Inovasi Pendidikan Fisika
p-ISSN: 2302-4496
e-ISSN: 2830-3881

Vol 15, No 01, 132-145
https://doi.org/10.26740/ipf
jipf@unesa.ac.id

Pengembangan Instrumen Keterampilan Pemecahan Masalah Berorientasi *Education for Sustainable Development* (ESD) untuk Peserta Didik SMA pada Materi Pemanasan Global

Tri Wulandari [1], Mukhayyarotin Niswati Rodliyatul Jauhariyah [1*]
[1] Pendidikan Fisika, Universitas Negeri Surabaya, Surabaya, Indonesia
[*] Email: mukhayyarotinjauhariyah@unesa.ac.id

Abstrak

Keterampilan pemecahan masalah merupakan salah satu kompetensi utama yang perlu dikembangkan dalam pembelajaran abad ke-21. Namun, hingga kini masih belum tersedia instrumen untuk mengukur keterampilan pemecahan masalah yang berorientasi Education for Sustainable Development (ESD). Penelitian ini bertujuan mengembangkan instrumen asesmen keterampilan pemecahan masalah berorientasi ESD pada materi pemanasan global untuk peserta didik SMA. Instrumen dikembangkan berdasarkan empat tahapan pemecahan masalah Polya yang diintegrasikan dengan tiga dimensi ESD, yaitu lingkungan, sosial, dan ekonomi. Penelitian menggunakan metode Research and Development (R&D) dengan model ADDIE, yang meliputi tahap analisis, perancangan, pengembangan, implementasi, dan evaluasi. Validasi isi melibatkan tiga validator, sedangkan uji coba empiris melibatkan 35 peserta didik kelas XI MAN 2 Gresik. Hasil validasi menunjukkan nilai Percentage of Agreement (PoA) sebesar 86% pada aspek isi dan konstruksi serta 90% pada aspek bahasa. Seluruh 20 butir memenuhi kriteria validitas empiris yang digunakan, dengan koefisien reliabilitas Cronbach's Alpha sebesar 0,969. Indeks tingkat kesukaran seluruh butir berada pada kategori sedang dan sebaran daya pembeda pada kategori lemah, cukup, hingga baik. Meskipun instrumen ini sudah dapat dimanfaatkan pendidik sebagai langkah awal, keterbatasan ESD yang sebatas konteks cerita dan keberadaan butir berdaya pembeda rendah mengisyaratkan perlunya pematangan lebih lanjut. Penyempurnaan melalui perbaikan butir, pendalaman ESD pada alur Polya, serta penguatan matematis menjadi krusial demi menghadirkan pengukuran yang jauh lebih utuh dan presisi.

Kata kunci: *Education for Sustainable Development* (ESD), instrumen penilaian, keterampilan pemecahan masalah, pemanasan global.

Development of an Education for Sustainable Development (ESD)-Oriented Assessment Instrument for High School Students' Problem-Solving Skills in Global Warming

Abstract

Problem-solving skills are a key competency that must be developed within the context of 21st-century education. However, there is currently a lack of instruments designed to measure problem-solving skills through the lens of Education for Sustainable Development (ESD). This study aimed to develop an assessment instrument for ESD-oriented problem-solving skills, focusing on the topic of global warming for high school students. The instrument was developed based on Polya's four-stage problem-solving model, integrated with the three dimensions of ESD: environmental, social, and economic. The study employed a Research and Development (R&D) approach using the ADDIE model, which encompasses the stages of analysis, design, development, implementation, and evaluation. Content validation was conducted by three experts, while an empirical trial involved 35 eleventh-grade students from MAN 2 Gresik. The validation results yielded a Percentage of Agreement (PoA) of 86% for content and

Copyright © The Author(s). This work is licensed under CC BY-NC 4.0

construct, and 90% for language. All 20 items met the established criteria for empirical validity, with a Cronbach's Alpha reliability coefficient of 0.969. The difficulty indices for all items fell into the moderate category, while the discrimination indices ranged from weak to good. Although this instrument can serve as an initial tool for educators, its limitations—specifically the restriction of ESD elements to narrative contexts and the presence of items with low discrimination—indicate a need for further refinement. Future improvements, including the revision of specific items, a deeper integration of ESD concepts into Polya's framework, and the strengthening of mathematical components, are essential to achieving a more comprehensive and precise assessment.
Keywords: *Education for Sustainable Development* (ESD)*, assessment instruments, problem-solving skills, global warming.*

Histori Naskah

Diserahkan: 21 Oktober 2026 Direvisi: 25 Oktober 2026 Diterima: 29 Oktober 2026

How to cite:

Wulandari, T. & Jauhariyah, M.N.R. (2026). Pengembangan Instrumen Keterampilan Pemecahan Masalah Berorientasi *Education for Sustainable Development* (ESD) untuk Peserta Didik SMA pada Materi Pemanasan Global. *Inovasi Pendidikan Fisika*, **15**(1), 132-145. DOI: https://doi.org/10.26740/ipf.vxx.nxx.xx-xx.

PENDAHULUAN

Pendidikan berkualitas merupakan salah satu tujuan utama dalam *Sustainable Development Goals* (SDGs), khususnya Tujuan 4 yang menekankan pentingnya penyediaan pendidikan yang inklusif, adil, dan mampu membekali peserta didik dengan pengetahuan, keterampilan, dan kompetensi untuk menghadapi berbagai tantangan global (United Nations, 2023; United Nations Educational Scientific and Cultural Organization, 2018, 2020). Salah satu pendekatan yang direkomendasikan untuk mendukung pencapaian tujuan tersebut adalah *Education for Sustainable Development* (ESD), yang mengintegrasikan dimensi lingkungan, sosial, dan ekonomi ke dalam proses pembelajaran sehingga peserta didik mampu berkontribusi terhadap pembangunan berkelanjutan (Kioupi & Voulvoulis, 2022; United Nations Educational Scientific and Cultural Organization, 2020). Berbagai penelitian menunjukkan bahwa implementasi ESD tidak hanya meningkatkan penguasaan konsep, tetapi juga mengembangkan keterampilan berpikir tingkat tinggi, termasuk kemampuan menganalisis permasalahan nyata, mengambil keputusan yang bertanggung jawab, dan memecahkan masalah dalam konteks kehidupan sehari-hari (Jauhariyah et al., 2019, 2021; Rieckmann, 2023). Salah satu konteks yang relevan untuk mengembangkan kompetensi tersebut dalam pembelajaran fisika adalah isu pemanasan global, karena fenomena ini merupakan permasalahan nyata yang menuntut peserta didik untuk mengintegrasikan konsep fisika dengan pertimbangan lingkungan, sosial, dan ekonomi dalam merumuskan solusi yang berkelanjutan (Olsson et al., 2022; Tilbury, 2014; United Nations, 2023).

Keterampilan pemecahan masalah (*problem-solving skills*) merupakan salah satu kompetensi utama yang perlu dikembangkan dalam pembelajaran abad ke-21 karena memungkinkan peserta didik mengidentifikasi permasalahan, menganalisis berbagai alternatif solusi, serta mengambil keputusan berdasarkan bukti dan penalaran ilmiah (OECD, 2022, 2023). Dalam pembelajaran fisika, keterampilan ini memiliki peran yang sangat penting karena peserta didik tidak hanya dituntut memahami konsep-konsep fisika, tetapi juga mampu menerapkannya untuk menjelaskan fenomena alam dan menyelesaikan permasalahan kontekstual yang dijumpai dalam kehidupan sehari-hari (Docktor & Mestre, 2014; Yuliantaningrum & Sunarti, 2020). Berbagai penelitian menunjukkan bahwa pembelajaran yang berorientasi pada pemecahan masalah mampu meningkatkan kemampuan berpikir kritis, penalaran ilmiah, dan pengambilan keputusan peserta didik, terutama ketika diintegrasikan dengan konteks permasalahan nyata yang berkaitan dengan keberlanjutan (Isac et al., 2025; Olsson et al., 2022). Oleh karena itu, pengembangan keterampilan pemecahan masalah dalam kerangka *Education for Sustainable Development* (ESD) menjadi semakin penting untuk membekali peserta didik dalam merespons isu-isu global, termasuk pemanasan global, secara kritis, ilmiah, dan bertanggung jawab (Tilbury, 2014).

