

Analisis Kualitas Instrumen Tes Literasi Sains Berbasis Etnosains Suku Dayak di Kalimantan Tengah

Rahmadyah Kusuma Putri ^{1*}, Rahmawati ², Rio Eka Desi Purwandari Hartanti ³, Bintang Sariyatno ⁴, Shanty Savitri ⁵

^{1, 2, 3, 4, 5} Universitas Palangka Raya, Indonesia

* rahmadyahkusumaputri@gmail.com

Abstrak

Permasalahan utama dalam studi ini adalah rendahnya capaian literasi sains suku dayak di Indonesia menunjukkan bahwa soal evaluasi yang ada belum sepenuhnya mengukur kemampuan berpikir ilmiah dan kritis melalui asesmen nasional dan internasional, yang menunjukkan perlunya penguatan kualitas instrumen berbasis etnosains di budaya local. Selama ini, sebagian besar soal literasi sains di sekolah masih mengacu pada model nasional atau internasional seperti PISA, yang belum sepenuhnya bersifat kontekstual. Penelitian ini bertujuan untuk menganalisis kualitas instrumen tes literasi sains berbasis etnosains suku dayak di Kalimantan Tengah sebagai upaya menghadirkan alat ukur yang kontekstual dan relevan dengan budaya lokal. Melalui pendekatan etnosains, konsep ilmiah dapat dikaitkan dengan pengetahuan lokal, sehingga siswa lebih mudah menalar fenomena sains yang dekat dengan lingkungannya. Penelitian ini merupakan penelitian kuantitatif yaitu menguji kualitas instrumen yang dikembangkan. Responden penelitian ini adalah siswa SMA kelas X, XI dan XII dari MAN 1 Pulang Pisau, SMAN 2 Kahayan Tengah dan SMAN 1 Pulang Pisau dengan total 257 siswa. Analisis data yang dilakukan adalah perhitungan uji validitas konstruk melalui *Multidimensional Scaling* (MDS), reliabilitas dengan *Cronbach's Alpha*, serta analisis tingkat kesukaran, daya beda, dan efektivitas distraktor. Hasil analisis menunjukkan nilai stres sebesar 0,00068 (0,68%) dan DAF 0,99 yang menandakan kecocokan model sangat baik. Berdasarkan model konfigurasi MDS diketahui 90% soal valid. Reliabilitas *Cronbach's Alpha* sebesar 0,753 menunjukkan bahwa instrumen memiliki reliabilitas tinggi. Berdasarkan tingkat kesukaran, 65% butir tergolong mudah, 25% sedang, dan 10% sukar. Daya beda menunjukkan 80% butir berkategori sangat baik, 10% baik, dan 10% jelek. Sementara itu, efektivitas distraktor menunjukkan 20% butir efektif dan 80% kurang efektif. Secara keseluruhan, instrumen ini dinyatakan valid dan reliabel, namun masih memerlukan penyempurnaan pada aspek distraktor dan proporsi tingkat kesukaran agar lebih representatif dalam mengukur kemampuan literasi sains siswa berbasis etnosains Kalimantan Tengah.

Kata kunci: Analisis Kualitas. Instrumen Tes, Literasi Sains, Etnosains Suku Dayak, *Multidimensional Scaling*

Pendahuluan

Pembelajaran berbasis etnosains di Indonesia semakin berkembang sebagai strategi untuk menghubungkan pengetahuan sains formal dengan kearifan lokal, sehingga pembelajaran menjadi lebih kontekstual, relevan, dan bermakna bagi peserta didik. Integrasi etnosains dapat diterapkan melalui pengembangan modul, media, dan model pembelajaran yang mengangkat praktik budaya lokal, seperti penggunaan tanaman obat, teknik pertanian tradisional, dan pengelolaan sumber daya alam, yang membantu siswa memahami konsep ilmiah dalam konteks kehidupan sehari-hari mereka (Putra et al, 2025). Kajian empiris menunjukkan bahwa

penerapan etnosains dalam pembelajaran IPA dapat meningkatkan pemahaman konsep, keterampilan proses sains, dan motivasi belajar siswa ketika dikombinasikan dengan model pembelajaran aktif seperti *project-based learning* dan *discovery learning* (Ningtyas et al, 2023).

Etnosains juga berperan penting dalam memperkuat profil Pelajar Pancasila serta menumbuhkan sikap ilmiah yang berpijak pada nilai budaya lokal, meskipun masih terdapat tantangan dalam penyediaan bahan ajar valid dan kesiapan guru dalam mengintegrasikan budaya lokal ke dalam pembelajaran (Ismail et al., 2024). Analisis bibliometrik terbaru menunjukkan tren penelitian etnosains di bidang pendidikan sains mengalami peningkatan signifikan sejak tahun 2020, baik dari segi jumlah publikasi maupun keragaman konteks budaya yang dikaji, yang menunjukkan semakin luasnya penerapan pendekatan ini di berbagai wilayah Indonesia (Rini et al., 2024). Dengan demikian, pembelajaran berbasis etnosains tidak hanya memperkaya makna belajar sains, tetapi juga memperkuat hubungan antara ilmu pengetahuan dan budaya lokal yang menjadi identitas bangsa Indonesia.

Etnosains tidak hanya berfungsi sebagai konteks pembelajaran, tetapi juga sebagai dasar pengukuran berbagai kemampuan ilmiah siswa. Melalui integrasi nilai-nilai budaya lokal dalam sains, etnosains dapat digunakan untuk menilai pemahaman konseptual, yaitu kemampuan siswa mengaitkan konsep ilmiah modern dengan praktik dan pengetahuan tradisional di lingkungannya (Restiani et al., 2023). Etnosains juga mengukur keterampilan proses sains, yang meliputi kemampuan mengamati, mengklasifikasi, menafsirkan data, dan merumuskan kesimpulan berdasarkan fenomena yang ada di masyarakat (Lestari et al., 2020). Siswa yang terbiasa mengaitkan konsep sains dengan praktik budaya lokal cenderung memiliki kesadaran ekologis dan sosial yang lebih tinggi karena mereka memahami bahwa pengetahuan ilmiah berperan dalam menjaga keseimbangan alam dan budaya (Hasibuan, 2023).

