

ANALISIS KUALITAS INSTRUMEN ASESMEN SOAL SAINS BERBASIS STEM BERORIENTASI PISA PADA SISWA KELAS X SMA NEGERI 1 SEMARAPURA

I Made Subrata^{1*}, I Gusti Ayu Rai², Ni Luh Putu Tejawati³

^{1,2,3} Universitas PGRI Mahadewa Indonesia

madesubrata@mahadewa.ac.id ; gustiayurai64@gmail.com ; tejawatiputu@gmail.com

ABSTRACT

This study aims to analyze the characteristics and quality of a STEM-based science assessment instrument oriented toward the Programme for International Student Assessment (PISA) for tenth-grade students at SMA Negeri 1 Semarang. This study used a mixed methods approach with a sequential explanatory design that integrates qualitative and quantitative analysis. The subjects were tenth-grade students in the 2025/2026 academic year. Qualitative data were collected through observation, interviews, and document studies to assess the suitability of the instrument's characteristics to the PISA scientific literacy framework, which includes aspects of context, content, and competency. Quantitative data were obtained through instrument trials to analyze item quality, including validity, reliability, difficulty level, and discriminatory power. The results showed that the developed instrument met the PISA assessment characteristics by incorporating real-world contexts, integration across STEM disciplines, and higher-order thinking skills. Quantitatively, most items had good validity, high reliability, moderate difficulty level, and adequate to good discriminatory power. However, some items needed revision to improve the instrument's quality. Thus, the developed assessment instrument is suitable for measuring students' scientific literacy and supporting the development of STEM-based assessments oriented towards international standards.

Keywords: science assessment, STEM, scientific literacy, PISA

ABSTRAK

Penelitian ini bertujuan untuk menganalisis karakteristik dan kualitas instrumen asesmen soal sains berbasis STEM berorientasi Programme for International Student Assessment pada siswa kelas X SMA Negeri 1 Semarang. Penelitian ini menggunakan pendekatan *mixed methods* dengan desain *sequential explanatory* yang mengintegrasikan analisis kualitatif dan kuantitatif. Subjek penelitian adalah siswa kelas X tahun pelajaran 2025/2026. Data kualitatif dikumpulkan melalui observasi, wawancara, dan studi dokumen untuk mengkaji kesesuaian karakteristik instrumen dengan kerangka literasi sains PISA, yang meliputi aspek konteks, konten, dan kompetensi. Data kuantitatif diperoleh melalui uji coba instrumen untuk menganalisis kualitas butir soal yang mencakup validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Hasil penelitian menunjukkan bahwa instrumen yang dikembangkan telah memenuhi karakteristik asesmen PISA dengan memuat konteks nyata, integrasi lintas disiplin STEM, serta kemampuan berpikir tingkat tinggi. Secara kuantitatif, sebagian besar butir soal memiliki validitas yang baik, reliabilitas tinggi, tingkat kesukaran kategori sedang, dan daya pembeda yang cukup hingga baik. Meskipun demikian, beberapa butir soal perlu direvisi untuk meningkatkan kualitas instrumen. Dengan demikian, instrumen asesmen yang dikembangkan layak digunakan untuk mengukur literasi sains siswa dan mendukung pengembangan asesmen berbasis STEM yang berorientasi pada standar internasional.

Kata kunci: asesmen sains, STEM, literasi sains, PISA

PENDAHULUAN

Sejak diluncurkan oleh *Organisation for Economic Co-operation and Development* melalui program *Programme for International Student Assessment* pada tahun 2000, penilaian terhadap kemampuan literasi sains siswa telah mengalami pergeseran dari sekadar penguasaan konten menuju kemampuan menerapkan pengetahuan dalam konteks kehidupan nyata. PISA secara berkala dilaksanakan setiap tiga tahun untuk mengukur kemampuan siswa berusia 15 tahun dalam bidang membaca, matematika, dan sains. Dalam konteks sains, PISA tidak hanya menilai penguasaan konsep, tetapi juga kemampuan berpikir ilmiah, penalaran, serta penerapan pengetahuan dalam menyelesaikan permasalahan kontekstual. Oleh karena itu, kerangka asesmen PISA menjadi rujukan penting dalam pengembangan instrumen penilaian yang berorientasi pada literasi sains (Subrata et al., 2026).

