Copyright © 2025 The Author(s)
Didaktika: Jurnal Kependidikan Vol. 14 No. 3 Agustus 2025

Pengembangan Instrumen Tes Literasi Sains Berbasis PISA pada Materi Sistem Pencernaan untuk Siswa Kelas V Sekolah Dasar

Enung Nurhasanah[1], Yuli Kurniawati Sugiyo Pranoto[2], Decky Avrilianda[3], Ellianawati Ellianawati [4]

[1,2,3,4]Universitas Negeri Semarang, Indonesia

[1]enungnurhasanah23@students.unnes.ac.id

Abstrak

Penelitian ini bertujuan untuk mengembangkan dan memvalidasi instrumen tes literasi sains berbasis kerangka PISA pada materi sistem pencernaan untuk siswa kelas V sekolah dasar. Latar belakang penelitian ini didasarkan pada kebutuhan akan instrumen penilaian yang tidak hanya mengukur penguasaan konsep, tetapi juga kemampuan berpikir ilmiah dalam konteks kehidupan nyata. Penelitian ini menggunakan metode Research and Development (R&D) dengan model pengembangan ADDIE yang dimodifikasi menjadi beberapa tahapan, yaitu: Analysis, Design, Development, Implementation, and Evaluation. Instrumen yang dikembangkan mengacu pada tiga domain literasi sains PISA, yaitu: konteks, pengetahuan, dan proses ilmiah. Validasi isi dilakukan oleh dua orang ahli menggunakan teknik Gregory dan menunjukkan tingkat kesesuaian yang tinggi antar penilai. Uji validitas empiris dilakukan menggunakan korelasi Point Biserial dengan bantuan SPSS 29.00, dan hasil menunjukkan bahwa sebagian besar butir soal memiliki nilai validitas tinggi. Selain itu, analisis konstruk dan bahasa menunjukkan bahwa instrumen telah disusun secara sistematis, menggunakan bahasa yang sesuai dengan kaidah EYD, serta mudah dipahami oleh siswa sekolah dasar. Dengan demikian, instrumen tes literasi sains yang dikembangkan dinyatakan valid secara isi dan empiris serta layak digunakan sebagai alat evaluasi kemampuan literasi sains siswa SD. Penelitian ini diharapkan dapat menjadi rujukan dalam pengembangan asesmen berbasis konteks dan mendukung peningkatan kemampuan literasi sains di jenjang pendidikan dasar.

Kata kunci: literasi sains, PISA, sistem pencernaan, instrumen penilaian, validitas isi

Pendahuluan

Perkembangan ilmu pengetahuan dan teknologi yang pesat di abad ke-21 menuntut kemampuan literasi sains yang kuat sejak dini. Literasi sains tidak hanya berarti menguasai konsep-konsep IPA secara faktual, tetapi juga membutuhkan keterampilan berpikir tingkat tinggi, seperti mengevaluasi, menganalisis, dan menciptakan solusi untuk masalah yang terjadi di lingkungan tertentu (Raharjo *et al*., 2023: Putri, 2020). Kemampuan ini sangat penting dalam membangun siswa yang berpikir kritis, peduli terhadap lingkungan, dan membuat keputusan berdasarkan bukti ilmiah (Zulfiana *et al*., 2023).

Literasi sains siswa Indonesia masih menjadi masalah besarc Indonesia sekarang berada di posisi 67 dari 81 negara, tetapi skornya masih jauh di bawah rata-rata OECD, yaitu 499 poin (OECD, 2023). Penurunan skor ini menunjukkan bahwa siswa di Indonesia belum mampu menerapkan pengetahuan sains ke dunia nyata, yang sebenernya merupakan dasar literasi sains (Thoriq, 2024; Azaly & Fitrihidajati, 2021).

Hasil penelitian lapangan memperkuat keadaan ini. Seperti yang ditunjukkan oleh temuan yang dilakukan pada guru kelas V SD di forum Kelompok Kerja Guru (KKG) Gugus VI Kecamatan Kalijati, siswa menghadapi kesulitan dalam menjawab pertanyaan sains yang membutuhkan penalaran dan pemahaman mendalam. Guru menyatakan bahwa sebagian besar soal dalam buku teks dan evaluasi harian masih hafalan dan tidak menyentuh aspek literai. Akibatnya, siswa tidak tertarik untuk belajar sains dan tidak memahami konsep-konsep sehari-hari seperti pencernaan manusia.

Salah satu topik penting dalam pembelajaran Ilmu Pengetahuan Alam (IPA) di kelas V SD adalah sistem pencernaan. Memahami sistem pencernaan tidak hanya penting untuk tujuan akademik, tetapi juga terkait langsung dengan kesadaran siswa akan pentingnya menjalani gaya hidup bersih dan pola makan sehat (Hampel *et al*., 2023). Evaluasi dalam pembelajarann sangat penting karena dapat memungkinkan guru untuk mengevaluasi kemampuan siswa dan pemahaman mereka tentang materi pelajaran (Risaad *et al*., 2025). Oleh karena itu, diperlukan alat evaluasi yang berkualitas alat ukur yang berkualitas tinggi untuk mengumpulkan data yang lebih akurat tentang kemampuan kognitif peserta didik (Dermawan *et al*., 2021).

Banyak upaya telah dilakukan di berbagai jenjang pendidikan untuk mengembangkan instrumen ujian yang berbasis literasi sains. Namun, penelitian yang dilakukan oleh Zulfiana *et al.*, (2023) menekankan pentingnya menggunakan instrumen evaluasi untuk meningkatkan kemampuan berpikir kritis, tetapi lebih umum. Penelitian Murti dan Sunarti (2021) menunjukkan bahwa pendekatan kontekstual dan berbasis budaya lokal efektif dalam meningkatkan literasi sains siswa; namun, materi dari sistem pencernaan tidak terkait secara langsung dengan kemampuan siswa untuk membaca dan memahami sains. Sementara itu, penelitian Anggreni *et al.*, (2022) mengembangkan instrumen yang terbukti valid dan reliabel dalam mengukur keterampilan literasi sains anak yang berkonsentrasi pada anak usia dini.