Meskipun pentingnya keterampilan pemecahan masalah telah banyak diakui dalam pembelajaran fisika, implementasinya di sekolah masih menghadapi berbagai kendala (Ayudha & Setyarsih, 2021; Makiyah et al., 2021). Berbagai penelitian melaporkan bahwa kemampuan pemecahan masalah peserta didik masih berada

pada kategori rendah karena proses pembelajaran cenderung berpusat pada guru (*teacher-centered*), berorientasi pada penyelesaian soal rutin, dan belum memberikan kesempatan kepada peserta didik untuk membangun strategi penyelesaian masalah secara mandiri (Ayumniyya & Setyarsih, 2021; Firmansyah et al., 2022; Franestian et al., 2020; Musengimana et al., 2025; Widiawati et al., 2022). Kondisi tersebut sejalan dengan hasil wawancara pra-penelitian dengan guru fisika di MAN 2 Gresik yang menunjukkan bahwa proses evaluasi masih didominasi oleh soal pilihan ganda dengan sistem penskoran benar–salah, sehingga belum mampu mengidentifikasi tahapan berpikir peserta didik selama menyelesaikan masalah fisika. Selain itu, guru juga belum memiliki instrumen asesmen yang dilengkapi rubrik penskoran untuk mengukur keterampilan pemecahan masalah secara komprehensif sesuai tahapan penyelesaian masalah (Meyer & Friege, 2026; Putri et al., 2020). Akibatnya, informasi yang diperoleh dari hasil penilaian lebih banyak menggambarkan jawaban akhir dibandingkan proses berpikir peserta didik dalam memecahkan masalah.

Berbagai penelitian telah mengembangkan instrumen untuk mengukur keterampilan pemecahan masalah dalam pembelajaran sains maupun fisika dan menunjukkan bahwa instrumen tersebut mampu memberikan informasi mengenai kemampuan peserta didik dalam menyelesaikan permasalahan ilmiah (Docktor & Mestre, 2014; Musengimana et al., 2025; Reinhard et al., 2022). Meskipun demikian , sebagian besar instrumen yang telah dikembangkan masih berfokus pada pengukuran kemampuan kognitif atau hasil akhir penyelesaian masalah dan belum secara eksplisit mengintegrasikan prinsip-prinsip *Education for Sustainable Development* (ESD) ke dalam konteks asesmen. Selain itu, penelitian-penelitian tersebut umumnya belum mengombinasikan indikator pemecahan masalah berdasarkan tahapan Polya dengan tiga dimensi ESD (lingkungan, sosial, dan ekonomi), serta belum dilengkapi rubrik analitik yang mampu mengevaluasi proses berpikir peserta didik secara sistematis. Kondisi tersebut menunjukkan masih adanya kebutuhan untuk mengembangkan instrumen asesmen yang tidak hanya valid dan reliabel, tetapi juga mampu mengukur keterampilan pemecahan masalah melalui konteks pembelajaran yang mendukung pencapaian kompetensi keberlanjutan.

Berdasarkan kesenjangan penelitian tersebut, penelitian ini menawarkan suatu kebaruan berupa pengembangan instrumen asesmen keterampilan pemecahan masalah yang mengintegrasikan indikator pemecahan masalah Polya dengan tiga dimensi *Education for Sustainable Development* (ESD), yaitu lingkungan, sosial, dan ekonomi, dalam konteks materi pemanasan global. Selain menghasilkan seperangkat butir soal, penelitian ini juga mengembangkan rubrik penskoran analitik yang memungkinkan guru mengevaluasi setiap tahapan proses pemecahan masalah peserta didik secara lebih sistematis dan komprehensif. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan instrumen asesmen keterampilan pemecahan masalah berorientasi *Education for Sustainable Development* (ESD) pada materi pemanasan global yang valid, reliabel, dan layak digunakan dalam pembelajaran fisika di tingkat SMA. Dengan demikian, rumusan masalah yang menjadi fokus penelitian ini ialah, ” Bagaimana kualitas (validitas, reliabilitas, tingkat kesukaran, dan daya pembeda) instrumen tes berorientasi *Education for Sustainable Development* (ESD) untuk mengukur keterampilan pemecahan masalah peserta didik SMA pada materi pemanasan global?”.

METODE

Desain Penelitian

Penelitian ini menggunakan pendekatan *Research and Development* (R&D) yang bertujuan mengembangkan instrumen asesmen keterampilan pemecahan masalah berbentuk soal esai berorientasi *Education for Sustainable Development* (ESD) pada materi pemanasan global. Proses pengembangan instrumen mengacu pada model ADDIE (*Analyze, Design, Development, Implementation,* dan *Evaluation*) yang dikembangkan oleh Robert Maribe Branch (2009). Model ini dipilih karena menyediakan tahapan yang sistematis dalam proses analisis kebutuhan, perancangan, pengembangan, implementasi, dan evaluasi produk sehingga menghasilkan instrumen yang valid, reliabel, dan layak digunakan dalam pembelajaran fisika. Produk yang dihasilkan berupa seperangkat instrumen asesmen yang terdiri atas kisi-kisi, butir soal esai, rubrik penskoran analitik, serta pedoman penilaian keterampilan pemecahan masalah berorientasi ESD. Produk yang dihasilkan berupa seperangkat instrumen asesmen yang selanjutnya dievaluasi melalui validitas isi, validitas empiris, reliabilitas, tingkat kesukaran, dan daya pembeda untuk memastikan kualitas instrumen sebelum digunakan dalam pembelajaran.

Partisipan

Penelitian ini melibatkan dua kelompok partisipan, yaitu validator ahli dan peserta didik. Validator ahli terdiri atas tiga orang pakar yang memiliki kompetensi dalam bidang pendidikan fisika, meliputi dua dosen Pendidikan Fisika Universitas Negeri Surabaya dan satu guru fisika MAN 2 Gresik. Ketiga validator berperan dalam mengevaluasi validitas isi instrumen berdasarkan aspek materi, konstruksi, dan bahasa. Selanjutnya, uji coba empiris dilakukan kepada 35 peserta didik kelas XI MAN 2 Gresik yang dipilih menggunakan teknik *purposive sampling* karena telah mempelajari materi pemanasan global sesuai dengan Kurikulum Merdeka. Data hasil uji coba digunakan untuk menganalisis validitas empiris, reliabilitas, tingkat kesukaran, dan daya pembeda butir soal sebagai dasar penyempurnaan instrumen.

Pengembangan Instrumen

Pengembangan instrumen diawali dengan penyusunan spesifikasi instrumen berdasarkan tujuan penelitian, indikator keterampilan pemecahan masalah, serta capaian pembelajaran pada materi pemanasan global dalam Kurikulum Merdeka. Indikator keterampilan pemecahan masalah mengacu pada tahapan Polya yang meliputi memahami masalah (*understanding the problem*), merencanakan penyelesaian (*devising a plan*), melaksanakan rencana (*carrying out the plan*), dan melakukan pengecekan kembali (*looking back*) (Polya, 1957). Selanjutnya, prinsip *Education for Sustainable Development* (ESD) diintegrasikan ke dalam penyusunan konteks soal melalui dimensi lingkungan, sosial, dan ekonomi sehingga setiap butir soal merepresentasikan permasalahan nyata yang berkaitan dengan isu pemanasan global. Berdasarkan indikator tersebut disusun kisi-kisi instrumen, butir soal esai, rubrik penskoran analitik, serta pedoman penilaian yang selanjutnya digunakan pada proses validasi ahli dan uji coba empiris. Setiap butir soal dikembangkan menggunakan stimulus berupa fenomena kontekstual terkait pemanasan global yang merepresentasikan permasalahan keberlanjutan sehingga peserta didik tidak hanya menerapkan konsep fisika, tetapi juga mempertimbangkan dampak lingkungan, sosial, dan ekonomi dalam proses pemecahan masalah. Tabel 1 menyajikan kisi-kisi instrumen keterampilan pemecahan masalah berorientasi ESD.

