

Pengembangan Instrumen Tes Digital Berbasis Literasi Sains pada Materi Suhu dan Kalor Terintegrasi Kearifan Lokal Batik Bojonegoro

Berliana Nur Aulia Putri^{1*}, Abu Zainuddin²

^{1, 2} Program Studi Pendidikan Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya,
Jl. Ketintang Selatan, Surabaya, Indonesia
berlianap@gmail.com

Abstract

Science literacy skills need to be supported by assessment instruments capable of measuring students' abilities objectively and within meaningful contexts. However, science literacy assessment instruments integrating local wisdom remain limited. This study aimed to develop a digital science literacy test instrument on temperature and heat by integrating the local wisdom of Bojonegoro batik-making and to evaluate its quality and effectiveness. The study employed a Research and Development (R&D) approach using the ADDIE model, consisting of the analysis, design, development, implementation, and evaluation stages. Instrument feasibility was evaluated through logical validity, Percentage of Agreement (PoA), and students' responses, while instrument quality was analyzed using the Rasch Model based on item fit, reliability, separation index, and item difficulty. Instrument effectiveness was examined through students' science literacy profiles using person measures and the Wright Map. The findings indicate that the instrument met the feasibility criteria, demonstrated good measurement quality based on the Rasch Model, and effectively mapped students' science literacy profiles objectively. The developed instrument has the potential to serve as an alternative to traditional digital science literacy assessments for physics learning.

Keywords: Digital Test Instrument, Science Literacy, Rasch Model, Temperature and Heat, Bojonegoro Batik

Abstrak

Kemampuan literasi sains peserta didik perlu didukung oleh instrumen tes yang mampu mengukur kemampuan secara objektif dan kontekstual. Namun, instrumen berbasis literasi sains yang mengintegrasikan kearifan lokal masih terbatas. Penelitian ini bertujuan untuk mengembangkan instrumen tes digital berbasis literasi sains pada materi suhu dan kalor dengan mengintegrasikan kearifan lokal pembuatan batik Bojonegoro serta menganalisis kualitas dan efektivitasnya. Penelitian ini menggunakan metode *Research and Development* (R&D) dengan model ADDIE yang meliputi tahap *analysis*, *design*, *development*, *implementation*, dan *evaluation*. Kelayakan instrumen dianalisis melalui validitas logis, *Percentage of Agreement* (PoA), dan respons peserta didik, sedangkan kualitas instrumen dianalisis menggunakan Model Rasch melalui parameter *item fit*, *reliability*, *separation index*, dan *item difficulty*. Efektivitas instrumen ditinjau berdasarkan profil kemampuan literasi sains peserta didik menggunakan *person measure* dan *Wright Map*. Hasil penelitian menunjukkan bahwa instrumen memenuhi aspek kelayakan, memiliki kualitas pengukuran yang baik berdasarkan analisis Model Rasch, serta mampu memetakan profil kemampuan literasi sains peserta didik secara objektif. Instrumen yang dikembangkan berpotensi menjadi alternatif asesmen literasi sains berbasis digital yang kontekstual pada pembelajaran fisika.

Kata kunci: Instrumen Tes Digital, Literasi Sains, Model Rasch, Suhu dan Kalor, Batik Bojonegoro

Copyright (c) 2026 Berliana Nur Aulia Putri, Abu Zainuddin

✉Corresponding author: Berliana Nur Aulia Putri

Email Address: berlianap@gmail.com (Jl. Ketintang Selatan, Surabaya, Indonesia)

Received 15 July 2026, Accepted 21 July 2026, Published 27 July 2026

PENDAHULUAN

Literasi sains menjadi salah satu kompetensi esensial yang perlu dimiliki oleh peserta didik untuk menghadapi tantangan pembelajaran abad ke-21 (Limiansih et al., 2024). Menurut kerangka kerja PISA (*Programme for International Student Assessment*) 2025, literasi sains tidak hanya berorientasi pada penguasaan konsep-konsep sains, tetapi juga mencakup kompetensi menjelaskan fenomena secara ilmiah, mengevaluasi dan menggunakan informasi ilmiah, serta mengambil keputusan berdasarkan bukti ilmiah dalam berbagai konteks kehidupan (OECD, 2023b). Penguasaan kemampuan literasi sains

berperan penting dalam membentuk peserta didik yang mampu berpikir kritis, memecahkan masalah, dan merespons berbagai isu yang berkaitan dengan sains serta teknologi (Aisyah et al., 2025). Dengan demikian, literasi sains tidak lagi dipandang sebatas penguasaan materi saja, tetapi juga sebagai kemampuan peserta didik dalam menghubungkan pengetahuan ilmiah dengan situasi nyata sehingga mampu mengambil keputusan secara bertanggung jawab. Oleh karena itu, pencapaian kompetensi literasi sains perlu didukung oleh sistem pembelajaran dan penilaian yang mampu mengukur kemampuan literasi sains peserta didik secara komprehensif.