Penemuan-penemuan ini menunjukkan betapa pentingnya melakukan penelitian yang mengembangkan instrumen tes literasi sains berbasis PISA yang berfokus pada materi yang relevan dengan kurikulum SD dan karakteristik perkembangan kognitif siswa kelas V. Diharapkan bahwa pengembangan instrumen ini tidak hanya akan menilai penguasaan konsep siswa, tetapi juga memberikan gambaran menyeluruh mengenai kemampuan mereka dalam memahami, menalar, dan menerapkan sains dalam kehidupan sehari-hari secara bermakna dan kontekstual.

Berdasarkan analisis kesenjangan tersebut, pertanyaan penelitian yang diangkat adalah: "Bagaimana efektivitas pengembangan instrumen tes literasi sains berbasis PISA pada materi sistem pencernaan manusia siswa kelas V?"

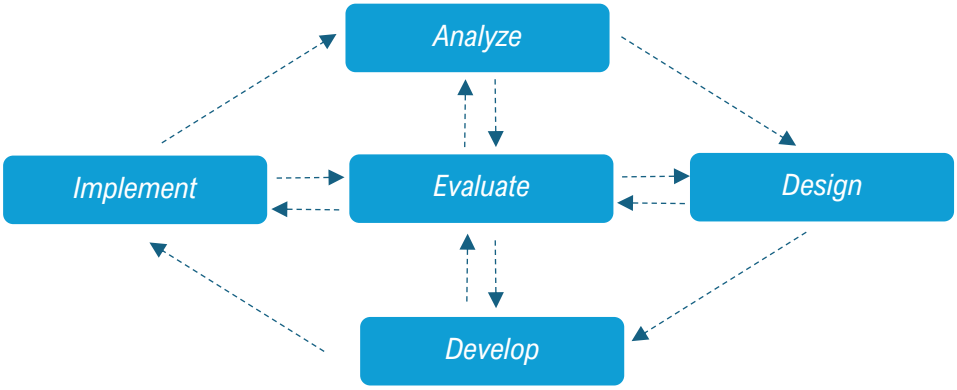
Tujuan penelitian ini adalah untuk membuat instrumen tes literasi sains berbasis PISA yang valid, dapat diandalkan, dan praktis untuk pembelajaran IPA SD. Penelitian ini baru karena mengintegrasikan pengembangan instrumen literasi sains berbasis PISA dengan konten pembelajaran tertentu dan berbasis kebutuhan riil guru dan siswa di sekolah dasar. Selain itu, penelitian ini dapat membantu mengembangkan instrumen penilaian formatif yang lebih kontekstual, fleksibel, dan sesuai dengan profil siswa Pancasila dalam Kurikulum Merdeka.

Metode

Penelitian ini menggunakan pendekatan *Research and Development* (R&D) dengan model pengembangan *Analysis, Design, Development, Implementation, and Evaluation* (ADDIE). Model ini dipilih karena memberikan struktur sistematis dan iteratif yang dapat digunakan untuk mengembangkan, menguji, dan menyempurnakan produk pendidikan secara berkelanjutan (Arni *et al.*, 2025; Adeoye *et al.,* 2024). Dalam konteks penelitian ini, produk yang dikembangkan

Copyright © 2025 The Author(s)

adalah instrumen tes literasi sains berbasis *Programme for International Student Assessment (*PISA) pada materi sistem pencernaan untuk siswa kelas V Sekolah Dasar.



Gambar 1 Tahap Pengembangan Instrumen Model ADDIE

Tahap analisis (*analysis*) melibatkan pengenalan masalah dan kebutuhan lapangan. Ini termasuk melakukan survei dan penelitian literatur dengan guru kelas V di Kecamatan Kalijati. Hasil analisis menunjukkan bahwa karena kebanyakan soal sains bersifat faktual, siswa kesulitan memahaminya. Selain itu, analisis kurikulum dan kompetensi dasar dilakukan untuk memastikan bahwa alat yang dikembangkan terkait dengan pembelajaran IPA di SD.

Pada tahap desain (*design*), dibuat instrumen tes literasi sains yang memiliki indikator yang berkaitan dengan sistem pencernaan. Agar lebih bermakna secara kontekstual, indikator soal disusun berdasarkan kompetensi literasi sains menurut PISA, yang meliputi: (1) menjelaskan fenomena secara ilmiah, (2) mengevaluasi dan merancang penyelidikan ilmiah, dan (3) menafsirkan data dan bukti secara ilmiah (OECD, 2023). Selain itu, pengetahuan yang diukur mencakup aspek konten seperti fungsi organ pencernaan dan proses yang terjadi di dalamnya, aspek prosedural seperti cara melakukan pengamatan atau eksperimen sederhana tentang pencernaan, serta aspek epistemik seperti pemahaman tentang bagaimana informasi ilmiah tentang sistem pencernaan diperoleh dan divalidasi. Konteks soal dirancang sedekat mungkin dengan kehidupan sehari-hari siswa, seperti kebiasaan makan, gangguan pencernaan, dan pentingnya makanan bergizi, sehingga mendorong siswa untuk berpikir kritis dan menerapkan pengetahuan ilmiah dalam situasi nyata.

Tahap pengembangan (*development*), instrumen yang telah dirancang kemudian dibentuk menjadi soal tertulis dalam berbagai jenis, seperti pilihan ganda, pilihan banyak, isian, uraian terbuka, dan menjodohkan konsep. Draft soal divalidasi oleh praktisi pendidikan dasar, ahli materi, dan ahli evaluasi pendidikan. *Content Validity Ratio* (CVR) digunakan untuk mengukur validitas isi, dan reliabilitas diuji dengan Alpha Cronbach menggunakan program statistik SPSS.