Di Indonesia, upaya peningkatan mutu pendidikan dilakukan melalui berbagai kebijakan, salah satunya adalah implementasi Asesmen Nasional yang bertujuan untuk mengevaluasi kualitas pembelajaran secara komprehensif. Asesmen ini menekankan pada

pengukuran kompetensi minimum, khususnya literasi membaca dan literasi sains, yang menjadi fondasi bagi pengembangan kemampuan berpikir kritis siswa. Namun demikian, hasil berbagai studi internasional, termasuk PISA, menunjukkan bahwa kemampuan literasi sains siswa Indonesia masih perlu ditingkatkan, terutama dalam hal penalaran ilmiah dan pemecahan masalah berbasis konteks (Subrata et al., 2026).

Salah satu pendekatan yang dinilai relevan untuk menjawab tantangan tersebut adalah pendekatan pembelajaran berbasis STEM (*Science, Technology, Engineering, and Mathematics*). Pendekatan ini mengintegrasikan berbagai disiplin ilmu untuk membangun pemahaman yang holistik serta mengembangkan keterampilan abad ke-21, seperti berpikir kritis, kreativitas, kolaborasi, dan komunikasi (Subrata et al., 2025). Dalam konteks pembelajaran sains, pendekatan STEM memungkinkan siswa untuk tidak hanya memahami konsep, tetapi juga mengaplikasikannya dalam situasi nyata melalui proses investigasi dan rekayasa. Dengan demikian, penerapan STEM memiliki potensi besar untuk meningkatkan kualitas literasi sains

siswa sesuai dengan tuntutan kerangka PISA.

Meskipun demikian, keberhasilan implementasi pembelajaran berbasis STEM tidak terlepas dari kualitas instrumen asesmen yang digunakan. Instrumen asesmen yang baik harus mampu mengukur capaian pembelajaran secara valid dan reliabel, serta mencerminkan karakteristik literasi sains sebagaimana yang diukur dalam PISA. Instrumen tersebut idealnya memuat konteks dunia nyata, menuntut kemampuan berpikir tingkat tinggi (*higher order thinking skills*), serta mengintegrasikan berbagai konsep lintas disiplin. Namun pada kenyataannya, masih banyak instrumen asesmen sains di sekolah yang cenderung berorientasi pada hafalan konsep dan belum sepenuhnya mengakomodasi karakteristik soal berbasis PISA.

Selain itu, aspek kualitas instrumen juga menjadi perhatian penting dalam pengembangan asesmen. Kualitas instrumen dapat ditinjau dari berbagai parameter psikometrik, seperti validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Validitas menunjukkan sejauh mana instrumen mampu mengukur apa yang seharusnya diukur, sedangkan reliabilitas berkaitan

dengan konsistensi hasil pengukuran. Tingkat kesukaran dan daya pembeda berperan dalam menentukan kemampuan butir soal dalam mengidentifikasi perbedaan kemampuan siswa. Analisis terhadap aspek-aspek ini menjadi krusial untuk memastikan bahwa instrumen asesmen yang dikembangkan benar-benar berkualitas dan layak digunakan dalam proses evaluasi pembelajaran (Subrata et al., 2026).