Pendekatan etnosains juga dapat mengukur keterampilan berpikir kritis, karena siswa diajak menganalisis kesesuaian antara pengetahuan tradisional dan prinsip ilmiah modern untuk menghasilkan pemahaman yang lebih holistik (Hastuti et al., 2022). Dengan berbagai dimensi kemampuan tersebut, etnosains berpotensi kuat untuk digunakan dalam pengukuran literasi sains. Melalui konteks budaya yang familiar, siswa dapat menunjukkan kemampuan mereka dalam menjelaskan fenomena secara ilmiah, mengevaluasi bukti, dan menerapkan sains dalam kehidupan sehari-hari, sesuai dengan tiga kompetensi utama literasi sains yang dikemukakan oleh OECD.

Berbagai studi menunjukkan bahwa pengembangan instrumen literasi sains berbasis etnosains telah dilakukan di beberapa daerah Indonesia dengan hasil yang beragam. Instrumen literasi sains berbasis etnosains Melayu Riau terbukti valid dan reliabel, namun analisisnya masih terbatas pada uji validitas isi dan reliabilitas internal tanpa pengujian terhadap validitas konstruk maupun fungsi distraktor (Amalia et al., 2024). Instrumen yang dikembangkan dengan konteks Hutan Larangan Kampung Adat Kuta telah melalui uji validitas, reliabilitas, tingkat kesukaran, dan daya beda, tetapi belum menerapkan pendekatan analisis multidimensional untuk memvisualisasikan hubungan antarbutir soal (Muslimin et al, 2024).

Penelitian di daerah Trenggalek dan Lamongan juga menghasilkan instrumen yang valid dan reliabel, namun sebagian butir masih memiliki daya pembeda rendah dan distraktor yang kurang berfungsi optimal (Murti et al, 2021; Maulida et al, 2022). Temuan serupa diperoleh pada pengembangan instrumen di Probolinggo, yang menunjukkan bahwa menjaga keseimbangan antara konteks budaya lokal dan kualitas psikometrik masih menjadi tantangan utama dalam pengembangan instrumen literasi sains berbasis etnosains (Ningsetyo et al, 2024). Berdasarkan hasil tersebut, terdapat gap penelitian berupa keterbatasan analisis

psikometrik yang komprehensif pada instrumen literasi sains berbasis etnosains. Sebagian besar penelitian hanya fokus pada validitas isi dan reliabilitas, tanpa menganalisis validitas konstruk secara multidimensional maupun fungsi distraktor secara detail. Selain itu, belum ditemukan penelitian yang secara spesifik mengembangkan instrumen berbasis etnosains dengan konteks masyarakat Kalimantan Tengah.

Faktanya instrumen tes literasi sains yang digunakan di sekolah, khususnya di Kalimantan Tengah mengacu pada standar internasional seperti *Programme for International Student Assessment* (PISA) yang dimodifikasi sesuai dengan materi tertentu (Yessi, 2019; Febryana et al., 2021). Instrumen tersebut memiliki keunggulan dari segi struktur dan indikator pengukuran, namun sering kali masih bersifat umum dan kurang memperhatikan konteks lokal. Oleh karena itu telah dikembangkan instrumen tes literasi sains berbasis etnosains Kalimantan Tengah (Muslimin et al, 2024). Namun kualitas instrumen tersebut belum diuji. Instrumen ini terdiri atas 20 butir soal pilihan ganda yang masing-masing mencerminkan indikator kemampuan literasi sains, seperti mengidentifikasi masalah ilmiah, menggunakan bukti ilmiah, dan menjelaskan fenomena secara ilmiah. Setiap butir disusun menggunakan skala penilaian obyektif dengan satu jawaban benar dan empat opsi pengecoh (*distractor*).

Instrumen tes literasi sains berbasis etnosains Kalimantan Tengah dikembangkan sebagai alat ukur untuk menilai kemampuan siswa dalam memahami konsep sains secara kontekstual dengan mengaitkan pengetahuan ilmiah modern dan pengetahuan tradisional masyarakat lokal. Dalam pengembangannya, konten dan konteks soal diambil dari berbagai bentuk etnosains masyarakat Kalimantan Tengah, seperti pengetahuan lokal tentang pemanfaatan tumbuhan obat oleh masyarakat dayak, sistem pertanian di lahan gambut, pengelolaan hutan secara tradisional, serta praktik konservasi sumber daya alam berbasis kearifan lokal. Setiap butir soal dirancang untuk menggambarkan situasi nyata yang sering ditemui masyarakat setempat, sehingga siswa dapat menalar fenomena sains melalui pendekatan yang dekat dengan kehidupan mereka. Melalui instrumen ini, diharapkan pengukuran literasi sains menjadi lebih kontekstual, relevan, dan mendukung pembelajaran yang berakar pada nilai-nilai budaya serta lingkungan lokal.

Pengujian kualitas instrumen dilakukan untuk memastikan bahwa setiap butir soal memiliki karakteristik yang sesuai dengan kriteria instrumen tes yang baik. Aspek yang dianalisis meliputi validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas distraktor. Kebaruan (*novelty*) penelitian ini terletak pada penerapan analisis *Multidimensional Scaling* (MDS) untuk memetakan validitas konstruk butir soal, yang belum digunakan dalam penelitian sejenis sebelumnya, serta penggunaan konteks etnosains di masyarakat Kalimantan Tengah sebagai dasar pengembangan instrumen literasi sains. Penelitian ini diharapkan dapat memberikan kontribusi teoretis dan praktis dalam pengembangan instrumen berbasis budaya lokal yang lebih valid, reliabel, dan kontekstual terhadap karakteristik pembelajaran sains di Indonesia.

Metode

Penelitian ini menggunakan pendekatan kuantitatif dengan jenis penelitian deskriptif analisis yang bertujuan untuk menguji kualitas instrumen tes literasi sains berbasis etnosains yang telah disusun. Instrumen yang diuji berupa 20 butir soal pilihan ganda dengan lima alternatif jawaban (a,b,c,d,e). Jawaban benar diberi skor 1, dan jawaban salah diberi skor 0. Uji coba instrumen dilaksanakan di SMAN 2 Kahayan Tengah, SMAN 2 Pulang Pisau, dan MAN 1 Pulang Pisau dengan jumlah responden 257 siswa yang terdiri atas siswa kelas X, XI, dan XII.