Tabel 1. Kisi-kisi instrumen keterampilan pemecahan masalah berorientasi ESD.

No	Aspek dan Komponen ESD	Indikator Keterampilan Pemecahan Masalah	Nomor Butir Soal
1	• Lingkungan: Perubahan Cuaca • Sosial-Budaya: Kesehatan	Memahami Masalah	1a
		Menyusun Rencana Penyelesaian	1b
		Melaksanakan Rencana Penyelesaian	1c
		Memeriksa Kembali Solusi yang Diperoleh	1d
2	• Lingkungan: Perubahan Cuaca • Sosial-Budaya: Tata Kelola	Memahami Masalah	2a
		Menyusun Rencana Penyelesaian	2b
		Melaksanakan Rencana Penyelesaian	2c
		Memeriksa Kembali Solusi yang Diperoleh	2d
3	• Lingkungan: Sumber Daya Alam dan Pencegahan & Penanganan Bencana	Memahami Masalah	3a
		Menyusun Rencana Penyelesaian	3b
		Melaksanakan Rencana Penyelesaian	3c
		Memeriksa Kembali Solusi yang Diperoleh	3d
4	• Lingkungan: Perubahan Cuaca dan Pencegahan & Penanganan Bencana	Memahami Masalah	4a
		Menyusun Rencana Penyelesaian	4b
		Melaksanakan Rencana Penyelesaian	4c
		Memeriksa Kembali Solusi yang Diperoleh	4d
5	• Lingkungan: Sumber Daya Alam • Ekonomi: Tanggungjawab Perusahaan	Memahami Masalah	5a
		Menyusun Rencana Penyelesaian	5b
		Melaksanakan Rencana Penyelesaian	5c
		Memeriksa Kembali Solusi yang Diperoleh	5d

Proses Validasi

Instrumen yang telah dikembangkan selanjutnya melalui dua tahap validasi, yaitu validasi isi (*content validation*) dan validasi empiris (*empirical validation*). Validasi isi dilakukan oleh tiga validator ahli yang terdiri atas dua dosen Pendidikan Fisika Universitas Negeri Surabaya dan satu guru fisika MAN 2 Gresik.

Proses validasi bertujuan untuk mengevaluasi kesesuaian instrumen berdasarkan aspek materi, konstruksi, dan bahasa. Masukan dan saran dari validator digunakan sebagai dasar revisi untuk menyempurnakan instrumen sebelum dilakukan uji coba kepada peserta didik.

Setelah direvisi, instrumen diujicobakan kepada 35 peserta didik kelas XI MAN 2 Gresik untuk memperoleh data empiris. Hasil uji coba digunakan sebagai dasar evaluasi kualitas butir soal melalui analisis validitas empiris, reliabilitas, tingkat kesukaran, dan daya pembeda sehingga diperoleh instrumen yang memenuhi kriteria kelayakan untuk digunakan dalam pembelajaran fisika. Revisi instrumen dilakukan berdasarkan komentar dan rekomendasi validator hingga seluruh aspek penilaian memenuhi kriteria kelayakan sebelum dilanjutkan pada tahap uji coba empiris.

Setelah uji coba empiris, kualitas setiap butir dievaluasi berdasarkan validitas empiris, reliabilitas, tingkat kesukaran, dan daya pembeda. Hasil analisis daya pembeda menunjukkan bahwa butir 1b, 1c, dan 1d memerlukan revisi. Namun, ketiga butir tersebut tidak direvisi dan tidak diuji ulang sebelum instrumen digunakan di kelas X. Dengan demikian, hasil analisis empiris dalam penelitian ini merepresentasikan kualitas butir berdasarkan uji coba awal, sedangkan efektivitas revisi terhadap ketiga butir tersebut belum dapat dievaluasi.

Analisis Data

Data hasil validasi ahli dianalisis secara deskriptif untuk menentukan tingkat validitas isi instrumen berdasarkan aspek materi, konstruksi, dan bahasa. Konsistensi penilaian antarvalidator dihitung menggunakan *Percentage of Agreement* (PoA) yang dikemukakan oleh Borich et al. (1994). Instrumen dinyatakan memiliki tingkat kesepakatan yang baik apabila nilai PoA mencapai ≥75%.

Selanjutnya, data hasil uji coba dianalisis untuk mengevaluasi kualitas psikometrik instrumen yang meliputi validitas empiris, reliabilitas, tingkat kesukaran, dan daya pembeda butir soal. Validitas empiris dianalisis menggunakan korelasi *Product Moment Pearson*, sedangkan reliabilitas instrumen dihitung menggunakan koefisien Cronbach's Alpha. Selain itu, setiap butir soal dianalisis berdasarkan indeks tingkat kesukaran dan daya pembeda untuk menentukan kelayakan butir sebelum digunakan sebagai instrumen akhir. Seluruh analisis data dilakukan menggunakan perangkat lunak IBM SPSS Statistics versi 27.

HASIL DAN PEMBAHASAN

Analisis (Analyze)

Fokus utama tahap ini mencakup analisis kebutuhan instrumen, kajian terhadap kurikulum yang berlaku, dan peninjauan materi pemanasan global. Tahap analisis kebutuhan diawali dengan kajian literatur dan wawancara untuk mengidentifikasi urgensi serta rendahnya keterampilan pemecahan masalah peserta didik. Berbagai studi menegaskan bahwa keterampilan tersebut merupakan kompetensi abad 21 yang krusial, namun tingkat penguasaannya di Indonesia masih tergolong rendah akibat metode pembelajaran yang belum optimal. Temuan ini memperkuat urgensi pengembangan instrumen tes khusus guna mengukur dan memetakan profil keterampilan pemecahan masalah peserta didik secara sistematis.

Analisis kurikulum menunjukkan bahwa MAN 2 Gresik telah menerapkan Kurikulum Merdeka yang mendukung prinsip Education for Sustainable Development (ESD) dan Profil Pelajar Pancasila pada mata pelajaran fisika. Guru fisika mengintegrasikan keterampilan pemecahan masalah melalui dimensi bernalar kritis, kreativitas, dan kemandirian untuk membentuk karakter siswa yang solutif. Analisis materi selanjutnya difokuskan pada penentuan topik pemanasan global yang sejalan dengan Kurikulum Merdeka, prinsip ESD UNESCO, dan target Sustainable Development Goals (SDGs). Penyelarasan ini mencakup capaian pembelajaran, indikator keterampilan pemecahan masalah, serta aspek lingkungan terkait perubahan cuaca untuk menghasilkan output berupa dokumen kebutuhan, peta kesesuaian CP/TP, dan batasan materi yang komprehensif.

Desain (Design)

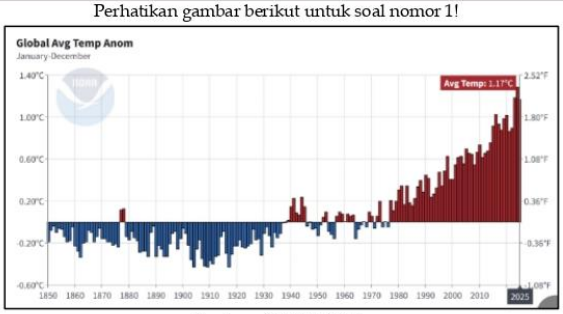
Produk yang dihasilkan dalam penelitian ini berupa instrumen asesmen keterampilan pemecahan masalah berorientasi ESD pada materi pemanasan global. Instrumen dikembangkan dalam bentuk tes esai yang terdiri atas 20 butir soal yang disusun berdasarkan lima konteks fenomena autentik terkait isu pemanasan global. Setiap konteks dirancang untuk mendorong peserta didik menganalisis permasalahan, merumuskan strategi penyelesaian, menerapkan konsep fisika, serta mengevaluasi solusi yang dihasilkan sesuai dengan

empat tahapan pemecahan masalah Polya, yaitu *understanding the problem, devising a plan, carrying out the plan,* dan *looking back*.