Di tingkat satuan pendidikan, seperti di SMA Negeri 1 Semarang, pengembangan instrumen asesmen sains berbasis STEM berorientasi PISA masih menghadapi berbagai tantangan. Tantangan tersebut meliputi keterbatasan referensi instrumen yang sesuai dengan kerangka PISA, kurangnya pengalaman guru dalam menyusun soal berbasis literasi sains, serta belum optimalnya analisis kualitas butir soal secara komprehensif. Kondisi ini menunjukkan adanya kebutuhan mendesak untuk mengembangkan dan menganalisis instrumen asesmen yang tidak hanya relevan dengan kurikulum, tetapi juga selaras dengan standar internasional.

Berdasarkan uraian tersebut, penelitian ini menjadi penting untuk dilakukan dengan tujuan menganalisis

karakteristik dan kualitas instrumen asesmen sains berbasis STEM berorientasi PISA. Analisis karakteristik instrumen difokuskan pada kesesuaian dengan kerangka literasi sains PISA, sedangkan analisis kualitas instrumen ditinjau dari aspek validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Hasil penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan instrumen asesmen yang lebih berkualitas serta mendukung peningkatan kemampuan literasi sains siswa, khususnya pada jenjang sekolah menengah atas.

Dengan demikian, penelitian ini tidak hanya relevan secara teoritis dalam pengembangan asesmen pendidikan, tetapi juga memiliki implikasi praktis bagi guru dalam merancang instrumen penilaian yang lebih efektif, kontekstual, dan berorientasi pada standar internasional seperti PISA.

METODE PENELITIAN

Penelitian ini menggunakan desain *mixed methods* dengan model *sequential explanatory*, yaitu mengintegrasikan pendekatan kualitatif dan kuantitatif secara bertahap untuk memperoleh pemahaman yang komprehensif mengenai karakteristik

dan kualitas instrumen asesmen sains berbasis STEM berorientasi *Programme for International Student Assessment*. Pendekatan kualitatif digunakan untuk menganalisis karakteristik instrumen berdasarkan kerangka PISA, sedangkan pendekatan kuantitatif digunakan untuk menguji kualitas instrumen secara empiris.

Subjek penelitian adalah siswa kelas X SMA Negeri 1 Semarang tahun pelajaran 2025/2026. Objek penelitian berupa instrumen asesmen soal sains berbasis STEM yang dikembangkan mengacu pada framework literasi sains PISA. Populasi penelitian adalah seluruh siswa kelas X, sedangkan sampel ditentukan menggunakan teknik *purposive sampling* dengan mempertimbangkan kesesuaian dengan tujuan penelitian.

Pada tahap kualitatif, pengumpulan data dilakukan melalui observasi, studi dokumen, dan wawancara. Observasi dilakukan untuk mengidentifikasi implementasi pembelajaran dan asesmen sains di kelas. Studi dokumen mencakup analisis kurikulum, kisi-kisi soal, serta contoh instrumen yang digunakan. Wawancara dilakukan terhadap guru mata pelajaran sains untuk memperoleh informasi

terkait kebutuhan dan karakteristik asesmen. Analisis data kualitatif dilakukan melalui tahapan reduksi data, penyajian data, dan penarikan kesimpulan, dengan penekanan pada analisis tematik untuk mengidentifikasi kesesuaian instrumen dengan kerangka PISA.

Pada tahap kuantitatif, penelitian diawali dengan penyusunan instrumen asesmen berbasis STEM yang mencakup indikator literasi sains PISA, meliputi konteks, konten, dan kompetensi. Instrumen yang dikembangkan kemudian divalidasi oleh ahli melalui *expert judgment* berdasarkan aspek materi, konstruksi, dan bahasa. Validitas isi dianalisis menggunakan *formula Gregory*. Selanjutnya, instrumen diuji cobakan kepada siswa untuk memperoleh data empiris.

Analisis kuantitatif dilakukan untuk mengukur kualitas instrumen yang meliputi: (1) validitas butir menggunakan korelasi *Pearson Product Moment*, (2) reliabilitas menggunakan koefisien Alpha Cronbach, (3) tingkat kesukaran soal, dan (4) daya pembeda butir soal. Data dianalisis menggunakan statistik deskriptif untuk menentukan kategori kualitas instrumen. Hasil analisis ini digunakan untuk menyeleksi

dan merevisi butir soal agar memenuhi kriteria instrumen yang baik.