Data hasil tes dianalisis untuk menentukan validitas isi, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas distraktor.

Uji validitas dianalisis dengan *Multidimensional Scaling* (MDS) merupakan sebuah teknik analisis statistik multivariat yang digunakan untuk menggambarkan struktur hubungan antar objek data berdasarkan kemiripannya secara grafis dalam suatu bidang multidimensi, untuk mendapatkan informasi dari data, melalui program SPSS *proxscal* untuk melihat pola kedekatan antar butir soal. Pemetaan tersebut membantu menentukan butir-butir yang konsisten maupun yang menyimpang dari struktur konseptual, sehingga MDS berperan penting dalam menilai validitas suatu instrumen (Zulham et al, 2024). Hasil analisis MDS dinilai berdasarkan indikator stress dan DAF. Nilai stress digunakan untuk melihat seberapa besar perbedaan data yang belum terwakili oleh model. Semakin kecil nilainya, semakin akurat dan baik hasil pemetaan MDS yang diperoleh (Tabel 1).

Tabel 1. Interpretasi Nilai Stress

100 x Stress (Percent)	Goodness of fit
Diatas 20%	Jelek
10% - 19,9%	Cukup
5% - 9,9%	Baik
2,5% - 4,9%	Sangat Baik
0% - 2.4%	Sempurna

Uji realibilitas dianalisis dengan *Cronbach's Alpha* dengan bantuan SPSS. Uji ini bertujuan melihat konsistensi internal antarbutir soal. Nilai *alpha* menentukan kriteria realibilitas instrumen (Tabel 2) (Taherdoost, 2016)

Tabel 2. Interpretasi Nilai Alpha

Nilai Alpha (α)	Kriteria
>0,90	Sangat reliabel
0,70 – 0,90	Reliabel tinggi
0,50 – 0,70	Cukup reliabel
< 0,50	Tidak reliabel

Uji tingkat kesukaran soal dihitung menggunakan rumus indeks kesukaran (P) (Elgadal et al, 2021). Uji tingkat kesukaran soal dilakukan untuk mengkategorikan soal sukar, sedang dan mudah (Tabel 3).

Tabel 3. Interpretasi Indeks Kesukaran

Indeks Kesukaran	Kriteria
< 0,30	Sukar
0,30 – 0,70	Sedang
> 0,70	Mudah

Uji daya beda soal dilakukan dengan rumus daya beda (D) (Quaigrain et al, 2017). Uji daya beda soal dilakukan untuk mengelompokkan siswa dengan prestasi belajar tinggi dengan siswa dengan prestasi belajar rendah (Tabel 4).

Tabel 4. Interpretasi Daya Beda

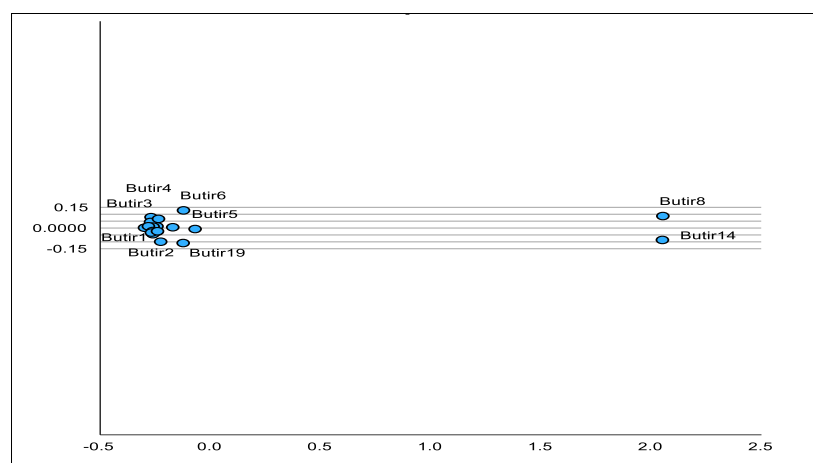
Daya Beda	Kriteria
$\geq 0,40$	Baik Sekali
0,30 – 0,39	Baik
0,20 – 0,29	Cukup
< 0,19	Jelek

Uji efektivitas distraktor ditentukan berdasarkan persentase siswa yang menjawab pilihan pengecoh, yaitu minimal sebanyak 5% siswa. Apabila persentase siswa yang menjawab pilihan pengecoh kurang dari 5%, maka pengecoh dikatakan tidak efektif. Sebaliknya apabila lebih dari 5%, maka pengecoh berfungsi dengan efektif (Elgadal et al, 2021).

Hasil

Validitas

Hasil uji validitas menggunakan MDS STRESS (*Standardized Residual Sum of Square*) merupakan kriteria atau pedoman yang digunakan untuk menguji validitas dari MDS. Berdasarkan hal tersebut dapat dilihat *stress* 0.00068 (0,68%) dan DAF 0,99. Nilai *stress* ini termasuk dalam kriteria *good of fit* yang sempurna dan DAF yang sangat baik, artinya konfigurasi MDS yang dihasilkan sangat akurat dan representatif terhadap data asli (Gambar 1).



Gambar 1. Konfigurasi MDS Instrumen

Gambar 1 menunjukkan konfigurasi MDS butir soal umumnya mengelompok di satu sisi yang sama (berdekatan), sementara butir nomor 8 dan nomor 14 berada di sisi yang berlawanan (jauh). Konfigurasi MDS menunjukkan butir-butir (atau subtest) yang memiliki korelasi tinggi muncul dalam posisi yang berdekatan dalam peta MDS, mencerminkan kesamaan konstruk atau domain yang diukur (Meyer et al, 2022). Sebaliknya, butir yang posisinya menjauh atau tidak sejajar dengan kelompoknya menunjukkan perbedaan makna konseptual, sehingga dapat diinterpretasikan sebagai butir yang tidak valid secara konstruk (Zulham et al, 2024). Butir 8 dan 14 menunjukkan posisi menyimpang dari grup butir soal lainnya, sehingga disimpulkan bahwa kedua butir ini tidak mampu secara konsisten mengukur kemampuan literasi sains berbasis etnosains yang diharapkan atau tidak valid secara konstruk. Dengan demikian, delapan belas soal dari instrumen yang diuji valid, artinya soal tersebut telah merepresentasikan isi atau domain konsep yang hendak diukur secara tepat dan relevan dengan indikator yang telah ditetapkan (Ihsan, 2015)

Realibilitas

Hasil uji realibilitas dengan analisis *Cronbach's Alpha* diperoleh nilai *Cronbach's Alpha* sebesar 0,753 (Tabel 5).

