



Prosiding

Seminar Nasional

Unit Kegiatan Mahasiswa Penalaran dan Riset

IKIP PGRI Bojonegoro

Tema “Eksplorasi Penalaran dalam Riset untuk Meningkatkan Kualitas Publikasi Ilmiah”



Kajian Statistik terhadap Validitas dan Reliabilitas Soal Materi Teks Prosedur untuk Peserta Didik Sekolah Menengah Pertama

David Luthfi Pratama^{1(✉)}, Cahyo Hasanudin²

^{1,2}Program Studi Pendidikan Bahasa dan Sastra Indonesia, IKIP PGRI Bojonegoro, Indonesia

davidluthfipratama@gmail.com¹, cahyo.hasanudin@ikippgribojonegoro.ac.id²

abstrak – Penelitian ini bertujuan mengevaluasi kualitas instrumen tes materi teks prosedur melalui pengujian validitas dan reliabilitas. Penelitian menggunakan pendekatan kuantitatif dengan desain evaluatif. Instrumen yang dikembangkan terdiri atas 30 butir, yaitu 25 soal pilihan ganda dan 5 soal uraian. Validitas isi dianalisis menggunakan indeks V Aiken berdasarkan penilaian dua validator ahli, sedangkan reliabilitas diuji dengan formula Kuder-Richardson 20 (KR-20). Hasil penelitian menunjukkan bahwa 3 butir soal berada pada kategori cukup, 15 butir berkategori tinggi, dan 12 butir termasuk sangat tinggi. Hasil uji coba menunjukkan bahwa sebagian besar soal memiliki tingkat kesukaran mudah, sedangkan kemampuan daya pembedanya masih rendah. Nilai reliabilitas yang diperoleh sebesar 0,48 menunjukkan bahwa konsistensi instrumen belum optimal. Berdasarkan hasil analisis validitas, tingkat kesukaran, daya pembeda, dan reliabilitas, hanya 5 butir soal yang memenuhi kriteria baik dan dinilai layak digunakan sebagai alat evaluasi pembelajaran. Oleh karena itu, butir soal yang belum memenuhi standar perlu diperbaiki dan disempurnakan agar instrumen mampu menghasilkan data pengukuran yang lebih akurat dan konsisten.

Kata kunci – validitas, reliabilitas, teks prosedur

Abstract – This study aims to evaluate the quality of a test instrument covering procedural text material by assessing its validity and reliability. A quantitative approach with an evaluative design was employed. The developed instrument consists of 30 items: 25 multiple-choice questions and 5 essay questions. Content validity was analyzed using Aiken's V index based on assessments by two expert validators, while reliability was tested using the Kuder-Richardson 20 (KR-20) formula. The results indicate that 3 items fell into the "sufficient" category, 15 into the "high" category, and 12 into the "very high" category. Trial results show that most questions were easy, yet their discrimination power remained low. The obtained reliability coefficient of 0.48 indicates that the instrument's consistency was not yet optimal. Based on the analysis of validity, difficulty level, discrimination power, and reliability, only 5 items met the criteria for quality and were deemed suitable for use as a learning assessment tool. Therefore, items failing to meet the standards require revision and refinement to ensure the instrument yields more accurate and consistent measurement data.

Keywords – validity, reliability, procedural text

PENDAHULUAN

Validitas adalah salah satu ukuran krusial yang menentukan kualitas alat penelitian dalam menilai variabel sesuai dengan tujuan yang telah ditentukan. Berdasarkan pernyataan Hastjarjo (2011), validitas merujuk pada kemampuan alat untuk dengan tepat merepresentasikan konstruk yang ingin diukur. Sejalan dengan pandangan tersebut, Rinaldi dan Albina (2025) mengungkapkan bahwa alat yang memiliki validitas tinggi akan mampu mengukur variabel penelitian dengan akurat sesuai dengan konsep yang diajukan. Di sisi lain, Rosa (2008) mengindikasikan bahwa presisi dan objektivitas hasil pengukuran sangat dipengaruhi oleh kualitas validitas alat yang digunakan. Dari penjelasan tersebut, validitas dapat dipahami sebagai ukuran ketepatan suatu alat dalam menilai variabel yang sedang diteliti. Alat yang mempunyai validitas tinggi akan menghasilkan data yang akurat, terpercaya, serta dapat dipertanggungjawabkan secara ilmiah.

Validitas dalam penelitian dapat dibedakan ke dalam beberapa bentuk berdasarkan aspek yang menjadi fokus pengukuran suatu instrumen. Mukhlisa (2023) menyatakan bahwa validitas mencakup tiga bentuk utama, yaitu validitas isi, validitas konstruk, dan validitas kriteria. Pendapat tersebut diperkuat oleh Soeharto (2017) yang menjelaskan bahwa suatu instrumen pengukuran dapat ditinjau melalui tiga tipe validitas, meliputi validitas isi, validitas konstruk, dan validitas kriteria. Di samping itu, Siswanto (2008) menambahkan bahwa validitas tampak dapat digunakan sebagai salah satu bentuk penilaian terhadap instrumen dengan melihat kesesuaian secara umum yang tampak dari segi bentuk, penampilan, serta relevansinya dengan hal yang hendak diukur. Berdasarkan beberapa pandangan tersebut, dapat dikemukakan bahwa validitas yang banyak digunakan dalam penelitian pendidikan terdiri atas validitas isi, konstruk, kriteria, dan tampak. Masing-masing jenis memiliki karakteristik dan kegunaan yang berbeda, namun secara keseluruhan berperan dalam memastikan bahwa instrumen penelitian mampu mengukur variabel yang ditetapkan secara akurat dan selaras dengan tujuan pengukuran.