Untuk menjamin keabsahan data kualitatif, digunakan teknik triangulasi sumber dan metode, *member check*, serta diskusi dengan rekan sejawat. Sementara itu, keandalan data kuantitatif dijamin melalui prosedur uji validitas dan reliabilitas instrumen.

Tahapan penelitian meliputi: (1) studi pendahuluan dan analisis kebutuhan, (2) pengembangan instrumen, (3) validasi ahli melalui kegiatan Focus Group Discussion (FGD), (4) uji coba terbatas, (5) analisis kualitas instrumen, dan (6) revisi serta finalisasi instrumen asesmen. Dengan tahapan tersebut, diharapkan instrumen yang dihasilkan memiliki karakteristik sesuai standar PISA dan kualitas yang layak digunakan dalam mengukur literasi sains siswa.

HASIL PENELITIAN

Hasil penelitian ini disajikan berdasarkan dua fokus utama, yaitu (1) karakteristik instrumen asesmen soal sains berbasis STEM berorientasi Programme for International Student Assessment, dan (2) kualitas instrumen yang ditinjau dari aspek validitas,

reliabilitas, tingkat kesukaran, dan daya pembeda.

1. Karakteristik Instrumen Asesmen Berbasis STEM Berorientasi PISA

Berdasarkan hasil analisis kualitatif terhadap 160 instrument yang dikembangkan, diperoleh bahwa karakteristik soal telah mengacu pada kerangka literasi sains PISA yang mencakup tiga aspek utama, yaitu konteks, konten, dan kompetensi. Dari aspek konteks, soal dirancang dengan mengangkat permasalahan nyata yang relevan dengan kehidupan sehari-hari siswa, seperti fenomena lingkungan, teknologi sederhana, dan isu 160 instrument. Hal ini menunjukkan bahwa 160 instrument telah memenuhi prinsip kontekstual sebagaimana yang ditekankan dalam PISA.

Pada aspek konten, 160 instrument mencakup materi lintas bidang ilmu sains, yaitu biologi, fisika, dan ilmu bumi dan 160 instrument. Integrasi konten ini mencerminkan pendekatan STEM yang menekankan keterpaduan antar disiplin ilmu. Selain itu, beberapa soal juga mengaitkan konsep sains dengan penerapan teknologi dan proses rekayasa sederhana, sehingga mendorong siswa untuk berpikir secara interdisipliner.

Dari aspek kompetensi, soal yang dikembangkan telah mengakomodasi tiga kompetensi utama literasi sains PISA, yaitu (1) menjelaskan fenomena ilmiah, (2) mengevaluasi dan merancang penyelidikan ilmiah, dan (3) menafsirkan data dan bukti ilmiah. Mayoritas butir soal berada pada level kognitif menengah hingga tinggi (HOTS), yang menuntut kemampuan analisis, evaluasi, dan penalaran. Dengan demikian, secara umum 160 instrument yang dikembangkan telah memiliki karakteristik yang sesuai dengan standar asesmen PISA dan prinsip pembelajaran berbasis STEM.

2. Kualitas Instrumen Asesmen

Analisis kuantitatif dilakukan untuk menguji kualitas instrumen berdasarkan parameter psikometrik, meliputi validitas, reliabilitas, tingkat kesukaran, dan daya pembeda.

a. Validitas Butir Soal

Hasil uji validitas menggunakan korelasi Pearson Product Moment menunjukkan bahwa sebagian besar butir soal memiliki koefisien korelasi dalam kategori valid. Dari seluruh butir yang diuji, lebih dari 80% dinyatakan valid dengan nilai koefisien di atas batas minimum

yang ditetapkan. Namun demikian, terdapat beberapa butir soal yang memiliki nilai korelasi rendah sehingga dinyatakan tidak valid dan perlu direvisi, terutama pada aspek kejelasan stimulus dan kesesuaian indikator.