Tahap implementasi (*implementation*), Instrumen diuji coba pada siswa kelas V dari empat sekolah dasar di Kecamatan Kalijati yang dipilih secara acak. Siswa berpartisipasi secara aktif dalam uji coba dan memberikan masukan tentang cara instrumen digunakan. Tahap akhir, yaitu evaluasi (*evaluation*) dilakukan secara formatif dan sumatif. Evaluasi formatif dilakukan sepanjang proses pengembangan, sedangkan evaluasi sumatif dilakukan setelah implementasi untuk menilai efektivitas, efisiensi, dan keterterapan instrumen dalam pembelajaran IPA.

Populasi penelitian adalah seluruh siswa kelas V Sekolah Dasar di Kecamatan Kalijati, Kabupaten Subang. Sampel penelitian terdiri dari seluruh siswa kelas V Sekolah Dasar di Kecamatan Kalijati, Kabupaten Subang. Sampel diambil secara acak sederhana (*simple random sampling*) dari empat SD dengan latar belakang yang beragam. Jumlah sampel yang

representatif diperoleh dengan menggunakan rumus Slovin (Sevilla, C., *et al* 2007) untuk menentukan jumlah sampel, dengan batas toleransi kesalahan (e) 5%.

Beberapa metode digunakan untuk mengumpulkan data; 1) Wawancara menyeluruh dengan guru dan siswa untuk menentukan kebutuhan dan kesulitan dalam pembelajaran literasi sains berbasis PISA; 2) Observasi di kelas untuk memantau pelaksanaan pembelajaran IPA dan penggunaan soal evaluasi; 3) Penyebaran alat ujian untuk mengevaluasi pemahaman dan kemampuan siswa dalam PISA.

Hasil tes diuji untuk validitas (daya beda dan tingkat kesukaran), statistik deskriptif (rerata, standar deviasi), dan reliabilitas menggunakan Alpha Cronbach. Untuk data kualitatiif dari observasi dan wawancara dianalisis secara tematik. Prosedur ini termasuk transkripsi, pengkodean, identifikasi tema, dan interpretasi makna.

Dengan metode ini, penelitian ini diharapkan menghasilkan instrumen tes yang valid dan dapat diandalkan untuk menilai kemampuan literasi sains siswa SD pada materi sistem pencernaan secara menyeluruh dan kontekstual melalui PISA.

Hasil

Tahap Analisis

Analisis dilakukan melalui observasi, wawancara, dan studi literatur secara sistematis untuk mengidentifikasi kebutuhan pengembangan instrumen literasi sains berbasis PISA pada materi sistem pencernaan di sekolah dasar. Hasil analisis menunjukkan bahwa soal-soal IPA yang digunakan di kelas sebagian besar bersifat faktual dan tidak menuntut siswa untuk berpikir kritis atau menerapkan pengetahuan yang mereka pelajari ke situasi dunia nyata. Karena tidak ada hubungan antara materi soal dan pengalaman sehari-hari siswa, siswa mungkin kesulitan memahaminya. Selain itu, asesmen yang dilakukan guru lebih berfokus pada hafalan konsep dan pemahaman dasar, belum menyentuh kemampuan literasi sains yang mencakup interpretasi data, pemecahan masalah, dan evaluasi informasi ilmiah secara kritis. literasi sains yang komprehensif diperlukan agar siswa dapat menghubungkan pengetahuan ilmiah dengan masalah dunia nyata. Oleh karena itu, diperlukan alat evaluasi yang dapat menilai kemampuan siswa dalam memahami dan menerapkan konsep-konsep ilmiah tentang sistem pencernaan melalui pendekatan PISA, seperti merencanakan penyelidikan, menafsirkan data, dan membuat keputusan berdasarkan bukti.

Tahap Perencanaan

Model ADDIE digunakan untuk membuat instrumen tes literasi sains berbasis PISA. Pada tahap desain, indikator soal disusun berdasarkan kompetensi literasi sains PISA, yaitu memberikan penjelasan tentang fenomena ilmiah, membuat laporan penyelidikan, dan menafsirkan data. Sebuah soal dapat dibuat dalam berbagai bentuk, yaitu pilihan ganda, pilihan banyak, uraian, dan menjodohkan. Mereka berkaitan dengan hal-hal seperti masalah pencernaan dan diet. Pada tahap pengembangan, reliabilitas diuji dengan menggunakan Alpha Cronbach dan validasi oleh ahli. Untuk memastikan kualitas instrumen, tingkat kesulitan dan daya beda soal dievaluasi. Selanjutnya, instrumen diujicobakan secara acak pada siswa kelas V di empat sekolah dasar. Untuk menilai efektivitas dan keterterapan instrumen dalam pembelajaran IPA kontekstual, evaluasi dilakukan secara formatif dan sumatif.

Tabel 1. Kisi-Kisi Instrumen Tes Literasi Sains

Domain Literasi Sains	Indikator Literasi Sains	Nomor Soal	Jumlah Soal
Konteks	Lingkungan	4, 10, 15, 17, 19,	5
Pengetahuan	Pengetahuan konten	1, 2, 3	3

Copyright © 2025 The Author(s)

Domain Literasi Sains	Indikator Literasi Sains	Nomor Soal	Jumlah Soal
Proses	Pengetahuan procedural	6,14,	2
	Menjelaskan fenomena secara ilmiah	5, 7, 18	3
	Mengevaluasi dan merancang penyelidikan ilmiah	11, 12, 13	3
	Menafsirkan data dan bukti ilmiah	8, 9, 16, 20	4
	Total Butir Soal		20

Tahap Pengembangan

Tahap pengembangan dilakukan dengan menyusun butir-butir soal berdasarkan indikator yang telah dirancang pada tahap desain. Menurut PISA, domain literasi sains—konteks, pengetahuan, dan proses ilmiah—digunakan dalam pembuatan soal. Setiap soal dirancang untuk menggambarkan situasi nyata dari kehidupan siswa, terutama yang berkaitan dengan sistem pencernaan. Alat ini digunakan untuk menilai berbagai aspek kemampuan siswa melalui soal pilihan ganda, pilihan banyak, uraian, dan menjodohkan.