Kriteria validitas berfungsi sebagai panduan utama dalam mengevaluasi sejauh mana instrumen penelitian dapat diterima. Utami dan rekan-rekannya (2024)

menjelaskan bahwa hasil dari pengujian validitas dapat dikelompokkan ke dalam beberapa kategori, mulai dari tidak valid hingga sangat valid, yang mempermudah peneliti dalam menilai kualitas instrumen tersebut. Dalam proses validasi isi, setiap elemen instrumen akan dinilai berdasarkan seberapa sesuai dengan aspek yang ingin diukur, yang mencakup kategori tidak relevan, kurang relevan, cukup relevan, dan relevan (An Nabil dkk, 2022). Di sisi lain, Kurniawati (2021) mengemukakan bahwa sebuah instrumen dapat dianggap valid jika mendapatkan koefisien validitas yang mencapai batas minimal yang telah ditentukan, seperti indeks validitas di atas 0,75. Oleh karena itu, kriteria validitas menjadi landasan dalam menentukan mutu instrumen sekaligus menjadi pedoman untuk memutuskan apakah suatu butir soal dapat dipakai, perlu diperbaiki, atau harus dihilangkan sebelum digunakan dalam penelitian.

Reliabilitas adalah elemen yang sangat penting yang mencerminkan kualitas suatu alat penelitian. Menurut Khumaedi (2012), semakin tinggi koefisien reliabilitas suatu alat, maka semakin besar pula konsistensi dari hasil pengukurannya. Pandangan ini diperkuat oleh Subhaktiyasa (2024) yang menyatakan bahwa reliabilitas berfungsi sebagai tolok ukur kualitas alat dalam mengukur variabel penelitian dengan cara yang konsisten. Selanjutnya, Muharromah dan rekan-rekan (2025) menjelaskan bahwa reliabilitas terkait dengan seberapa andal dan stabilnya suatu alat ukur ketika diterapkan dalam kondisi yang serupa. Berdasarkan penjelasan tersebut, reliabilitas dapat diartikan sebagai kapasitas alat untuk memberikan hasil pengukuran yang stabil, konsisten, dan dipercaya, sehingga data yang dihasilkan memiliki tingkat keajegan yang baik dan dapat dijadikan sebagai dasar dalam penelitian.

Pengelompokan reliabilitas suatu instrumen dapat dilakukan berdasarkan cara atau prosedur yang diterapkan untuk mengetahui kestabilan hasil pengukurannya. Menurut Yusup (2018), tingkat keandalan instrumen dapat diperiksa melalui beberapa metode, seperti pengukuran berulang (*test-retest*), penggunaan dua bentuk instrumen yang setara (*equivalent*), serta pendekatan konsistensi internal (*internal consistency*). Pada pendekatan konsistensi internal sendiri terdapat sejumlah teknik analisis yang dapat diterapkan, di antaranya metode split-half, KR-20, KR-21, dan

Alpha Cronbach. Sementara itu, Sappaile (2007) memandang reliabilitas dari aspek kestabilan jawaban yang diberikan serta keteraturan hubungan yang terbentuk di antara butir-butir dalam suatu instrumen. Pendapat tersebut sejalan dengan Busnawir (2023) yang menjelaskan bahwa pendekatan konsistensi internal banyak dimanfaatkan dalam penelitian karena mampu menggambarkan keselarasan setiap butir ketika digunakan untuk mengukur konstruk atau konsep yang sama. Berdasarkan uraian dari beberapa ahli tersebut, bentuk reliabilitas yang umum diterapkan dalam penelitian meliputi metode test-retest, bentuk ekuivalen, konsistensi internal, konsistensi respons, serta konsistensi antarbutir. Keseluruhan pendekatan tersebut memiliki tujuan utama untuk memastikan bahwa instrumen yang digunakan mampu menghasilkan data pengukuran yang stabil, ajeg, dan memiliki tingkat kepercayaan yang memadai.

Koefisien reliabilitas menjadi salah satu ukuran yang digunakan untuk mengetahui sejauh mana sebuah instrumen penelitian dapat diandalkan dalam menghasilkan data yang konsisten. Hakim (2011) menjelaskan bahwa tinggi rendahnya koefisien reliabilitas tidak semata-mata ditentukan oleh kualitas dan konsistensi instrumen, tetapi juga dapat dipengaruhi oleh karakteristik responden atau sampel yang terlibat dalam proses pengukuran. Sementara itu, Yusup (2018) mengemukakan bahwa reliabilitas berkaitan dengan kemampuan suatu instrumen dalam memberikan hasil pengukuran yang dapat dipercaya apabila digunakan dalam penelitian. Sejalan dengan pandangan tersebut, Naga (2015) menyatakan bahwa koefisien reliabilitas dipengaruhi oleh keterkaitan atau hubungan antarbutir dalam instrumen serta pola skor yang diperoleh dari hasil pengukuran. Berdasarkan beberapa pendapat tersebut, koefisien reliabilitas dapat dipahami sebagai nilai yang menunjukkan tingkat keajegan dan keterandalan suatu instrumen dalam mengukur suatu variabel. Semakin baik tingkat reliabilitas yang dimiliki, semakin dapat dipercaya pula hasil pengukuran yang diperoleh untuk mendukung proses penelitian.

Teks prosedur adalah salah satu jenis teks yang berfungsi memberikan arahan atau petunjuk kepada pembaca dalam melaksanakan suatu aktivitas melalui tahapan yang tersusun secara teratur. Putri dkk. (2024) menjelaskan bahwa teks prosedur

memuat informasi faktual yang disampaikan melalui rangkaian tahapan atau langkah secara terperinci, runtut, dan sistematis. Sejalan dengan pendapat tersebut, Setiawati dkk. (2023) mengemukakan bahwa teks prosedur disusun dengan tujuan membantu pembaca memahami tata cara melakukan suatu kegiatan sehingga dapat melaksanakannya dengan tepat. Selain itu, Riyanti dkk. (2025) menekankan bahwa penguasaan materi teks prosedur penting bagi peserta didik, sebab pembelajaran tidak hanya berfokus pada pemahaman konsep, tetapi juga mencakup kemampuan mengenali berbagai jenis teks prosedur dan menghasilkan teks sesuai dengan ketentuan kebahasaan yang berlaku. Dengan demikian, berdasarkan beberapa pendapat tersebut, teks prosedur dapat dipahami sebagai teks yang menyajikan serangkaian tahapan secara runtut, logis, dan sistematis untuk memberikan panduan kepada pembaca dalam melakukan suatu aktivitas atau mencapai hasil yang telah ditentukan secara tepat dan efektif.

