



Prosiding Seminar Nasional

Unit Kegiatan Mahasiswa Penalaran dan Riset
IKIP PGRI Bojonegoro

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



Uji Validitas Butir dan Reliabilitas Tes Teks Narasi untuk Peserta Didik Sekolah Menengah Pertama

Azmi Laila Azzahro^{1(□)}, Cahyo Hasanuddin²

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

azmilaila106@gmail.com¹, cahyo.hasanudin@ikip PGRI Bojonegoro.ac.id²

abstrak – Kemampuan siswa memahami teks narasi hanya dapat terukur secara akurat apabila menggunakan instrumen yang telah teruji kevalidan dan keandalannya. Penelitian ini bertujuan menguji validitas butir soal sekaligus reliabilitas tes teks narasi bagi peserta didik sekolah menengah pertama. Pendekatan yang digunakan bersifat kuantitatif dengan desain evaluatif, melibatkan peserta didik SMP yang telah menerima materi teks narasi sebagai subjek, yang dipilih melalui teknik purposive sampling. Instrumen yang dikembangkan terdiri atas 30 butir soal, mencakup 25 soal pilihan ganda dan 5 soal uraian. Kevalidan isi diperiksa lewat penilaian dua validator ahli dengan skala Likert yang kemudian dihitung menggunakan rumus V Aiken, sementara keandalannya dianalisis melalui koefisien Alpha Cronbach. Temuan penelitian memperlihatkan bahwa dari 30 butir yang diujikan, 9 butir berada pada kategori validitas sangat tinggi, 15 butir pada kategori tinggi, dan 6 butir pada kategori cukup. Hasil analisis daya pembeda mengungkap 10 butir soal yang memenuhi kriteria baik, yakni nomor 2, 3, 7, 9, 13, 17, 18, 24, 26, dan 28. Adapun perhitungan reliabilitas menghasilkan nilai 0,75 yang mengindikasikan konsistensi tinggi. Berdasarkan temuan tersebut, instrumen tes yang dikembangkan terbukti valid, reliabel, dan layak dipakai untuk mengukur kemampuan siswa dalam memahami teks narasi.

Kata kunci – Peserta Didik SMP, Teks Narasi, Reliabilitas.

Abstract – A trustworthy measurement of students' narrative-text comprehension can only be obtained through a test instrument that has been proven valid and reliable. This study set out to examine the item validity and reliability of a narrative-text test designed for junior high school students. A quantitative, evaluative-design approach was adopted, with junior high school students who had already covered narrative-text material as the subjects, selected through purposive sampling. The instrument comprised 30 items: 25 multiple-choice questions and 5 essay questions. Content validity was checked by two expert validators using a Likert scale and computed with the Aiken's V formula, while reliability was determined through the Cronbach Alpha coefficient. Findings show that among the 30 items tested, 9 fell into the very-high validity category, 15 into the high category, and 6 into the adequate category. Item-discrimination analysis identified 10 items meeting the good criterion, namely items 2, 3, 7, 9, 13, 17, 18, 24, 26, and 28. The reliability coefficient reached 0.75, reflecting a high degree of consistency. These results confirm that the developed test instrument is valid, reliable, and appropriate for measuring junior high school students' ability to comprehend narrative texts.

Keywords – Junior High School Students, Narrative Text, Reliability.

PENDAHULUAN

Di jenjang sekolah menengah pertama (SMP), teks narasi menempati posisi penting dalam pembelajaran bahasa Indonesia karena menyuguhkan rangkaian peristiwa yang disusun secara kronologis untuk menghibur sekaligus memberikan pengalaman estetik bagi pembaca (Ramadhan et al., 2024). Materi ini menuntut peserta didik memahami pengertian, ciri, dan struktur teks narasi, sekaligus mampu memproduksinya sesuai kompetensi dasar yang berlaku dalam kurikulum. Struktur tersebut tersusun atas orientasi, komplikasi, resolusi, dan koda secara sistematis (Nandini et al., 2024). Karena cakupan materi yang harus dikuasai peserta didik SMP tergolong kompleks, diperlukan instrumen evaluasi yang valid dan reliabel agar kemampuan mereka dapat terukur secara akurat (Subhaktiyasa, 2024).

Fase perkembangan kognitif peserta didik SMP menuntut pendekatan evaluasi yang terstandar dan dapat diukur secara objektif, sehingga kemampuan mereka memahami teks narasi perlu diuji dengan instrumen yang disusun secara sistematis. Yusup (2018) mengartikan validitas sebagai ukuran yang memperlihatkan sejauh mana suatu instrumen dinyatakan valid. Ramadhan dkk. (2024) menambahkan bahwa validitas menggambarkan ketepatan instrumen dalam mengukur konstruk yang dituju. Adapun Subhaktiyasa (2024) memandang validitas instrumen sebagai landasan untuk memastikan data yang terkumpul benar-benar merepresentasikan variabel yang diteliti, terutama dalam konteks kemampuan berbahasa peserta didik SMP.