Tabel 2. Pengembangan Soal Berdasarkan Indikator Literasi Sains

Indikator Literasi Sains	Bentuk Soal
Lingkungan	Doni sering makan terlalu cepat dan tidak mengunyah makanannya dengan baik. Kebiasaan ini dapat menyebabkan gangguan pencernaan karena.... A. Makanan langsung diserap tubuh tanpa dicerna B. Proses pengunyahan penting untuk mencerna makanan secara mekanik C. Perut akan menjadi lebih kenyang D. Makanan terasa lebih enak jika dikunyah cepat
Pengetahuan konten	Langkah pertama yang perlu dilakukan untuk mengetahui pengaruh makanan pedas terhadap lambung adalah.... A. Makan makanan pedas setiap hari B. Membaca buku tentang pencernaan C. Melakukan pengamatan terhadap orang yang makan pedas D. Menyusun rencana percobaan dan menentukan variabelnya
Pengetahuan procedural	Sebutkan dan jelaskan secara berurutan organ-organ yang dilalui makanan sejak masuk ke mulut hingga sisa makanan dikeluarkan dari tubuh!

<table>
<tr><th>Indikator Literasi Sains</th><th>Bentuk Soal</th></tr>
<tr><td>Menjelaskan fenomena secara ilmiah</td><td>Pasangkan fenomena di Kolom A dengan penjelasan ilmiah di Kolom B yang sesuai!
<table>
<tr><th>Fenomena</th><th></th><th>Penjelasan Ilmiah</th></tr>
<tr><td>1. Makanan terasa manis di mulut</td><td></td><td>A. Sistem pencernaan bekerja terlalu berat</td></tr>
<tr><td>2. Perut terasa kenyang</td><td></td><td>B. Enzim amilase memecah karbohidrat</td></tr>
<tr><td>3. Mual karena makan berlebihan</td><td></td><td>C. Lambung telah terisi penuh</td></tr>
</table></td></tr>
<tr><td>Mengevaluasi dan merancang penyelidikan ilmiah</td><td>Kani ingin mengetahui pengaruh makan pagi terhadap semangat belajar di sekolah. Langkah yang tepat dalam menyusun penyelidikan ilmiah adalah....
□ A. Menentukan pertanyaan penelitian
□ B. Mengamati teman yang tidak sarapan
□ C. Menulis hasil observasi dalam jurnal
□ D. Mewawancarai guru olahraga</td></tr>
</table>

Uji ahli yang melibatkan dua validator—ahli materi dan ahli evaluasi (penilaian)—digunakan untuk memastikan validitas isi penelitian ini. Tujuan uji ini adalah untuk mengevaluasi seberapa sesuai materi soal dengan indikator dalam kisi-kisi instrumen literasi sains. Setiap validator menilai instrumen menggunakan angket yang disusun berdasarkan tiga aspek utama: (1) kesesuaian materi dengan domain dan indikator literasi sains; (2) konstruksi soal (ketepatan bentuk soal, kejelasan pilihan jawaban, dan ketepatan kata kunci); dan (3) kejelasan bahasa (ketepatan struktur bahasa, ketepatan struktur soal, dan ketepatan kata kunci).

Tabel 3. Skala Likert

Skor	Kategori Penilaian	Keterangan
4	Sangat Baik	Pernyataan/butir sangat relevan/sesuai, tidak perlu perbaikan
3	Baik	Pernyataan/butir relevan/sesuai, perlu sedikit perbaikan
2	Cukup	Pernyataan/butir kurang relevan/sesuai, perlu perbaikan signifikan
1	Tidak Baik	Pernyataan/butir tidak relevan/tidak sesuai, perlu direvisi total

Hasil validasi menunjukkan bahwa ahli materi mendapatkan skor rata-rata 3,67, dan ahli evaluasi mendapatkan skor rata-rata 3,75. Instrumen yang dikembangkan sangat layak untuk digunakan dalam uji coba lapangan, menurut kriteria validitas Riduwan (2016), dengan skor antara 3,26 dan 4,00, masuk dalam kategori "sangat valid".

Copyright © 2025 The Author(s)

Tabel 4. Rekapitulasi Hasil Validasi Isi oleh Ahli Materi dan Ahli Evaluasi

Aspek yang Dinilai	Skor Ahli Materi	Kategori	Skor Ahli Evaluasi	Kategori
Kesesuaian Konten	3,75	Sangat Valid	3,8	Sangat Valid
Konstruksi Soal	3,5	Sangat Valid	3,7	Sangat Valid
Kejelasan Bahasa	3,75	Sangat Valid	3,75	Sangat Valid
Rata-Rata Total	3,67	Sangat Valid	3,75	Sangat Valid

Dengan demikian, instrumen tes literasi sains yang dikembangkan telah memenuhi syarat validitas isi dan dinyatakan layak untuk digunakan pada tahap uji coba empiris. Validasi yang diberikan oleh ahli ini menjadi landasan penting untuk menguji instrumen secara lebih lanjut untuk validitas empiris dan reliabilitas.

Tahap Implementasi

Implementasi dilakukan di SDN Alun-alun, Kecamatan Kalijati. Siswa kelas V digunakan sebagai uji coba. Alat literasi sains diberikan kepada siswa untuk mengetahui sejauh mana alat tersebut dapat mengukur kemampuan mereka dalam literasi sains sesuai dengan domain dan indikator yang telah ditetapkan. Uji coba dilakukan selama kegiatan pembelajaran IPA. Siswa mengerjakan semua soal dalam waktu yang ditetapkan.