Struktur dalam teks prosedur dapat dipahami sebagai rangkaian bagian yang membentuk keseluruhan teks dan berfungsi mengarahkan pembaca dalam melaksanakan suatu aktivitas. Ayunisyah dkk. (2020) mengemukakan bahwa teks prosedur memiliki beberapa bagian utama, yaitu judul, tujuan, alat atau bahan yang diperlukan, serta tahapan kegiatan. Sejalan dengan pendapat tersebut, Nur'aini dkk. (2023) menjelaskan bahwa penyusunan teks prosedur mencakup judul, tujuan, alat dan bahan, kemudian diikuti oleh langkah-langkah kegiatan yang disusun secara teratur. Sementara itu, Meltrianda dan Zulfikarni (2024) menyatakan bahwa keempat bagian tersebut menjadi komponen penting yang mendasari pembentukan sebuah teks prosedur. Oleh karena itu, dapat disimpulkan bahwa teks prosedur pada umumnya dibangun oleh judul, tujuan, alat dan bahan, serta rangkaian tahapan kegiatan. Seluruh bagian tersebut disusun secara teratur dan berurutan agar pembaca dapat memahami serta melaksanakan setiap proses sesuai petunjuk hingga tujuan kegiatan dapat tercapai dengan tepat.

Kaidah kebahasaan pada teks prosedur berkaitan dengan penggunaan unsur-unsur bahasa tertentu yang digunakan untuk menyampaikan tata cara atau petunjuk dalam melaksanakan suatu aktivitas. Priyatno dkk. (2023) menjelaskan bahwa teks prosedur memiliki karakteristik bahasa yang ditandai dengan penggunaan kalimat

perintah, kata penghubung yang menunjukkan hubungan waktu, serta ungkapan yang menerangkan cara melakukan suatu tindakan. Penggunaan unsur tersebut berfungsi membantu pembaca memahami keterkaitan dan tahapan kegiatan yang disampaikan. Sejalan dengan hal tersebut, Fitriyani dan Mukhlis (2021) mengungkapkan bahwa kalimat imperatif banyak ditemukan dalam teks prosedur karena digunakan untuk menyampaikan instruksi atau mengarahkan pembaca agar melakukan tindakan tertentu. Di sisi lain, Arifah (2025) menyatakan bahwa penggunaan kaidah kebahasaan dalam teks prosedur diperlukan untuk menyusun petunjuk secara berurutan dan bertahap sehingga informasi yang diberikan dapat dipahami serta diterapkan dengan mudah. Berdasarkan beberapa pendapat tersebut, dapat disimpulkan bahwa teks prosedur memanfaatkan berbagai unsur kebahasaan, seperti kalimat imperatif, konjungsi temporal, dan keterangan cara, serta bentuk bahasa lainnya yang mendukung penyampaian tahapan kegiatan secara terstruktur, jelas, dan berurutan agar pembaca dapat melaksanakan kegiatan sesuai tujuan yang hendak dicapai.

METODE PENELITIAN

Jenis dan Desain Penelitian

Penelitian ini mengadopsi pendekatan kuantitatif dengan desain evaluatif yang bertujuan untuk menilai kualitas instrumen tes pada materi teks prosedur bagi peserta didik sekolah menengah pertama. Kajian penelitian difokuskan pada pengujian validitas setiap butir soal serta reliabilitas tes sebagai tolak ukur kelayakan instrumen yang digunakan dalam evaluasi pembelajaran. Pelaksanaan penelitian mencakup beberapa tahap, yaitu penyusunan instrumen, pelaksanaan uji coba soal, analisis validitas butir, dan pengujian reliabilitas guna memastikan kualitas instrumen yang dikembangkan.

Subjek Penelitian

Partisipan penelitian ini terdiri atas siswa sekolah menengah pertama yang sebelumnya telah mendapatkan pembelajaran mengenai teks prosedur. Sampel penelitian ditetapkan melalui teknik purposive sampling dengan

mempertimbangkan kesesuaian pengalaman belajar peserta didik terhadap materi yang menjadi dasar penyusunan instrumen tes. Banyaknya peserta didik yang dijadikan sampel ditentukan berdasarkan kebutuhan penelitian, khususnya untuk memenuhi keperluan analisis statistik dalam proses pengujian validitas dan reliabilitas instrumen. Penelitian ini juga melibatkan dua guru Bahasa Indonesia yang berperan sebagai validator ahli untuk memberikan penilaian terhadap instrumen yang telah disusun. Penilaian dilakukan dengan memperhatikan sejumlah aspek, meliputi keselarasan kisi-kisi dengan kompetensi yang hendak diukur, kesesuaian substansi materi, kualitas konstruksi soal, ketepatan penggunaan bahasa, dan keterkaitan butir soal dengan kompetensi pada materi teks prosedur. Masukan serta hasil penilaian dari para validator selanjutnya digunakan sebagai bahan evaluasi untuk merevisi dan menyempurnakan instrumen sebelum diterapkan dalam tahap uji coba kepada peserta didik.

Instrumen Penelitian

Penelitian ini memanfaatkan instrumen berupa tes objektif yang dikembangkan dengan mengacu pada kompetensi materi teks prosedur. Perumusan setiap butir soal didasarkan pada sejumlah indikator pembelajaran, meliputi kemampuan peserta didik dalam mengenali struktur teks prosedur, memahami isi dan tujuan yang disampaikan, mengidentifikasi karakteristik kebahasaan, serta menentukan dan memahami tahapan kegiatan yang tersusun dalam teks. Sebelum digunakan sebagai alat pengumpulan data, rancangan kisi-kisi instrumen terlebih dahulu diperiksa dan dinilai melalui proses validasi yang melibatkan ahli di bidang pembelajaran Bahasa Indonesia dan evaluasi pendidikan. Tahap validasi ini dilakukan untuk menilai kesesuaian antara butir soal dengan kompetensi yang menjadi sasaran pengukuran sekaligus memastikan bahwa instrumen telah memenuhi kriteria kelayakan untuk digunakan dalam penelitian.

Teknik Pengumpulan Data

Pengumpulan data dalam penelitian ini dilakukan melalui dua tahap, yaitu validasi instrumen dan uji coba instrumen tes. Pada tahap validasi, data diperoleh

menggunakan lembar validasi yang diisi oleh dua orang guru Bahasa Indonesia sebagai validator ahli. Para validator melakukan penilaian terhadap beberapa aspek, meliputi kesesuaian butir soal dengan indikator pembelajaran, ketepatan materi, kualitas konstruksi soal, penggunaan bahasa, serta tingkat keterbacaan soal. Hasil penilaian tersebut dijadikan dasar dalam merevisi dan menyempurnakan instrumen agar layak digunakan pada tahap berikutnya.

