



Analisis Butir Soal Materi Rantai Makanan Meliputi Validitas, Reliabilitas, Tingkat Kesukaran, dan Daya Beda

Mu'ah Sharoh Suprihatin^{1*}, Fina Fakhriyah²

^{1,2}Program Studi Pendidikan Guru Sekolah Dasar, Universitas Muria Kudus, Indonesia

Email: muahsharoh95@gmail.com¹, fina.fakhriyah@umk.ac.id²

*Penulis korespondensi: muahsharoh95@gmail.com

Abstract. This research is motivated by the importance of the availability of quality evaluation instruments to accurately measure student learning outcomes, especially in the Natural and Social Sciences (IPAS) subject matter. The quality of a test instrument greatly determines the validity of the learning outcome data obtained, but in practice, it is still often found that test items have not gone through the qualitative and quantitative parameter analysis stages. This study aims to analyze and describe the quality of the IPS test items on the food chain material in grade V of elementary school in terms of validity, reliability, level of difficulty, and discrimination. The research method used is descriptive quantitative with 19 fifth grade students of Growong Kidul 02 Elementary School as test subjects. Data were collected through an objective test instrument in the form of 25 multiple choice questions which were then processed using the Microsoft Excel program. The research findings show that of the total 25 test items tested, 18 items were declared valid and had an adequate level of reliability to be used as a measuring tool. The results of the difficulty level analysis show a varied distribution of questions, including easy, medium, and difficult categories. Meanwhile, in terms of discriminatory power, a diversity of question quality was found, ranging from poor, adequate, good, to excellent. The implication of this study is the availability of a standardized assessment instrument for the food chain topic that can serve as a reference for teachers at Elementary School in conducting learning evaluations. However, the results of this study also provide recommendations for improvements or revisions to items with low discriminatory power to optimize the overall quality of the instrument for future use.

Keywords: Discriminatory Power; Level of Difficulty; Reliability; Test Items; Validity.

Abstrak. Penelitian ini dilatarbelakangi oleh pentingnya ketersediaan instrumen evaluasi yang berkualitas untuk mengukur capaian belajar siswa secara akurat, khususnya pada muatan pelajaran Ilmu Pengetahuan Alam dan Sosial (IPAS). Kualitas suatu instrumen tes sangat menentukan validitas data hasil belajar yang diperoleh, namun pada praktiknya masih sering ditemukan butir soal yang belum melalui tahapan analisis parameter kualitatif maupun kuantitatif. Penelitian ini bertujuan untuk menganalisis dan mendeskripsikan kualitas butir soal IPAS materi rantai makanan di kelas V sekolah dasar ditinjau dari aspek validitas, reliabilitas, tingkat kesukaran, dan daya beda. Metode penelitian yang digunakan adalah deskriptif kuantitatif dengan subjek uji coba sebanyak 19 siswa kelas V SD Negeri Growong Kidul 02. Data dikumpulkan melalui instrumen tes objektif berupa 25 butir soal pilihan ganda yang kemudian diolah menggunakan bantuan program Microsoft Excel. Temuan penelitian menunjukkan bahwa dari total 25 butir soal yang diuji, sebanyak 18 butir soal dinyatakan valid dan memiliki tingkat reliabilitas yang memadai untuk digunakan sebagai alat ukur. Hasil analisis tingkat kesukaran menunjukkan sebaran soal yang variatif, mencakup kategori mudah, sedang, hingga sukar. Sementara itu, pada aspek daya beda, ditemukan keberagaman kualitas soal yang merentang dari kategori kurang baik, cukup, baik, hingga sangat baik. Implikasi dari penelitian ini adalah tersedianya instrumen asesmen yang terstandarisasi untuk materi rantai makanan yang dapat menjadi rujukan bagi guru di SD dalam melaksanakan evaluasi pembelajaran. Meskipun demikian, hasil penelitian ini juga memberikan rekomendasi untuk melakukan perbaikan atau revisi terhadap butir soal yang memiliki daya beda rendah agar kualitas instrumen secara keseluruhan menjadi lebih optimal pada penggunaan di masa mendatang.

Kata kunci: Butir Soal; Daya Beda; Reliabilitas; Tingkat Kesukaran; Validitas.

1. LATAR BELAKANG

Penilaian hasil belajar merupakan elemen fundamental dalam sistem pendidikan yang berfungsi sebagai instrumen untuk memantau sejauh mana peserta didik telah mencapai standar kompetensi yang ditetapkan (Muhadi et al., 2025). Salah satu instrumen asesmen formatif maupun sumatif yang paling konvensional adalah ujian tulis. Tes ini disusun dalam format butir soal yang spesifik dan berfungsi vital untuk mengevaluasi kedalaman pemahaman konseptual peserta didik terhadap bahan ajar yang telah disampaikan (Sa'adah et al., 2025). Guna mengidentifikasi kompetensi berpikir tingkat tinggi (HOTS) siswa SD pada muatan pelajaran IPAS, diperlukan alat evaluasi yang dirancang secara sistematis agar dapat mengukur ketajaman analisis siswa (Susanti et al., 2025). Ironisnya, instrumen evaluasi yang diimplementasikan dalam praktik pengajaran sering kali belum melewati proses analisis kualitas butir soal yang komprehensif. Akibatnya, interpretasi hasil penilaian terhadap kapabilitas aktual peserta didik menjadi kurang akurat (Jendriadi et al., 2025).