Pengembangan instrumen tes teks narasi bagi peserta didik SMP perlu memperhatikan tujuan uji validitas secara utuh, yakni memastikan setiap butir soal benar-benar mengukur kemampuan yang ditargetkan sesuai kompetensi peserta didik SMP. Wiranata dkk. (2024) menyebutkan bahwa uji validitas pada dasarnya diarahkan untuk menghasilkan instrumen yang mampu mengukur pemahaman peserta didik secara tepat. Nandini dkk. (2024) menambahkan bahwa uji validitas bertujuan mengembangkan alat ukur yang valid dan reliabel guna menilai capaian belajar peserta didik. Sementara itu, Hasanudin dkk. (2024) menekankan bahwa uji

validitas dimaksudkan untuk memastikan kesesuaian antara isi instrumen dan kompetensi yang hendak diukur pada peserta didik SMP.

Di samping validitas, reliabilitas juga merupakan unsur krusial dalam pengembangan instrumen tes teks narasi bagi peserta didik SMP, sebab reliabilitas menunjukkan seberapa konsisten hasil pengukuran suatu instrumen apabila digunakan berulang kali. Ramadhan dkk. (2024) menyatakan bahwa reliabilitas menjadi salah satu penanda utama instrumen pengukuran yang baik, yaitu kemampuannya memberikan hasil konsisten pada kelompok yang sama meski diujikan pada waktu berbeda. Subhaktiyasa (2024) memandang reliabilitas berkaitan dengan sejauh mana hasil pengukuran dapat dipercaya dan tetap konsisten. Yusup (2018) pun menegaskan bahwa instrumen dikategorikan reliabel jika hasil pengukurannya tidak berubah meskipun digunakan berkali-kali. Oleh karena itu, reliabilitas menjadi jaminan mutu instrumen evaluasi bagi peserta didik SMP pada materi teks narasi.

Pengujian reliabilitas pada instrumen tes teks narasi bagi peserta didik SMP dimaksudkan untuk memastikan keandalan sekaligus konsistensi alat ukur yang dipakai. Nandini dkk. (2024) menjelaskan bahwa uji reliabilitas berfungsi mengetahui apakah instrumen yang dikembangkan memiliki konsistensi internal memadai sehingga layak dipakai sebagai alat evaluasi bagi peserta didik SMP. Wiranata dkk. (2024) menambahkan bahwa uji reliabilitas bertujuan memastikan instrumen dapat diandalkan dalam memberikan hasil yang stabil. Adapun Hasanudin dkk. (2023) berpendapat bahwa reliabilitas instrumen yang tinggi mencerminkan kualitas pengukuran yang baik sehingga data yang dihasilkan dapat dipertanggungjawabkan secara ilmiah dalam evaluasi pembelajaran peserta didik SMP.

Teks narasi yang menjadi objek instrumen dalam penelitian ini adalah teks yang memaparkan rangkaian peristiwa secara kronologis, dengan tujuan menghibur maupun menyampaikan pengalaman kepada pembaca, khususnya dalam konteks pembelajaran peserta didik SMP. Alimah dan Indihadi (2022) memaknai teks narasi sebagai karangan yang menceritakan suatu kejadian sedemikian rupa sehingga pembaca seolah turut mengalaminya. Munjayana (2024) menambahkan bahwa teks narasi berfungsi menyampaikan cerita secara berurutan dari awal hingga akhir

dengan menghadirkan konflik beserta penyelesaiannya. Sementara itu, Alber dkk. (2023) memandang teks narasi sebagai teks yang memuat unsur tokoh, latar, dan alur yang saling terkait untuk membangun cerita bermakna bagi peserta didik SMP.

Dalam pembelajaran di SMP, teks narasi memiliki fungsi menghibur, menyampaikan pengalaman estetis, sekaligus menanamkan nilai moral kepada peserta didik. Alimah dan Indihadi (2022) menyatakan bahwa tujuan utama teks narasi adalah menghibur pembaca sembari menyampaikan pesan lewat rangkaian peristiwa yang dikisahkan. Munjayana (2024) berpendapat bahwa teks narasi turut mengembangkan kemampuan berbahasa peserta didik SMP melalui aktivitas menulis pengalaman atau cerita. Alber dkk. (2023) menambahkan bahwa teks narasi juga melatih kreativitas peserta didik SMP dalam mengungkapkan ide dan gagasan secara tertulis. Struktur teks narasi sendiri tersusun atas orientasi, komplikasi, resolusi, dan koda secara sistematis, yang wajib dikuasai peserta didik SMP (Nandini et al., 2024).

Urgensi pengujian validitas dan reliabilitas instrumen evaluasi bagi peserta didik SMP semakin banyak ditegaskan dalam kajian-kajian mutakhir. Sobiri dan Nashrullah (2026) menerangkan bahwa instrumen yang tidak valid maupun tidak reliabel berisiko memunculkan kesalahan pengukuran (*measurement error*) yang dapat memicu bias hasil penelitian, sehingga uji validitas dan reliabilitas menjadi tahap yang tidak boleh dilewatkan dalam penelitian kuantitatif, khususnya penelitian pendidikan yang banyak mengangkat variabel abstrak seperti kemampuan memahami teks narasi pada peserta didik SMP. Sejalan dengan pandangan tersebut, Zayrin dkk. (2025) menegaskan bahwa keandalan dan ketepatan instrumen merupakan faktor krusial dalam menghasilkan data berkualitas, sehingga pengembangan instrumen yang baik semestinya diawali dengan perumusan konstruk yang jelas, penyusunan kisi-kisi, validasi ahli, serta uji coba empiris melalui teknik statistik yang sesuai.