Setelah instrumen dinyatakan layak untuk diuji, tahap berikutnya dilakukan dengan mengimplementasikan tes kepada siswa sekolah menengah pertama yang sebelumnya telah memperoleh pembelajaran tentang teks prosedur. Dalam pelaksanaannya, setiap peserta didik mengerjakan seluruh soal yang tersedia dengan mengikuti arahan pengerjaan serta batas waktu yang telah ditetapkan. Jawaban yang diperoleh dari pelaksanaan tes selanjutnya dihimpun dan digunakan sebagai sumber data penelitian. Data tersebut kemudian diolah melalui analisis statistik untuk menentukan tingkat validitas pada masing-masing butir soal dan mengidentifikasi tingkat konsistensi atau reliabilitas instrumen secara keseluruhan. Tahapan pengumpulan data melalui pelaksanaan uji coba instrumen ini memiliki kesamaan dengan prosedur penelitian Hasanudin dan Asror (2017), yang menggunakan proses pengujian instrumen sebagai bagian dari pengembangan alat ukur untuk mengetahui prestasi belajar peserta didik.

Teknik Analisis Data

Uji Validitas Butir Soal

Pengujian validitas bertujuan untuk memastikan tingkat ketepatan instrumen dalam mengukur kompetensi peserta didik pada materi teks prosedur sesuai dengan sasaran pengukuran yang telah dirumuskan. Pada penelitian ini, validitas yang dikaji adalah validitas isi, yang penilaiannya melibatkan dua guru Bahasa Indonesia sebagai ahli atau validator. Penilaian terhadap instrumen dilakukan dengan menggunakan skala Likert berskala 1 sampai 5. Melalui skala tersebut, validator memberikan penilaian terhadap tingkat keselarasan dan relevansi setiap butir soal dengan aspek kompetensi yang menjadi fokus pengukuran. Skala tersebut juga telah digunakan

oleh Hasanudin dkk. (2025) dalam penelitian yang berkaitan dengan pengembangan instrumen tes membaca.

Hasil penilaian yang diberikan oleh para ahli selanjutnya diolah menggunakan indeks Aiken's V guna mengetahui tingkat keabsahan masing-masing butir dalam instrumen. Besarnya nilai koefisien yang semakin mendekati 1,00 menunjukkan adanya tingkat kesepakatan yang tinggi di antara validator mengenai kesesuaian serta mutu butir soal yang telah dievaluasi. Butir instrumen yang memperoleh nilai koefisien tinggi dapat dinyatakan telah memenuhi kriteria validitas yang ditetapkan. Sementara itu, butir yang memperoleh nilai dengan kategori validitas rendah atau sangat rendah perlu diperbaiki sebelum digunakan pada tahap berikutnya. Proses penyempurnaan dilakukan berdasarkan berbagai masukan, koreksi, maupun saran yang disampaikan oleh para validator. Setelah dilakukan perbaikan sesuai hasil evaluasi ahli, instrumen kemudian dapat dilanjutkan ke tahap pengujian kepada peserta didik. Metode analisis menggunakan rumus V Aiken ini juga pernah diterapkan oleh Hasanudin dkk. (2024) untuk menilai hasil validasi yang diberikan oleh ahli materi dan ahli media.

Uji Reliabilitas Tes

Analisis reliabilitas dalam penelitian ini bertujuan untuk mengetahui konsistensi instrumen dalam memberikan hasil pengukuran serta mengevaluasi kualitas tes yang digunakan untuk mengukur pemahaman peserta didik mengenai materi teks prosedur. Kualitas instrumen dianalisis berdasarkan beberapa aspek, yaitu tingkat kesukaran soal, daya pembeda, dan reliabilitas tes secara keseluruhan. Analisis tingkat kesukaran dilakukan dengan melihat proporsi peserta didik yang mampu menjawab setiap butir soal secara benar. Berdasarkan hasil tersebut, masing-masing soal diklasifikasikan ke dalam tiga kelompok tingkat kesukaran, yaitu mudah, sedang, dan sulit. Selanjutnya, analisis daya pembeda digunakan untuk mengetahui kemampuan setiap butir soal dalam membedakan peserta didik yang memiliki penguasaan materi tinggi dari peserta didik yang tingkat penguasaan materinya lebih rendah.

Pada bagian berikutnya, keandalan perangkat tes dianalisis melalui penerapan koefisien *Alpha Cronbach* dengan menggunakan data yang diperoleh setelah instrumen diujicobakan kepada peserta didik. Angka koefisien yang dihasilkan dapat dijadikan dasar untuk mengetahui sejauh mana keterkaitan dan kestabilan antarbutir dalam instrumen tersebut. Apabila nilai reliabilitas yang diperoleh semakin tinggi, maka tingkat konsistensi instrumen dalam menghasilkan skor pengukuran juga semakin baik. Kondisi tersebut menunjukkan bahwa instrumen memiliki kelayakan yang lebih tinggi untuk dimanfaatkan sebagai sarana penilaian dalam proses pembelajaran. Setelah itu, butir-butir pertanyaan yang belum mencapai batas kelayakan yang telah ditetapkan akan diperiksa kembali. Pemeriksaan dilakukan dengan mempertimbangkan tingkat kesulitan, kemampuan soal dalam membedakan peserta didik berdasarkan penguasaan materi, serta pengaruh setiap butir terhadap tingkat keandalan instrumen secara keseluruhan. Soal yang tidak memenuhi ketentuan dapat mengalami penyempurnaan atau tidak disertakan dalam perangkat tes akhir. Melalui tahapan tersebut, diharapkan tersusun instrumen yang dapat menghasilkan informasi pengukuran secara tepat, ajeg, dan memiliki dasar pertanggungjawaban ilmiah.