b. Reliabilitas Instrumen

Uji reliabilitas menggunakan koefisien Alpha Cronbach menunjukkan bahwa instrumen memiliki tingkat reliabilitas yang tinggi. Nilai koefisien reliabilitas berada pada kategori baik, yang mengindikasikan bahwa instrumen memiliki konsistensi internal yang memadai dalam mengukur kemampuan literasi sains siswa. Hal ini menunjukkan bahwa instrumen dapat digunakan secara berulang dengan hasil yang relatif stabil.

c. Tingkat Kesukaran Soal

Analisis tingkat kesukaran menunjukkan bahwa sebagian besar butir soal berada pada kategori sedang, dengan proporsi yang seimbang antara soal mudah, sedang, dan sukar. Distribusi tingkat kesukaran yang proporsional ini menunjukkan bahwa instrumen mampu mengakomodasi variasi kemampuan siswa. Soal dengan

tingkat kesukaran sedang mendominasi, yang sesuai dengan prinsip penyusunan tes yang baik.

d. Daya Pembeda Butir Soal

Hasil analisis daya pembeda menunjukkan bahwa sebagian besar butir soal memiliki daya pembeda dalam kategori cukup hingga baik. Hal ini berarti bahwa instrumen mampu membedakan siswa yang memiliki kemampuan tinggi dan rendah secara efektif. Namun, beberapa butir soal dengan daya pembeda rendah ditemukan pada soal yang memiliki tingkat kesukaran terlalu tinggi atau terlalu rendah, sehingga perlu dilakukan perbaikan.

3. Sintesis Hasil Analisis

Berdasarkan hasil analisis kualitatif dan kuantitatif, dapat disimpulkan bahwa instrumen asesmen soal sains berbasis STEM berorientasi PISA yang dikembangkan telah memenuhi karakteristik utama asesmen literasi sains dan memiliki kualitas yang cukup baik. Meskipun demikian, beberapa butir soal masih memerlukan revisi untuk meningkatkan validitas dan daya pembeda.

Secara keseluruhan, instrumen ini layak digunakan sebagai alat ukur literasi sains siswa kelas X, khususnya dalam konteks pembelajaran berbasis STEM. Hasil penelitian ini juga menunjukkan bahwa pengembangan instrumen yang mengacu pada kerangka PISA dapat menjadi alternatif strategis dalam meningkatkan kualitas asesmen sains di sekolah menengah.

PEMBAHASAN

Hasil penelitian menunjukkan bahwa instrumen asesmen soal sains berbasis STEM berorientasi Programme for International Student Assessment yang dikembangkan telah memenuhi karakteristik utama dalam kerangka literasi sains PISA, yaitu mencakup aspek konteks, konten, dan kompetensi. Temuan ini mengindikasikan bahwa integrasi pendekatan STEM dalam pengembangan instrumen asesmen mampu menghasilkan soal yang tidak hanya mengukur penguasaan konsep, tetapi juga kemampuan berpikir tingkat tinggi siswa dalam menghadapi permasalahan nyata.

Dari segi karakteristik, instrumen yang dikembangkan telah mengadopsi konteks kehidupan sehari-hari sebagai stimulus soal. Hal ini sejalan dengan

prinsip asesmen PISA yang menekankan pada penerapan pengetahuan dalam situasi autentik. Penggunaan konteks nyata, seperti isu lingkungan dan teknologi, terbukti mampu meningkatkan keterlibatan kognitif siswa dalam memahami permasalahan yang diberikan. Selain itu, integrasi lintas disiplin dalam pendekatan STEM memungkinkan siswa untuk menghubungkan konsep biologi, fisika, dan ilmu lainnya secara holistik. Dengan demikian, instrumen tidak hanya bersifat multidisipliner, tetapi juga mendorong terbentuknya pola pikir sistematis dan terintegrasi.