Kualitas suatu instrumen evaluasi ditentukan oleh empat parameter esensial yang meliputi validitas, reliabilitas, indeks kesukaran, serta daya beda setiap butir soal (Rachmawati & Pradana, 2025). Derajat validitas mengacu pada taraf akurasi sebuah instrumen evaluasi dalam merepresentasikan variabel atau kompetensi yang menjadi sasaran pengukuran secara tepat dan objektif (Buka et al., 2025). Sementara itu, parameter reliabilitas merepresentasikan derajat stabilitas dan konsistensi hasil penilaian yang dihasilkan oleh instrumen tersebut dalam berbagai kondisi pengujian yang berulang (Arbeni et al., 2025). Proporsi peserta yang menjawab benar menjadi tolok ukur dalam menentukan derajat kesukaran soal. Di sisi lain, daya pembeda digunakan sebagai instrumen untuk memverifikasi kemampuan suatu butir tes dalam mendiferensiasi kapasitas akademik antara siswa yang unggul dengan siswa yang memiliki kemampuan di bawah rata-rata (Aprilia & S, 2025). Evaluasi mendalam terhadap aspek validitas, reliabilitas, kesukaran, dan daya beda menjadi prasyarat bagi guru dalam mengonstruksi alat ukur yang transparan dan akuntabel, sehingga mampu merefleksikan capaian belajar siswa secara autentik (Hajiriah, 2025).

Pada muatan pelajaran IPAS, terutama dalam topik rantai makanan, penguasaan siswa terhadap prinsip fundamental seperti perpindahan energi dan sirkulasi nutrisi menjadi fondasi krusial yang melandasi pemahaman fenomena saintifik yang lebih abstrak pada tingkat pendidikan selanjutnya (Bambang et al., 2025). Dengan demikian, penguasaan materi ini berfungsi sebagai fondasi konseptual esensial dalam pembelajaran sains. Konsekuensinya, instrumen evaluasi yang diaplikasikan pada materi rantai makanan memerlukan analisis

mendalam guna menjamin alat ukur tersebut merefleksikan capaian kompetensi siswa secara akurat.

Evaluasi terhadap butir tes mencakup parameter validitas, reliabilitas, tingkat kesukaran, serta daya beda. Dalam hal ini, validitas berfungsi sebagai instrumen untuk memverifikasi akurasi soal dalam merepresentasikan penguasaan siswa terhadap materi rantai makanan. Sedangkan parameter reliabilitas merujuk pada derajat stabilitas dan keterandalan instrumen dalam memberikan hasil penilaian yang serupa meskipun dilakukan pengukuran berulang terhadap materi yang sama (Subhaktiyasa, 2024). Ketidaktepatan pada aspek validitas dan reliabilitas akan mengakibatkan bias dalam perolehan data, sehingga informasi mengenai kapabilitas aktual peserta didik tidak dapat terpotret secara objektif. Selain itu, analisis indeks kesukaran sangat esensial untuk memetakan distribusi tingkat kesulitan materi rantai makanan, guna mengidentifikasi apakah soal tersebut tergolong dalam kategori mudah, sedang, atau sukar. Di sisi lain, indeks daya pembeda mencerminkan sensitivitas suatu butir soal dalam mengklasifikasikan peserta didik berdasarkan gradasi pemahamannya, terutama pada materi yang membutuhkan penalaran mendalam seperti rantai makanan (Fitraynsyah & Hilmiyati, 2024). Temuan penelitian tersebut mengindikasikan bahwa implementasi format soal pilihan ganda memiliki keunggulan dalam mengoptimalkan indeks daya beda serta memperkuat relevansi kontekstual butir soal.

Konsekuensinya, evaluasi mutu butir soal pada materi rantai makanan tidak hanya krusial bagi peningkatan standar asesmen, namun juga menjadi jaminan atas kedalaman pemahaman siswa. Perangkat tes yang telah divalidasi secara saksama memungkinkan pendidik untuk memformulasikan umpan balik yang konstruktif, mengonstruksi strategi instruksional yang adaptif, serta menetapkan kebijakan pedagogis berbasis data. Oleh sebab itu, penelitian ini memosisikan analisis butir soal bukan sekadar sebagai prosedur administratif, melainkan instrumen strategis untuk memperkuat sistem evaluasi pembelajaran IPAS. Melalui penyediaan alat ukur yang valid dan reliabel, guru dapat mengekstraksi data yang akurat guna mendasari pengambilan keputusan pendidikan yang lebih presisi.

Penelitian ini diorientasikan untuk melakukan analisis mendalam terhadap butir soal materi rantai makanan, mengingat masih terbatasnya ketersediaan instrumen evaluasi di tingkat SD yang telah teruji secara empiris melalui parameter validitas, reliabilitas, tingkat kesukaran, dan daya beda. Melalui kajian ini, diharapkan tenaga pendidik memperoleh rujukan valid dalam mengonstruksi serta menyempurnakan alat ukur asesmen, sehingga proses asesmen menjadi lebih representatif, bermakna, dan mampu mendasari strategi pembelajaran yang lebih adaptif.

2. KAJIAN TEORITIS

Hakikat Evaluasi dan Analisis Butir Soal dalam IPAS

Evaluasi dalam sistem pendidikan modern dipandang sebagai proses sistematis untuk menentukan sejauh mana tujuan instruksional telah dicapai oleh peserta didik (Boroallo et al., 2025). Dalam muatan pelajaran Ilmu Pengetahuan Alam dan Sosial (IPAS) di jenjang sekolah dasar, instrumen penilaian harus mampu memotret kemampuan kognitif siswa secara akurat agar keputusan pedagogis yang diambil tidak mengalami bias. Analisis butir soal merupakan prosedur krusial untuk memverifikasi kualitas perangkat tes melalui pendekatan kuantitatif (Zayrin et al., 2025). Tanpa proses analisis yang kredibel, hasil belajar yang dilaporkan sering kali gagal mencerminkan penguasaan konsep siswa yang sebenarnya, terutama pada materi kompleks seperti interaksi makhluk hidup dalam ekosistem.