Meskipun literasi sains menjadi salah satu kompetensi esensial yang perlu dikembangkan dalam pembelajaran, kemampuan literasi sains peserta didik Indonesia masih menunjukkan capaian yang relatif rendah berdasarkan hasil penilaian PISA. Pada penilaian PISA 2022, Indonesia memperoleh skor literasi sains sebesar 383. Nilai tersebut masih berada di bawah rata-rata OECD yang mencapai 485 (OECD, 2023a). Kondisi tersebut menunjukkan bahwa penguatan kemampuan literasi sains tidak hanya membutuhkan proses pembelajaran yang efektif, tetapi juga didukung oleh sistem penilaian yang mampu mengukur kemampuan literasi sains peserta didik secara tepat. Sementara itu, penilaian yang diterapkan di sekolah masih didominasi oleh instrumen yang berorientasi pada penguasaan konsep dan penyelesaian soal secara matematis, sehingga belum sepenuhnya mengukur kemampuan literasi sains peserta didik sesuai dengan indikator yang dikembangkan dalam kerangka kerja PISA (Yuliasih & Sarwi, 2021). Oleh karena itu, diperlukan instrumen penilaian yang mampu mengukur kemampuan literasi sains peserta didik secara komprehensif sesuai dengan kompetensi yang ada pada kerangka kerja PISA.

Instrumen penilaian memiliki peran penting dalam proses evaluasi pembelajaran karena menjadi dasar untuk memperoleh informasi yang valid mengenai pencapaian kemampuan peserta didik (Yuniarti & Dwikurnaningsih, 2022). Sejalan dengan hal tersebut, penelitian oleh Wulandari et al. (2024) menyatakan bahwa instrumen yang digunakan dalam pembelajaran perlu disesuaikan dengan tujuan pembelajaran agar mampu mengukur kemampuan peserta didik secara tepat, objektif, dan komprehensif. Dalam konteks literasi sains, instrumen penilaian perlu dirancang berdasarkan kerangka kerja PISA agar selaras dengan kompetensi yang hendak diukur. Selain memperhatikan kesesuaian dengan kompetensi yang diukur, instrumen juga harus memiliki kualitas yang baik sehingga mampu menghasilkan informasi yang akurat, konsisten, dan dapat dipercaya sebagai dasar pengambilan keputusan dalam pembelajaran (Triyanti et al., 2025). Dengan demikian, pengembangan instrumen tes tidak hanya berfokus pada penyusunan butir soal, tetapi juga memastikan bahwa instrumen yang dihasilkan memenuhi karakteristik sebagai alat ukur yang berkualitas untuk mengukur kemampuan literasi sains peserta didik.

Untuk menghasilkan instrumen tes yang berkualitas, pada proses pengembangan instrumen perlu dilakukan penyesuaian antara kompetensi yang diukur, karakteristik materi pembelajaran, serta konteks yang dekat dengan kehidupan peserta didik. Materi suhu dan kalor merupakan salah satu topik fisika yang erat kaitannya dengan berbagai fenomena dalam kehidupan sehari-hari. Meskipun demikian,

materi suhu dan kalor masih sering menimbulkan kesulitan konseptual bagi peserta didik, sehingga diperlukan penyajian konteks yang relevan untuk membantu menghubungkan konsep sains dengan fenomena nyata (Utama et al., 2025). Salah satu konteks yang dapat diintegrasikan adalah kearifan lokal pembuatan batik Bojonegoro karena pada tahapan produksinya terdapat berbagai konsep fisika, khususnya materi suhu dan kalor. Kondisi tersebut dapat dimanfaatkan untuk mendukung stimulus dalam penyusunan instrumen berbasis literasi sains. Selain itu, pada era kemajuan teknologi, penyajian instrumen dalam bentuk digital juga dapat mendukung pelaksanaan asesmen yang lebih fleksibel sekaligus memungkinkan penyajian stimulus kontekstual dalam berbagai bentuk, seperti gambar, tabel data, grafik, maupun video sehingga mendukung pengukuran kemampuan literasi sains peserta didik (Pratiwi et al., 2024). Dengan demikian, pengembangan instrumen tes digital berbasis literasi sains yang mengintegrasikan kearifan lokal pembuatan batik Bojonegoro pada materi suhu dan kalor diharapkan mampu menghasilkan instrumen penilaian yang lebih kontekstual, relevan, dan sesuai dengan kompetensi yang hendak diukur.

Beberapa penelitian terdahulu menunjukkan bahwa pengembangan instrumen tes berbasis literasi sains telah dilakukan pada berbagai materi, konteks pembelajaran, dan media yang berbeda. Wardi & Jauhariyah (2023) mengembangkan instrumen berbasis literasi sains untuk mengukur kemampuan peserta didik pada materi inti atom dan radioaktivitas. Pada materi suhu dan kalor, Muizzah & Jauhariyah (2025) mengembangkan instrumen literasi sains dengan mengintegrasikan konteks Batik Karang khas Tuban, sedangkan Royani & Setyarsih (2022) mengembangkan instrumen tes digital berbantuan Google Form pada materi termodinamika. Beberapa penelitian tersebut menunjukkan bahwa pengembangan instrumen literasi sains telah dilakukan melalui berbagai pendekatan, namun kajian yang mengintegrasikan kearifan lokal pembuatan batik Bojonegoro pada materi suhu dan kalor serta mengevaluasi kualitas instrumen secara komprehensif menggunakan Model Rasch masih sangat terbatas. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan instrumen tes digital berbasis literasi sains pada materi suhu dan kalor yang mengintegrasikan kearifan lokal pembuatan batik Bojonegoro serta menganalisis kualitas instrumen yang dihasilkan berdasarkan parameter Model Rasch sebagai alat ukur kemampuan literasi sains peserta didik.