Pada aspek kompetensi, dominasi soal dengan kategori *higher order thinking skills* (HOTS) menunjukkan bahwa instrumen telah dirancang untuk mengembangkan kemampuan analisis, evaluasi, dan penalaran ilmiah siswa. Hal ini penting mengingat salah satu kelemahan utama siswa dalam studi PISA adalah rendahnya kemampuan dalam menafsirkan data dan menyelesaikan masalah berbasis konteks. Oleh karena itu, keberadaan instrumen dengan karakteristik tersebut menjadi langkah strategis dalam meningkatkan kualitas literasi sains siswa.

Ditinjau dari kualitas instrumen, hasil analisis menunjukkan bahwa sebagian besar butir soal telah memenuhi kriteria validitas yang baik. Hal ini menunjukkan bahwa instrumen mampu mengukur konstruk literasi sains yang sesuai dengan tujuan pembelajaran. Namun, masih ditemukannya beberapa butir yang tidak valid mengindikasikan bahwa proses pengembangan instrumen perlu dilakukan secara lebih cermat, terutama dalam merumuskan indikator, stimulus, dan pilihan jawaban. Dengan kata lain, validitas instrumen sangat dipengaruhi oleh kejelasan konstruksi soal dan kesesuaian dengan kompetensi yang diukur.

Selanjutnya, nilai reliabilitas yang tinggi menunjukkan bahwa instrumen memiliki konsistensi internal yang baik. Hal ini mengindikasikan bahwa instrumen dapat digunakan sebagai alat ukur yang stabil dalam mengevaluasi kemampuan siswa. Reliabilitas yang tinggi juga mencerminkan bahwa butir soal memiliki keterkaitan yang baik satu sama lain dalam mengukur konstruk yang sama, yaitu literasi sains berbasis STEM.

Dari aspek tingkat kesukaran, distribusi soal yang didominasi kategori

sedang menunjukkan bahwa instrumen telah dirancang secara proporsional. Soal dengan tingkat kesukaran sedang dinilai ideal karena mampu mengakomodasi berbagai tingkat kemampuan siswa. Namun demikian, keberadaan soal dengan tingkat kesukaran terlalu tinggi atau terlalu rendah perlu diperhatikan karena dapat memengaruhi akurasi pengukuran. Soal yang terlalu mudah tidak mampu membedakan kemampuan siswa, sedangkan soal yang terlalu sulit dapat menurunkan motivasi siswa dalam mengerjakan tes.

Sementara itu, daya pembeda butir soal secara umum berada pada kategori cukup hingga baik, yang berarti instrumen mampu membedakan siswa berdasarkan tingkat kemampuannya. Temuan ini memperkuat bahwa instrumen yang dikembangkan memiliki kualitas yang memadai untuk digunakan dalam evaluasi pembelajaran. Meskipun demikian, butir soal dengan daya pembeda rendah perlu direvisi agar dapat berfungsi secara optimal dalam mengidentifikasi perbedaan kemampuan siswa.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa pengembangan instrumen asesmen berbasis STEM berorientasi PISA

memiliki potensi besar dalam meningkatkan kualitas asesmen sains di sekolah. Instrumen yang dihasilkan tidak hanya relevan dengan tuntutan kurikulum nasional, tetapi juga selaras dengan standar internasional. Hal ini menjadi penting dalam upaya mempersiapkan siswa menghadapi tantangan global yang menuntut kemampuan berpikir kritis, kreatif, dan adaptif terhadap perkembangan ilmu pengetahuan dan teknologi.