Validitas dan Reliabilitas

Kualitas sebuah instrumen evaluasi ditentukan oleh dua pilar utama, yakni validitas dan reliabilitas. Validitas merujuk pada derajat ketepatan alat ukur dalam mengevaluasi variabel yang seharusnya diukur, sehingga data yang dihasilkan bersifat sah (Rambe et al., 2025). Sementara itu, reliabilitas berkaitan dengan aspek keajekan atau konsistensi hasil pengukuran ketika instrumen tersebut diujikan kembali pada subjek yang sama di waktu yang berbeda (Bawono & Wibowo, 2025). Dalam konteks materi rantai makanan, validitas memastikan bahwa soal benar-benar mengukur pemahaman saintifik tentang aliran energi, sedangkan reliabilitas menjamin kestabilan skor yang diperoleh siswa sebagai dasar pengambilan keputusan akademik.

Tingkat Kesukaran dan Daya Pembeda

Indeks tingkat kesukaran merupakan indikator yang menggambarkan proporsi keberhasilan peserta didik dalam menyelesaikan sebuah butir soal (Muldia et al., 2025). Soal yang ideal dalam pembelajaran IPAS harus memiliki distribusi tingkat kesulitan yang seimbang antara kategori mudah, sedang, dan sukar guna menjaga motivasi sekaligus tantangan intelektual siswa.

Di samping itu, daya pembeda atau daya beda sangat diperlukan untuk mengevaluasi sensitivitas soal dalam memisahkan kelompok siswa berkemampuan tinggi dengan siswa berkemampuan rendah (Surbakti, 2025). Melalui kombinasi analisis tingkat kesukaran dan daya beda, guru dapat mengidentifikasi butir soal mana yang perlu dipertahankan, direvisi, atau bahkan dieliminasi dari bank soal. Selanjutnya, penelitian oleh Handayani et al. (2025) menekankan bahwa banyak soal IPAS di sekolah dasar belum memenuhi kriteria daya beda

yang optimal, sehingga memerlukan analisis mendalam untuk meningkatkan mutu asesmen. Relevansi dari berbagai kajian tersebut memberikan pijakan bagi penelitian ini untuk memfokuskan analisis pada materi rantai makanan di kelas V, dengan tujuan menghasilkan alat evaluasi yang tidak hanya valid secara teoretis tetapi juga andal secara empiris sebagai rujukan dalam proses penilaian di sekolah dasar.

3. METODE PENELITIAN

Penelitian ini menggunakan metode kuantitatif yang bertujuan untuk menganalisis validitas, reliabilitas, tingkat kesukaran, dan daya beda pada soal-soal IPAS materi rantai makanan. Penelitian dilakukan di SD Negeri Growong Kidul 02, salah satu sekolah dasar unggulan di Kecamatan Juwana yang sudah menerapkan Kurikulum Merdeka.

Penelitian ini dilakukan di SD Negeri Growong Kidul 02. SD Negeri Growong Kidul 02 dipilih karena merupakan salah satu sekolah dasar yang termasuk dalam sekolah percontohan atau sekolah model di Kabupaten Pati. SD Negeri Growong Kidul 02 juga merupakan sekolah yang sudah menerapkan Kurikulum Merdeka. Faktor lain yang menjadikan SD Negeri Growong Kidul 02 dijadikan studi kasus penelitian adalah karena kemudahan akses dan koordinasi antara sekolah dan peneliti. Penelitian yang dilakukan di SD Negeri Growong Kidul 02, dengan jumlah 60 peserta didik kelas V. Sampel penelitian diambil dengan teknik purposive sampling, dimana kelas VC dipilih sebagai sampel utama dengan jumlah 19 peserta didik. Kelas VC dipilih berdasarkan kriteria representatif yaitu memiliki karakteristik heterogenitas kemampuan akademik yang mewakili variasi keseluruhan populasi kelas V di SD Negeri Growong Kidul 02. Pemilihan ini bertujuan agar hasil pengujian instrumen dapat menggambarkan kondisi yang realistik dari populasi kelas V.

Instrumen penelitian berupa tes objektif yang terdiri dari 25 butir soal pilihan ganda, yang dikembangkan berdasarkan kisi-kisi soal yang mengacu pada indikator pembelajaran dari kurikulum yang berlaku, khususnya pada materi rantai makanan. Kisi-kisi tersebut memastikan setiap butir soal relevan dengan capaian pembelajaran, tujuan pembelajaran, dan alur tujuan pembelajaran sehingga keterkaitan antar soal dengan indikator pembelajaran dapat dipertanggungjawabkan secara sistematis. Proses pembuatan, perancangan, dan penyusunan 25 butir soal dikerjakan melalui tiga tahapan. Tahap pertama adalah perencanaan, bagian ini dimulai dari menentukan tujuan, membuat kisi-kisi soal, dan menyiapkan bahan pembuatan soal. Pada tahap pertama, perencanaan berdasarkan pada

kurikulum yang berlaku, Capaian Pembelajaran (CP), Tujuan Pembelajaran (TP), Alur Tujuan Pembelajaran (ATP), dan materi rantai makanan.