Karakteristik utama instrumen ini terletak pada integrasi prinsip ESD ke dalam setiap stimulus soal melalui dimensi lingkungan, sosial, dan ekonomi. Dengan demikian, peserta didik tidak hanya dituntut memahami konsep fisika mengenai pemanasan global, tetapi juga mempertimbangkan implikasi keberlanjutan dalam proses pemecahan masalah. Selain itu, instrumen dilengkapi dengan rubrik penskoran analitik yang dirancang untuk mengevaluasi setiap tahapan proses berpikir peserta didik secara sistematis, sehingga hasil asesmen tidak hanya menggambarkan jawaban akhir, tetapi juga kualitas proses pemecahan masalah yang dilakukan. Gambar 1 menyajikan karakteristik instrumen pemecahan masalah berorientasi ESD.

Stimulus Soal Nomor 1

Perhatikan gambar berikut untuk soal nomor 1!



Sumber : NOAA, 2025

Grafik anomali suhu global dari NOAA menampilkan perubahan suhu bumi sejak tahun 1850 hingga beberapa tahun terakhir dengan membandingkan suhu setiap tahun terhadap rata-rata suhu acuan. Perubahan nilai anomali suhu tersebut menjadi salah satu data penting untuk memahami kondisi pemanasan global dalam jangka panjang. Ketika suhu bumi mengalami perubahan, keseimbangan energi di atmosfer juga dapat terganggu sehingga memengaruhi pola musim, curah hujan, dan kejadian cuaca ekstrem di berbagai wilayah. Dampak perubahan tersebut tidak hanya berkaitan dengan lingkungan, tetapi juga dapat memengaruhi kehidupan manusia, seperti ketersediaan air, kegiatan pertanian, kenyamanan hidup, serta risiko gangguan kesehatan dan penyebaran penyakit. *(Sumber : NOAA, 2025; The Lancet, 2024).*

Aspek dan Komponen ESD

- Lingkungan: Perubahan Cuaca
- Sosial-Budaya: Kesehatan

Indikator Keterampilan Pemecahan Masalah	Ranah Kognitif	Indikator Soal	Soal	Rambu-Rambu Jawaban	No Soal	Skor Maks
Memahami Masalah (*Understanding the Problem*)	C4	Disajikan grafik anomali suhu global, peserta didik dapat menganalisis masalah utama yang timbul akibat peningkatan anomali suhu global serta menjelaskan alasan fenomena tersebut tidak dapat dipahami hanya sebagai kenaikan suhu biasa.	Berdasarkan pola pada grafik, analisislah masalah utama yang muncul jika anomali suhu global terus meningkat, serta jelaskan mengapa masalah tersebut tidak cukup dipahami hanya sebagai kenaikan suhu biasa!	Berdasarkan grafik, pada periode awal banyak nilai anomali suhu berada di bawah atau mendekati rata-rata suhu acuan, sedangkan pada periode akhir grafik semakin banyak nilai anomali suhu positif yang ditunjukkan dengan batang merah. Hal ini menunjukkan bahwa suhu global cenderung meningkat dalam jangka panjang. Masalah ini tidak cukup dipahami sebagai kenaikan suhu biasa, karena perubahan suhu global dapat mengganggu keseimbangan energi di atmosfer. Akibatnya, pola musim, curah hujan, dan intensitas cuaca ekstrem dapat berubah. Jika kondisi tersebut terus terjadi, manusia dapat menghadapi masalah yang lebih luas, seperti kekurangan air, gagal panen, menurunnya kenyamanan hidup, serta meningkatnya risiko penyakit akibat perubahan suhu dan lingkungan.	1a	5

Gambar 1. Karakteristik instrumen pemecahan masalah berorientasi ESD.

Penggunaan soal esai dipilih karena mampu mengungkap proses berpikir peserta didik secara lebih komprehensif dibandingkan soal pilihan ganda, terutama dalam mengukur kemampuan pemecahan masalah dan penalaran ilmiah (Docktor & Mestre, 2014; Musengimana et al., 2025). Selain itu, integrasi konteks ESD ke dalam stimulus soal memberikan kesempatan kepada peserta didik untuk menghubungkan konsep fisika dengan permasalahan keberlanjutan sehingga asesmen tidak hanya berorientasi pada penguasaan konsep, tetapi juga pada pengembangan kompetensi keberlanjutan sebagaimana direkomendasikan United Nations Educational Scientific and Cultural Organization (2020) dan Kioupi & Voulvoulis (2022).

Pengembangan (Develop)

Validitas isi instrumen dievaluasi oleh tiga validator yang terdiri atas dua dosen Pendidikan Fisika Universitas Negeri Surabaya dan satu guru fisika MAN 2 Gresik. Penilaian dilakukan terhadap tiga aspek, yaitu isi, konstruksi, dan bahasa. Hasil validasi menunjukkan bahwa seluruh aspek memperoleh modus skor sebesar 4 dengan kategori *sangat valid*. Tingkat kesepakatan antarvalidator yang dihitung menggunakan *Percentage of Agreement* (PoA) juga menunjukkan hasil yang sangat baik, yaitu sebesar 86% pada aspek isi, 86% pada aspek konstruksi, dan 90% pada aspek bahasa. Hasil tersebut menunjukkan bahwa instrumen yang dikembangkan telah memenuhi kriteria validitas isi dan layak untuk dilanjutkan ke tahap uji coba empiris.

Tingginya tingkat validitas isi menunjukkan bahwa setiap butir soal telah sesuai dengan indikator keterampilan pemecahan masalah, capaian pembelajaran materi pemanasan global, serta prinsip ESD yang menjadi dasar pengembangan instrumen. Selain itu, tingginya nilai PoA mengindikasikan adanya konsistensi penilaian di antara para validator, sehingga instrumen memiliki representasi isi yang baik dan mampu mengukur konstruk yang dituju secara tepat. Kesepakatan antarvalidator merupakan salah satu indikator penting dalam memastikan kualitas instrumen sebelum dilakukan pengujian empiris (AERA et al., 2014; Borich et al., 1994).

Temuan penelitian ini sejalan dengan penelitian Maulita et al. (2019) yang menegaskan bahwa validitas isi merupakan prasyarat utama dalam pengembangan instrumen asesmen karena memastikan kesesuaian antara indikator, materi, dan tujuan pengukuran. Hasil ini juga mendukung temuan Sherliyanti & Jauhariyah (2024) bahwa validitas konstruksi berperan penting dalam menghasilkan instrumen yang mampu mengungkap kemampuan berpikir peserta didik secara akurat. Selain itu, penggunaan bahasa yang jelas dan tidak menimbulkan penafsiran ganda memperkuat keterbacaan instrumen sehingga peserta didik dapat memusatkan perhatian pada proses pemecahan masalah, bukan pada interpretasi redaksi soal. Dengan demikian, instrumen yang dikembangkan memiliki dasar validitas isi yang kuat sebagai prasyarat untuk melanjutkan evaluasi kualitas psikometrik pada tahap berikutnya. Tabel 2 menyajikan hasil validasi instrumen pemecahan masalah berorientasi ESD.

Tabel 2. Hasil validasi instrumen pemecahan masalah berorientasi ESD

Ranah	Modus	PoA (%)	Kriteria
Konstruksi	4	86	Sangat Valid
Isi	4	86	Sangat Valid
Bahasa	4	90	Sangat Valid

Tingginya validitas isi menunjukkan bahwa setiap butir soal telah merepresentasikan konstruk keterampilan pemecahan masalah yang ingin diukur. Hal ini menunjukkan adanya kesesuaian antara indikator pembelajaran, indikator keterampilan pemecahan masalah berdasarkan tahapan Polya, tujuan pengukuran, serta konteks pemanasan global yang digunakan sebagai stimulus soal. Dengan demikian, setiap butir instrumen tidak hanya mengukur penguasaan konsep fisika, tetapi juga mampu memfasilitasi peserta didik untuk mengidentifikasi permasalahan, merancang strategi penyelesaian, menerapkan konsep fisika secara ilmiah, dan mengevaluasi solusi yang diperoleh sesuai dengan tahapan pemecahan masalah. Kesesuaian antara indikator dan tujuan pengukuran merupakan salah satu prinsip utama dalam validitas isi karena memastikan bahwa setiap butir instrumen benar-benar merepresentasikan domain kompetensi yang hendak diukur (AERA et al., 2014; DeVellis, 2017; DeVellis & Thorpe, 2022).