METODE

Penelitian ini merupakan penelitian dan pengembangan (*Research and Development*) dengan menggunakan model ADDIE (*Analysis, Design, Development, Implementation, dan Evaluation*). Model ADDIE dipilih karena memiliki tahapan pengembangan yang sistematis serta memberikan kesempatan untuk melakukan penyempurnaan produk pada setiap tahap pengembangan sehingga sesuai apabila digunakan dalam pengembangan instrumen tes (Cheung, 2016). Produk yang dikembangkan pada penelitian ini berupa instrumen tes digital berbasis literasi sains pada materi suhu dan kalor dengan mengintegrasikan kearifan lokal pembuatan batik Bojonegoro. Instrumen tes disusun berdasarkan kerangka kerja PISA 2025 dalam bentuk 15 butir soal uraian yang dilengkapi dengan rubrik penskoran.

Penelitian ini melibatkan tiga validator dan 13 peserta didik untuk uji coba terbatas pada tahap *development*, serta melibatkan 45 peserta didik pada tahap implementasi.

Data pada penelitian ini dikumpulkan melalui metode wawancara, metode angket, dan metode tes sesuai dengan tahapan pengembangan instrumen. Wawancara dilakukan kepada guru fisika kelas XI dan pemilik Batik'e Seto Bojonegoro untuk memperoleh informasi mengenai kebutuhan pengembangan instrumen serta proses pembuatan batik Bojonegoro sebagai konteks kearifan lokal yang diintegrasikan dalam instrumen. Selanjutnya, metode angket digunakan untuk memperoleh data validitas isi atau kelayakan instrumen melalui penilaian oleh tiga validator. Selain itu, metode angket juga digunakan pada pelaksanaan uji coba terbatas untuk mengetahui respon peserta didik terhadap instrumen yang dikembangkan. Adapun metode tes digunakan pada tahap implementasi untuk memperoleh data kemampuan literasi sains peserta didik sebagai dasar analisis kualitas instrumen dan profil literasi sains peserta didik.

Data pada proses pengembangan instrumen dianalisis menggunakan pendekatan deskriptif kualitatif dan deskriptif kuantitatif sesuai dengan karakteristik data yang diperoleh. Data hasil wawancara pada tahap analisis serta perancangan instrumen dideskripsikan secara kualitatif sebagai dasar pengembangan instrumen. Selanjutnya, data hasil validasi ahli dianalisis untuk menentukan kelayakan instrumen berdasarkan aspek isi, konstruk, dan bahasa. Sementara itu, tingkat kesepakatan antar validator dianalisis menggunakan *Percentage of Agreement* (PoA). Setelah instrumen direvisi berdasarkan saran masukan ahli, maka dilakukan uji coba terbatas yang kemudian peserta didik diberikan angket respon. Angket tersebut dianalisis secara deskriptif kuantitatif menggunakan persentase serta didukung analisis deskriptif kualitatif terhadap tanggapan peserta didik sebagai bahan evaluasi dan penyempurnaan instrumen tes sebelum diimplementasikan.

Data hasil implementasi dianalisis menggunakan Model Rasch berbantuan perangkat lunak Winstep. Analisis kualitas instrumen dilakukan berdasarkan empat parameter, yaitu *item fit*, *reliability*, *separation index*, dan *item difficulty*. Kesesuaian butir soal ditentukan berdasarkan kriteria *item fit* yang meliputi nilai *Outfit Mean Square* (MNSQ) sebesar 0,50-1,50, *Outfit Z-Standard* (ZSTD) sebesar -2,0 hingga +2m0, dan *Point Measure Correlation* (Pt. Measure Corr.) sebesar 0,40-0,85 (Sumintono & Widhiarso, 2015). *Reliability* dianalisis berdasarkan nilai *Cronbach Alpha* untuk menunjukkan konsistensi internal instrumen, sedangkan *person reliability* dan *item reliability* diinterpretasikan berdasarkan kategori reliabilitas menurut Sumintono & Widhiarso (2015). *Separation Index* dianalisis berdasarkan nilai *person separation* dan *item separation* untuk menentukan kemampuan instrumen dalam membedakan tingkat kemampuan peserta didik serta tingkat kesukaran butir. Selanjutnya, *item difficulty* ditentukan berdasarkan nilai mean logit dan standar deviasi (SD) sehingga butir soal dikelompokkan ke dalam kategori sangat sukar, sukar, mudah, dan sangat mudah menurut Sumintono & Widhiarso (2015). Seluruh hasil analisis tersebut diinterpretasikan secara terpadu untuk menentukan kualitas instrumen tes yang dikembangkan.

Selain menghasilkan informasi mengenai kualitas instrumen, analisis Model Rasch juga

digunakan untuk memperoleh profil kemampuan literasi sains peserta didik. Profil kemampuan literasi sains dianalisis berdasarkan parameter *person measure* dan *Wright Map*. Parameter *person measure* digunakan untuk mengukur kemampuan literasi sains setiap peserta didik dalam satuan logit, yang selanjutnya diinterpretasikan ke dalam kategori sangat tinggi, tinggi, sedang, dan rendah berdasarkan nilai rata-rata (*mean*) dan standar deviasi (SD). Selanjutnya, *Wright Map* digunakan untuk memvisualisasikan distribusi kemampuan peserta didik dan tingkat kesukaran butir soal pada skala logit yang sama sehingga dapat menggambarkan hubungan antara kemampuan peserta didik dengan karakteristik butir soal secara komprehensif. Hasil analisis kedua parameter tersebut digunakan untuk memperoleh gambaran profil kemampuan literasi sains peserta didik sebagai indikator efektivitas instrumen tes yang dikembangkan.