Tahap Evaluasi

Data dari siswa kelas V SDN Alun-alun, Kecamatan Kalijati, digunakan untuk menilai instrumen. Instrumen terdiri dari dua puluh soal dengan berbagai bentuk, termasuk pilihan ganda, pilihan banyak, menjodohkan, dan uraian. Soal-soal ini menggabungkan metrik dari domain indikator literasi sains PISA. Untuk mengevaluasi validitas setiap soal, daya beda dan tingkat kesukaran dihitung. Hasilnya menunjukkan bahwa sebagian besar soal berada dalam kategori daya beda sedang hingga tinggi dan memiliki tingkat kesukaran sedang. Oleh karena itu, soal-soal ini dianggap baik dan proporsional untuk mengukur kemampuan siswa.

Tabel 4. Rekapitulasi Hasil Validasi Isi oleh Ahli Materi dan Ahli Evaluasi

Butir	r-hitung	r-tabel	Validitas	Daya Pembeda	Keterangan Daya Pembeda	Tingkat Kesukaran	Keterangan Kesukaran
Soal 1	0,728	0,444	Valid	0,516	Baik	0,65	Sedang
Soal 2	0,621	0,444	Valid	0,544	Baik	0,62	Sedang
Soal 3	0,563	0,444	Valid	0,516	Baik	0,66	Sedang
Soal 4	0,459	0,444	Valid	0,372	Baik diperbaiki	0,58	Sedang
Soal 5	0,690	0,444	Valid	0,710	Baik	0,74	Mudah
Soal 6	0,650	0,444	Valid	0,690	Baik	0,69	Sedang
Soal 7	0,675	0,444	Valid	0,670	Baik	0,67	Sedang
Soal 8	0,648	0,444	Valid	0,491	Baik	0,65	Sedang
Soal 9	0,680	0,444	Valid	0,690	Baik	0,66	Sedang
Soal 10	0,461	0,444	Valid	0,630	Baik	0,61	Sedang
Soal 11	0,490	0,444	Valid	0,332	Baik diperbaiki	0,55	Sedang
Soal 12	0,728	0,444	Valid	0,305	Baik diperbaiki	0,60	Sedang

Butir	r-hitung	r-tabel	Validitas	Daya Pembeda	Keterangan Daya Pembeda	Tingkat Kesukaran	Keterangan Kesukaran
Soal 13	0,563	0,444	Valid	0,446	Baik	0,63	Sedang
Soal 14	0,766	0,444	Valid	0,442	Baik	0,69	Sedang
Soal 15	0,683	0,444	Valid	0,516	Baik	0,66	Sedang
Soal 16	0,507	0,444	Valid	0,534	Baik	0,65	Sedang
Soal 17	0,690	0,444	Valid	0,660	Baik	0,28	Sukar
Soal 18	0,710	0,444	Valid	0,730	Baik	0,76	Mudah
Soal 19	0,675	0,444	Valid	0,650	Baik	0,29	Sukar
Soal 20	0,705	0,444	Valid	0,700	Baik	0,27	Sukar

Hasil uji validitas yang dilakukan dengan SPSS 29 menunjukkan bahwa empat soal (nomor 6, 7, 9, dan 10) tidak valid. Ini karena nilai korelasi Pearson yang diperoleh tidak memenuhi kriteria validitas jika dibandingkan dengan nilai r-tabel. Untuk langkah berikutnya, alat yang digunakan hanya terdiri dari soal-soal yang telah teruji dan valid. Ini membuat pengukuran pemahaman peserta didik tentang materi yang diujikan lebih akurat.

Hasil analisis menunjukkan bahwa sebagian besar butir soal berada dalam kategori baik, yang berarti mereka dapat secara efektif membedakan siswa berkemampuan tinggi dan rendah (soal 1, 2, 3, 8, 13, 14, 15, dan 16). Soal lain, seperti soal 4, 11, 12, 17, 18, 19, dan 20, berada dalam kategori "baik diperbaiki", yang berarti mereka masih dapat digunakan dengan beberapa revisi. Namun, beberapa soal tidak layak digunakan karena daya pembedanya sangat rendah. Daya pembeda yang baik sangat penting karena menunjukkan seberapa baik suatu soal dapat membedakan siswa yang mampu dan siswa yang tidak mampu dalam pelajaran (Anastasi, A., & Urbina, 1997). Oleh karena itu, kualitas daya pembeda ini menjadi salah satu indikator penting dalam pembuatan alat evaluasi yang sah dan dapat diandalkan.

Tabel 5. Uji Reliabilitas

Reability Statistic Cronbach's Alpha	N of Items
0,738	21

Uji reliabilitas instrumen dilakukan dengan menggunakan program SPSS versi 29. Nilai Alpha Cronbach adalah 0,738. Nilai ini berada di atas ambang batas minimum sebesar 0,70 yang disarankan oleh Nunnally, J. C., & Bernstein (1994) menunjukkan bahwa instrumen memiliki tingkat konsistensi internal yang cukup tinggi. Artinya, soal-soal dalam instrumen ini saling berkorelasi secara positif dan stabil dalam mengukur pengetahuan sains peserta didik tentang bahan-bahan yang berkaitan dengan sistem pencernaan.

Setelah instrumen ditunjukkan valid secara isi, analisis daya pembeda dilakukan untuk setiap soal. Tujuan dari analisis ini adalah untuk menentukan seberapa jauh setiap soal dapat membedakan siswa dengan kemampuan literasi sains rendah dan tinggi. Daya pembeda dihitung dengan melihat korelasi antara jawaban benar dan skor total, yang merupakan teknik statistik yang umum digunakan dalam evaluasi instrumen tes. Menurut Kerlinger (2000), kriteria untuk menilai daya pembeda suatu butir soal dapat diklasifikasikan sebagai berikut: nilai lebih dari 0,40 menunjukkan daya pembeda yang baik; antara 0,30 hingga 0,39 cukup; 0,20 hingga 0,29 kurang; dan di bawah 0,20 menunjukkan daya pembeda yang buruk.