Selain mencerminkan indikator keterampilan pemecahan masalah, setiap butir soal juga dirancang menggunakan konteks autentik yang berkaitan dengan isu pemanasan global sebagai implementasi prinsip ESD. Pendekatan ini memungkinkan instrumen mengukur kemampuan peserta didik dalam menerapkan konsep fisika pada permasalahan nyata yang memiliki dimensi lingkungan, sosial, dan ekonomi. Penggunaan konteks autentik dalam asesmen sejalan dengan rekomendasi United Nations Educational Scientific and Cultural Organization (2020) yang menekankan bahwa asesmen ESD seharusnya tidak hanya mengukur penguasaan konsep, tetapi juga kemampuan peserta didik dalam menggunakan pengetahuan tersebut untuk mengambil keputusan yang bertanggung jawab terhadap isu keberlanjutan.

Temuan ini sejalan dengan penelitian Sherliyanti & Jauhariyah (2024) yang menunjukkan bahwa kesesuaian antara indikator, tujuan pengukuran, dan karakteristik butir soal merupakan faktor utama yang menentukan tingginya validitas isi suatu instrumen asesmen. Hasil serupa juga dilaporkan oleh Maulita et al. (2019), yang menegaskan bahwa keterlibatan validator ahli dari berbagai latar belakang keahlian berkontribusi

terhadap penyempurnaan isi, konstruksi, dan bahasa instrumen sebelum dilakukan uji empiris. Dengan demikian, validitas isi yang tinggi pada penelitian ini menunjukkan bahwa instrumen telah memiliki representasi konstruk yang memadai dan siap untuk dievaluasi lebih lanjut melalui pengujian empiris.

Dengan demikian, validitas isi pada penelitian ini tidak hanya menunjukkan kesesuaian antara indikator dan tujuan pengukuran, tetapi juga memperlihatkan keberhasilan integrasi prinsip *Education for Sustainable Development* ke dalam desain asesmen fisika. Hal ini menjadikan instrumen yang dikembangkan tidak sekadar mengukur kemampuan pemecahan masalah, tetapi juga relevan untuk mendukung pembelajaran fisika yang kontekstual dan berorientasi pada pembangunan berkelanjutan.

Hasil uji validitas empiris menunjukkan bahwa seluruh butir soal memenuhi kriteria valid berdasarkan koefisien korelasi *Product Moment Pearson*. Nilai koefisien korelasi setiap butir berada di atas nilai r_{tabel} (0,334) sehingga seluruh butir dinyatakan valid dan dapat digunakan pada tahap analisis selanjutnya. Selain itu, hasil uji reliabilitas menunjukkan bahwa instrumen memperoleh koefisien Cronbach's Alpha sebesar 0,969, yang menunjukkan tingkat konsistensi internal yang sangat tinggi. Hasil tersebut mengindikasikan bahwa instrumen mampu memberikan hasil pengukuran yang konsisten ketika digunakan untuk mengukur keterampilan pemecahan masalah peserta didik. Tabel 3 menyajikan data hasil validasi butir soal ESD yang dikembangkan dalam penelitian ini.

Tabel 3. Hasil Uji Validitas Butir Soal

No Soal	r_{hitung}	Keterangan	Kriteria
1a	0,716	Valid	Tinggi
1b	0,570	Valid	Cukup
1c	0,583	Valid	Cukup
1d	0,583	Valid	Cukup
2a	0,777	Valid	Tinggi
2b	0,657	Valid	Tinggi
2c	0,757	Valid	Tinggi
2d	0,809	Valid	Sangat Tinggi
3a	0,871	Valid	Sangat Tinggi
3b	0,830	Valid	Sangat Tinggi
3c	0,869	Valid	Sangat Tinggi
3d	0,877	Valid	Sangat Tinggi
4a	0,846	Valid	Sangat Tinggi
4b	0,905	Valid	Sangat Tinggi
4c	0,873	Valid	Sangat Tinggi
4d	0,873	Valid	Sangat Tinggi
5a	0,867	Valid	Sangat Tinggi
5b	0,824	Valid	Sangat Tinggi
5c	0,871	Valid	Sangat Tinggi
5d	0,848	Valid	Sangat Tinggi
Cronbach's Alpha = 0.969 (Reliabilitas Sangat Tinggi)			

Hasil analisis validitas empiris menunjukkan bahwa seluruh butir memenuhi kriteria berdasarkan korelasi *Product Moment Pearson* yang digunakan dalam penelitian ini. Koefisien *Cronbach's Alpha* sebesar 0,969 menunjukkan konsistensi internal yang tinggi pada data uji coba. Meskipun demikian, hasil tersebut perlu diinterpretasikan bersama dengan hasil analisis daya pembeda. Kondisi ini mengindikasikan bahwa proses penyusunan indikator, pengembangan kisi-kisi, serta penyusunan rubrik penskoran telah menghasilkan representasi konstruk yang baik. Menurut *Standards for Educational and Psychological Testing* (AERA et al., 2014). Pemenuhan kriteria validitas empiris dan tingginya koefisien reliabilitas tidak secara otomatis menunjukkan bahwa setiap butir telah memiliki kemampuan diskriminasi yang memadai. Dalam penelitian ini, butir 1b, 1c, dan 1d tetap menunjukkan daya pembeda yang rendah. Oleh karena itu, evaluasi kualitas instrumen perlu mempertimbangkan berbagai bukti empiris secara terpadu, termasuk hasil analisis pada tingkat butir.

Koefisien Cronbach's Alpha sebesar 0,969 menunjukkan bahwa instrumen memiliki konsistensi internal yang sangat tinggi. Nilai tersebut mengindikasikan bahwa seluruh butir soal bekerja secara harmonis dalam mengukur konstruk yang sama, yaitu keterampilan pemecahan masalah. Tingginya reliabilitas juga menunjukkan bahwa variasi skor yang diperoleh peserta didik lebih banyak dipengaruhi oleh perbedaan kemampuan mereka daripada oleh kesalahan pengukuran (*measurement error*). Oleh karena itu, instrumen ini memiliki tingkat keandalan yang memadai untuk digunakan sebagai alat asesmen dalam pembelajaran fisika (DeVellis & Thorpe, 2022).

Temuan penelitian ini sejalan dengan penelitian Musengimana et al. (2025) yang melaporkan bahwa instrumen dengan validitas empiris dan reliabilitas tinggi mampu menghasilkan pengukuran yang lebih akurat terhadap kemampuan pemecahan masalah dalam pembelajaran sains. Hasil ini juga mendukung penelitian Hebebci & Usta (2022) yang menegaskan bahwa kualitas psikometrik instrumen merupakan prasyarat utama sebelum instrumen digunakan untuk mengevaluasi kompetensi berpikir tingkat tinggi peserta didik. Dengan demikian, hasil validitas empiris dan reliabilitas pada penelitian ini memperkuat bukti bahwa instrumen yang dikembangkan telah memenuhi karakteristik sebagai instrumen asesmen yang layak digunakan dalam pembelajaran fisika berorientasi ESD.

Menariknya, tingginya validitas empiris pada penelitian ini sejalan dengan hasil validitas isi yang telah diperoleh sebelumnya. Hal ini menunjukkan bahwa kesesuaian indikator, tujuan pengukuran, dan konteks ESD yang telah dikonfirmasi melalui validasi ahli berhasil dipertahankan pada saat instrumen diujikan kepada peserta didik. Dengan demikian, terdapat konsistensi antara validitas teoritis dan validitas empiris, yang semakin memperkuat kualitas psikometrik instrumen yang dikembangkan.

Secara keseluruhan, temuan validitas dan reliabilitas empiris menunjukkan bahwa instrumen yang dikembangkan memiliki sifat psikometrik yang kuat, menunjukkan bahwa instrumen ini mampu menghasilkan pengukuran keterampilan pemecahan masalah siswa yang akurat dan konsisten dalam konteks ESD.