Tingkat kesukaran instrumen untuk soal pilihan ganda dihitung menggunakan rumus *indek kesukaran* (P) = $\frac{\text{banyak siswa yang menjawab benar (B)}}{\text{jumlah seluruh peserta tes (JS)}}$, Sementara itu tingkat kesukaran pada soal uraian diperoleh dengan *tingkat kesukaran* = $\frac{\text{rata-rata skor}}{\text{skor maksimum soal}}$. Selain itu, analisis daya beda dilakukan untuk mengetahui kemampuan setiap butir soal dalam membedakan peserta didik yang memiliki kemampuan tinggi dan rendah. Perhitungan daya beda menggunakan rumus $\text{Daya beda (D)} = \frac{BA}{JA} - \frac{BB}{JB} = PA - PB$.

Keterangan :

D	= Indeks daya beda.
BA	= Banyak peserta kelompok atas yang menjawab benar.
BB	= Banyak peserta kelompok bawah yang menjawab benar.
JA	= Jumlah peserta kelompok atas.
JB	= Jumlah peserta kelompok bawah.
PA	= Rasio peserta kelompok atas yang menjawab benar.
PB	= Rasio peserta kelompok bawah yang menjawab benar.

Kriteria Interpretasi

Hasil pengujian validitas dan reliabilitas kemudian ditafsirkan berdasarkan kriteria statistik yang lazim digunakan dalam penelitian pendidikan. Butir soal yang belum memenuhi standar validitas akan diperbaiki atau dikeluarkan dari instrumen, sedangkan instrumen dapat dinyatakan reliabel apabila memperoleh koefisien reliabilitas yang berada pada kategori tinggi. Dengan demikian, instrumen yang telah melalui proses pengujian diharapkan memiliki tingkat ketepatan dan konsistensi yang baik sehingga mampu digunakan sebagai alat ukur yang efektif dalam menilai kemampuan peserta didik pada materi teks prosedur.

HASIL DAN PEMBAHASAN

Untuk mengukur kemampuan siswa dalam memahami teks narasi dapat digunakan instrumen tes berikut.

1. Tabel Tabel nilai

No	Instrumen Soal	Bentuk Soal	Soal Ke-
1.	Mengidentifikasi pengertian teks prosedur.	Pilihan Ganda	Soal 1
2.	Menentukan tujuan teks prosedur.	Pilihan Ganda	Soal 2
3.	Mengidentifikasi ciri teks prosedur.	Pilihan Ganda	Soal 3
4.	Menentukan struktur teks prosedur.	Pilihan Ganda	Soal 4
5.	Mengidentifikasi bagian tujuan.	Pilihan Ganda	Soal 5
6.	Mengidentifikasi langkah.	Pilihan Ganda	Soal 6
7.	Menentukan kata kerja imperatif.	Pilihan Ganda	Soal 7
8.	Menentukan kalimat perintah.	Pilihan Ganda	Soal 8
9.	Menentukan urutan langkah.	Pilihan Ganda	Soal 9
10.	Mengidentifikasi konjungsi.	Pilihan Ganda	Soal 10
11.	Menganalisis struktur.	Pilihan Ganda	Soal 11
12.	Menganalisis bahasa.	Pilihan Ganda	Soal 12
13.	Membedakan kalimat efektif.	Pilihan Ganda	Soal 13
14.	Menyimpulkan isi.	Pilihan Ganda	Soal 14
15.	Menentukan alat.	Pilihan Ganda	Soal 15
16.	Menentukan urutan benar.	Pilihan Ganda	Soal 16
17.	Mengaitkan Direct Instruction.	Pilihan Ganda	Soal 17
18.	Membandingkan teks.	Pilihan Ganda	Soal 18
19.	Menilai judul.	Pilihan Ganda	Soal 19
20.	Menilai bahasa.	Pilihan Ganda	Soal 20

21.	Menilai kualitas.	Pilihan Ganda	Soal 21
22.	Analisis teks prosedur.	Pilihan Ganda	Soal 22
23.	Mengidentifikasi fakta.	Pilihan Ganda	Soal 23
24.	Mengidentifikasi proses	Pilihan Ganda	Soal 24
25.	Menilai langkah.	Pilihan Ganda	Soal 25
26.	Menjelaskan pengertian.	Esai	Soal 26
27.	Menjelaskan struktur.	Esai	Soal 27
28.	Menganalisis langkah.	Esai	Soal 28
29.	Menyusun teks.	Esai	Soal 29
30.	Perbandingan.	Esai	Soal 30

Berdasarkan instrumen di atas, soal tes dapat disusun dalam dua bentuk, yaitu pilihan ganda sebanyak 25 dan esai sebanyak 5. Setelah soal tersusun, kemudian dilakukan uji validitas dengan melibatkan dua validator. Adapun hasil validasi dapat dilihat pada tabel berikut.

Tabel 2. Tabel nilai

No Butir	Validator		S1	S2	$\sum S$	V	Ketegori
	1	2					
1	3	4	2	3	7	0,625	Tinggi
2	2	4	1	3	7	0,5	Cukup
3	4	4	3	3	7	0,75	Tinggi
4	3	4	2	3	6	0,625	Tinggi
5	5	5	4	4	8	1	Sangat Tinggi
6	4	4	3	3	7	0,75	Tinggi
7	5	4	4	3	8	0,875	Sangat Tinggi
8	5	5	4	4	7	1	Sangat Tinggi
9	4	4	3	3	6	0,75	Tinggi
10	5	5	4	4	7	1	Sangat Tinggi
11	5	4	4	3	8	0,875	Sangat Tinggi
12	4	4	3	3	8	0,75	Tinggi
13	5	5	4	4	8	1	Sangat Tinggi
14	5	5	4	4	7	1	Sangat Tinggi
15	4	4	3	3	8	0,75	Tinggi
16	2	4	1	3	7	0,5	Cukup
17	3	5	2	4	8	0,75	Tinggi
18	3	4	2	3	6	0,625	Tinggi
19	2	4	1	3	8	0,5	Cukup
20	4	4	3	3	8	0,75	Tinggi
21	4	4	3	3	8	0,75	Tinggi
22	5	5	4	4	7	1	Sangat Tinggi
23	3	4	2	3	7	0,625	Tinggi