Rumus V Aiken menjadi salah satu teknik yang lazim dipakai untuk menguji validitas isi instrumen teks narasi bagi peserta didik SMP. Dalam penelitiannya terhadap instrumen *Scientific Habits of Mind*, Utami dkk. (2024) menerangkan bahwa indeks validitas Aiken dihitung dari penilaian sejumlah validator ahli berdasarkan kriteria skala tertentu, yang hasilnya kemudian dikelompokkan mulai dari kategori

tidak valid hingga sangat valid. Indeks Aiken dinilai efektif karena mampu memotret tingkat kesepakatan antarvalidator terhadap kualitas tiap butir instrumen secara objektif dan terukur, sehingga relevan diterapkan dalam pengembangan instrumen tes teks narasi bagi peserta didik SMP.

Berbagai penelitian mengenai uji validitas dan reliabilitas instrumen pemahaman teks narasi bagi peserta didik SMP telah dilakukan pada beragam jenis teks. Riana dan Hasanudin (2024), misalnya, menguji validitas dan reliabilitas instrumen pemahaman materi teks berita pada peserta didik SMP kelas VII dan menemukan bahwa mayoritas butir soal berada pada kategori validitas sangat tinggi dengan nilai reliabilitas Alpha Cronbach sebesar 0,822. Temuan tersebut menjadi salah satu acuan penting dalam menyusun prosedur pengujian instrumen pemahaman teks narasi pada jenjang SMP. Bertolak dari berbagai temuan mutakhir tersebut, penelitian ini berupaya mengembangkan instrumen tes teks narasi yang teruji validitas dan reliabilitasnya, sehingga dapat dipakai sebagai alat ukur kemampuan peserta didik SMP secara akurat dan dapat dipertanggungjawabkan secara ilmiah.

Sampai saat ini, koefisien reliabilitas Alpha Cronbach tetap menjadi metode paling lazim untuk mengukur konsistensi internal instrumen tes, terutama pada instrumen berbentuk skala atau butir dikotomis (Tavakol & Dennick, 2011). Haladyna dan Rodriguez (2013) menekankan bahwa penyusunan butir soal yang berkualitas tidak dapat dilepaskan dari proses validasi konten serta analisis statistik yang cermat agar instrumen benar-benar mengukur konstruk yang dituju.

Validitas isi instrumen evaluasi pembelajaran Bahasa Indonesia juga banyak disoroti dalam penelitian-penelitian terbaru. Kaaffah dkk. (2021) menemukan bahwa validitas isi pada alat evaluasi buku teks pelajaran Bahasa Indonesia SMA perlu ditelaah secara saksama agar selaras dengan kompetensi dasar yang ditetapkan kurikulum. Senada dengan itu, Taneo dan Henukh (2026) menegaskan bahwa uji validitas isi dan konstruk merupakan tahap krusial dalam pengembangan instrumen evaluasi pembelajaran agar instrumen tersebut benar-benar sahih untuk digunakan. Rofida dan Anggraini (2026), dalam penelitiannya terhadap instrumen penilaian Bahasa Indonesia berbantuan kecerdasan buatan generatif, turut membuktikan

bahwa validasi ahli beserta uji validitas empiris memberi kontribusi signifikan terhadap mutu instrumen penilaian kemampuan berpikir kritis peserta didik.

Selain validitas dan reliabilitas, tingkat kesukaran dan daya pembeda butir soal juga menjadi indikator penting kualitas instrumen tes. Hanifah (2014) menjelaskan bahwa tingkat kesukaran dan daya pembeda butir soal turut menentukan keandalan tes secara keseluruhan, sementara Yusrizal (2016) mendapati bahwa proporsi tingkat kesukaran butir soal yang seimbang memengaruhi kualitas hasil ujian. Qomariyah (2022) juga menyoroti pentingnya analisis daya pembeda dalam mengenali butir soal yang mampu membedakan peserta didik berkemampuan tinggi dan rendah. Temuan senada dikemukakan Muluki dkk. (2020) serta Ndiung dan Jediut (2020), yang menyebutkan bahwa instrumen tes hasil belajar yang baik mesti melalui tahap uji tingkat kesukaran dan daya pembeda sebelum dipakai sebagai alat evaluasi resmi.

Kajian tentang kemampuan peserta didik memahami maupun memproduksi teks narasi turut memperkuat landasan penelitian ini. Adlani dkk. (2021) mendapati bahwa penggunaan media video teks narasi mampu meningkatkan keterampilan menulis teks narasi peserta didik. Hutabarat (2021) menambahkan bahwa penerapan model pembelajaran Think Talk Write efektif meningkatkan kemampuan peserta didik menyusun cerita pendek berbasis struktur narasi. Subekti (2019) turut membuktikan bahwa media gambar membantu peserta didik menyusun rangkaian peristiwa dalam teks narasi secara lebih runtut. Sejalan dengan itu, Ilmiarah dkk. (2025) dan Hanifa dkk. (2024) menegaskan bahwa kemampuan menulis narasi peserta didik masih beragam dan dipengaruhi strategi pembelajaran yang diterapkan guru, sehingga diperlukan instrumen evaluasi yang valid dan reliabel untuk mengukur capaian tersebut secara objektif.