HASIL DAN DISKUSI

Hasil dan diskusi penelitian disajikan berdasarkan tiga fokus utama, yaitu pengembangan instrumen tes digital berbasis literasi sains, kualitas instrumen berdasarkan analisis Model Rasch, serta profil kemampuan literasi sains peserta didik. Setiap temuan disajikan secara ringkas dan diikuti diskusi yang mengaitkan hasil penelitian dengan teori maupun penelitian terdahulu.

Pengembangan Instrumen Tes Digital Berbasis Literasi Sains

Pengembangan instrumen tes digital berbasis literasi sains pada materi suhu dan kalor yang mengintegrasikan kearifan lokal pembuatan batik Bojonegoro dilaksanakan menggunakan model ADDIE yang meliputi tahap *Analysis*, *Design*, *Development*, *Implementation*, dan *Evaluation*. Setiap tahap menghasilkan temuan yang menjadi dasar penyempurnaan instrumen hingga diperoleh produk akhir yang layak digunakan untuk mengukur kemampuan literasi sains peserta didik.

Tahap analisis menunjukkan bahwa materi suhu dan kalor pada Kurikulum Merdeka memiliki keterkaitan dengan kompetensi literasi sains dalam kerangka kerja PISA 2025 sehingga berpotensi dikembangkan menjadi instrumen penilaian berbasis kontekstual. Hasil wawancara dengan guru fisika menunjukkan bahwa instrumen yang digunakan di sekolah belum sepenuhnya berbasis literasi sains karena terdapat keterbatasan waktu untuk mengembangkan instrumen sehingga instrumen di sekolah masih didominasi soal pilihan ganda dan uraian yang berorientasi pada penguasaan konsep dan perhitungan matematis. Selain itu, analisis terhadap proses pembuatan batik Bojonegoro menunjukkan adanya keterkaitan antara tahapan membatik dengan konsep suhu, kalor, perubahan wujud zat, dan perpindahan kalor sehingga dipilih sebagai konteks kearifan lokal dalam pengembangan instrumen. Temuan tersebut menunjukkan adanya kebutuhan pengembangan instrumen yang mampu mengintegrasikan konteks kehidupan nyata dengan kompetensi literasi sains sehingga penilaian tidak hanya berfokus pada penguasaan konsep, tetapi juga kemampuan peserta didik dalam menjelaskan fenomena secara ilmiah, mengevaluasi dan menggunakan informasi ilmiah, serta mengambil keputusan berdasarkan bukti ilmiah sesuai kerangka kerja PISA (OECD, 2023b).

Berdasarkan hasil analisis, disusun rancangan instrumen berupa 15 soal uraian berbasis literasi

sains yang mengacu pada kompetensi dan jenis pengetahuan literasi sains dalam kerangka kerja PISA 2025 serta diintegrasikan dengan konteks kearifan lokal pembuatan batik Bojonegoro. Selain penyusunan kisi-kisi, butir soal, dan rubrik penskoran, pada tahap ini juga dikembangkan angket respon peserta didik yang terdiri atas lima aspek, yaitu ketertarikan terhadap instrumen, kejelasan tampilan dan penyajian soal, kemudahan penggunaan, manfaat instrumen, serta keterlaksanaan instrumen. Selama proses perancangan dilakukan evaluasi formatif melalui telaah dosen pembimbing sehingga diperoleh rancangan awal instrumen yang siap memasuki tahap pengembangan. Karakteristik instrumen tersebut menunjukkan bahwa pengembangan instrumen telah disesuaikan dengan kerangka kerja PISA 2025 sehingga butir soal dirancang untuk mengukur kemampuan literasi sains peserta didik melalui konteks yang autentik dan dekat dengan kehidupan sehari-hari.

Tahap pengembangan meliputi validasi logis oleh tiga validator serta uji coba terbatas kepada 13 peserta didik. Hasil validitas logis dan *Percentage of Agreement* (PoA) disajikan pada Tabel 1.

Tabel 1. Hasil Validitas Logis dan *Percentage of Agreement* (PoA)

Aspek	Persentase Validitas (%)	Kategori	PoA (%)
Isi	93,33	Sangat valid	94,29
Konstruk	88,89	Sangat valid	92,06
Bahasa	93,75	Sangat valid	93,75
Rata-rata	91,99	Sangat valid	93,37

Instrumen memperoleh rata-rata validitas sebesar 91,99% dengan kategori sangat valid, sedangkan nilai PoA sebesar 93,37% menunjukkan bahwa tingkat kesepakatan antarvalidator sangat tinggi. Berdasarkan hasil tersebut, dilakukan penyempurnaan terhadap rubrik penskoran, kisi-kisi, dan redaksi pada beberapa butir soal sebelum dilakukan uji coba terbatas. Hasil ini menunjukkan bahwa instrumen telah memenuhi kelayakan aspek isi, konstruk, dan bahasa sehingga layak digunakan pada tahap berikutnya.

Setelah direvisi, instrumen tes diujicobakan secara terbatas kepada 13 peserta didik untuk memperoleh respon terhadap instrumen yang dikembangkan. Hasil analisis angket respon peserta didik disajikan pada Tabel 2.