Analisis Item (Item Analysis)

Analisis kualitas butir dilakukan melalui penghitungan indeks tingkat kesukaran dan daya pembeda. Hasil analisis menunjukkan bahwa seluruh 20 butir soal memiliki indeks tingkat kesukaran antara 0,451 dan 0,640, yang termasuk dalam kategori sedang berdasarkan kriteria yang digunakan dalam penelitian ini. Sementara itu, indeks daya pembeda berkisar antara 0,177 dan 0,488. Sebanyak 17 butir memiliki indeks daya pembeda antara 0,222 dan 0,488, sedangkan tiga butir, yaitu 1b, 1c, dan 1d, masing-masing memiliki indeks sebesar 0,177 dan memerlukan revisi. Temuan ini menunjukkan bahwa meskipun seluruh butir memiliki tingkat kesukaran dalam kategori sedang dan memenuhi kriteria validitas empiris yang digunakan, kualitas daya pembeda belum merata pada seluruh butir. Oleh karena itu, interpretasi kelayakan instrumen perlu mempertimbangkan hasil analisis setiap butir secara terpadu. Tabel 4 menyajikan hasil analisis tingkat kesukaran dan daya pembeda sebagai dasar penentuan tindak lanjut terhadap setiap butir soal.

Tabel 4. Hasil Analisis Butir Soal

No Butir Soal	Tingkat Kesukaran	Kriteria	Daya Pembeda	Kriteria	Keputusan
1a	0,594	Sedang	0,244	Cukup	Dipertahankan
1b	0,502	Sedang	0,177	Lemah	Direkomendasikan untuk direvisi
1c	0,485	Sedang	0,177	Lemah	Direkomendasikan untuk direvisi
1d	0,451	Sedang	0,177	Lemah	Direkomendasikan untuk direvisi
2a	0,617	Sedang	0,266	Cukup	Dipertahankan
2b	0,560	Sedang	0,222	Cukup	Dipertahankan
2c	0,514	Sedang	0,288	Cukup	Dipertahankan
2d	0,474	Sedang	0,311	Cukup	Dipertahankan
3a	0,628	Sedang	0,333	Cukup	Dipertahankan
3b	0,565	Sedang	0,311	Cukup	Dipertahankan
3c	0,508	Sedang	0,355	Cukup	Dipertahankan
3d	0,485	Sedang	0,377	Cukup	Dipertahankan

No Butir Soal	Tingkat Kesukaran	Kriteria	Daya Pembeda	Kriteria	Keputusan
4a	0,640	Sedang	0,422	Baik	Dipertahankan
4b	0,571	Sedang	0,488	Baik	Dipertahankan
4c	0,571	Sedang	0,400	Baik	Dipertahankan
4d	0,497	Sedang	0,400	Baik	Dipertahankan
5a	0,628	Sedang	0,422	Baik	Dipertahankan
5b	0,548	Sedang	0,355	Cukup	Dipertahankan
5c	0,502	Sedang	0,422	Baik	Dipertahankan
5d	0,491	Sedang	0,422	Baik	Dipertahankan

Distribusi tingkat kesukaran yang didominasi kategori sedang menunjukkan bahwa instrumen memiliki tingkat tantangan yang proporsional bagi peserta didik. Butir dengan tingkat kesukaran sedang umumnya lebih mampu memberikan informasi mengenai variasi kemampuan peserta didik dibandingkan butir yang terlalu mudah maupun terlalu sukar. Dalam konteks asesmen keterampilan pemecahan masalah, distribusi tingkat kesukaran yang seimbang juga memberikan kesempatan kepada peserta didik untuk menunjukkan proses berpikirnya secara optimal melalui setiap tahapan pemecahan masalah. Oleh karena itu, komposisi tingkat kesukaran pada instrumen ini mendukung fungsi asesmen sebagai alat untuk mengidentifikasi kemampuan problem solving secara lebih akurat (AERA et al., 2014; DeVellis & Thorpe, 2022).

Hasil analisis daya pembeda menunjukkan adanya variasi kemampuan butir dalam membedakan peserta didik berdasarkan tingkat keterampilan pemecahan masalah. Sebanyak 17 butir memiliki indeks daya pembeda antara 0,222 dan 0,488, sedangkan butir 1b, 1c, dan 1d memperoleh indeks sebesar 0,177. Berdasarkan kriteria yang digunakan, ketiga butir tersebut memerlukan revisi karena kemampuan diskriminasinya masih lemah. Temuan ini menunjukkan bahwa hasil validasi ahli, validitas empiris, dan reliabilitas instrumen perlu dilengkapi dengan evaluasi kualitas setiap butir. Dengan demikian, tindak lanjut terhadap butir yang memiliki daya pembeda rendah diperlukan untuk meningkatkan kualitas instrumen sebelum digunakan secara lebih luas. Karakteristik ini merupakan salah satu indikator penting dalam evaluasi kualitas psikometrik suatu instrumen asesmen (Ebel & Friesbie, 1991; Engelhardt, 2009; Kılıç et al., 2023)

Jika dikaitkan dengan hasil validitas isi, validitas empiris, dan reliabilitas yang telah diperoleh sebelumnya, analisis tingkat kesukaran dan daya pembeda semakin memperkuat bukti bahwa instrumen yang dikembangkan memiliki kualitas psikometrik yang baik. Kombinasi karakteristik tersebut menunjukkan bahwa instrumen tidak hanya mampu mengukur keterampilan pemecahan masalah secara valid dan konsisten, tetapi juga memiliki kemampuan yang memadai dalam mengidentifikasi variasi kemampuan peserta didik. Dengan demikian, instrumen ini berpotensi digunakan sebagai asesmen autentik untuk mengevaluasi keterampilan pemecahan masalah dalam pembelajaran fisika berorientasi ESD.

Menariknya, butir-butir yang dikembangkan dalam penelitian ini menggunakan konteks autentik pemanasan global yang mengintegrasikan dimensi lingkungan, sosial, dan ekonomi. Meskipun konteks soal lebih kompleks dibandingkan soal fisika konvensional, hasil analisis menunjukkan bahwa tingkat kesukaran dan daya pembeda tetap berada pada kategori yang memadai. Temuan ini mengindikasikan bahwa integrasi konteks *Education for Sustainable Development* tidak mengurangi kualitas psikometrik instrumen, tetapi justru memungkinkan pengukuran keterampilan pemecahan masalah dalam situasi yang lebih kontekstual dan bermakna.

Secara keseluruhan, hasil analisis menunjukkan bahwa instrumen memiliki bukti awal validitas empiris dan konsistensi internal yang tinggi. Namun, kemampuan diskriminasi belum memadai pada seluruh butir. Indeks daya pembeda yang rendah pada butir 1b, 1c, dan 1d menunjukkan perlunya pemeriksaan lebih lanjut terhadap karakteristik butir, termasuk kesesuaian tuntutan soal, rubrik penskoran, dan respons peserta didik. Karena ketiga butir tersebut tidak direvisi dan diuji ulang, penelitian ini belum memberikan bukti mengenai apakah perbaikan butir akan meningkatkan daya pembedanya.

Produk Akhir (Final Product)

Berdasarkan hasil validasi isi, pengujian validitas empiris, analisis reliabilitas, serta analisis kualitas butir, penelitian ini menghasilkan instrumen asesmen keterampilan pemecahan masalah berorientasi Education for Sustainable Development (ESD) pada materi pemanasan global. Instrumen terdiri atas 20 butir soal esai yang disusun berdasarkan empat tahapan pemecahan masalah Polya dan dilengkapi dengan rubrik penskoran analitik.

Hasil pengujian awal menunjukkan bahwa seluruh butir memenuhi kriteria validitas empiris yang digunakan, dengan koefisien *Cronbach's Alpha* sebesar 0,969. Seluruh butir juga memiliki indeks tingkat kesukaran dalam kategori sedang. Namun, analisis daya pembeda menunjukkan bahwa butir 1b, 1c, dan 1d memiliki indeks sebesar 0,177 dan memerlukan revisi. Karena ketiga butir tersebut tidak direvisi dan tidak diuji ulang, kualitasnya setelah perbaikan belum dapat ditentukan. Dengan demikian, instrumen yang dikembangkan merupakan produk awal yang telah memperoleh bukti validitas empiris dan konsistensi internal, tetapi masih memerlukan penyempurnaan dan evaluasi lanjutan sebelum direkomendasikan untuk penggunaan yang lebih luas.