Praktik evaluasi pembelajaran di berbagai jenjang pendidikan turut memperlihatkan pentingnya pengujian kualitas instrumen secara berkala. Muhadi dkk. (2025) menekankan bahwa evaluasi instrumen hasil pembelajaran mesti dilakukan berkesinambungan agar selaras dengan desain pembelajaran yang direncanakan. Sabrina dkk. (2025) dan Saputra dkk. (2025) memperlihatkan bahwa mutu butir soal try out maupun penilaian hasil belajar pada Kurikulum Merdeka sangat bergantung pada hasil analisis validitas, reliabilitas, tingkat kesukaran, dan

daya pembeda. Rachmawati dan Bayu Pradana (2025) memperkuat temuan tersebut lewat analisis butir soal mata pelajaran ekonomi yang menunjukkan keempat aspek itu saling berkaitan dalam menentukan kelayakan instrumen. Di samping itu, transformasi taksonomi Bloom dalam evaluasi pembelajaran berbasis HOTS (Listiani & Rachmawati, 2022) serta analisis soal berdasarkan taksonomi Bloom (Lestari dkk., 2022) turut menjadi rujukan dalam menyusun butir soal teks narasi yang mengukur berbagai level kognitif peserta didik SMP.

METODE PENELITIAN

Jenis dan desain penelitian

Penelitian ini mengadopsi pendekatan kuantitatif dengan desain evaluatif, yang difokuskan pada pengujian kualitas instrumen tes materi teks narasi bagi peserta didik sekolah menengah pertama. Pengujian ditempuh melalui analisis validitas dan reliabilitas instrumen guna mengetahui tingkat kelayakan soal sebagai alat evaluasi pembelajaran. Tahapan penelitian mencakup penyusunan kisi-kisi dan instrumen soal, validasi oleh ahli, uji coba instrumen, hingga analisis data hasil pengujian.

Subjek Penelitian

Subjek penelitian ini adalah peserta didik sekolah menengah pertama yang telah menerima pembelajaran materi teks narasi. Penentuan subjek dilakukan dengan teknik purposive sampling, dengan mempertimbangkan kesesuaian materi yang telah dipelajari peserta didik terhadap instrumen yang digunakan. Jumlah subjek disesuaikan dengan kebutuhan analisis statistik dalam pengujian instrumen. Selain peserta didik, penelitian ini juga melibatkan dua orang guru sebagai validator ahli, yang bertugas menilai kesesuaian isi soal dengan indikator pembelajaran, ketepatan materi, konstruksi soal, penggunaan bahasa, serta keterbacaan soal. Hasil penilaian validator selanjutnya dijadikan dasar revisi instrumen sebelum diujicobakan.

Instrumen Penelitian

Instrumen penelitian berupa tes objektif pada materi teks narasi. Penyusunan soal mengacu pada indikator pembelajaran yang meliputi pemahaman pengertian

teks narasi, identifikasi dan penyimpulan jenis teks narasi, analisis struktur teks narasi, evaluasi teks narasi, serta penulisan dan penghasilan teks narasi. Sebelum digunakan, instrumen lebih dulu divalidasi oleh ahli pembelajaran bahasa Indonesia dan ahli evaluasi pendidikan guna memastikan kesesuaiannya dengan kompetensi yang hendak diukur.

Teknik Pengumpulan Data

Data dikumpulkan melalui dua tahap, yaitu validasi instrumen dan uji coba soal. Pada tahap validasi, data diperoleh dari lembar penilaian yang diisi oleh dua validator ahli, yakni guru bahasa Indonesia, yang menilai aspek materi, konstruksi soal, penggunaan bahasa, dan kesesuaian soal dengan indikator pembelajaran.

Setelah tahap validasi rampung, instrumen direvisi mengikuti saran validator sebelum diujicobakan kepada peserta didik. Peserta didik kemudian mengerjakan soal sesuai petunjuk yang telah disediakan, dan hasil pengerjaannya dijadikan data utama dalam analisis validitas dan reliabilitas instrumen. Prosedur uji coba ini mengacu pada penelitian Nandini dkk. (2024) dalam pengembangan instrumen pemahaman materi teks laporan hasil observasi.

Teknik Analisis Data

Uji Validitas Butir Soal

Uji validitas ditempuh untuk mengetahui ketepatan instrumen dalam mengukur kemampuan peserta didik pada materi teks narasi. Jenis validitas yang dipakai dalam penelitian ini adalah validitas isi melalui penilaian dua validator ahli, dengan penilaian menggunakan skala Likert berentang skor 1 hingga 5. Skala serupa pernah dipakai Wiranata dkk. (2024) dalam penelitian instrumen tes pemahaman teks prosedur.