24	4	4	3	3	7	0,75	Tinggi
25	3	4	2	3	6	0,625	Tinggi
26	5	5	4	4	7	1	Sangat Tinggi
27	5	4	4	3	7	0,875	Sangat Tinggi
28	3	4	2	3	6	0,625	Tinggi
29	5	4	4	3	5	0,875	Sangat Tinggi
30	5	4	4	3	8	0,875	Sangat Tinggi

Berdasarkan hasil perhitungan yang telah dilakukan, validator menyatakan bahwa instrumen tes yang terdiri atas 30 butir soal, yaitu 25 soal pilihan ganda dan 5 soal uraian, layak untuk diujikan kepada peserta didik. Kelayakan tersebut didasarkan pada hasil penilaian validator yang menunjukkan bahwa seluruh butir soal memperoleh kategori cukup, tinggi, atau sangat tinggi. Pada tabel terlihat bahwa soal yang memperoleh kategori cukup yaitu nomor 2, 16, dan 19. Selanjutnya, soal yang termasuk dalam kategori tinggi adalah nomor 1, 3, 4, 6, 9, 12, 15, 17, 18, 20, 21, 23, 24, 25, dan 28. Adapun soal yang memperoleh kategori sangat tinggi meliputi nomor 5, 7, 8, 10, 11, 13, 14, 22, 26, 27, 29, dan 30.

Tahap berikutnya adalah pelaksanaan uji coba instrumen tes pada kelas uji coba yang melibatkan siswa sekolah menengah pertama. Hasil pelaksanaan uji coba tersebut disajikan pada tabel berikut.

Tabel 3. Tabel nilai

No Butir	P	Kriteria	D	Kriteria	Variasi butir	pi	qi	pi qi
1	1	Mudah	0	Jelek	0	1	0	0
2	1	Mudah	0	Jelek	0	1	0	0
3	0,5	Sedang	1	Baik	0,33	0,5	0,5	0,25
4	1	Mudah	0	Jelek	0	1	0	0
5	0,75	Mudah	-0,5	Jelek	0,25	0,75	0,25	0,19
6	1	Mudah	0	Jelek	0	1	0	0
7	0,25	Sukar	-0,5	Jelek	0,25	0,25	0,75	0,19
8	1	Mudah	0	Jelek	0	1	0	0
9	1	Mudah	0	Jelek	0	1	0	0
10	0,5	Sedang	0	Jelek	0,33	0,5	0,5	0,25
11	1	Mudah	0	Jelek	0	1	0	0
12	1	Mudah	0	Jelek	0	1	0	0
13	0,25	Sukar	0,5	Baik	0,25	0,25	0,75	0,19
14	1	Mudah	0	Jelek	0	1	0	0
15	1	Mudah	0	Jelek	0	1	0	0

16	1	Mudah	0	Jelek	0	1	0	0
17	1	Mudah	0	Jelek	0	1	0	0
18	1	Mudah	0	Jelek	0	1	0	0
19	0,25	Sukar	0,5	Baik	0,25	0,25	0,75	0,19
20	0,75	Mudah	-0,5	Jelek	0,25	0,75	0,25	0,19
21	1	Mudah	0	Jelek	0	1	0	0
22	0,75	Mudah	0,5	Baik	0,25	0,75	0,25	0,19
23	0	Sukar	0	Jelek	0	0	1	0
24	0,5	Sedang	0	Jelek	0,33	0,5	0,5	0,25
25	1	Mudah	0	Jelek	0	1	0	0
26	1	Mudah	0	Jelek	0	1	0	0
27	1	Mudah	0	Jelek	0	1	0	0
28	1	Mudah	0	Jelek	0	1	0	0
29	1	Mudah	0	Jelek	0	1	0	0
30	0,5	Sedang	1	Baik	0,33	0,5	0,5	0,25

Berdasarkan pelaksanaan uji coba lapangan, diperoleh data hasil tes dengan variasi skor sebagaimana disajikan pada tabel sebelumnya. Analisis terhadap tingkat kesukaran memperlihatkan beberapa butir soal tergolong dalam kategori mudah. Sementara itu, hasil perhitungan daya pembeda memperlihatkan bahwa sebanyak 25 butir soal berada pada kategori jelek, sedangkan 5 butir soal lainnya termasuk dalam kategori baik.

Tahap berikutnya adalah pengujian reliabilitas instrumen menggunakan rumus *Kuder-Richardson 20* (KR-20). Berdasarkan hasil perhitungan, diperoleh koefisien reliabilitas sebesar 0,48. Nilai tersebut menunjukkan bahwa reliabilitas instrumen masih berada pada kategori rendah, sehingga tingkat konsistensi instrumen belum dapat dikatakan memadai karena koefisien yang diperoleh belum mendekati nilai ideal, yaitu 1,00. Dengan mempertimbangkan hasil analisis tingkat kesukaran, daya pembeda, dan reliabilitas, hanya 5 butir soal yang memenuhi kriteria baik sehingga dinyatakan layak digunakan sebagai instrumen penelitian. Adapun butir-butir soal yang memenuhi kriteria tersebut setelah melalui tahap uji coba disajikan pada bagian berikut.

Tabel 4. Tabel soal

No Butir	Soal Tes Teks Narasi
3	Ciri utama teks prosedur adalah... A. Imajinatif B. Mengandung langkah

	C. Subjektif
	D. Bebas
13	Kalimat efektif adalah.... A. Masukan gula lalu kemudian B. Masukkan gula C. Gula dimasukkan oleh kita D. Kita memasukkan gula sekali
19	Judul tidak sesuai jika... A. Sama isi B. Berbeda isi C. Singkat D. Panjang
22	Perhatikan teks berikut! "Untuk membuat minuman teh manis, pertama siapkan gelas, teh, gula, dan air panas. Masukkan teh ke dalam gelas, lalu tuangkan air panas. Setelah itu, tambahkan gula sesuai selera dan aduk hingga larut." Berdasarkan teks tersebut, analisislah pernyataan yang paling tepat! A. Teks tersebut tidak memiliki tujuan yang jelas B. Langkah-langkah sudah sistematis dan mudah diikuti C. Teks tersebut termasuk teks narasi D. Bahasa yang digunakan bersifat opini
30	Bandingkan teks prosedur sederhana dan kompleks, jelaskan perbedaanya!