Tabel 2. Hasil Respon Peserta Didik pada Uji Coba Terbatas

Aspek	Persentase (%)	Kategori
Ketertarikan terhadap instrumen	80,12	Baik
Kejelasan tampilan dan penyajian soal	77,56	Baik
Kemudahan penggunaan	76,92	Baik
Manfaat instrumen	76,28	Baik
Keterlaksanaan instrumen	71,15	Baik
Rata-rata	76,41	Baik

Sebagaimana ditunjukkan pada Tabel 2, respon peserta didik berada pada kategori baik dengan rata-rata persentase sebesar 76,41%. Aspek ketertarikan memperoleh nilai tertinggi, sedangkan aspek keterlaksanaan memperoleh nilai terendah, tetapi seluruh aspek tetap berada pada kategori baik. Hasil tersebut menunjukkan bahwa instrumen tes berbasis digital yang diujikan mudah dipahami, mudah digunakan, serta memiliki keterlaksanaan yang baik. Selain butir pernyataan pada beberapa aspek di

atas, terdapat beberapa pertanyaan yang menuntut peserta didik untuk memberikan jawaban sesuai pengalaman yang dimiliki. Hasil menunjukkan bahwa terdapat beberapa saran dan masukan dari peserta didik, yang selanjutnya digunakan sebagai dasar penyempurnaan instrumen sebelum diimplementasikan.

Instrumen yang telah direvisi kemudian diimplementasikan kepada 45 peserta didik kelas XI SMA Negeri 1 Gondang Bojonegoro. Hasil pengerjaan pada tahap implementasi tersebut digunakan sebagai dasar evaluasi sumatif terhadap kualitas instrumen dan profil kemampuan literasi sains peserta didik sebagai indikator efektivitas instrumen yang dikembangkan. Proses evaluasi dilakukan secara berkelanjutan melalui evaluasi formatif pada setiap tahap pengembangan dan evaluasi sumatif setelah tahap implementasi dengan menggunakan Model Rasch berbantuan Winstep. Proses tersebut menunjukkan bahwa model ADDIE memungkinkan penyempurnaan instrumen secara sistematis sehingga menghasilkan instrumen yang siap dianalisis kualitasnya berdasarkan Model Rasch.

Kualitas Instrumen Berdasarkan Analisis Model Rasch

Kualitas instrumen tes digital berbasis literasi sains dianalisis menggunakan Model Rasch berbantuan Winstep berdasarkan empat parameter, yaitu *item fit*, *reliability*, *separation index*, dan *item difficulty*. Keempat parameter tersebut memberikan informasi yang saling melengkapi mengenai kesesuaian butir soal terhadap model pengukuran, konsistensi pengukuran, kemampuan instrumen dalam membedakan kemampuan peserta didik, serta variasi tingkat kesukaran butir soal. Oleh karena itu, interpretasi terhadap seluruh parameter digunakan sebagai dasar dalam menentukan kualitas instrumen yang dikembangkan.

Hasil analisis *item fit* menunjukkan bahwa seluruh butir soal memenuhi minimal dua dari tiga indikator yang digunakan, yaitu *Outfit MNSQ*, *Outfit ZSTD* dan *Point Measure Correlation*, sehingga seluruh butir dinyatakan fit dan layak digunakan dalam proses pengukuran kemampuan literasi sains. Hasil analisis item fit disajikan pada Tabel 3.

Tabel 3. Hasil Analisis Item Fit

No.	Butir Soal	Outfit MNSQ	Outfit ZSTD	Pt. Measure Corr.	Keputusan
1.	BS1	0,75	-1,4	0,52	Fit
2.	BS2	0,89	-0,5	0,69	Fit
3.	BS3	0,95	-0,2	0,63	Fit
4.	BS4	1,45	2,1	0,52	Fit
5.	BS5	0,99	0,0	0,55	Fit
6.	BS6	1,06	0,4	0,62	Fit
7.	BS7	1,38	1,8	0,75	Fit
8.	BS8	0,86	-0,6	0,64	Fit
9.	BS9	1,04	0,3	0,56	Fit
10.	BS10	0,68	-1,8	0,66	Fit
11.	BS11	1,23	1,2	0,57	Fit
12.	BS12	1,08	0,5	0,75	Fit
13.	BS13	0,89	-0,5	0,68	Fit
14.	BS14	0,58	-2,3	0,78	Fit
15.	BS15	1,06	0,4	0,62	Fit

Berdasarkan Tabel 3, seluruh butir soal memiliki nilai *Outfit MNSQ* pada rentang 0,58-1,45 dan nilai *Point Measure Correlation* pada rentang 0,52-0,78, sehingga memenuhi kriteria *item fit*. Meskipun terdapat dua butir yang memiliki *Outfit ZSTD* di luar rentang ideal, kedua butir tersebut tetap memenuhi indikator *item fit* lainnya sehingga masih dapat dipertahankan. Temuan tersebut menunjukkan bahwa setiap butir soal telah bekerja sesuai konstruk literasi sains yang diukur dan mampu menghasilkan pengukuran yang konsisten. Hasil ini sejalan dengan Ibrahim (2025) dan Ilhami et al. (2025) yang menyatakan bahwa butir soal yang memenuhi kriteria *item fit* mampu menggambarkan konstruk yang diukur secara konsisten sehingga layak digunakan dalam proses pengukuran.

Selanjutnya, analisis reliabilitas menunjukkan nilai *Cronbach Alpha* sebesar 0,89, *person reliability* sebesar 0,89, dan *item reliability* sebesar 0,90. Nilai tersebut menunjukkan bahwa instrumen memiliki konsistensi internal yang baik, baik dari interaksi antara peserta didik dan butir soal maupun konsistensi peserta didik dalam memberikan respons, serta konsistensi butir soal dalam mengukur konstruk yang sama. Dengan demikian, instrumen yang dikembangkan mampu menghasilkan pengukuran yang stabil dan dapat dipercaya. Temuan ini selaras dengan Bond et al. (2020) yang menyatakan bahwa semakin tinggi nilai reliabilitas, semakin baik tingkat konsistensi instrumen tersebut dalam menghasilkan pengukuran yang objektif.