Penilaian validator selanjutnya dianalisis dengan rumus V Aiken untuk mengetahui tingkat validitas tiap butir soal. Nilai V yang semakin mendekati 1,00 menunjukkan tingkat validitas yang semakin tinggi. Butir soal dengan validitas rendah direvisi berdasarkan masukan validator sebelum diujicobakan. Rumus V

Aiken serupa juga pernah dipakai Nandini dkk. (2024) dalam validasi instrumen tes teks laporan hasil observasi.

Uji Reliabilitas Tes

Uji reliabilitas dilakukan guna mengetahui konsistensi instrumen tes, mencakup analisis tingkat kesukaran soal, daya beda soal, serta reliabilitas keseluruhan tes. Tingkat kesukaran dipakai untuk mengelompokkan soal ke kategori mudah, sedang, dan sukar berdasarkan proporsi jawaban benar peserta didik, sementara daya beda dipakai untuk mengetahui kemampuan soal membedakan peserta didik berkemampuan tinggi dan rendah.

Reliabilitas instrumen dianalisis memakai koefisien Alpha Cronbach berdasarkan hasil uji coba. Nilai reliabilitas yang tinggi menandakan instrumen memiliki konsistensi internal yang baik sehingga layak dipakai dalam penelitian, sedangkan butir soal yang tidak memenuhi kriteria akan direvisi atau dieliminasi sebelum instrumen dipakai lebih lanjut.

Kriteria Interpretasi

Hasil analisis validitas dan reliabilitas diinterpretasikan mengikuti kategori statistik yang lazim dipakai dalam penelitian pendidikan. Instrumen dinyatakan reliabel apabila koefisien Alpha Cronbach berada pada kategori tinggi, sehingga layak dipakai sebagai alat evaluasi kemampuan peserta didik pada materi teks narasi.

HASIL DAN PEMBAHASAN

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

Tabel 1. Tabel nilai

No	Instrumen Soal	Bentuk Soal	Soal Ke-
1.	Menjelaskan pengertian dan tujuan teks narasi	Pilihan Ganda	Soal 1-8
2.	Mengidentifikasi & menyimpulkan jenis teks narasi	Pilihan Ganda	Soal 9-16

3.	Menelaah struktur, unsur, dan isi teks narasi	Pilihan Ganda	Soal 17-25
4.	Menilai kualitas dan kelengkapan teks narasi	Uraian	Soal 26-28
5.	Menulis dan menghasilkan teks narasi	Uraian	Soal 29-30

Merujuk pada instrumen di atas, soal tes disusun dalam dua bentuk, yakni 25 soal pilihan ganda dan 5 soal esai. Setelah tersusun, soal tersebut diuji validitasnya dengan melibatkan dua validator, dan hasil validasi tersebut disajikan pada tabel berikut.

Tabel 2. Tabel hasil validasi

No Butir	Validator		S1	S2	$\sum S$	V	Ketegori
	1	2					
1	3	3	2	2	4	0.5	Cukup
2	3	3	2	2	4	0.5	Cukup
3	3	3	2	2	4	0.5	Cukup
4	3	3	2	2	4	0.5	Cukup
5	3	3	2	2	4	0.5	Cukup
6	3	3	2	2	4	0.5	Cukup
7	4	3	3	2	5	0.625	Tinggi
8	4	3	3	2	5	0.625	Tinggi
9	4	3	3	2	5	0.625	Tinggi
10	4	4	3	3	6	0.75	Tinggi
11	4	4	3	3	6	0.75	Tinggi
12	4	4	3	3	6	0.75	Tinggi
13	4	4	3	3	6	0.75	Tinggi
14	4	4	3	3	6	0.75	Tinggi
15	4	4	3	3	6	0.75	Tinggi
16	5	4	4	3	7	0.875	Sangat Tinggi
17	5	4	4	3	7	0.875	Sangat Tinggi
18	5	4	4	3	7	0.875	Sangat Tinggi
19	5	4	4	3	7	0.875	Sangat Tinggi

20	3	5	2	4	6	0.75	Tinggi
21	3	5	2	4	6	0.75	Tinggi
22	3	5	2	4	6	0.75	Tinggi
23	3	5	2	4	6	0.75	Tinggi
24	3	5	2	4	6	0.75	Tinggi
25	3	5	2	4	6	0.75	Tinggi
26	4	5	3	4	7	0.875	Sangat Tinggi
27	4	5	3	4	7	0.875	Sangat Tinggi
28	4	5	3	4	7	0.875	Sangat Tinggi
29	5	5	4	4	8	1	Sangat Tinggi
30	5	5	4	4	8	1	Sangat Tinggi

Hasil uji validitas terhadap 30 butir soal menggunakan rumus V Aiken menghasilkan tiga kategori, yaitu cukup, tinggi, dan sangat tinggi. Butir nomor 1, 2, 3, 4, 5, dan 6 berada pada kategori cukup dengan nilai V sebesar 0,5. Butir nomor 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 21, 22, 23, 24, dan 25 berada pada kategori tinggi sehingga dinyatakan layak dipakai dalam penelitian, sementara kategori sangat tinggi diraih oleh butir nomor 16, 17, 18, 19, 26, 27, 28, 29, dan 30. Karena tidak ditemukan butir soal berkategori rendah maupun sangat rendah, seluruh butir soal dinyatakan valid dan dapat dipakai dalam penelitian tanpa perlu digugurkan.