Karakteristik instrumen ini terletak pada integrasi konteks pemanasan global dengan dimensi lingkungan, sosial, dan ekonomi dalam kerangka *Education for Sustainable Development* (ESD). Integrasi tersebut dirancang untuk mendorong peserta didik menerapkan konsep fisika dalam menganalisis permasalahan keberlanjutan. Meskipun demikian, kualitas psikometrik instrumen tetap perlu dievaluasi pada tingkat butir. Temuan mengenai daya pembeda yang rendah pada tiga butir menunjukkan bahwa relevansi konteks ESD perlu disertai dengan tuntutan soal dan rubrik penskoran yang mampu membedakan tingkat keterampilan pemecahan masalah peserta didik. Instrumen ini tidak hanya berfungsi sebagai alat ukur keterampilan pemecahan masalah, tetapi juga mendukung implementasi pembelajaran fisika yang kontekstual dan berorientasi pada pembangunan berkelanjutan (Kioupi & Voulvoulis, 2022; United Nations Educational Scientific and Cultural Organization, 2020).

Meskipun beberapa butir soal masih memerlukan penyempurnaan berdasarkan hasil analisis daya pembeda, secara keseluruhan instrumen telah memenuhi kriteria kelayakan sebagai instrumen asesmen pendidikan. Revisi yang dilakukan terhadap butir-butir tersebut diharapkan semakin meningkatkan kualitas instrumen tanpa mengubah representasi konstruk yang diukur. Oleh karena itu, instrumen yang dikembangkan berpotensi digunakan sebagai alternatif asesmen autentik untuk mengukur keterampilan pemecahan masalah peserta didik pada pembelajaran fisika yang mengintegrasikan prinsip-prinsip *Education for Sustainable Development*.

Dengan karakteristik tersebut, instrumen yang dikembangkan tidak hanya memenuhi persyaratan psikometrik sebagai alat ukur keterampilan pemecahan masalah, tetapi juga merepresentasikan prinsip-prinsip asesmen autentik yang relevan dengan implementasi *Education for Sustainable Development* dalam pembelajaran fisika.

KESIMPULAN

Penelitian ini menghasilkan instrumen asesmen keterampilan pemecahan masalah berorientasi *Education for Sustainable Development* (ESD) pada materi pemanasan global yang valid, reliabel, dan layak digunakan melalui model pengembangan ADDIE. Instrumen terdiri atas 20 butir soal esai yang disusun berdasarkan empat tahapan pemecahan masalah Polya dan dilengkapi dengan rubrik penskoran analitik. Hasil validasi isi menunjukkan nilai *Percentage of Agreement* (PoA) sebesar 86% pada aspek isi dan konstruksi serta 90% pada aspek bahasa. Pada uji coba terhadap 35 peserta didik kelas XI MAN 2 Gresik, seluruh butir memenuhi kriteria validitas empiris yang digunakan, dengan koefisien Cronbach's Alpha sebesar 0,969. Seluruh butir memiliki tingkat kesukaran dalam kategori sedang, dan sebaran daya pembeda pada kategori lemah, cukup, hingga baik.

Penelitian ini memiliki keterbatasan substansi karena integrasi komponen ESD masih sebatas stimulus cerita dan belum menyatu ke dalam indikator pemecahan masalah Polya. Dampaknya, domain ESD belum terintegrasi secara merata serta belum menyatu mendalam ke dalam alur penalaran peserta didik. Selain itu, analisis daya pembeda menunjukkan bahwa butir 1b, 1c, dan 1d memiliki indeks sebesar 0,177 sehingga memerlukan revisi sebelum digunakan di kelas X. Dengan demikian, penelitian selanjutnya perlu mencakup penelaahan dan revisi terhadap butir yang memiliki daya pembeda rendah.

Implikasi penelitian ini adalah para pendidik dapat mengadopsi instrumen ini sebagai perangkat awal untuk mengukur keterampilan pemecahan masalah pada materi pemanasan global, namun perlu menyadari bahwa aspek *Education for Sustainable Development* (ESD) yang dihadirkan masih sebatas stimulus konteks

cerita. Bagi peneliti dan pengembang instrumen selanjutnya, temuan ini menuntut integrasi domain ESD yang lebih substansial ke dalam alur penalaran Polya serta penyertaan konsep-konsep inti fisika dengan perhitungan matematis yang lebih kompleks agar instrumen dapat mengukur kemampuan siswa secara kuantitatif dan komprehensif.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Program Studi Pendidikan Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya serta MAN 2 Gresik atas dukungan dan fasilitasi selama pelaksanaan penelitian. Penulis juga menyampaikan apresiasi kepada para validator ahli dan seluruh peserta didik yang telah berpartisipasi dalam penelitian ini.

KONTRIBUSI PENULIS

Tri Wulandari: Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Writing – Original Draft Preparation. **Mukhayyarotin Niswati Rodliyatul Jauhariyah:** Conceptualization, Methodology, Validation, Supervision, Writing – Review & Editing, Project Administration.

PERNYATAAN BEBAS KONFLIK KEPENTINGAN

Para penulis menyatakan bahwa tidak ada konflik kepentingan finansial maupun hubungan pribadi yang dapat mempengaruhi hasil yang dilaporkan dalam naskah ini.

PERNYATAAN ETIKA PENELITIAN DAN PUBLIKASI

Para penulis menyatakan bahwa penelitian dan penulisan naskah ini telah mematuhi standar etika penelitian dan publikasi, sesuai dengan prinsip ilmiah, serta bebas dari plagiasi.

PERNYATAAN PEMANFAATAN TEKNOLOGI ASISTIF

Para penulis menyatakan bahwa Kecerdasan Buatan Generatif (*Generative Artificial Intelligence*) digunakan secara terbatas sebagai teknologi asistif dalam proses penyuntingan dan penyempurnaan penulisan naskah ini. Secara khusus, ChatGPT (OpenAI) digunakan untuk membantu perbaikan struktur artikel, penyuntingan bahasa, peningkatan kejelasan dan koherensi narasi ilmiah, serta penyempurnaan redaksi berdasarkan hasil penelitian yang tersedia. Seluruh saran yang dihasilkan oleh AI telah ditinjau dan dievaluasi oleh para penulis. Para penulis bertanggung jawab penuh atas keakuratan data, interpretasi hasil, validitas sitasi dan referensi, serta isi akhir naskah, dan memastikan bahwa penggunaan teknologi asistif tersebut tidak menggantikan tanggung jawab akademik maupun penilaian ilmiah para penulis.

DAFTAR PUSTAKA

AERA, APA, & NCME. (2014). *Standards for Educational and Psychological Testing*. American Educational Research Assosiation.

Ayudha, C. F. H., & Setyarsih, W. (2021). Studi Literatur : Analisis Praktik Pembelajaran Fisika Di Sma Untuk Melatih Keterampilan Pemecahan Masalah. *Jurnal Pendidikan Fisika Undiksha*, *11*(1), 16. https://doi.org/10.23887/jjpf.v11i1.33427

Ayumniyya, L., & Setyarsih, W. (2021). Profil Kemampuan Berpikir Tingkat Tinggi Siswa SMA dalam Pemecahan Masalah pada Materi Hukum Newton. *IPF: Inovasi Pendidikan Fisika*, *10*(1), 50–58.

Borich, G. D., Godbout, R. C., Peck, R. F., Kash, M. M., & Poynor, L. H. (1994). An Evaluation of the Personalized Model of Teacher. *Research and Development Center for Teacher Education The University of Texas at Austin*.

DeVellis, R. F. (2017). *Scale Development: Theory and Applications Fourth Edition*. SAGE Publications Inc.

DeVellis, R. F., & Thorpe, C. T. (2022). Scale Development Theory and Applications (Applied Social Research Methods). In *SAGE Publication, Inc*. SAGE Publication, Inc.