Tabel 5. Hasil Uji Realibilitas

<i>Cronbach's Alpha</i>	<i>N of Items</i>
.753	20

Tabel 5 menyajikan hasil uji realibilitas. Nilai ini berada dalam rentang 0,70 – 0,79, artinya termasuk dalam kriteria reliabel. Dengan demikian disimpulkan instrumen tes memiliki konsistensi internal yang baik, sehingga dapat dinyatakan reliabel dan layak digunakan dalam pengumpulan data.

Tingkat Kesukaran

Hasil uji tingkat kesukaran, diketahui bahwa instrumen terdiri atas soal mudah, sedang dan sukar (Tabel 6).

Tabel 6. Hasil Uji Tingkat Kesukaran

Kriteria Kesukaran	Jumlah Butir	Nomor Butir
Mudah	13	1,2,3,4,7,9,10,11,12,13,15,17,18
Sedang	5	5,6,16,19,20
Sukar	2	8,14

Tabel 6 menyajikan hasil analisis tingkat kesukaran terhadap dua puluh butir soal yang diuji. Berdasarkan hasil perhitungan, terdapat tiga belas butir soal (65%) yang tergolong mudah, yaitu butir nomor 1, 2, 3, 4, 7, 9, 10, 11, 12, 13, 15, 17, dan 18. Sebanyak lima butir soal (25%) termasuk dalam kategori sedang, yaitu butir nomor 5, 6, 16, 19, dan 20. Sementara itu, dua butir soal (10%) berada pada kategori sukar, yaitu butir nomor 8 dan 14. Hasil ini menunjukkan bahwa sebagian besar butir soal berada pada kategori mudah, sehingga secara keseluruhan tingkat kesukaran tes cenderung rendah. Kondisi ini menunjukkan bahwa instrumen tes yang dikembangkan belum memiliki proporsi tingkat kesukaran yang ideal. Proporsi ideal instrumen soal terdiri dari 25% soal sukar, 50% soal sedang, dan 25% soal mudah (Arifin, 2007).

Daya Beda

Hasil uji daya beda diketahui instrumen terdiri dari butir yang memiliki kriteria baik sekali, baik, dan jelek (Tabel 7).

Tabel 7. Hasil Uji Daya Beda

Kriteria Daya beda	Jumlah Butir	Nomor Butir
Baik Sekali	16	1,2,4,6,7,9,10,11,12,13,15,16,17,18,19,20
Baik	2	3, 5
Jelek	2	8,14

Tabel 7 menunjukkan hasil analisis daya beda terhadap dua puluh butir soal. Berdasarkan hasil perhitungan, terdapat enam belas butir soal (80%) yang memiliki daya beda baik sekali, yaitu butir nomor 1, 2, 4, 6, 7, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, dan 20. Selanjutnya, dua butir soal (10%) tergolong baik, yaitu butir nomor 3 dan 5. Sementara itu, dua butir soal (10%) lainnya memiliki daya beda jelek, yakni butir nomor 8 dan 14. Hasil ini menunjukkan bahwa sebagian besar butir soal memiliki daya beda baik sekali dan baik. Daya beda yang baik menunjukkan bahwa siswa berkemampuan tinggi cenderung menjawab benar, sedangkan siswa berkemampuan rendah cenderung salah (Arikunto, 2014; Sudijono, 2017).

Efektivitas Distraktor

Hasil analisis menunjukkan bahwa butir soal memiliki distraktor yang efektif dan tidak efektif (Tabel 8).

Tabel 8. Hasil Uji Efektivitas Distraktor

Butir	Frekuensi Pilihan Jawaban					Kesimpulan Fungsi Distraktor
	A	B	C	D	E	
1	9	7	81*	2	1	Tidak efektif
2	4	6	87*	2	1	Tidak efektif
3	78	5	7	5	4	Efektif
4	14	77*	2	5	2	Tidak efektif
5	12	16	9	61*	2	Tidak efektif
6	19	6	10	5	59*	Efektif
7	77*	4	11	2	7	Tidak efektif
8	7*	4	82	3	3	Tidak efektif
9	10	4	76*	4	5	Tidak efektif
10	11	76*	4	6	3	Tidak efektif
11	5	81*	7	1	6	Tidak efektif
12	91*	4	3	3	0	Tidak efektif
13	5	9	81*	2	2	Tidak efektif
14	3	7	4*	2	82	Tidak efektif
15	4	7	79*	7	2	Tidak efektif
16	17	63*	13	5	2	Tidak efektif
17	79*	6	7	5	2	Tidak efektif
18	77*	4	8	2	9	Tidak efektif
19	14	9	11	57*	7	Efektif
20	14	69*	6	6	5	Efektif

Tabel 8 menampilkan hasil analisis efektivitas distraktor (pengecoh) pada dua puluh butir soal pilihan ganda. Berdasarkan hasil analisis, diketahui bahwa hanya empat butir soal (20%) yang memiliki distraktor efektif, yaitu butir nomor 3, 6, 19, dan 20. Sementara itu, enam belas butir soal (80%) lainnya dinyatakan tidak efektif, karena satu atau lebih pengecohnya tidak berfungsi sebagaimana mestinya. Distraktor yang efektif harus bersifat plausible, yaitu tampak masuk akal bagi peserta yang belum memahami konsep dengan benar. Sementara itu, distraktor yang terlalu jelas salah atau tidak relevan akan cenderung diabaikan, sehingga gagal berfungsi sebagai distraktor yang mampu membedakan peserta dengan tingkat pemahaman yang berbeda (Raina et al., 2023).