Setelah itu, instrumen tes dilakukan uji pada kelas uji coba yang melibatkan siswa sekolah menengah pertama. Adapun hasil uji coba dapat dilihat pada tabel berikut.

Tabel 3. Tabel nilai

No Butir	P	Kriteria	D	Kriteria	Variasi butir	P_1	Q_2	P_1Q_1
1	0,25	Sukar	-0,5	Jelek	0,1875	0,25	0,75	0,1875
2	0,5	Sedang	0,5	Baik	0,25	0,5	0,5	0,25
3	0,75	Mudah	0,5	Baik	0,1875	0,75	0,25	0,1875
4	1	Mudah	0	Jelek	0	1	0	0
5	0,5	Sedang	0	Jelek	0,25	0,5	0,5	0,25
6	1	Mudah	0	Jelek	0	1	0	0
7	0,75	Mudah	0,5	Baik	0,1875	0,75	0,25	0,1875

8	1	Mudah	0	Jelek	0	1	0	0
9	0,5	Sedang	0,5	Baik	0,25	0,5	0,5	0,25
10	1	Mudah	0	Jelek	0	1	0	0
11	0,25	Sukar	0	Jelek	0,1875	0,25	0,75	0,1875
12	1	Mudah	0	Jelek	0	1	0	0
13	0,75	Mudah	0,5	Baik	0,1875	0,75	0,25	0,1875
14	0,5	Sedang	0	Jelek	0,25	0,5	0,5	0,25
15	1	Mudah	0	Jelek	0	1	0	0
16	1	Mudah	0	Jelek	0	1	0	0
17	0,75	Mudah	0,5	Baik	0,1875	0,75	0,25	0,1875
18	0,5	Sedang	0,5	Baik	0,25	0,5	0,5	0,25
19	1	Mudah	0	Jelek	0	1	0	0
20	1	Mudah	0	Jelek	0	1	0	0
21	0,25	Sukar	-0,5	Jelek	0,1875	0,25	0,75	0,1875
22	0,5	Sedang	0	Jelek	0,25	0,5	0,5	0,25
23	1	Mudah	0	Jelek	0	1	0	0
24	0,75	Mudah	0,5	Baik	0,1875	0,75	0,25	0,1875
25	1	Mudah	0	Jelek	0	1	0	0
26	0,5	Sedang	0,5	Baik	0,25	0,5	0,5	0,25
27	1	Mudah	0	Jelek	0	1	0	0
28	0,75	Mudah	0,5	Baik	0,1875	0,75	0,25	0,1875
29	0,5	Sedang	0	Jelek	0,25	0,5	0,5	0,25
30	0,25	Sukar	0	Jelek	0,1875	0,25	0,75	0,1875

Hasil analisis daya pembeda mengungkap 10 butir soal berkategori baik, yaitu nomor 2, 3, 7, 9, 13, 17, 18, 24, 26, dan 28. Butir-butir tersebut terbukti mampu membedakan peserta didik berkemampuan tinggi dan rendah sehingga dinyatakan layak dipakai dalam penelitian. Sebaliknya, 20 butir soal lain, yaitu nomor 1, 4, 5, 6, 8, 10, 11, 12, 14, 15, 16, 19, 20, 21, 22, 23, 25, 27, 29, dan 30, berada pada kategori jelek pada daya pembeda sehingga perlu diperbaiki agar kualitas instrumen meningkat. Adapun berdasarkan tingkat kesukaran (P), dari 30 butir soal terdapat 12 butir

berkategori mudah ($P = 1$), 6 butir berkategori mudah ($P = 0,75$), 8 butir berkategori sedang ($P = 0,5$), dan 4 butir berkategori sukar ($P = 0,25$).

Reliabilitas instrumen turut dihitung menggunakan rumus Alpha Cronbach berdasarkan jumlah variansi butir dan nilai piqi tiap soal, dengan k selaku jumlah butir soal sebanyak 30, jumlah hasil perkalian proporsi jawaban benar dan salah ($\Sigma p_i q_i$) sebesar 3,875, serta variansi total instrumen (σ_i^2) sebesar 14,09. Perhitungan tersebut menghasilkan nilai reliabilitas sebesar 0,75 yang termasuk kategori tinggi (0,60–0,79), sehingga instrumen dinyatakan layak dipakai dalam penelitian.

Adapun butir soal yang memenuhi kategori baik pada daya pembeda dinyatakan layak digunakan dalam instrumen penelitian. Rincian butir soal tersebut disajikan pada tabel berikut.