Analisis *separation index* menunjukkan bahwa nilai *person separation* sebesar 2,78 mampu mengelompokkan peserta didik ke dalam 4 strata kemampuan, sedangkan nilai *item separation* sebesar 2,97 mampu mengelompokkan butir soal ke dalam empat strata tingkat kesukaran. Hasil tersebut menunjukkan bahwa instrumen memiliki kemampuan diskriminasi yang baik dalam membedakan kemampuan peserta didik maupun karakteristik tingkat kesukaran butir soal. Kemampuan tersebut penting karena memungkinkan instrumen memberikan informasi pengukuran yang lebih rinci mengenai variasi kemampuan peserta didik sekaligus memastikan bahwa setiap butir soal memiliki tingkat kesukaran yang beragam. Temuan ini sejalan dengan Sumintono & Widhiarso (2015) yang menyatakan bahwa nilai *separation index* yang tinggi menunjukkan kemampuan instrumen dalam membedakan responden dan butir soal secara lebih objektif sehingga kualitas pengukuran menjadi lebih baik.

Variasi tingkat butir soal dianalisis menggunakan parameter *item difficulty* berdasarkan nilai logit setiap butir soal. Nilai logit masing-masing butir soal disajikan pada Tabel 4.

Tabel 4. Nilai logit

No.	Butir Soal	Measure (Logit)
1.	BS5	1,34
2.	BS8	1,05
3.	BS15	0,67
4.	BS14	0,51
5.	BS12	0,47
6.	BS3	0,15
7.	BS11	0,00
8.	BS9	-0,12
9.	BS6	-0,36
10.	BS10	-0,44

No.	Butir Soal	Measure (Logit)
11.	BS4	-0,48
12.	BS1	-0,52
13.	BS13	-0,56
14.	BS2	-0,76
15.	BS7	-0,97

Berdasarkan nilai logit pada Tabel 4, butir soal dikelompokkan ke dalam kategori sangat mudah, mudah, sukar, dan sangat sukar sesuai batas logit yang digunakan pada analisis Model Rasch. Hasil tersebut menunjukkan bahwa instrumen memiliki distribusi tingkat kesukaran yang bervariasi sehingga tidak didominasi oleh butir soal dengan tingkat kesukaran tertentu. Variasi tersebut mengindikasikan bahwa instrumen mampu mengukur kemampuan peserta didik pada berbagai rentang kemampuan serta mencerminkan penyusunan butir soal yang proporsional berdasarkan kompetensi literasi sains PISA 2025, level kognitif, dan konteks kearifan lokal pembuatan batik Bojonegoro. Temuan ini selaras dengan Qirom & Nurlaelah (2023) yang menyatakan bahwa distribusi tingkat kesukaran yang beragam merupakan salah satu indikator penting kualitas instrumen karena memungkinkan pengukuran kemampuan peserta didik secara lebih akurat pada berbagai tingkatan kemampuan.

Berdasarkan keseluruhan hasil analisis Model Rasch, instrumen tes digital berbasis literasi sains yang dikembangkan memenuhi parameter *item fit*, *reliability*, *separation index*, dan *item difficulty*, sehingga memiliki kualitas pengukuran yang baik dan layak untuk digunakan sebagai alat ukur kemampuan literasi sains peserta didik. Keempat parameter tersebut saling melengkapi dalam menunjukkan bahwa instrumen tidak hanya sesuai dengan Model Rasch, tetapi juga mampu menghasilkan pengukuran yang konsisten, membedakan kemampuan peserta didik secara objektif serta memiliki tingkat kesukaran butir yang proporsional. Kualitas instrumen tersebut menjadi dasar untuk menganalisis profil kemampuan literasi sains peserta didik melalui parameter *person measure* dan visualisasi *Wright Map*.

Profil Kemampuan Literasi Sains Peserta Didik sebagai Indikator Efektivitas Instrumen

Efektivitas instrumen tes digital berbasis literasi sains ditunjukkan melalui kemampuannya dalam menghasilkan informasi mengenai profil kemampuan literasi sains peserta didik. Profil tersebut dianalisis menggunakan analisis menggunakan parameter *person measure* dan divisualisasikan melalui *Wright Map* berdasarkan analisis Model Rasch berbantuan perangkat lunak Winstep. Analisis ini memberikan gambaran mengenai tingkat kemampuan peserta didik sekaligus menunjukkan sejauh mana instrumen mampu menjalankan fungsi pengukurannya.

Hasil analisis *person measure* menunjukkan bahwa nilai rata-rata kemampuan peserta didik sebesar -0,60 logit dengan standar deviasi 1,11 logit. Berdasarkan nilai tersebut, kemampuan literasi sains peserta didik dikelompokkan ke dalam empat kategori, yaitu seperti disajikan pada Tabel 5.