Tahap kedua adalah merumuskan butir soal, pada tahap ini dilakukan pembuatan butir soal pilihan ganda sebanyak 25 butir soal. Setelah 25 butir soal sudah disiapkan, maka dilakukan penyusunan butir soal dengan bahasa yang jelas, baik, dan mudah dipahami peserta didik. Perumusan butir soal harus sudah sesuai dengan apa yang telah ditentukan pada Capaian Pembelajaran, Tujuan Pembelajaran, Alur Tujuan Pembelajaran, dan materi yang telah dipilih. Tahap ketiga yaitu uji coba dan analisis butir soal, 25 butir soal yang telah disusun dengan sistematika bahasa yang baik selanjutnya dilakukan uji soal. Hasil dari uji soal kemudian akan dianalisis menggunakan statistika untuk mengetahui tingkat validitas, reliabilitas, tingkat kesukaran, dan daya beda soal. Setelah dianalisis, maka soal yang tidak valid, memiliki reliabilitas rendah, terlalu mudah, dan memiliki daya beda rendah akan dievaluasi untuk selanjutnya digunakan dalam tes penilaian.

Uji instrumen dilakukan dengan empat tahapan utama, yaitu uji validitas, uji reliabilitas, tingkat kesukaran, dan daya beda soal. Uji validitas digunakan untuk mengetahui sejauh mana soal-soal pilihan ganda dan uraian benar-benar dapat mengukur kemampuan kognitif siswa dengan objektif dan menyeluruh (Pramudita et al., 2025). Uji validitas dilakukan dengan rumus koefisien korelasi biserial. Berdasarkan analisis statistik, soal yang valid adalah soal yang memiliki r hitung $>$ r tabel. Sedangkan soal yang tidak valid adalah soal yang memiliki r hitung $<$ r tabel.

Tahapan uji yang kedua adalah uji reliabilitas, uji reliabilitas digunakan untuk mengukur sejauh mana soal-soal tersebut mampu memberikan hasil yang konsisten apabila digunakan secara berulang kali. Atau dengan kata lain, uji reliabilitas digunakan untuk mengukur bagaimana kemampuan suatu alat uji berupa soal pilihan ganda dan uraian dapat digunakan dalam waktu yang berbeda, penguji yang berbeda, tetapi dengan kondisi pengukuran yang masih sama (Ridha et al., 2025). Uji reliabilitas diukur dengan rumus KR-20, soal yang dianggap reliable adalah jenis soal yang memiliki nilai KR-20 $>$ 0,90.

Tahapan uji yang ketiga adalah tingkat kesukaran soal, tingkat kesukaran soal digunakan untuk mengetahui tingkat proporsi soal tersebut dengan menggolongkan soal yang mudah, sedang, dan sulit atau sukar. Tingkat kesukaran soal perlu diketahui karena digunakan untuk menentukan proporsi soal yang baik untuk mengukur kemampuan kognitif siswa secara menyeluruh terutama siswa yang heterogen (Putra et al., 2025). Dalam pengukuran uji tingkat kesukaran dapat dilakukan dengan cara menganalisis skor rata-rata yang diperoleh oleh siswa, analisis tersebut kemudian digolongkan pada jenis soal mudah,

sedang, dan sukar. Hasil yang didapatkan dari analisis uji tingkat kesukaran soal dapat digunakan sebagai dasar oleh penguji atau pendidik dalam melakukan revisi soal yang kurang dari standar pembelajaran yang telah ditentukan.

Tahapan uji yang terakhir adalah dengan daya beda soal, daya beda soal digunakan sebagai alat uji untuk mengetahui sebuah soal dapat membedakan tingkat kemampuan kognitif siswa. Daya beda soal dibedakan menjadi soal dengan daya beda kurang baik, cukup, baik, dan sangat baik, yang ditentukan dengan menghitung indeks diskriminasi soal tersebut. Soal yang baik adalah soal yang mampu memberikan daya pembeda kemampuan kognitif siswa secara optimal, dengan kata lain soal yang memiliki indeks diskriminasi tinggi adalah soal yang mampu membedakan siswa berkemampuan kognitif tinggi dan siswa dengan kemampuan kognitif rendah (Gunawan et al., 2025). Hasil dari uji daya beda soal akan dapat digunakan oleh penguji atau pendidik sebagai dasar dalam mengevaluasi bentuk soal yang tidak sesuai kriteria dalam standar pembelajaran.

Data yang diperoleh dari tes kemudian akan dilakukan analisis validitas, reliabilitas, tingkat kesukaran, dan daya beda menggunakan metode statistik deskriptif. Analisis dalam penelitian ini dilakukan dengan menggunakan perangkat lunak uji statistik *Microsoft Excel* untuk menguji dan menganalisis empat tahapan tersebut. Landasan penelitian ini diperkuat oleh beberapa temuan terdahulu yang relevan. Studi yang dilakukan oleh Basri et al. (2025) menunjukkan bahwa analisis butir soal menggunakan bantuan perangkat lunak seperti *Microsoft Excel* terbukti efektif dalam mempermudah guru melakukan evaluasi mandiri terhadap instrumen tes buatannya.