Tabel 4. Tabel Butir Soal Berkategori Baik

No Butir	Soal Tes Teks Narasi
2	(Bacalah teks berikut untuk soal nomor 3–4) "Pagi itu, langit Jakarta masih berselimut kabut tipis. Budi berlari tergesa-gesa menuruni tangga sambil menyandang tas ranselnya yang penuh. Ibunya berteriak dari dapur, 'Budi! Sarapan dulu!' Namun Budi sudah lenyap di balik pintu pagar, meninggalkan aroma nasi goreng yang mengepul sendirian." Ciri teks narasi yang paling tampak dalam kutipan tersebut adalah... a. Bersifat argumentatif b. Menggambarkan urutan kejadian dengan latar waktu dan tempat yang jelas c. Berisi langkah-langkah kerja sistematis d. Menyajikan data dan fakta ilmiah
3	Bagian "orientasi" dalam teks narasi berfungsi untuk... a. Menyelesaikan masalah tokoh b. Memperkenalkan tokoh, latar, dan situasi awal c. Menampilkan konflik yang memuncak d. Menyampaikan pesan moral
7	Urutan struktur teks narasi yang benar adalah... a. Kompleksi → Orientasi → Resolusi → Koda

	b. Koda → Orientasi → Komplikasi → Resolusi c. Orientasi → Komplikasi → Resolusi → Koda d. Resolusi → Komplikasi → Orientasi → Koda
9	Cerita pendek dimulai langsung dengan konflik tanpa memperkenalkan tokoh/latar. Evaluasi yang paling tepat adalah... a. Sudah baik, langsung menarik minat pembaca b. Kurang lengkap, tidak ada orientasi sehingga pembaca bingung c. Terlalu panjang karena banyak konflik d. Sudah memenuhi syarat teks narasi yang baik
13	Kalimat pembuka orientasi terbaik untuk fenomena "anak menemukan dompet di jalan" adalah... a. Dompet adalah benda yang sering hilang di jalan b. Sore itu, langkah Dani terhenti saat ujung sepatunya menyentuh benda tebal di pinggir trotoar c. Menemukan dompet adalah pengalaman menyenangkan d. Di dalam dompet biasanya terdapat uang dan kartu identitas
17	Kalimat dengan kata kerja aktif yang tepat untuk teks narasi adalah... a. Buku itu dibaca oleh Santi dengan penuh konsentrasi b. Santi membaca buku itu dengan penuh konsentrasi sambil menunggu hujan reda c. Oleh Santi, buku tersebut sedang dalam proses pembacaan d. Kegiatan membaca dilakukan oleh Santi
18	Prinsip "AMBAK" (Apa Manfaatnya BAgiKu) dalam quantum learning diterapkan dalam teks narasi dengan cara... a. Menambahkan data statistik di setiap paragraf b. Menyisipkan amanat/pelajaran hidup yang relevan bagi pembaca c. Menggunakan bahasa ilmiah yang kompleks d. Membuat cerita sepanjang mungkin
26	Manfaat prinsip "Tulis dengan Cepat, Perbaiki Kemudian" dalam quantum learning adalah... a. Menghasilkan teks sempurna sejak pertama b. Mencegah

	hambatan kreativitas agar ide mengalir bebas c. Membuat tulisan lebih panjang dan detail d. Menghindari kesalahan ejaan sejak awal
--	--

	Tujuan teknik "Mind Mapping" sebelum menulis teks narasi adalah...
28	a. Langsung menulis tanpa persiapan b. Memetakan ide, tokoh, konflik, dan alur secara visual c. Membuat daftar kosakata sulit d. Menghitung jumlah kata

	Kalimat koda yang paling efektif menurut prinsip quantum learning adalah...
24	a. Mengulang kejadian sebelumnya b. Memberikan penutupan bermakna, refleksi, atau pesan yang menginspirasi c. Kalimat paling panjang dalam cerita d. Menjelaskan latar cerita secara lengkap

SIMPULAN

Merujuk pada hasil penelitian yang telah dipaparkan, dapat ditarik simpulan sebagai berikut. Pertama, uji validitas terhadap 30 butir soal dengan rumus V Aiken memperlihatkan 6 butir berkategori cukup, 15 butir berkategori tinggi, dan 9 butir berkategori sangat tinggi, sehingga seluruh butir soal dinyatakan valid dan layak dipakai sebagai instrumen tes teks narasi bagi peserta didik sekolah menengah pertama. Kedua, uji reliabilitas dengan koefisien Alpha Cronbach menghasilkan nilai 0,75 yang tergolong kategori tinggi, sehingga instrumen dinyatakan memiliki konsistensi internal yang baik dan layak dipakai sebagai alat ukur penelitian. Ketiga, analisis daya pembeda menunjukkan 10 butir soal berkategori baik, yaitu nomor 2, 3, 7, 9, 13, 17, 18, 24, 26, dan 28, sehingga butir-butir tersebut dinyatakan paling layak dipakai sebagai instrumen penelitian yang mampu membedakan peserta didik berkemampuan tinggi dan rendah pada materi teks narasi.