Pembahasan

Hasil analisis menunjukkan bahwa instrumen tes literasi sains berbasis etnosains Kalimantan Tengah memiliki kualitas yang cukup baik berdasarkan lima aspek pengujian, yaitu validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas distraktor. Uji validitas konstruk menggunakan *Multidimensional Scaling* menghasilkan nilai *stress* sebesar 0,068% dan DAF 0,99, yang menandakan kesesuaian model sangat baik. Model menunjukkan bahwa struktur hubungan antarbutir sangat baik dan sesuai dengan konstruk literasi sains yang diukur. Hampir seluruh butir soal dinyatakan valid secara konstruk, kecuali butir nomor 8 dan 14 yang menunjukkan penyimpangan posisi dari kelompok butir lainnya. Hasil ini menunjukkan bahwa sebagian besar butir telah merepresentasikan dimensi literasi sains berbasis etnosains Kalimantan Tengah secara konsisten dan bermakna. Validitas menjadi aspek penting karena menunjukkan sejauh mana instrumen benar-benar mengukur apa yang seharusnya diukur, sehingga hasil pengukuran dapat dipercaya untuk merepresentasikan konstruk yang diteliti (Ihsan, 2015). Validitas yang baik juga menjadi dasar bagi penyusunan dan interpretasi hasil tes agar sesuai dengan tujuan pengukuran (Taherdoost, 2016).

Selain validitas, reliabilitas menjadi faktor penentu kualitas instrumen. Hasil uji reliabilitas menunjukkan nilai *Cronbach's Alpha* sebesar 0,753, yang termasuk kategori reliabel.

Reliabilitas yang tinggi mengindikasikan konsistensi internal antarbutir soal dan kestabilan hasil pengukuran ketika digunakan dalam kondisi berbeda (Quaigrain et al, 2017). Hal ini menunjukkan bahwa instrumen yang dikembangkan memiliki tingkat keandalan yang memadai untuk digunakan dalam konteks pengukuran kemampuan literasi sains berbasis etnosains. Dengan kata lain, siswa dengan kemampuan literasi sains serupa akan cenderung memperoleh hasil yang konsisten ketika mengikuti tes ini.

Berdasarkan daya beda, tingkat kesukaran dan efektivitas distraktor kualitas soal dapat dikategorikan menjadi empat kategori, yaitu: 1) Baik, 2) Revisi, 3) Cukup, 4) Tidak Baik (Adawiah, 2020). Soal dinyatakan baik apabila memiliki daya pembeda baik atau cukup baik, tingkat kesukaran sedang, dan seluruh distraktornya berfungsi efektif. Soal perlu direvisi apabila daya pembeda dan tingkat kesukarannya memadai, namun terdapat distraktor yang tidak berfungsi dengan baik. Soal dikategorikan cukup baik apabila daya pembedanya baik namun tingkat kesukarannya terlalu mudah atau terlalu sukar. Soal tidak baik, yaitu daya pembeda jelek.

Kategori soal dinyatakan baik apabila memiliki daya pembeda baik atau cukup baik, tingkat kesukaran sedang, dan seluruh distraktornya berfungsi efektif baik. Soal baik pada instrumen ini adalah butir nomor 6, 19, dan 20. Ketiga butir ini memiliki daya pembeda baik, tingkat kesukaran sedang, serta distraktor yang berfungsi efektif. Dengan demikian, ketiga butir ini dapat dikategorikan sebagai soal yang baik dan layak digunakan tanpa revisi. Salah satu contoh cuplikan soal kategori baik dapat dilihat berikut ini.

19. Perhatikan tabel berikut yang menunjukkan persentase populasi hewan endemik Kalimantan Tengah yang diamati di kawasan konservasi selama satu tahun terakhir.

Tabel 9. Persentase Populasi Hewan Endemik Kalimantan

Hewan Endemik	Persentase Populasi
Orangutan Kalimantan	35%
Owa-owa Kalimantan	25%
Bekantan	20%
Burung enggang (Rangkong)	10%
Lainnya	10%

Dari tabel di atas, seorang siswa diminta mengevaluasi kondisi populasi hewan endemik tersebut dan menyusun rekomendasi penelitian. Evaluasi manakah yang paling tepat berdasarkan data tersebut?

- Populasi orangutan yang paling besar menunjukkan bahwa tidak ada ancaman terhadap kelangsungan hidup spesies tersebut.
- Burung enggang yang memiliki populasi terkecil sebaiknya dikeluarkan dari daftar hewan yang dipantau karena populasinya tidak signifikan.
- Owa-owa dan bekantan yang menempati urutan kedua dan ketiga menunjukkan bahwa ekosistem hutan Kalimantan Tengah masih sangat stabil.
- Proporsi populasi orangutan yang dominan menunjukkan kebutuhan untuk fokus konservasi pada spesies lain yang lebih sedikit populasinya.
- Karena hewan kategori "lainnya" hanya 10%, maka tidak perlu dimasukkan ke dalam kebijakan pelestarian jangka panjang.

Butir nomor 19 tersebut termasuk dalam tingkat kesukaran sedang, ini menunjukkan bahwa sebagian besar siswa (57%) dapat menjawab soal dengan benar. Butir ini juga mampu membedakan siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah. Siswa berkemampuan rendah umumnya akan memilih distraktor, ini dibuktikan dengan empat pilihan jawaban lainnya dipilih oleh siswa dengan persentase lebih dari 5%. Selanjutnya kategori soal revisi apabila daya pembeda dan tingkat kesukarannya memadai, namun terdapat

distraktor yang tidak berfungsi dengan baik. Soal revisi adalah butir nomor 5 dan 16. Revisi dapat dilakukan dengan memperbaiki pilihan pengecoh agar lebih menarik bagi peserta dengan kemampuan rendah. Salah satu contoh cuplikan soal kategori revisi dapat dilihat berikut ini.