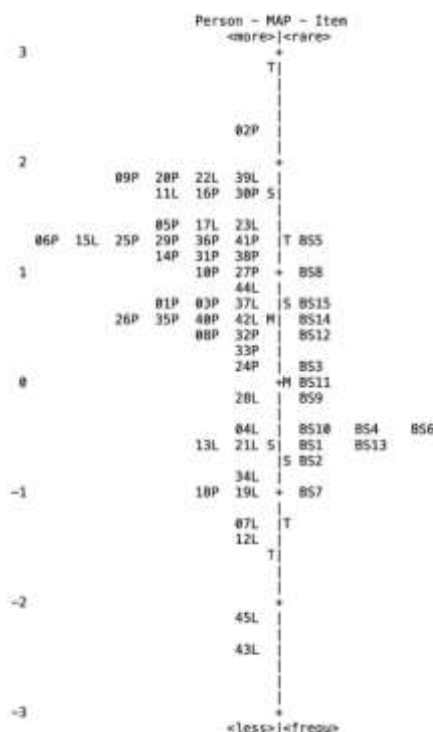
Tabel 5. Kategori Kemampuan Literasi

Kategori	Jumlah Peserta Didik
Sangat tinggi	8
Tinggi	18

Kategori	Jumlah Peserta Didik
Sedang	10
Rendah	9
Total	45

Berdasarkan Tabel 5, sebagian besar peserta didik berada pada kategori tinggi, diikuti kategori sedang, rendah, dan sangat tinggi. Sebaran tersebut menunjukkan bahwa kemampuan literasi sains peserta didik tidak bersifat homogen, melainkan bervariasi pada beberapa tingkat kemampuan. Variasi tersebut mengindikasikan bahwa instrumen mampu membedakan kemampuan literasi sains peserta didik secara objektif serta memberikan informasi yang dapat dimanfaatkan untuk mengidentifikasi kebutuhan pembelajaran sesuai karakteristik masing-masing peserta didik. Temuan ini sejalan dengan Yudiar et al. (2022) yang menyatakan bahwa *person measure* memberikan estimasi kemampuan peserta didik secara lebih objektif dibandingkan skor mentah karena kemampuan diperkirakan berdasarkan model probabilistik.

Hubungan antara kemampuan peserta didik dengan tingkat kesukaran butir soal selanjutnya divisualisasikan melalui Wright Map sebagaimana ditunjukkan pada Gambar 1.



Gambar 1. Wright Map

Berdasarkan Gambar 1, kemampuan peserta didik tersebar pada rentang logit -2,42 hingga 2,23 logit, sedangkan tingkat kesukaran butir soal berada pada rentang -0,97 hingga 1,34 logit. Rata-rata kemampuan peserta didik berada di atas rata-rata tingkat kesukaran butir soal (0,00 logit), yang menunjukkan bahwa secara umum peserta didik mampu menyelesaikan sebagian besar butir soal yang diberikan. Selain itu, sebagian besar peserta didik berada pada rentang kemampuan yang beririsan dengan tingkat kesukaran butir soal sehingga menunjukkan kesesuaian antara karakteristik instrumen

dengan kemampuan peserta didik sebagai sasaran pengukuran. Visualisasi tersebut memperlihatkan bahwa instrumen mampu mengidentifikasi peserta didik dengan kemampuan tinggi maupun rendah sekaligus menyusun hierarki tingkat kesukaran butir soal secara objektif pada satu skala logit. Hasil ini menunjukkan bahwa *Wright Map* tidak hanya memberikan visualisasi distribusi kemampuan peserta didik, tetapi juga memperkuat bahwa instrumen telah berfungsi secara efektif dalam menghasilkan informasi pengukuran yang komprehensif.

Secara keseluruhan, hasil penelitian menunjukkan bahwa instrumen tes digital berbasis literasi sains yang dikembangkan tidak hanya memenuhi aspek pengembangan dan kualitas pengukuran berdasarkan Model Rasch, tetapi juga mampu menghasilkan profil kemampuan literasi sains peserta didik secara objektif melalui analisis *person measure* dan *Wright Map*. Temuan tersebut menunjukkan bahwa instrumen layak digunakan sebagai alternatif asesmen literasi sains pada materi suhu dan kalor yang mengintegrasikan kearifan lokal pembuatan batik Bojonegoro serta memberikan informasi yang bermanfaat sebagai dasar penyusunan tindak lanjut pembelajaran.

KESIMPULAN

Penelitian ini berhasil mengembangkan instrumen tes berbasis literasi sains pada materi suhu dan kalor yang mengintegrasikan kearifan lokal pembuatan batik Bojonegoro menggunakan model ADDIE. Instrumen yang dikembangkan memenuhi aspek kelayakan serta memiliki kualitas pengukuran yang baik berdasarkan analisis Model Rasch, sehingga mampu memetakan profil kemampuan literasi sains peserta didik secara objektif melalui parameter *person measure* dan *Wright Map*. Temuan pada penelitian ini berkontribusi terhadap pengembangan asesmen literasi sains berbasis digital yang mengintegrasikan kearifan lokal sebagai konteks pengukuran yang autentik serta memperkuat pemanfaatan Model Rasch dalam mengevaluasi kualitas instrumen secara komprehensif. Instrumen yang dikembangkan berpotensi digunakan sebagai alternatif asesmen pada pembelajaran fisika, sedangkan penelitian selanjutnya disarankan untuk menguji instrumen pada cakupan sampel yang lebih luas, materi fisika yang berbeda, serta konteks kearifan lokal lainnya untuk memperoleh bukti empiris yang lebih komprehensif mengenai kualitas dan efektivitas instrumen.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada Program Studi S1-Pendidikan Fisika Universitas Negeri Surabaya, validator ahli, guru fisika, serta peserta didik SMA Negeri 1 Gondang Bojonegoro yang telah memberikan dukungan dan berpartisipasi dalam pelaksanaan penelitian ini.