4. HASIL DAN PEMBAHASAN

Uji Validitas

Uji validitas merupakan metode penting dalam analisis butir soal untuk memastikan setiap item pertanyaan dalam tes benar-benar mengukur apa yang seharusnya diukur, sesuai dengan tujuan pembelajaran atau kompetensi yang ingin dicapai (Baihaqi et al., 2025). Dalam penelitian ini, uji validitas dilakukan terhadap 25 butir soal pilihan ganda yang dianalisis menggunakan rumus koefisien korelasi biserial, dengan batas nilai r tabel sebesar 0,456 sebagai acuan. Berdasarkan hasil uji validitas, 18 dari 25 butir soal dinyatakan valid karena memiliki nilai r hitung lebih besar dari r tabel (0,456). Jika ditinjau dari materi pada kisi-kisi, butir soal yang valid umumnya merepresentasikan materi rantai makanan dan aliran energi secara jelas dan sesuai dengan indikator pembelajaran.

Sebagai contoh, soal nomor 1 dan 2 yang mengukur kemampuan siswa dalam mengidentifikasi peran makhluk hidup dalam rantai makanan (produsen, konsumen, dan dekomposer) menunjukkan validitas yang baik. Hal ini menandakan bahwa materi tersebut mudah dipahami siswa dan disajikan dengan konteks yang konkret, seperti ekosistem sawah dan laut. Demikian pula soal nomor 6, 8, dan 9 yang berkaitan dengan sumber energi utama serta peran konsumen primer dan sekunder, terbukti mampu mengukur pemahaman konsep dasar aliran energi secara tepat.

Sebaliknya, beberapa soal yang dinyatakan tidak valid, seperti nomor 3, 4, 5, 7, dan 12, meskipun masih berada dalam lingkup materi rantai makanan, cenderung memiliki redaksi atau konteks materi yang kurang spesifik. Hal ini menyebabkan soal belum sepenuhnya merepresentasikan indikator pembelajaran yang diharapkan. Dengan demikian, ketidakvalidan butir soal tersebut lebih disebabkan oleh ketidaktepatan pengemasan materi, bukan karena materi yang diujikan tidak relevan.

Keabsahan setiap butir soal menjadi penentu utama dalam menghasilkan data penilaian yang objektif, sehingga gambaran penguasaan konsep siswa terhadap materi pelajaran benar-benar terukur secara tepat (Manasikana et al., 2025). Ketika suatu soal dinyatakan valid, ini berarti bahwa siswa yang memiliki pemahaman lebih baik terhadap materi cenderung memperoleh skor lebih tinggi pada soal tersebut, dibandingkan dengan siswa yang pemahamannya masih rendah.

Sebaliknya, jika suatu soal tidak valid, maka besar kemungkinan soal tersebut tidak selaras dengan indikator pembelajaran, terlalu membingungkan, atau justru tidak menantang sama sekali. Soal seperti ini tidak hanya gagal memberikan informasi yang akurat tentang penguasaan siswa, tetapi juga berpotensi menciptakan ketidakadilan dalam penilaian. Oleh karena itu, soal-soal tersebut telah disusun dengan mempertimbangkan keterkaitan antara tujuan pembelajaran, materi, dan konstruk kognitif yang tepat.

Guna menghasilkan instrumen evaluasi yang berkualitas, pendidik perlu menjamin bahwa setiap butir soal IPAS dirancang untuk menguji kedalaman pemahaman serta keterampilan berpikir kritis siswa, melampaui sekadar kemampuan kognitif tingkat rendah seperti menghafal (Muhadi et al., 2025). Instrumen yang memiliki tingkat validitas tinggi memungkinkan pendidik memantau perkembangan kompetensi siswa secara akurat, sehingga umpan balik yang diberikan menjadi lebih efektif dan mampu menciptakan proses pembelajaran IPAS yang lebih bermakna. Berikut adalah hasil uji validitas terhadap 25 butir soal pilihan ganda.

Tabel 1. Hasil Uji Validitas.

No. Soal	r tabel	r hitung	Keterangan
1	0,456	0,483	Valid
2	0,456	0,576	Valid
3	0,456	0,260	Tidak Valid
4	0,456	0,231	Tidak Valid
5	0,456	0,401	Tidak Valid
6	0,456	0,709	Valid
7	0,456	0,138	Tidak Valid
8	0,456	0,598	Valid
9	0,456	0,819	Valid
10	0,456	0,934	Valid
11	0,456	0,934	Valid
12	0,456	0,413	Tidak Valid
13	0,456	0,934	Valid
14	0,456	0,934	Valid
15	0,456	0,934	Valid
16	0,456	0,934	Valid
17	0,456	0,934	Valid
18	0,456	0,934	Valid
19	0,456	0,934	Valid
20	0,456	0,934	Valid
21	0,456	0,349	Tidak Valid
22	0,456	0,461	Valid
23	0,456	0,349	Tidak Valid
24	0,456	0,461	Valid
25	0,456	0,461	Valid

Sumber: Output Ms. Excel, diolah oleh Peneliti (2025)

Uji Reliabilitas

Tujuan utama dari pengujian reliabilitas adalah untuk memverifikasi sejauh mana perangkat evaluasi mampu mempertahankan konsistensi temuan datanya ketika diujikan kembali pada parameter lingkungan yang setara (Waruwu et al., 2025). Salah satu teknik yang umum digunakan adalah perhitungan KR-20, dimana nilai KR-20 $> 0,90$ menunjukkan reliabilitas. Hasil uji reliabilitas menunjukkan nilai KR-20 sebesar 0,950 yang termasuk kategori sangat reliabel. Tingginya reliabilitas ini tidak terlepas dari keseragaman materi dalam kisi-kisi soal, yaitu fokus pada topik Harmoni dalam Ekosistem dengan materi utama rantai makanan dan aliran energi.