5. *Dalam berbagai upacara adat suku Dayak di Kalimantan Tengah, digunakan berbagai jenis daun-daunan yang dianggap memiliki kekuatan simbolik maupun fungsional. Misalnya, dalam ritual penyembuhan tradisional, daun sangkuang atau daun sirih hutan digunakan untuk "menarik penyakit" dari tubuh seseorang. Daun tersebut biasanya direndam dalam air atau diusapkan langsung ke kulit sambil diiringi mantra. Bagaimana kamu menjelaskan secara ilmiah mengapa daun-daunan tersebut bisa membantu proses penyembuhan dalam praktik ritual?*

- a) *Karena mantra yang dibacakan membuat tubuh lebih siap menerima penyembuhan secara spiritual.*
- b) *Karena semua tanaman yang digunakan dalam budaya pasti memiliki efek farmakologis kuat.*
- c) *Karena penggunaan daun dalam budaya Dayak sudah turun-temurun dan tidak perlu dibuktikan secara ilmiah.*
- d) *Karena daun-daunan tersebut mengandung senyawa aktif yang secara biologis dapat meredakan gejala penyakit.*
- e) *Karena tubuh manusia bisa sembuh sendiri selama ritual dilakukan*

Butir nomor 5 tersebut sudah baik pada daya beda dan tingkat kesukaran sedang, namun perlu direvisi pada distraktor, ini dibuktikan dengan pilihan jawaban E dipilih oleh siswa dengan persentase 2%, sementara pilihan jawaban lainnya sudah memenuhi persentase minimal untuk efektivitas distraktor yaitu minimal 5%. Jika dianalisis narasi pilihan E kontras dari segi keilmiahan dibandingkan pilihan jawaban lainnya. Oleh karena itu, pilihan E kurang menarik bagi siswa, sehingga pilihan jawaban ini perlu direvisi.

Butir yang dikategorikan cukup baik adalah soal yang memiliki daya pembeda baik atau cukup baik, namun tingkat kesukarannya terlalu mudah atau terlalu sukar. Berdasarkan hasil, sebagian besar soal, yaitu 13 butir (nomor 1, 2, 3, 4, 7, 9, 10, 11, 12, 13, 15, 17, dan 18) termasuk dalam kategori ini. Walaupun memiliki daya pembeda baik, sebagian besar butir tergolong mudah dan memiliki distraktor yang tidak berfungsi, sehingga tes cenderung kurang menantang bagi peserta dengan kemampuan tinggi. Salah satu contoh cuplikan soal kategori revisi dapat dilihat berikut ini:

4. *Hutan tropis Kalimantan Tengah merupakan rumah bagi berbagai jenis pohon bernilai ekonomi dan budaya, seperti ulin, meranti, tengkawang, dan bangkirai. Masyarakat Dayak memiliki pengetahuan turun-temurun tentang pemanfaatan dan pelestarian pohon-pohon ini. Misalnya, pohon tengkawang hanya diambil buahnya saat musim tertentu dan tidak ditebang karena diyakini sebagai "pohon pemberi kehidupan". Namun, banyak pohon tengkawang dan meranti kini mulai jarang ditemukan di sekitar desa karena penebangan liar dan perubahan fungsi hutan. Jika kamu diminta menjadi ketua tim penelitian sekolah untuk mengkaji persebaran pohon tengkawang berdasarkan pengetahuan lokal dan pendekatan ilmiah, desain penelitian mana yang paling sesuai?*

- a) *Melakukan eksperimen di laboratorium untuk menumbuhkan bibit tengkawang dan mengukur pertumbuhannya.*
- b) *Wawancara tokoh adat dan mengamati langsung lokasi-lokasi yang disebutkan, lalu mencatat jumlah dan ukuran pohon yang ditemukan.*
- c) *Mengambil foto pohon-pohon besar di sekitar sekolah dan membuat klasifikasi morfologi daunnya.*
- d) *Menghitung jumlah pohon secara acak di hutan, tanpa melibatkan informasi dari masyarakat setempat.*
- e) *Menebang beberapa pohon untuk dianalisis umur dan kualitas kayunya di laboratorium kampus terdekat.*

Butir nomor 4 tersebut termasuk dalam tingkat kesukaran mudah, ini menunjukkan bahwa hampir seluruh siswa (77%) dapat menjawab soal dengan benar. Namun, distraktor pada butir ini tidak efektif, ini dilihat dari hanya 2% siswa memilih pilihan jawaban C dan E. Pilihan jawaban tersebut dinilai kurang menarik bagi siswa. Adapun butir yang dikategorikan tidak baik adalah yang memiliki daya pembeda rendah. Dalam hal ini, butir nomor 8 dan 14 termasuk kategori tidak baik. Kedua butir ini juga memiliki tingkat kesukaran tinggi dan distraktor yang tidak efektif. Salah satu contoh cuplikan soal kategori revisi dapat dilihat berikut ini

8. Masyarakat Dayak Kalimantan Tengah mempraktikkan ladang berpindah (*huma*), dengan membakar lahan sebelum ditanami padi. Mereka percaya bahwa abu sisa pembakaran dapat menyuburkan tanah. Untuk membuktikan hal ini, dilakukan pengukuran kandungan hara tanah dari dua jenis lahan:

Tabel 10. Pengukuran Kandungan Hara Tanah

Unsur Hara	Lahan A (baru dibakar)	Lahan B (tidak dibakar 2 tahun)
Nitrogen(N)	45ppm	22ppm
Fosfor(P)		17ppm
Kalium(K)	50ppm	25ppm

Berdasarkan data di atas, manakah simpulan paling tepat secara ilmiah yang mendukung kearifan lokal pertanian Dayak?

- Tanah pada lahan yang tidak dibakar mengandung lebih banyak unsur hara karena lebih stabil.
- Pembakaran lahan menyebabkan hilangnya unsur hara penting dalam tanah dan menurunkan kesuburan.
- Abu hasil pembakaran berperan meningkatkan kandungan nitrogen, fosfor, dan kalium dalam tanah.
- Ladang berpindah menyebabkan tanah tidak subur dan tidak cocok untuk ditanami kembali.
- Penggunaan pupuk sintetis lebih efektif dibanding abu dalam meningkatkan kesuburan tanah

Butir seperti ini tidak mampu mengukur kemampuan peserta secara akurat dan disarankan untuk dihapus atau diganti, karena keberadaannya dapat menurunkan reliabilitas dan validitas instrumen secara keseluruhan (Rezigalla et al., 2024). Penghapusan butir dengan tingkat kesukaran ekstrem, terutama yang disertai distraktor tidak berfungsi, dapat meningkatkan kualitas psikometrik instrumen, termasuk daya pembeda dan konsistensi konstruk. Secara umum, hasil analisis menunjukkan bahwa hanya sebagian kecil butir yang benar-benar memenuhi kriteria soal yang baik. Sebagian besar masih perlu direvisi, terutama pada aspek distraktor dan tingkat kesukaran agar komposisi soal menjadi lebih proporsional antara kategori mudah, sedang, dan sukar. Penelitian serupa juga menunjukkan bahwa tidak mudah menghasilkan instrumen tes literasi sains berbasis etnosains yang memenuhi semua kriteria kualitas soal.