Copyright © 2025 The Author(s)

Tingkat kesukaran suatu soal menunjukkan seberapa besar proporsi peserta tes yang mampu menjawab soal tersebut dengan benar. Soal dikategorikan sebagai sulit jika nilai indeksnya kurang dari 0,30, sedang jika berada di antara 0,31 dan 0,70, dan mudah jika nilai indeksnya lebih dari 0,70. Hasil analisis yang dilakukan terhadap 20 butir soal dalam instrumen tes literasi sains menunjukkan bahwa sebagian besar soal, yaitu 16 soal, atau 80 persen, berada dalam kategori sedang. Soal-soal nomor 1–4, 6–16, dan 19 termasuk dalam kategori sedang, dengan indeks kesukaran antara 0,55 dan 0,69. Soal-soal ini sempurna karena mampu membedakan kemampuan siswa secara proporsional dan memberikan tantangan yang sesuai.

Sementara itu, tiga soal (15%) yang termasuk dalam kategori sulit adalah soal nomor 17, 19, dan 20, yang memiliki indeks kesukaran masing-masing 0,28, 0,29, dan 0,27. Karena hanya sebagian kecil siswa yang mampu menjawab dengan benar, soal-soal ini harus ditinjau lagi untuk memastikan bahwa masalah yang dihadapi siswa tidak disebabkan oleh redaksi yang ambigu atau kerancuan. Satu soal (5 persen dari soal-soal ini), soal nomor 18, termasuk dalam kategori mudah dengan indeks kesukaran 0,76. Untuk memberikan rasa percaya diri dan kenyamanan awal bagi siswa dengan kemampuan rendah, tes harus tetap menggunakan soal yang mudah.

Secara keseluruhan, tingkat kesukaran instrumen ini cukup proporsional dan seimbang. Ini menunjukkan bahwa perangkat telah memenuhi salah satu syarat kualitas soal yang baik dari segi distribusi tingkat kesukaran. Sebagian besar soal berkategori sedang, dengan beberapa soal sulit dan mudah yang digunakan untuk menyebarkan tingkat pengukuran kemampuan siswa.

Pembahasan

Penelitian ini mengembangkan dan memvalidasi alat tes literasi sains yang didasarkan pada kerangka kerja PISA 2015, yang berfokus pada materi yang dipelajari siswa di kelas V sekolah dasar tentang kesehatan pencernaan. Terdapat tiga kategori utama literasi sains: konteks, pengetahuan, dan proses ilmiah. Instrumen yang dibuat terdiri dari dua puluh lima soal. Validasi isi instrumen oleh dua ahli dengan metode Gregory menunjukkan bahwa sebagian besar soal sangat relevan dengan indikator literasi sains yang telah ditetapkan. Tingkat kesepakatan antara para ahli dalam mengevaluasi kelayakan isi instrumen diukur dengan metode Gregory. Tingkat kesesuaian yang tinggi antara penilai menunjukkan bahwa instrumen memiliki validitas isi yang baik (Gregory, 2000). Hasil uji validitas butir dengan korelasi Point Biserial juga menunjukkan bahwa sebagian besar pertanyaan memiliki nilai validitas yang tinggi. Hal ini sejalan dengan Anastasi, A., & Urbina (1997), yang menyatakan bahwa validitas butir yang tinggi menunjukkan seberapa representatif dan konsisten suatu item dapat mengukur konstruk yang dimaksud. Hasil ini mendukung tujuan utama penelitian: membuat alat evaluasi yang memungkinkan siswa menilai penguasaan konten dan kemampuan berpikir ilmiah mereka dalam situasi dunia nyata.

Selain validitas isi dan validitas butir, aspek daya pembeda dan tingkat kesukaran juga menjadi perhatian dalam evaluasi kualitas instrumen. Daya pembeda didefinisikan sebagai kemampuan sebuah soal untuk membedakan siswa yang memiliki kemampuan tinggi dari siswa yang memiliki kemampuan rendah. Hasil analisis menunjukkan bahwa semua butir soal memiliki daya pembeda dalam kategori baik, dengan nilai 0,33 hingga 0,73. Soal-soal dengan daya pembeda yang tinggi, seperti soal nomor 5, 6, dan 18, menunjukkan bahwa soal tersebut sangat efektif dalam menentukan perbedaan tingkat penguasaan siswa. Ini meningkatkan kualitas instrumen sebagai alat ukur yang akurat dan adil.

Tingkat kesukaran soal juga merupakan indikator penting yang menunjukkan sejauh mana peserta dapat memberikan jawaban yang tepat. Hasil penelitian menunjukkan bahwa 80% soal

berada dalam kategori sedang dengan indeks 0,55–0,69. Ini menunjukkan bahwa soal-soal tersebut memiliki tingkat kesulitan yang seimbang dan memiliki kemampuan untuk mengukur kemampuan siswa secara optimal. Soal nomor 17, 19, dan 20 termasuk dalam kategori sukar (15%), dengan indeks kesukaran kurang dari 0,30. Soal nomor 18, dengan indeks kesukaran 0,76, termasuk dalam kategori mudah (5%). Komposisi ini menunjukkan bahwa soal-soal telah disusun dengan memperhatikan distribusi tingkat kesukaran ideal, dengan sebagian besar soal memiliki tingkat kesulitan sedang, dengan beberapa soal yang mudah dan sulit untuk menjamin cakupan pengukuran yang luas terhadap kemampuan siswa.