Seluruh butir soal dirancang berdasarkan dua tujuan pembelajaran utama, yaitu identifikasi peran makhluk hidup dan penjelasan aliran energi. Konsistensi materi ini membuat setiap soal saling berkaitan dalam mengukur konstruk yang sama, sehingga instrumen mampu memberikan hasil yang stabil dan konsisten. Dengan demikian, reliabilitas yang tinggi menunjukkan bahwa materi yang diujikan telah tersusun secara sistematis dan tidak menyimpang dari capaian pembelajaran IPAS fase C kelas V. Berikut adalah hasil uji reliabilitas terhadap 25 butir soal pilihan ganda.

Tabel 2. Hasil Uji Reliabilitas.

KR-20	Jumlah Butir Soal	Keterangan
0,950	25	Reliabel

Sumber: Output Ms. Excel, diolah oleh Peneliti (2025)

Tingkat reliabilitas yang signifikan pada 25 butir soal membuktikan bahwa perangkat evaluasi tersebut memiliki stabilitas yang baik dalam mengukur kompetensi siswa secara konsisten. Keajekan ini krusial untuk memastikan bahwa skor yang diperoleh murni merepresentasikan kemampuan peserta didik tanpa distorsi unsur kebetulan.

Uji Tingkat Kesukaran

Indikator tingkat kesukaran soal menggambarkan kemampuan peserta didik dalam menyelesaikan tugas secara tepat, yang dikalkulasi melalui rasio rata-rata skor terhadap skor maksimal 1. Hasil uji tingkat kesukaran menunjukkan bahwa sebagian besar soal berada pada kategori sedang, khususnya soal-soal yang menguji pemahaman dan aplikasi konsep aliran energi (misalnya soal nomor 6 hingga 20). Hal ini menunjukkan bahwa materi aliran energi menuntut kemampuan berpikir yang lebih dari sekadar mengingat, sehingga tingkat kesukarannya menjadi lebih seimbang.

Soal-soal yang tergolong mudah, seperti nomor 1, 2, 3, 5, 8, dan 9, umumnya berkaitan dengan pengenalan peran makhluk hidup dalam rantai makanan, yang bersifat konseptual dasar dan sering dijumpai dalam pembelajaran sehari-hari. Sebaliknya, soal-soal dengan kategori sukar (nomor 21–25) didominasi oleh materi evaluasi dan kreasi aliran energi, seperti menilai efisiensi transfer energi, merancang rantai makanan alternatif, dan menyusun piramida energi. Hal ini wajar karena materi tersebut menuntut kemampuan berpikir tingkat tinggi (C5–C6).

Dari hasil ini, dapat disimpulkan bahwa secara umum tingkat kesukaran soal menunjukkan variasi tingkat kesukaran sehingga dapat digunakan untuk mengukur kemampuan siswa secara komprehensif. Berikut adalah hasil uji tingkat kesukaran terhadap 25 butir soal pilihan ganda.

Tabel 3. Uji Tingkat Kesukaran.

No. Butir Soal	Indeks Kesukaran	Keterangan
1	0,84	Mudah
2	0,79	Mudah
3	0,89	Mudah
4	0,63	Sedang
5	0,84	Mudah
6	0,68	Sedang
7	0,63	Sedang
8	0,74	Mudah
9	0,74	Mudah
10	0,68	Sedang
11	0,68	Sedang
12	0,68	Sedang
13	0,68	Sedang
14	0,68	Sedang
15	0,68	Sedang
16	0,68	Sedang
17	0,68	Sedang
18	0,68	Sedang
19	0,68	Sedang
20	0,68	Sedang
21	0,11	Sukar
22	0,16	Sukar
23	0,11	Sukar
24	0,16	Sukar
25	0,16	Sukar

Sumber: Output Ms. Excel, diolah oleh Peneliti (2025)

Uji Daya Beda

Analisis daya pembeda diterapkan untuk mengevaluasi efektivitas suatu butir soal dalam mengklasifikasikan peserta didik yang memiliki kompetensi akademik tinggi dan rendah terhadap materi yang diujikan. Ditinjau dari materi pada kisi-kisi, soal-soal yang memiliki daya beda sangat baik umumnya berkaitan dengan materi analisis dampak perubahan dalam rantai makanan dan aliran energi, seperti soal nomor 6 dan 8 hingga 20. Materi ini efektif membedakan siswa berkemampuan tinggi dan rendah karena menuntut pemahaman konseptual yang mendalam serta kemampuan mengaitkan sebab akibat dalam ekosistem.

Sebaliknya, soal nomor 4 yang memiliki daya beda rendah berkaitan dengan identifikasi konsumen puncak, yang meskipun relevan dengan materi rantai makanan, disajikan secara relatif sederhana sehingga dapat dijawab dengan benar oleh sebagian besar siswa tanpa membedakan tingkat penguasaan konsep. Soal dengan daya beda cukup, seperti nomor 2, 3, 21, dan 23, menunjukkan bahwa materi sudah sesuai, namun masih memerlukan penyempurnaan redaksi atau konteks agar lebih menantang kemampuan berpikir siswa. Berikut adalah hasil uji daya beda terhadap 25 butir soal pilihan ganda.

Tabel 4. Uji Daya Beda.