Penelitian berjudul Pengembangan Instrumen Tes Literasi Sains Berbasis Kearifan Lokal Di Trenggalek Pengembangan Instrumen Tes Literasi Sains Berbasis Kearifan Lokal Di Trenggalek tidak memiliki komposisi soal yang proporsional karena hanya terdiri dari soal sukar dan sedang, tanpa soal mudah (Murti et al, 2021). Sejalan dengan itu, penelitian Pengembangan Instrumen Tes Literasi Sains Berbasis Kearifan Lokal Di Kabupaten Lamongan juga menunjukkan hal yang sama, yaitu komposisi soal yang didominasi soal sedang (73%), soal sukar (23%) dan mudah (4%) (Maulida et al, 2022). Penelitian Pengembangan Instrumen Tes Literasi Sains Berbasis Kearifan Lokal Di Probolinggo hanya terdiri dari soal sedang (100%) (Hamdu et al, 2025). Temuan penelitian – penelitian ini menunjukkan bahwa integrasi konteks etnosains dapat mengukur literasi sains siswa, tetapi tidak menjamin semua butir mencapai kualitas baik. Dengan demikian, dibandingkan penelitian-penelitian serupa, hasil ini

mengonfirmasi bahwa proses pengembangan instrumen literasi sains berbasis etnosains memerlukan tahapan uji dan revisi berulang agar menghasilkan butir soal yang valid, reliabel, dan berfungsi optimal.

Kesimpulan

Hasil analisis menunjukkan bahwa instrumen tes literasi sains berbasis etnosains Kalimantan Tengah memiliki kualitas yang cukup baik berdasarkan lima aspek pengujian, yaitu validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas distraktor. Uji validitas konstruk menggunakan MDS menghasilkan nilai stress sebesar 0,068% dan DAF 0,99, yang menunjukkan bahwa model memiliki kecocokan yang sangat baik. Hampir semua butir dinyatakan valid secara konstruk, kecuali butir nomor 8 dan 14 yang posisinya menyimpang dari kelompok lainnya. Uji reliabilitas menghasilkan nilai *Cronbach's Alpha* sebesar 0,753, yang menunjukkan konsistensi internal antarkategori butir soal tergolong baik. Berdasarkan tingkat kesukaran, 65% soal tergolong mudah, 25% sedang, dan 10% sukar. Komposisi ini belum ideal karena dominasi soal mudah dapat membatasi kemampuan instrumen dalam mengukur variasi kemampuan siswa. Dari hasil uji daya beda, 80% butir menunjukkan daya pembeda sangat baik, 10% baik, dan 10% jelek. Butir nomor 8 dan 14 termasuk kategori tidak baik karena tidak mampu membedakan siswa berkemampuan tinggi dan rendah. Pada aspek efektivitas distraktor, hanya 4 butir (20%) yang memiliki distraktor berfungsi efektif, sedangkan 16 butir (80%) dinilai kurang efektif karena ada opsi yang tidak dipilih minimal oleh 5% peserta. Secara keseluruhan, instrumen ini tergolong valid dan reliabel, namun masih memerlukan revisi, terutama pada efektivitas distraktor dan distribusi tingkat kesukaran agar lebih seimbang.

Implikasi penelitian ini menunjukkan bahwa pengembangan instrumen berbasis etnosains dapat menjadi pendekatan yang relevan untuk mengukur literasi sains secara kontekstual sesuai budaya lokal Kalimantan Tengah. Keterbatasan penelitian terletak pada jumlah sampel uji coba yang masih terbatas dan belum mencakup keragaman sekolah. Oleh karena itu, penelitian selanjutnya disarankan memperluas uji coba di wilayah yang lebih luas serta melakukan item analisis lanjutan menggunakan pendekatan *Rasch Model* untuk memperkuat validitas konstruk dan fungsi distraktor.

Acknowledgment

-

Daftar Pustaka

- Adawiah, R. (2020). The Application-Based Analysis Of Questions Item Quality In Junior High School. *JISAE (Journal of Indonesian Student Assessment and Evaluation)*, 6(2), 140–148. <https://doi.org/10.21009/jisae>
- Amalia, D. V., Ilhami, A., Fuadiyah, S., & Kusumanegara, A. (2024). Development of a Scientific Literacy Instrument Based on Riau Malay Ethnoscience in Science Subjects. *Pedagogik: Jurnal Pendidikan*, 11(1), 1–18. <https://doi.org/10.33650/pjp.v11i1.6382>
- Arifin, Z. (2007). *Evaluasi Pembelajaran*. Remaja Rosdakarya. Jawa Barat
- Arikunto, S. (2014). *Proseur Penelitian: Suatu Pendekatan Praktik*. Rineka Cipta. Jakarta Timur
- Elgadal, A. H., & Mariod, A. A. (2021). Item Analysis of Multiple-choice Questions (MCQs): Assessment Tool For Quality Assurance Measures. *Sudan Journal of Medical Sciences*, 16(3), 334–346. <https://doi.org/10.18502/sjms.v16i3.9695>