Docktor, J. L., & Mestre, J. P. (2014). Synthesis of discipline-based education research in physics. *Physical Review Special Topics - Physics Education Research*, *10*(020119), 1–58. https://doi.org/10.1103/PhysRevSTPER.10.020119

Ebel, R. L., & Friesbie, D. A. (1991). Essentials of Educational Measurement Fifth Edition. In *Prentice Hall, Inc.* Prentice Hall, Inc. https://ebookppsunp.files.wordpress.com/2016/06/robert_l-ebel_david_a-

_frisbie_essentials_of_edbookfi-org.pdf
Engelhardt, P. V. (2009). An Introduction to Classical Test Theory as Applied to Conceptual Multiple-choice Tests. *Reviews in Physics Education Research*, *2*(1), 1–40. https://doi.org/10.1119/revperv2.1.2
Firmansyah, F., Sukarno, S., Kafrita, N., & Al Farisi, S. (2022). Pengaruh Model Pembelajaran Problem Based Learning (PBL) Terhadap Kemampuan Pemecahan Masalah Fisika Siswa SMA Negeri 11 Muaro Jambi. *Physics and Science Education Journal (PSEJ)*, *2*(April), 75–82. https://e-journal.lp2m.uinjambi.ac.id/ojp/index.php/psej/article/view/1447
Franestian, I. D., Suyanta, & Wiyono, A. (2020). Analysis problem solving skills of student in Junior High School. *Journal of Physics: Conference Series*, *1440*(1). https://doi.org/10.1088/1742-6596/1440/1/012089
Hebebci, M. T., & Usta, E. (2022). The Effects of Integrated STEM Education Practices on Problem Solving Skills, Scientific Creativity, and Critical Thinking Dispositions. *Participatory Educational Research*, *9*(6), 358–379. https://doi.org/10.17275/per.22.143.9.6
Isac, M. M., Sandoval-Hernández, A., & Sass, W. (2025). *Knowledge and Willingness to Act Pro-Environmentally: Perspectives from IEA TIMSS 2019 and ICCS 2016 data*. Springer. https://library.oapen.org/handle/20.500.12657/97008
Jauhariyah, M. N. R., Hariyono, E., Abidin, E. N., & Prahani, B. K. (2019). Fostering Prospective Physics Teachers' Creativity in Analysing Education for Sustainable Development Based Curricula. *Journal of Physics: Conference Series*, *1417*(1). https://doi.org/10.1088/1742-6596/1417/1/012086
Jauhariyah, M. N. R., Prahani, B. K., Syahidi, K., Deta, U. A., Lestari, N. A., & Hariyono, E. (2021). ESD for physics: How to infuse education for sustainable development (ESD) to the physics curricula? *Journal of Physics: Conference Series*, *1747*(1), 012032. https://doi.org/10.1088/1742-6596/1747/1/012032
Kioupi, V., & Voulvoulis, N. (2022). Education for Sustainable Development as the Catalyst for Local Transitions Toward the Sustainable Development Goals. *Frontiers in Sustainability*, *3*(July), 1–18. https://doi.org/10.3389/frsus.2022.889904
Kılıç, A. F., Koyuncu, İ., & Uysal, İ. (2023). Scale Development Based on Item Response Theory: A Systematic Review. *International Journal of Psychology and Educational Studies*, *10*(1), 209–223. https://doi.org/10.52380/ijpes.2023.10.1.982
Makiyah, Y. S., Mahmudah, I. R., Sulistyaningsih, D., & Susanti, E. (2021). Hubungan Keterampilan Komunikasi Abad 21 Dan Keterampilan Pemecahan Masalah Mahasiswa Pendidikan Fisika. *Journal of Teaching and Learning Physics*, *6*(1), 1–10. https://doi.org/10.15575/jotalp.v6i1.9412
Maulita, S. R., Sukarmin, S., & Marzuki, A. (2019). The Content Validity: Two-Tier Multiple Choices Instrument to Measure Higher-Order Thinking Skills. *Journal of Physics: Conference Series*, *1155*(1). https://doi.org/10.1088/1742-6596/1155/1/012042
Meyer, A., & Friege, G. (2026). An atomized approach to assessing energy problem solving in physics using multidimensional item response theory. *Frontiers in Education*, *10*(February), 01–12. https://doi.org/10.3389/feduc.2026.1759878
Musengimana, T., Yadav, L. L., Uwamahoro, J., & Nizeyimana, G. (2025). Assessing physics students ' problem-solving skills : a baseline investigation. *Discover Education*, *4*(196), 1–15. https://doi.org/10.1007/s44217-025-00640-1
OECD. (2022). *PISA 2022 Assessment and Analytical Framework*. OECD Publishing.
OECD. (2023). PISA 2022 Results (Volume II): Learning During – and From – Disruption. In *OECD Publishing: Vol. II*. OECD Publising. https://www.oecd-ilibrary.org/education/pisa-2022-results-volume-ii_a97db61c-en
Olsson, D., Gericke, N., & Boeve-de Pauw, J. (2022). The effectiveness of education for sustainable development revisited – a longitudinal study on secondary students ' action competence for sustainability. *Environmental Education Research*, *28*(3), 405–429. https://doi.org/10.1080/13504622.2022.2033170
Polya, G. (1957). *How to Solve It: A New Aspect of Mathematical Method* (p. 284). Princeton University Press.
Putri, R., Sari, R., Jumadi, J., & Ariswan, A. (2020). Pengembangan dan Validasi Instrumen Tes untuk Mengukur Keterampilan Menyelesaikan Masalah Peserta Didik SMA pada Pelajaran Fisika. *Jurnal Penelitian Pembelajaran Fisika*, *11*(1), 17–26. https://doi.org/10.26877/jp2f.v11i1.3993
Reinhard, N., Schubert, F. K., Bertolini, E., Hagedorn, N., Manoli, G., Sekiguchi, M., Yoshii, T., Rieger, D., & Helfrich-Förster, C. (2022). The Neuronal Circuit of the Dorsal Circadian Clock Neurons in Drosophila melanogaster. *Frontiers in Physiology*, *13*(April), 1–29. https://doi.org/10.3389/fphys.2022.886432

Rieckmann, M. (2023). *Education for Sustainable Development in Teacher Education and Schools – Transformative and Disruptive Education Summer School ' Teaching for Sustainable*. *August*.

Robert Maribe Branch. (2009). Approach, Instructional Design: The ADDIE. In *Department of Educational Psychology and Instructional Technology University of Georgia* (Vol. 53, Issue 9).

Sherliyanti, Y., & Jauhariyah, M. N. R. (2024). Validitas Instrumen Penilaian Berbasis Kelas untuk Mengukur Kompetensi Literasi Sains Peserta Didik pada Materi Pemanasan Global. *IPF: Inovasi Pendidikan Fisika*, *13*(3), 168–176. https://doi.org/10.26740/ipf.v13n3.p168-176

Tilbury, D. (2014). *Education for sustainable development: an expert review of processes and learning; 2011*. *May*.

United Nations. (2023). Education for sustainable development in the framework of the 2030 genda for Sustainable Development. *General Assembly*, *14449*(July), 1–19.

United Nations Educational Scientific and Cultural Organization. (2018). *Issues and trends in Education for Sustainable Development*. UNESCO Publishing.

United Nations Educational Scientific and Cultural Organization. (2020). Education for sustainable development: a roadmap. In *Education for sustainable development: a roadmap*. UNESCO. https://doi.org/10.54675/yfre1448

Widiawati, R., Hikmawati, H., & 'Ardhuha, J. (2022). Pengembangan Perangkat Pembelajaran Berbasis Model Problem Based Learning untuk Meningkatkan Kemampuan Pemecahan Masalah Fisika Peserta Didik pada Materi Fluida Dinamis. *Jurnal Ilmiah Profesi Pendidikan*, *7*(3c), 1803–1810. https://doi.org/10.29303/jipp.v7i3c.857

Yuliantaningrum, L., & Sunarti, T. (2020). Pengembangan Instrumen Soal HOTS untuk Mengukur Keterampilan Berpikir Kritis, Berpikir Kreatif, dan Pemecahan Masalah Materi Gerak Lurus pada Peserta Didik SMA. *IPF: Inovasi Pendidikan Fisika*, *09*(02), 76–82.