No. Butir Soal	Daya Pembeda	Keterangan
1	0,33	Baik
2	0,23	Cukup
3	0,22	Cukup
4	0,14	Kurang Baik
5	0,33	Baik
6	0,67	Sangat Baik
7	0,36	Baik
8	0,56	Sangat Baik
9	0,56	Sangat Baik
10	0,67	Sangat Baik
11	0,67	Sangat Baik
12	0,58	Sangat Baik
13	0,67	Sangat Baik
14	0,67	Sangat Baik
15	0,67	Sangat Baik
16	0,67	Sangat Baik
17	0,67	Sangat Baik
18	0,67	Sangat Baik
19	0,67	Sangat Baik
20	0,67	Sangat Baik
21	0,20	Cukup
22	0,30	Baik
23	0,20	Cukup
24	0,30	Baik
25	0,30	Baik

Sumber: Output Ms. Excel, diolah oleh Peneliti (2025)

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan instrumen asesmen materi rantai makanan yang terdiri dari 25 soal pilihan ganda yang secara keseluruhan telah memenuhi standar validitas dan reliabilitas yang baik, dengan hanya tujuh soal yang dinyatakan tidak valid. Secara umum, instrumen ini memiliki variasi dalam tingkat kesukaran karena terdapat soal mudah, sedang, dan sukar, mayoritas butir soal efektif dalam membedakan kemampuan kognitif siswa (daya pembeda baik). Temuan ini memberikan kontribusi signifikan berupa referensi soal teruji yang dapat dimanfaatkan oleh guru kelas V di SD Negeri Growong Kidul 02.

Meskipun demikian, disarankan agar kajian lanjutan berfokus pada pengembangan instrumen serupa untuk materi muatan pelajaran IPAS lainnya karena sangat diperlukan guna memperkaya alat ukur asesmen guru secara komprehensif.

UCAPAN TERIMA KASIH

Penulis menyampaikan apresiasi dan terima kasih kepada Kepala Sekolah serta Bapak/Ibu Guru SD Negeri Growong Kidul 02 yang telah memberikan izin penelitian, bantuan fasilitas, serta dukungan penuh selama proses pengambilan data di kelas V. Ucapan terima kasih juga ditujukan kepada para siswa kelas V yang telah berpartisipasi dengan antusias dalam pengisian instrumen penelitian ini. Terakhir, terima kasih kepada rekan-rekan sejawat yang telah memberikan bantuan dalam proses ulasan naskah sehingga artikel ini dapat terselesaikan dengan baik.

DAFTAR REFERENSI

- Aprilia, R. P., & S, E. M. (2025). ANALISIS MISKONSEPSI SISWA PADA MATERI GEOMETRI DENGAN MENGGUNAKAN INSTRUMEN DIAGNOSTIK FOUR TIER DI KELAS IV MIS MADINATUSSALAM. *Jurnal Inovasi Pembelajaran*, 6(2), 549–561.
- Arbeni, W., Windiani, A., Sihotang, D. S. B., Anggraini, N., Wulandari, S., & Nugroho, A. (2025). Test Reliability Analysis in Educational Evaluation : A Quantitative Approach to Consistency and Validity. *Jurnal Nasional Holistic Science*, 5(1), 59–64.
- Baihaqi, M., Mahmudah, Mubarak, A. J. Al, Rabbaniyah4, N., & Susanto, D. (2025). Studi Validasi dan Reliabilitas Butir Soal Tes Sumatif dalam Pendidikan Bahasa Arab. *Jurnal Ilmiah Indonesia*, 5(4), 1530–1542.
- Bambang, Suharini, E., & Widyatmoko, A. (2025). PEMAHAMAN MENDALAM TENTANG EKOSISTEM UNTUK SISWA KELAS 5 DALAM MATA PELAJARAN IPAS. *JOURNAL OF ISLAMIC PRIMARY EDUCATION*, 6(1), 47–59.
- Basri, M. B., Sultan, Rapi, M., Baharman, & Sakaria. (2025). Meningkatkan Kompetensi Evaluasi Pembelajaran melalui Pelatihan Analisis Butir Soal bagi Mahasiswa Calon Guru. *Journal of Community Service*, 5(2), 633–639.
- Bawono, Y., & Wibowo, W. P. (2025). Validitas dan Reliabilitas Alat Ukur Intensitas Pemberian Dongeng untuk Anak Usia Sekolah Dasar. *Jurnal Psikologi*, 2(3), 1–9.
- Boroallo, R. P., Purnamasari, D. I., Kasmawati, & Mas'adi. (2025). Pentingnya Evaluasi Pembelajaran Dalam Meningkatkan Kualitas Pengajaran Di Era Modern. *Jurnal Pengabdian Masyarakat Dan Riset Pendidikan*, 3(4), 2632–2638.
- Buka, N., Jayamin, M. A., Kholis, N., & Rasyid, N. A. (2025). Peran Validitas dan Rehabilitas Dalam Evaluasi Pembelajaran. *Jurnal Penelitian Ilmu-Ilmu Sosial*, 02(11), 224–227.
- Fitraynsyah, A., & Hilmiyati, F. (2024). Peran Tingkat Kesukaran dan Daya Pembeda dalam Analisis Butir Tes : Kajian Literatur untuk Pendidikan Menengah. *Jurnal Riset Dan Evaluasi Pendidikan*, 1(4), 252–262.