Kualitas reliabilitas sebuah instrumen dapat diketahui melalui nilai koefisien yang diperoleh setelah dilakukan pengujian dan dibandingkan dengan kriteria yang telah ditentukan. Setiyawan (2020) mengemukakan bahwa tingkat reliabilitas suatu instrumen tidak hanya ditentukan oleh satu aspek, tetapi dapat dipengaruhi oleh berbagai komponen, seperti karakteristik butir soal, keragaman kemampuan peserta didik yang menjadi kelompok pengukuran, objektivitas proses pemberian skor, tingkat kesulitan tes, serta keseragaman isi instrumen. Rendahnya koefisien reliabilitas dapat menunjukkan bahwa masih terdapat kelemahan dalam instrumen, misalnya jumlah butir yang belum mencukupi atau adanya soal yang kurang relevan dengan kompetensi yang hendak diukur. Sejalan dengan hal tersebut, Muharromah dkk. (2025) menjelaskan bahwa konsistensi instrumen juga berkaitan dengan jumlah soal yang digunakan, kejelasan instruksi pengerjaan, bentuk atau format tes, serta

kondisi peserta didik ketika proses pengukuran berlangsung. Faktor lain yang turut mempengaruhi hasil reliabilitas dikemukakan oleh Setiyorini dkk. (2022), meliputi tingkat pemahaman peserta didik terhadap materi, kesiapan sebelum mengikuti tes, kondisi fisik, kejelasan penyusunan kalimat dalam soal, situasi saat tes dilaksanakan, dan ketersediaan waktu pengerjaan. Dengan demikian, seluruh aspek tersebut perlu dipertimbangkan secara cermat dalam penyusunan dan pelaksanaan instrumen agar data yang dihasilkan tetap konsisten, tepat, dan dapat dipercaya sebagai hasil pengukuran.

SIMPULAN

Hasil evaluasi terhadap kelayakan instrumen tes pada materi teks prosedur menunjukkan bahwa dari keseluruhan 30 soal yang dianalisis, diperoleh variasi tingkat validitas pada setiap butir. Sebanyak 3 soal termasuk dalam kategori cukup, kemudian 15 soal berada pada tingkat validitas tinggi, sedangkan 12 soal lainnya mencapai kategori sangat tinggi. Berbeda dengan hasil validitas tersebut, analisis reliabilitas instrumen memperoleh nilai koefisien sebesar 0,48. Nilai tersebut mengindikasikan bahwa konsistensi instrumen masih tergolong rendah karena koefisien yang dihasilkan belum mencapai tingkat yang mendekati nilai 1. Setelah seluruh aspek penilaian diperhitungkan, yang mencakup validitas, tingkat kesukaran, daya pembeda, dan reliabilitas, diketahui bahwa hanya terdapat 5 butir soal yang memenuhi kriteria kualitas yang ditetapkan. Dengan demikian, kelima butir tersebut dapat dipertimbangkan untuk digunakan sebagai instrumen dalam kegiatan evaluasi pembelajaran.

REFERENSI

- An Nabil, N. R., Wulandari, I., Yamtinah, S., Ariani, S. R. D., & Ulfa, M. (2022). Analisis indeks Aiken untuk mengetahui validitas isi instrumen asesmen kompetensi minimum berbasis konteks sains kimia. *Paedagogia*, 25(2), 184–195.
- Arifah, S. (2025). Karakteristik kaidah kebahasaan teks prosedur dalam aplikasi Cookpad: Resep masakan harian. *Secondary: Jurnal Inovasi Pendidikan Menengah*, 6(1). <https://doi.org/10.51878/secondary.v6i1.8759>

- Ayunisyah, S. D., Arifin, M., & Yulistio, D. (2020). Analisis struktur teks prosedur siswa kelas VII SMPN 7 Kota Bengkulu. *Jurnal Ilmiah KORPUS*, 4(1), 118–127. <https://doi.org/10.33369/jik.v4i1.8346>
- Busnawir. (2023). Konsistensi internal instrumen tes: Perbandingan beberapa metode estimasi berdasarkan variasi ukuran sampel. *Jurnal Pendidikan Matematika*, 14(2). <https://doi.org/10.36709/jpm.v14i2.100>
- Fitriyani, K., & Mukhlis, M. (2021). Kalimat imperatif dalam teks prosedur. *Deiksis*, 13(3), 241–252. <https://doi.org/10.30998/deiksis.v13i3.7024>
- Hakim, M. A. (2011). Koefisien reliabilitas sebagai karakteristik sampel: Pengujian dengan instrument *Portrait Values Questionnaire (PVQ)* dari Schwartz. *Psikoislamika: Jurnal Psikologi dan Psikologi Islam*, 8(2). <https://doi.org/10.18860/psi.v0i0.1550>
- Hasanudin, C., & Asror, A. G. (2017). Efektivitas model pembelajaran quantum learning dengan media aplikasi bamboomedia bmgames apps terhadap keterampilan membaca permulaan pada siswa kelas I MI Se-Kecamatan Kedungadem. *Pedagogia: Jurnal Pendidikan*, 6(2), 150–159. <https://doi.org/10.21070/pedagogia.v6i2.907>
- Hasanudin, C., Fitrianiingsih, A., Fitriyana, N., & Ulfaida, N. (2024). Design and validity of local-wisdom-based reading apps using Adobe Animate CC 2022. *International Journal of Information and Education Technology*, 14(1), 1-11. <https://doi.org/10.18178/ijiet.2024.14.1.2017>
- Hasanudin, C., Fitrianiingsih, A., Zulaeha, I., Fitriyana, N., & Saddhono, K. (2025). Si Raca App in Quantum Learning, Is It Effective to be Implemented in Early Reading Material for Primary School. *International Research Journal of Multidisciplinary Scope (IRJMS)*, 6(1), 383-394. <https://doi.org/10.47857/irjms.2025.v06i01.01875>
- Hastjarjo, T. D. (2011). Validitas eksperimen. *Buletin Psikologi*, 19(2), 70–80. <https://doi.org/10.22146/bps.11558>
- Khumaedi, M. (2012). Reliabilitas instrumen penelitian pendidikan. *Jurnal Pendidikan Teknik Mesin*, 12(1). <https://doi.org/10.15294/jptm.v12i1.5273>
- Kurniawati, K. (2021). Analisis validitas isi instrumen tes berpikir kritis IPS kelas V SD Kota Yogyakarta. *Pelita: Jurnal Penelitian dan Karya Ilmiah*, 21(1), 130–140.
- Meltrianda, F., & Zulfikarni, Z. (2024). Struktur dan kebahasaan teks prosedur siswa kelas VII MTsN 1 Kota Padang. *Jurnal Pendidikan Tambusai*, 8(3), 41587–41597. <https://doi.org/10.30743/bahastra.v8i1.7096>
- Muharromah, F., Al Fani, F., & Ridho, A. R. (2025). Reliabilitas alat ukur: Jenis-jenis, cara pengukuran, dan faktor-faktor yang mempengaruhinya dalam