REFERENSI

Aisyah, S. N., Rahayu, N. C. E., & Zainuddin, A. (2025). Increasing scientific literacy in the inheritance and biotechnology chapter by implementing problem-based learning. *Edu Sains Jurnal Pendidikan Sains & Matematika*, 12(2), 172–181. <https://doi.org/10.23971/eds.v12i2.8277>

- Bond, T., Yan, Z., & Heene, M. (2020). *Applying the Rasch Model*. Routledge.
<https://doi.org/10.4324/9780429030499>
- Cheung, L. (2016). Using the ADDIE model of instructional design to teach chest radiograph interpretation. *Journal of Biomedical Education*, 2016, 1–6.
<https://doi.org/10.1155/2016/9502572>
- Ibrahim, N. L. (2025). Analisis literatur penggunaan Model Rasch untuk validasi instrumen asesmen pendidikan : Tren, temuan, dan implikasi. *Lencana: Jurnal Inovasi Ilmu Pendidikan*, 3(2), 46–60. <https://doi.org/10.55606/lencana.v3i2.5012>
- Ilhami, A., Hidayat, T., Riandi, R., & Sriyati, S. (2025). Psychometric validation of a research skills assessment for pre-service science teachers using the Rasch Model. *THABIEA : JOURNAL OF NATURAL SCIENCE TEACHING*, 8(2), 149. <https://doi.org/10.21043/thabiea.v8i2.34359>
- Limiansih, K., Sulistyani, N., & Melissa, M. M. (2024). Persepsi guru SMP terhadap literasi sains dan implikasinya pada pembelajaran sains di sekolah. *Jurnal Pendidikan MIPA*, 14(3), 786–796. <https://doi.org/10.37630/jpm.v14i3.1858>
- Muizzah, N. A., & Jauhariyah, M. N. R. (2025). Development of science literacy test instruments on temperature and heat materials by integrating local wisdom in the production of Karang Batik, a speciality of Tuban Regency. *Jurnal Pijar Mipa*, 20(6), 1072–1079. <https://doi.org/10.29303/jpm.v20i6.9817>
- OECD. (2023a). *PISA 2022 Results (Volume I): The State of Learning And Equity In Education*, PISA, OECD Publishing, Paris,
- OECD. (2023b). *PISA 2025 Science Framework (draft)*. OECD Publishing.
- Pratiwi, I. T., Viyanti, & Permadi, D. (2024). The development of the two-tier diagnostic test instrument with Google Form to measure student misconceptions on energy and energy forms. *Impulse: Journal of Research and Innovation in Physics Education*, 4(1), 22–30. <https://doi.org/10.14421/impulse.2024.41-03>
- Qirom, M. S., & Nurlaelah, E. (2023). Rasch Model: Analysing the items quality of mathematics higher-order thinking skill instrument. *JPP (Jurnal Pendidikan Dan Pembelajaran)*, 30(1), 8. <https://doi.org/10.17977/um047v30i12023p8-16>
- Royani, A., & Setyarsih, W. (2022). Development of google form-based five-tier e-diagnostic test to identify conception levels and track students' misconceptions on thermodynamics materials. *Prisma Sains : Jurnal Pengkajian Ilmu Dan Pembelajaran Matematika Dan IPA IKIP Mataram*, 10(3), 450. <https://doi.org/10.33394/j-ps.v10i3.5209>
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Model Rasch untuk Penelitian Ilmu-ilmu Sosial*. Tim Komunikata Publishing House.
- Triyanti, A., Yuniarti, A., & Afandi, A. (2025). Analisis instrumen tes pengambilan keputusan taksonomi presseisen dan abilitas peserta didik menggunakan model rasch. *BIODIK*, 11(1), 46–58. <https://doi.org/10.22437/biodik.v11i1.36941>

- Utama, R., Toifur, M., Ishafit, & Okimustava. (2025). Investigating senior high school students' conceptual understanding of heat and temperature: A qualitative analysis in physics education. *JPF: Jurnal Pendidikan Fisika FKIP UM Metro*, 13(2), 2442–4838. <https://doi.org/10.24127/jpf.v13i2.14137>
- Wardi, L. Z., & Jauhariyah, M. N. R. (2023). Analisis profil kompetensi literasi sains siswa SMA pada materi inti atom dan radioaktivitas. *Inovasi Pendidikan Fisika*, 12(2), 74–80.
- Wulandari, A., Zuhro, M. A., Al Ghifari, M. A. R., Marzuki, M., & Wilantara, P. (2024). Penggunaan instrumen evaluasi tes dan non tes di SMA Al-Islam Surakarta. *PANDU: Jurnal Pendidikan Anak Dan Pendidikan Umum*, 2(1), 52–57. <https://doi.org/10.59966/pandu.v2i1.1046>
- Yudiar, A., Supriatna, M., & Yudha, E. S. (2022). Grit instrument development on high school students: Rasch Model analysis. *Jurnal Kependidikan: Jurnal Hasil Penelitian Dan Kajian Kepustakaan Di Bidang Pendidikan, Pengajaran Dan Pembelajaran*, 8(2), 258. <https://doi.org/10.33394/jk.v8i2.4794>
- Yuliasih, F., & Sarwi, S. (2021). Instrumen penilaian berbasis keterampilan abad ke-21 untuk mengukur kemampuan literasi sains siswa SMA. *Unnes Physics Education Journal*, (3). <https://doi.org/10.15294/upej.v9i3.45876>
- Yuniarti, S. I., & Dwikurnaningsih, Y. (2022). Online instrument of elementary school teacher's performance assessment. *Journal of Education Technology*, 6(3), 398–409. <https://doi.org/10.23887/jet.v6i3.46155>