- Gunawan, R., Supianto, A. A., & Herlambang, A. D. (2025). KUALITAS BUTIR-BUTIR SOAL UNTUK RANAH KOGNITIF. *Jurnal Edukasi Dan Teknologi Pembelajaran*, 06(01), 33–47.
- Hajiriah, T. L. (2025). ANALISIS EVALUASI MISKONSEPSI DAN PEMAHAMAN KONSEPTUAL DALAM PEMBELAJARAN IPA: TINJAUAN SISTEMATIS. *Panthera: Jurnal Ilmiah Pendidikan Sains Dan Terapan*, 5(2), 162–182.
- Handayani, R. D., Prayoga, J. D., & Putri, N. M. (2025). PERUBAHANNYABERBASIS PROGRAM ITEMAN. *Jurnal Natural Science Educational Research*, 8(2), 149–160.
- Jendriadi, Andriasari, S., Pikri, H., Islamidar, & Helena, E. (2025). ANALISIS KEMAMPUAN BERPIKIR TINGKAT TINGGI SISWA DALAM PEMBELAJARAN MATEMATIKA: EVALUASI DAN TINDAK LANJUT SISWA KELAS 5 SD DI SDN 03 KOTO BALINGKA. *Jurnal Ilmiah PGSD FKIP UNIVERSITAS MANDIRI*, 11(02), 287–303.
- Manasikana, A., Harani, D. A., Nabila, M., Azyati, N., Mumtaza, U., & Sukma, Z. (2025). Telaah Literatur tentang Efektivitas Evaluasi Sumatif dalam Menilai Pencapaian Kompetensi Peserta Didik. *Journal of Education and Islamic Studies*, 1(2).
- Muhadi, Jarir, Khairina, Rajuna, & Prasetyo, E. (2025). Evaluasi Perencanaan Desain Pembelajaran, Pelaksanaan Proses Kegiatan Pembelajaran, dan Evaluasi Instrumen Hasil Pembelajaran. *Edu Society: Jurnal Pendidikan, Ilmu Sosial, Dan Pengabdian Kepada Masyarakat*, 5(2), 156–165.
- Muldia, H., Sulistyarini, & Asriati, N. (2025). Analisis Butir Soal Berbasis Level Kognitif pada Ujian Akhir Semester Mata Pelajaran PPKn Kelas VII SMP Negeri 2 Pemangkat. *Jurnal Pendidikan Kewarganegaraan*, 9(2), 438–451.
- Pramudita, A., Prihantoro, W. K., Fikriyah, N., Falasifah, L., & Nursolihah, S. (2025). Implementasi Instrumen Tes Objektif Berbasis HOTS dalam Evaluasi Pembelajaran PAI di SMA Negeri 1 Jetis Bantul. *Islamic Education Journal*, 2(4).
- Putra, D. D., Nindiasari, H., & Fathurrohman, M. (2025). Pengembangan Instrumen Tes Kemampuan Pemecahan Masalah Matematis Siswa Sekolah Menengah Kejuruan. *Jurnal Inovasi Pendidikan ISSN*, 4(3), 967–983.
- Rachmawati, D., & Pradana, A. B. (2025). Analisis butir soal mata pelajaran ekonomi: validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. *Jurnal Pendidikan Ekonomi (JUPE)*, 13(3), 273–284.
- Rambe, A. R., Bandawaty, E., & Achmad, F. S. (2025). PENGARUH KEPEMIMPINAN DAN DISIPLIN TERHADAP KINERJA PEGAWAI DENGAN MOTIVASI KERJA SEBAGAI PERAN MEDIASI (STUDI KASUS PADA BADAN PENGAWAS PEMILIHAN UMUM RI). *Jurnal Bisnis Dan Manajemen*, 30(3), 51–68.
- Ridha, A. R., Sudarto, & Ghifari, F. Al. (2025). RELIABILITAS ALAT UKUR, JENIS-JENIS, CARA PENGUKURAN DAN FAKTOR-FAKTOR. *Jurnal Ilmiah Pengembangan Pendidikan*, 4(1), 121–131.

- Sa'adah, N., Jihana, N., Furqan, M., Asril, Z., & Fitriza, R. (2025). Implementasi Penilaian Sumatif Pada Ranah Kognitif Dalam Mengukur Pemahaman Peserta Didik Kelas VII Pada Mata Pelajaran SKI di MTs. *Jurnal Pendidikan Islam Dan Isu-Isu Sosial*, 23(2), 164–179.
- Subhaktiyyasa, P. G. (2024). Evaluasi Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif: Sebuah Studi Pustaka. *Journal of Education Research Evaluasi*, 5(4), 5599–5609.
- Surbakti, I. W. (2025). ANALISIS KUALITAS BUTIR SOAL PADA UJI COBA EVALUASI PEMBELAJARAN MATEMATIKA TINGKAT SMA. *Jurnal Ilmu Pengetahuan*, 5(1).
- Susanti, N. I., Rozy, F. S., Munawaroh, L., & Alfiyah, A. S. (2025). Analisis Tingkat Keterampilan Berpikir Kritis Siswa Kelas VI Sekolah Dasar pada Mata Pelajaran IPAS. *Jurnal Pemikiran Pendidikan*, 31(2), 325–336. <https://doi.org/10.30587/didaktika.v31i2.9847>
- Waruwu, M., Natijatul, S., Utami, P. R., Yanti, E., & Rusydiana, M. (2025). Metode Penelitian Kuantitatif: Konsep, Jenis, Tahapan dan Kelebihan. *Jurnal Ilmiah Profesi Pendidikan*, 10(1), 917–932.
- Zayrin, A. A., Nupus, H., Maizia, K. K., Marsela, S., Hidayatullah, R., & Harmonedi. (2025). Analisis Instrumen Penelitian Pendidikan (Uji Validitas Dan Reliabilitas Instrumen Penelitian). *QOSIM: Jurnal Pendidikan, Sosial & Humaniora*, 3(2), 780–789.