- pembelajaran. *JIPDAS (Jurnal Ilmiah Pendidikan Dasar)*, 4(1). <https://doi.org/10.65946/232k8198>
- Muharromah, F., Al Fani, F., & Ridho, A. R. (2025). Reliabilitas alat ukur: Jenis-jenis, cara pengukuran, dan faktor-faktor yang mempengaruhinya dalam pembelajaran. *JIPDAS (Jurnal Ilmiah Pendidikan Dasar)*, 4(1). <https://doi.org/10.65946/232k8198>
- Mukhlisa, N. (2023). Validitas tes. *Juara SD: Jurnal Pendidikan dan Pembelajaran Sekolah Dasar*, 2(1). <https://doi.org/10.26858/jppsd.v2i1.46314>
- Naga, D. S. (2015). Peranan interkorelasi butir terhadap koefisien reliabilitas Cronbach Alpha dan Kuder-Richardson. *Jurnal Kependidikan*, 34(1). <https://doi.org/10.21831/jk.v34i1.7283>
- Nur'aini, A., Septiana, I., & Mukhlis, M. (2023). Analisis deskriptif struktur teks prosedur hasil pembelajaran peserta didik SMP kelas VIIF. *Bahastra: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 8(1), 15–20. <https://doi.org/10.30743/bahastra.v8i1.7096>
- Priyatno, T., Marni, S., & Yulianti, U. (2023). Analisis kaidah kebahasaan teks prosedur siswa kelas VII SMPN 3 Bayang. *Alinea: Jurnal Bahasa, Sastra dan Pengajaran*, 3(2), 434–441. <https://doi.org/10.58218/alinea.v3i2.522>
- Putri, N. K., Sutri, S., & Syafroni, R. N. (2024). Analisis struktur teks prosedur dalam aplikasi Lemon8. *Bahtera Indonesia: Jurnal Penelitian Bahasa dan Sastra Indonesia*, 9(1), 227–233. <https://doi.org/10.31943/bi.v9i1.599>
- Rinaldi, R., & Albina, M. (2025). Analisis konsep, jenis, dan kelayakan instrumen penelitian pendidikan. *QAZI: Journal of Islamic Studies*, 2(1), 225–232.
- Riyanti, A., Tarigan, S. A., Jannah, A. M., Mulawarman, N., & Selviana, A. N. (2025). Analisis pemahaman peserta didik kelas VIII SMP Negeri 1 Kota Tarakan terhadap pembelajaran menulis teks prosedur. *IdeBahasa*, 7(1). <https://doi.org/10.37296/idebahasa.v7i1.308>
- Rosa, Y. (2008). Validitas instrumen ukur variabel sosial bidang permukiman. *Jurnal Permukiman*, 3(4), 264–279. <https://doi.org/10.31815/jp.2008.3.264-279>
- Sappaile, B. I. (2007). Konsep instrumen penelitian pendidikan. *Jurnal Pendidikan dan Kebudayaan*, 13(66), 379–391. <https://doi.org/10.24832/jpnk.v13i66.356>
- Setiawati, S., Hilaliyah, H., & Lapasau, M. (2023). Teks prosedur pada kemasan produk makanan dan implikasinya terhadap pembelajaran Bahasa Indonesia. *Hortatori: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 7(2). <https://doi.org/10.30998/jh.v7i2.2490>
- Setiyawan, A. (2020). Faktor-faktor yang mempengaruhi reliabilitas tes. *AN NUR: Jurnal Studi Islam*, 6(2). Retrieved from <https://mail.jurnalannur.ac.id/index.php/An-Nur/article/view/53>

- Setiyorini, T. J., Jaelani, Z. R., & Ngafif, A. (2022). Analisis faktor-faktor yang mempengaruhi reliabilitas tes tata bahasa Inggris di universitas di Indonesia. *Didaktis: Jurnal Pendidikan dan Ilmu Pengetahuan*, 22(3). <https://doi.org/10.30651/didaktis.v22i3.11286>
- Siswanto, S. (2008). Validitas sebagai alat penentuan kehandalan tes hasil belajar. *Jurnal Pendidikan Akuntansi Indonesia*, 6(1). <https://doi.org/10.21831/jpai.v6i1.1795>
- Soeharto. (2017). *Validitas konstruk dalam penelitian pendidikan*. <https://doi.org/10.13140/RG.2.2.12356.22401>
- Subhaktiyasa, P. G. (2024). Evaluasi validitas dan reliabilitas instrumen penelitian kuantitatif: Sebuah studi pustaka. *Journal of Education Research*, 5(4), 5599–5609. <https://doi.org/10.37985/jer.v5i4.1747>
- Utami, L., Festiyed, Ilahi, D. P., Ratih, A., Yenti, E., & Lazulva. (2024). Analisis indeks Aiken untuk mengetahui validitas isi instrumen *Scientific Habits of Mind*. *Journal of Research and Education Chemistry*, 6(1). [https://doi.org/10.25299/jrec.2024.vol6\(1\).17430](https://doi.org/10.25299/jrec.2024.vol6(1).17430)
- Yusup, F. (2018). Uji validitas dan reliabilitas instrumen penelitian kuantitatif. *Tarbiyah: Jurnal Ilmiah Kependidikan*, 7(1), 17–23. <https://doi.org/10.18592/tarbiyah.v7i1.2100>
- Yusup, F. (2018). Uji validitas dan reliabilitas instrumen penelitian kuantitatif. *Tarbiyah: Jurnal Ilmiah Kependidikan*, 7(1), 17–23. <https://doi.org/10.18592/tarbiyah.v7i1.2100>