Hasil ini sejalan dengan penelitian yang dilakukan oleh Anggreni *et al*., (2022) yang mengembangkan alat literasi sains untuk anak usia dini menggunakan pendekatan Research and Development (R&D). Dalam penelitian mereka, mereka menekankan bahwa hasil penilaian keterampilan literasi sains harus benar-benar mencerminkan kemampuan siswa. Instrumen yang mereka buat telah terbukti memiliki tingkat kepercayaan dan validitas yang tinggi setelah melalui tahapan uji ahli dan uji lapangan. Hal ini sejalan dengan temuan penelitian kami yang menunjukkan bahwa alat literasi sains yang dibuat berdasarkan kerangka kerja PISA 2015 untuk siswa kelas V SD juga memiliki kualitas yang baik, dengan hasil validasi ahli yang menunjukkan bahwa sebagian besar soal sangat relevan dan reliabilitas sebesar 0,738 yang tergolong tinggi. Kesesuaian hasil ini menunjukkan bahwa pembuatan alat literasi sains yang sistematis dan berbasis indikator yang jelas sangat penting untuk memastikan pengukuran kompetensi peserta didik secara tepat dan konsisten. Dengan kata lain, dapat disimpulkan bahwa alat yang dikembangkan memiliki kredibilitas yang tinggi dan layak digunakan untuk mengukur kemampuan literasi sains karena respon peserta didik terhadap item-item tes cukup

Penelitian yang dilakukan oleh Putri (2020) juga mendukung bahwa membangun alat literasi sains yang sah dan dapat diandalkan sangat penting untuk mengukur kemampuan siswa. Dalam penelitiannya, Putri membuat

instrumen tes literasi sains tentang keanekaragaman makhluk hidup dengan menggunakan model pengembangan ADDIE. Instrumen tersebut diuji validitas butir dengan membandingkan nilai korelasi butir soal dengan skor total dan reliabilitas dengan menggunakan rumus Kuder Richardson 20 (KR-20). Hasil menunjukkan bahwa 18 item tes adalah valid dan reliabel dengan koefisien r11 sebesar 0,859. Hasil ini sejalan dengan temuan penelitian kami sebelumnya yang membuat instrumen literasi sains tentang sistem pencernaan. Hasilnya menunjukkan bahwa validitas isi menunjukkan bahwa mayoritas bagian soal sangat relevan, dan reliabilitas termasuk dalam kategori tinggi berdasarkan nilai Cronbach's Alpha sebesar 0,738. Kesamaan ini menunjukkan bahwa metode sistematis dalam desain dan pengujian instrumen dapat menghasilkan alat evaluasi yang dapat diandalkan untuk mengevaluasi literasi sains siswa sekolah dasar pada berbagai topik IPAS.

Hasil penelitian ini memiliki dampak signifikan pada pendidikan, terutama dalam hal pembuatan alat evaluasi yang berfokus pada literasi sains di sekolah dasar. Instrument ini dapat digunakan sebagai alat ukur alternatif yang lebih luas yang menilai kompetensi ilmiah kontekstual dan aspek kognitif. Secara teoretis, penelitian ini menegaskan bahwa struktur PISA sangat penting untuk evaluasi pembelajaran yang menekankan keterampilan kontemporer seperti berpikir kritis, pemecahan masalah, dan literasi data. Secara praktis, alat ini dapat digunakan oleh pendidik untuk menilai pemahaman siswa terhadap sains dalam kehidupan sehari-hari.

Meskipun hasil penelitian menunjukkan keberhasilan dalam pengembangan instrumen, terdapat beberapa keterbatasan yang perlu disadari. Pertama, validasi empiris dilakukan pada skala yang terbatas, sehingga tidak mungkin untuk menggunakan hasil untuk generalisasi. Kedua, beberapa bagian soal menunjukkan nilai validitas yang rendah; ini mungkin disebabkan oleh redaksi yang tidak jelas atau ketidaksesuaian dengan pengalaman siswa. Ketiga, karena

Copyright © 2025 The Author(s)

elemen reliabilitas instrumen belum dianalisis dalam penelitian ini, tidak mungkin untuk memastikan bahwa hasil akan konsisten jika digunakan dalam berbagai situasi dan waktu.

Berdasarkan temuan dan keterbatasan penelitian ini, disarankan agar penelitian mengembangkan dan menguji instrumen dengan sampel yang lebih luas dan bervariasi secara geografis. Untuk memastikan bahwa instrumen akan tetap stabil dalam berbagai kondisi, pengujian reliabilitas juga harus dilakukan. Selain itu, perlu dilakukan pengujian reliabilitas untuk memastikan bahwa instrumen stabil dalam berbagai kondisi. Untuk meningkatkan partisipasi siswa, butir soal harus dibuat yang lebih sesuai dengan lingkungan atau berbasis kearifan lokal. Penelitian lanjutan juga dapat mengeksplorasi efektivitas instrumen ini dalam meningkatkan hasil belajar dan literasi sains siswa jika digunakan dalam proses pembelajaran secara langsung.

Kesimpulan

Penelitian ini telah berhasil mengembangkan instrumen tes literasi sains berbasis kerangka kerja PISA 2015 yang difokuskan pada materi sistem pencernaan untuk siswa kelas V sekolah dasar. Instrumen ini dirancang untuk mencakup tiga domain utama literasi sains, yaitu konteks, pengetahuan, dan proses ilmiah. Hasil validasi oleh dua ahli menggunakan teknik Gregory menunjukkan bahwa sebagian besar butir soal tergolong sangat relevan terhadap indikator yang telah ditetapkan. Selain itu, uji validitas empiris melalui analisis Point Biserial menunjukkan bahwa sebagian besar soal memiliki validitas yang memadai untuk mengukur kompetensi literasi sains siswa secara efektif. Secara umum, penelitian ini telah memberikan kontribusi terhadap pengembangan asesmen berbasis literasi sains di tingkat sekolah dasar. Instrumen yang dikembangkan dapat menjadi alat ukur alternatif yang tidak hanya menilai penguasaan konsep, tetapi juga kemampuan berpikir ilmiah siswa dalam konteks kehidupan nyata. Penelitian ini membuka peluang bagi pengembangan instrumen serupa di jenjang lain dan dalam topik pembelajaran yang berbeda, serta menjadi pijakan awal untuk membangun asesmen sains yang lebih kontekstual dan berorientasi pada keterampilan abad ke-21.