- Febryana, N. E., Septiana, N., & Rohmadi, M. (2021). Literasi Sains Siswa Kelas IX dengan Implementasi Media Pembelajaran Berbasis eXe Learning Pada Materi Pewarisan Sifat. *Bioeduca: Journal of Biology Education*, 3(1), 60-70. <https://doi.org/10.21580/bioeduca.v3i1.6638>
- Hamdu, G., & Putri, A. R. (2025). Pengembangan Instrumen Tes Literasi Sains Berbasis Etnosains Hutan Larangan Kampung Adat Kuta. *Jurnal Lensa Pendas*, 10(2), 326-339 <https://doi.org/10.33222/jlp.v10i2.4967> h.
- Hastuti, P. W., Anjarsari, P., & Yamtinah, S. (2022). Scientific Literacy Assessment Instrument on Addictive Substances Topic in Ethnoscience Integrated Inquiry Based-Learning. *Journal of Science Education Research*, 6(1), 31-36.
- Ihsan, H. (2015). Validitas isi alat ukur penelitian: Konsep dan panduan penilaiannya. *Pedagogia*, 13(3), 173-179. <https://doi.org/10.17509/pedagogia.v13i3.6004>
- Ismail, I. A., Weriza, J., Mawardi, L., Usmeldi, F., & Handri, S. (2024). Tinjauan Sistematis Analisis Integrasi Etnosains dalam Pembelajaran IPA dan Dampaknya terhadap Kompetensi Era Modern dan Nilai-Nilai Pancasila. *Vol*, 4, 207-219. <https://doi.org/10.52436/1.jpti.478>
- Lestari, H., Setiawan, W., & Siskandar, R. (2020). Science Literacy Ability of Elementary Students Through Nature of Science-based Learning with the Utilization of the Ministry of Education and Culture's " Learning House". *Jurnal Penelitian Pendidikan IPA*, 6(2), 215-220. <https://doi.org/10.29303/jppipa.v6i2.410>
- Maulida, F., & Sunarti, T. (2022). Pengembangan Instrumen Tes Literasi Sains Berbasis Kearifan Lokal Di Kabupaten Lamongan. *Orbita: Jurnal Hasil Kajian, Inovasi Dan Aplikasi Pendidikan Fisika*, 8(1), 59-68. <https://doi.org/10.61132/arjuna.v2i2.612>
- Meyer, E. M., & Reynolds, M. R. (2022). Multidimensional Scaling of Cognitive Ability and Academic Achievement Scores. *Journal of Intelligence*, 10(4). <https://doi.org/10.3390/jintelligence10040117>
- Murti, W., & Sunarti, T. (2021). Pengembangan Instrumen Tes Literasi Sains Berbasis Kearifan Lokal Di Trenggalek. *Orbita: Jurnal Hasil Kajian, Inovasi Dan Aplikasi Pendidikan Fisika*, 7(1), 33-43. <https://doi.org/10.61132/arjuna.v2i2.612>
- Muslimin, T. P., & Fatimah, A. A. B. (2024). Peningkatan Kemampuan Literasi Siswa UPT SPF SDN Sangir melalui Modul Ajar Digital Berbasis Budaya. *Jurnal Dieksis ID*, 4(2), 101-113. <https://doi.org/10.54065/dieksis.4.2.2024.529>
- Ningsetyo, M., & Sunarti, T. (2024). Pengembangan Instrumen Tes Literasi Sains Berbasis Kearifan Lokal Di Probolinggo. *Jurnal Arjuna: Publikasi Ilmu Pendidikan, Bahasa Dan Matematika*, 2(2), 59-68. <https://doi.org/10.61132/arjuna.v2i2.612>
- Ningtyas, S. A., & Setiawan, B. (2023). Penerapan Discovery Learning Berbasis Etnosains untuk Meningkatkan Hasil Belajar Peserta Didik pada Materi Getaran dan Gelombang. *Jurnal Pendidikan MIPA*, 13(3), 628-637. <https://doi.org/https://doi.org/10.37630/jpm.v13i3.1133>
- Putra, B., & Wahyuni, S. (2025). Integrasi Kearifan Lokal dalam Pembelajaran IPA untuk Meningkatkan Literasi Sains Siswa: Kajian Literatur. *Kalam Cendikia*, 13(2). <https://doi.org/10.20961/jkc.v13i2.103518>

- Quaigrain, K., & Arhin, A. K. (2017). Using reliability and item analysis to evaluate a teacher-developed test in educational measurement and evaluation. *Cogent Education*, 4(1). <https://doi.org/10.1080/2331186X.2017.1301013>
- Raina, V., Liusie, A., & Gales, M. (2023). Assessing Distractors in Multiple-Choice Tests. Proceedings of the 4th Workshop on Evaluation and Comparison of NLP Systems, 12–22. <https://doi.org/10.18653/v1/2023.eval4nlp-1.2>
- Restiani, N. L. D., Margunayasa, I. G., & Paramita, M. V. A. (2023). Improving scientific literacy of elementary school students through problem-based learning model with balinese local wisdom. *Jurnal Ilmiah Sekolah Dasar*, 7(4), 590-598. <https://doi.org/10.23887/jisd.v7i4.62080>
- Rezigalla, A. A., Eleragi, A. M. E. S. A., Elhussein, A. B., Alfaifi, J., ALGhamdi, M. A., Al Ameer, A. Y., ... & Adam, M. I. E. (2024). Item analysis: the impact of distractor efficiency on the difficulty index and discrimination power of multiple-choice items. *BMC Medical Education*, 24(1), 445. <https://doi.org/10.1186/s12909-024-05433-y>
- Rini, E. F. S., Rahardjo, S. B., & Bramastia, B. (2024). Bibliometric analysis from 2014-2024: Integration of ethnoscience in science learning. *Evaluation In Education*, 6(1). <https://doi.org/10.37251/jee.v6i1.1262>
- Sudijono, A. (2017). Pengantar Evaluasi Pendidikan. PT. Raja Grafindo Persada. Jawa Barat
- Taherdoost, H. (2016). Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. *International journal of academic research in management (Ijarm)*, 5. <https://doi.org/10.2139/ssrn.3205040>
- Yessi, M. (2019). Pembelajaran Berbasis Guided Inquiry Untuk Meningkatkan Literasi Sains Siswa Kelas X MIPA Pada Materi Larutan Elektrolit Dan Nonelektrolit. *Jurnal Ilmiah Kanderang Tingang*, 10(1), 27-37. <https://doi.org/10.37304/jikt.v10i1.18>
- Zulham, M., Sukmawati, S., Fitriani A , F. A., & Taufiq, T. (2024). Penguatan Literasi Numerasi Siswa SDN 15 Salolo Melalui Implementasi Kurikulum Merdeka Berbasis Teknologi Informasi. *Jurnal IPMAS*, 4(3), 181–192. <https://doi.org/10.54065/ipmas.4.3.2024.492>