Lebih lanjut, penelitian ini menghasilkan suatu model asesmen sains berbasis STEM yang sesuai dengan karakteristik sekolah target PISA di SMA Negeri 1 Semarang. Model ini diharapkan dapat menjadi acuan dalam pengembangan instrumen asesmen yang lebih kontekstual dan aplikatif. Selain itu, penerapan model asesmen ini berpotensi untuk mengidentifikasi kesenjangan (*gap*) dalam pencapaian literasi sains siswa, sehingga dapat menjadi dasar dalam perbaikan strategi pembelajaran.

Implikasi dari penelitian ini tidak hanya terbatas pada pengembangan instrumen, tetapi juga memberikan rekomendasi praktis bagi guru dan pemangku kebijakan pendidikan dalam merancang asesmen yang lebih

berkualitas. Dengan demikian, penelitian ini memberikan kontribusi signifikan dalam mendukung peningkatan mutu pendidikan sains, khususnya dalam mempersiapkan siswa agar lebih siap menghadapi tantangan di era global yang semakin berbasis teknologi.

SIMPULAN

Berdasarkan hasil penelitian dan pembahasan, dapat disimpulkan bahwa instrumen asesmen soal sains berbasis STEM berorientasi *Programme for International Student Assessment* yang dikembangkan telah memenuhi karakteristik utama kerangka literasi sains PISA. Instrumen tersebut dirancang dengan mengintegrasikan aspek konteks, konten, dan kompetensi secara seimbang. Dari aspek konteks, soal disajikan dalam bentuk permasalahan nyata yang relevan dengan kehidupan sehari-hari siswa. Dari aspek konten, instrumen mengakomodasi keterpaduan berbagai disiplin ilmu sains, seperti biologi, fisika, dan ilmu bumi, yang mencerminkan pendekatan STEM. Sementara itu, dari aspek kompetensi, instrumen telah mampu mengukur kemampuan berpikir tingkat tinggi (*higher order thinking skills*), termasuk kemampuan menjelaskan fenomena

ilmiah, mengevaluasi penyelidikan, serta menafsirkan data dan bukti ilmiah.

Ditinjau dari kualitas instrumen, hasil analisis menunjukkan bahwa sebagian besar butir soal telah memenuhi kriteria validitas yang baik, sehingga layak digunakan untuk mengukur kemampuan literasi sains siswa. Uji reliabilitas menunjukkan bahwa instrumen memiliki tingkat konsistensi internal yang tinggi, yang menandakan keandalan dalam penggunaan sebagai alat evaluasi. Selain itu, distribusi tingkat kesukaran soal tergolong proporsional dengan dominasi kategori sedang, sehingga mampu mengakomodasi variasi kemampuan siswa. Daya pembeda butir soal juga berada pada kategori cukup hingga baik, yang menunjukkan bahwa instrumen efektif dalam membedakan siswa berdasarkan tingkat kemampuannya.

Meskipun demikian, masih terdapat beberapa butir soal yang perlu direvisi, terutama pada aspek validitas dan daya pembeda, agar kualitas instrumen menjadi lebih optimal. Secara keseluruhan, instrumen asesmen yang dikembangkan telah memenuhi kriteria sebagai alat ukur yang baik dan layak digunakan dalam pembelajaran sains berbasis STEM.

Dengan demikian, penelitian ini memberikan kontribusi dalam pengembangan instrumen asesmen yang tidak hanya selaras dengan kurikulum nasional, tetapi juga mengacu pada standar internasional PISA. Instrumen ini diharapkan dapat mendukung peningkatan kualitas asesmen sains di sekolah, serta membantu siswa dalam mengembangkan kemampuan literasi sains yang relevan dengan tuntutan abad ke-21.