References

Adeoye, M. A., Wirawan, K. A. S. I., Made Shania Satya Pradnyani, & Nyoman Intan Septiarini. (2024). Revolutionizing Education: Unleashing the Power of the ADDIE Model for Effective Teaching and Learning. *JPI (Jurnal Pendidikan Indonesia)*, *13*(1), 202–209. https://doi.org/10.23887/jpiundiksha.v13i1.68624

Anastasi, A., & Urbina, S. (1997). *Psychological Testing (7th ed.)*. Prentice Hall.

Anggreni, R. A., Putu Aditya Antara, & Putu Rahayu Ujianti. (2022). Pengembangan Instrumen Literasi Sains Pada Anak. *Jurnal Pendidikan Anak Usia Dini Undiksha*, *10*(2), 291–301. https://doi.org/10.23887/paud.v10i2.49303

Arni, K. J., Ariana, L., & Rahayu, S. (2025). *PENERAPAN MODEL DESAIN PEMBELAJARAN ADDIE DALAM PEMBELAJARAN IPA : Kajian Literatur dan Implementasi APPLICATION OF THE ADDIE LEARNING DESIGN MODEL IN SCIENCE LEARNING : A Literature Review and Implementation*. *2*(November 2024), 173–178.

Azaly, Q. R., & Fitrihidajati, H. (2021). Pengembangan Media Pembelajaran Berbasis Microsoft Office Sway pada Materi Perubahan Lingkungan untuk Melatihkan Kemampuan Literasi Sains Siswa Kelas X SMA. *Berkala Ilmiah Pendidikan Biologi (BioEdu)*, *11*(1), 218–227. https://doi.org/10.26740/bioedu.v11n1.p218-227

Dermawan, D. D., Wardani, S., Kurniawati, Y., & Pranoto, S. (2021). Development Using Higher Order Thinking Skills Assisted by Quizizz Applications in Science Learning. *Proceeding ISET (2021) Universitas Negeri Semarang*, *7*, 89–97.

Ferdiansyah, I. (2023). Pengembangan Instrumen Penilaian Berbasis Hots Pada Mata Pelajaran Ipa Kelas V Sd. *Jurnal Teknologi Pendidikan*, *12*(1), 47–59. https://doi.org/10.32832/tek.pend.v12i1.7760

Gregory, R. J. (2000). *Psychological Testing: History, Principles, and Applications*. Allyn & Bacon.

Hampel, K., Ucieklak-Jeż, P., Bem, A., & Hajduova, Z. (2023). Health Promotion in Schools - A Key to Improving Young People's Health. *Finanse i Prawo Finansowe*, *1*, 89–108. https://doi.org/10.18778/2391-6478.s1.2023.05

Kerlinger, F. N. (2000). *Foundations of Behavioral Research (4th ed.)*. Holt, Rinehart and Winston.

Murti, W. W., & Sunarti, T. (2021). Pengembangan Instrumen Tes Literasi Sains Berbasis Kearifan Lokal Di Trenggalek. *ORBITA: Jurnal Kajian, Inovasi Dan Aplikasi Pendidikan Fisika*, *7*(1), 33. https://doi.org/10.31764/orbita.v7i1.4386

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory (3rd ed.)*. McGraw-Hill.

OECD. (2023). PISA 2022 Results (Volume I): The State of Learning and Equity in Education. In *OECD Publishing* (Vol. 46, Issue 183). https://doi.org/https://doi.org/10.1787/ee4d0a78-en

Putri, R. K. (2020). Pengembangan Instrumen Tes Literasi Sains Siswa Pada Topik Keanekaragaman Makhluk Hidup. *Diklabio: Jurnal Pendidikan Dan Pembelajaran Biologi*, *4*(1), 71–78. https://doi.org/10.33369/diklabio.4.1.71-78

Raharjo, T. J., Rusdarti, R., Subali, B., Suminar, T., Harianingsih, H., & Rahmawati, S. (2023). Pelatihan Penguatan Literasi Sains bagi Guru Sekolah Indonesia-Jeddah, Saudi Arabia. *Journal of Community Empowerment*, *3*(1), 1–6. https://journal.unnes.ac.id/sju/index.php/jce

Riduwan. (2016). *Skala Pengukuran Variabel-variabel Penelitian*. Alfabeta.

Risaad, R., Sarwi, M., Raharjo, T. J., Avrilianda, D., & Subali, B. (2025). *Pengaruh keajaiban*. *10*.

Sevilla, C. G., Ochave, J. A., Punsalan, T. G., Regala, B. P., & Uriarte, G. G. (2007). *Research Methods*. Rex Book Store.

Thoriq, T. (2024). Implementasi Percobaan Kimia Sederhana dalam meningkatkan Literasi Sains Siswa SMPN 1 dan SMAN 1 Sumedang. *Jurnal Abdimas Kartika Wijayakusuma*, *5*(2), 451–465. https://doi.org/10.26874/jakw.v5i2.491

Zulfiana, S., Gunamantha, I. M., & Putrayasa, I. B. (2023). Pengembangan Instrumen Kemampuan Berpikir Tingkat Tinggi Dan Literasi Sains Pada Pembelajaran Ipa Kelas V Sd. *PENDASI: Jurnal Pendidikan Dasar Indonesia*, *7*(1), 13–24.