DAFTAR PUSTAKA

- Friedrich, A., Schreiter, S., Vogel, M., Becker-Genschow, S., Brücken, R., Kuhn, J., & Lehmann, J. (2024). What shapes statistical and data literacy research in K-12 STEM education? *International Journal of STEM Education*, 11(58), 1–20.
- Hewi, L., & Shaleh, M. (2020). Refleksi Hasil PISA (The Programme For International Student Assessment): Upaya Perbaikan Bertumpu Pada Pendidikan Anak Usia Dini. *Jurnal Golden Age*, 4(01), 30–41. <https://doi.org/10.29408/jga.v4i01.2018>
- Liu, X., Hansen, K. Y., & Valcke, M. (2024). A decade of PISA: Student perceived instructional quality and mathematics achievement. *ZDM – Mathematics Education*, 56, 859–891.
- Mashuri, S., & Jahring. (2024). Test instrument development of PISA-based mathematical problem-solving ability. *Scientific Research*

- Journal of Education and Literature*, 4(2), 1–6.
- Prabawa, A. F., Indreswari, H., Setiyowati, A. J., & Mahalle, S. (2024). Enhancing PISA ranking through effective learning methods: A systematic literature review. *Jurnal Pendidikan Progresif*, 14(2), 100–110.
- Pursitasari, I. D., Permanasari, A., & Jaenudin, D. (2022). Pelatihan Penyusunan E-Asesmen Literasi Sains Berbasis AKM Bagi Guru IPA SMP di Kabupaten Bogor. *Jurnal Pemberdayaan Masyarakat*, 1(1), 26–33. <https://doi.org/10.46843/jmp.v1i1.265>
- Siska, & Samsul Maarif. (2023). The Science, Technology, Engineering, Mathematics (STEM)-based Problem-Based Learning (PBL) Learning Model on the Numeracy Skills. *MIMBAR PGSD Undiksha*, 11(2), 209–215. <https://doi.org/10.23887/jjsgsd.v11i2.64836>
- Seeon, U., Masantiah, C., & Tippayakulpairroj, D. (2024). Readiness to implement STEM education in accordance with PISA 2025 framework. *Interdisciplinary Academic and Research Journal*, 4(1), 45–60.
- Subrata, I. Made., Dharmadewi, A. A. I. M., & Lium, H. (2025). Pengaruh Pembelajaran Contextual Teaching And Learning Berbasis Lingkungan Sekolah Terhadap Kreativitas Dan Hasil Belajar Biologi Peserta Didik Kelas XI SMA Negeri 9 DenpasarR. 26(2). <https://doi.org/10.59672/widyadari.v26i2.5569>
- Subrata, I. Made., Fridayanthi, Putu Dessy., Darmada, I. Made. (2026). *Konsep Asesmen Literasi dan Numerasi Berbasis Pendekatan STEM Di Sekolah Menengah Atas*. 15(1).
- Sukendra, I Komang, Widana, I. W., & Juwana, I. D. P. (2023). Senior High School Mathematics E-Module Based on STEM Orienting to Higher Order Thinking Skills Questions. *JPI (Jurnal Pendidikan Indonesia)*, 12(4), 647–657. <https://doi.org/10.23887/jpiundiksha.v12i4.61042>
- Volante, L., & Mattei, P. (2024). The politicization of PISA in evidence-based policy discourses. *Educational Policy Studies*, 22(8), 1–15.
- Yayuk, E., Restian, A., & Ekowati, D. W. (2023). Literasi Numerasi dalam Kerangka Kurikulum Merdeka Berbasis Art Education. *International Journal of Community Service Learning*, 7(2), 228–238. <https://doi.org/10.23887/ijcs.v7i2.56278>
- Yusmar, F., & Fadilah, R. E. (2023). Analisis Rendahnya Literasi Sains Peserta Didik Indonesia: Hasil PISA Dan Faktor Penyebab. *LENSA (Lentera Sains): Jurnal Pendidikan IPA*, 13(1), 11–19. <https://doi.org/10.24929/lensa.v13i1.283>
- Zukhruf, A., & Nurwahidah. (2024). Reflection on Indonesia's PISA scores and teacher competency challenges. *Jurnal Pendidikan IPS*, 15(1), 1–10.