REFERENSI

Adlani, S., Agustina, R. T., & Muchtar, M. (2021). Implementasi Video Teks Narasi untuk Meningkatkan Keterampilan Menulis Teks Narasi Siswa Kelas V. *Jurnal*

- Pembelajaran, Bimbingan, dan Pengelolaan Pendidikan*, 1(1), 72–77.
<https://doi.org/10.17977/um065v1i12021p72-77>.
- Alber, Mukhlis, M., Hermaliza, Gadink, M., & Widyawati, K. (2023). Pengembangan Bahan Ajar Menulis Teks Narasi Berbasis Local Wisdom bagi Siswa Sekolah Menengah. *Jurnal Sastra Indonesia*, 12(2), 169–176.
<https://doi.org/10.15294/jsi.v12i2.70488>.
- Alimah, M., & Indihadi, D. (2022). Analisis Teks Narasi Implementasi Strategi Mind Mapping Peserta Didik di Sekolah Dasar. *Edukatif: Jurnal Ilmu Pendidikan*, 4(4), 5512–5519. <https://doi.org/10.31004/edukatif.v4i4.3245>.
- Dhiya, S. R., & Arianti, R. (2024). Analisis Majas Metafora, Personifikasi dan Simile dalam Kumpulan Puisi Jalan Menuju Rumahmu Karya Acep Zamzam Noor. *Jurnal Pendidikan Tambusai*, 8(2), 1–12.
<https://doi.org/10.31004/jptam.v8i2.17989>.
- Haladyna, T. M., & Rodriguez, M. C. (2013). Developing and Validating Test Items. Routledge. <https://doi.org/10.4324/9780203850381>.
- Hanifa, A. P., Putri, E. N., & Jacky, S. M. (2024). Kemampuan Menulis Narasi. *GURUKU: Jurnal Pendidikan dan Sosial Humaniora*, 2(3), 138–148.
<https://doi.org/10.59061/guruku.v2i3.701>.
- Hanifah, N. (2014). Perbandingan Tingkat Kesukaran, Daya Pembeda Butir Soal dan Reliabilitas Tes Bentuk Pilihan Ganda Biasa dan Pilihan Ganda Asosiasi Mata Pelajaran Ekonomi. *SOSIO E-KONS*, 6(1), 41–55.
<https://doi.org/10.30998/sosioekons.v6i1.1715>.
- Hasanudin, C., Fitriyaningsih, 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., Subyantoro, Zulaeha, I., & Pristiwati, R. (2023). Learning materials and their prototypes for academic writing skills: the needs of Indonesian lecturers in the post-COVID-19 era. *European Journal of Educational Research*, 12(1), 435–453.
<https://doi.org/10.12973/eu-jer.12.1.435>.
- Hutabarat, L. (2021). The Application Think Talk Write Learning Model to Improve Composing Short Stories. *Hortatori: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 4(2), 122–126. <https://doi.org/10.30998/jh.v4i2.530>.
- Ilmiarah, F. A., Setyawan, A., & Widayati, S. (2025). Analisis Keterampilan Menulis Karangan Narasi Siswa Kelas V SDN Tanjung Jati 2 Tahun Ajar 2024/2025. *Maktab: Jurnal Pendidikan dan Teknologi*, 3(2).
<https://doi.org/10.56480/maktab.v3i2.1173>.

- Kaaffah, R. R. S., Wijiyono, A. W., & Rahmayanti, I. (2021). Validitas Isi pada Alat Evaluasi Buku Teks Pelajaran Bahasa Indonesia Kelas X SMA. *Imajeri: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 3(2), 158–167. <https://doi.org/10.22236/imajeri.v3i2.6572>.
- Lestari, G. P., Yensy, N. A., Hanifah, H., Agustinsa, R., & Susanto, E. (2022). Analisis Soal Latihan Buku Matematika Kurikulum 2013 Kelas VIII Materi Lingkaran Berdasarkan Taksonomi Bloom. *Jurnal Penelitian Pembelajaran Matematika Sekolah*, 6(1), 23–31. <https://doi.org/10.33369/jp2ms.6.1.23-31>.
- Listiani, W., & Rachmawati, R. (2022). Transformasi Taksonomi Bloom dalam Evaluasi Pembelajaran Berbasis HOTS. *Jurnal Jendela Pendidikan*, 2(3), 397–402. <https://doi.org/10.57008/jjp.v2i03.266>.
- Muhadi, M., Jarir, J., Khairina, K., Rajuna, R., & Prasetyo, E. (2025). Evaluasi Perencanaan Desain Pembelajaran, Pelaksanaan Proses Kegiatan Pembelajaran, dan Evaluasi Instrumen Hasil Pembelajaran. *Jurnal Pendidikan, Ilmu Sosial, dan Pengabdian kepada Masyarakat*, 5(2), 156–165. <https://doi.org/10.56832/edu.v5i2.1084>.
- Muluki, A., Bundu, P., & Sukmawati, S. (2020). Analisis Kualitas Butir Tes Semester Ganjil Mata Pelajaran IPA Kelas IV MI Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar*, 4(1), 86–96. <https://doi.org/10.23887/jisd.v4i1.23335>.
- Munjayana, S. (2024). Pembelajaran Menulis Teks Narasi dengan Metode Mind Mapping pada Siswa MI. *FONDATIA*, 8(1), 66–76. <https://doi.org/10.36088/fondatia.v8i1.4536>.
- Nandini, F. G., Cahyawati, R. S., & Hasanudin, C. (2024). Langkah-Langkah Uji Validitas dan Reliabilitas Instrumen Pemahaman Materi Teks Laporan Hasil Observasi. *Seminar Nasional dan Gelar Karya Produk Hasil Pembelajaran*, 6(2), 378–388. <https://doi.org/10.70118/prosidingukmpr.v2i2.2977>.
- Ndiung, S., & Jediut, M. (2020). Pengembangan Instrumen Tes Hasil Belajar Matematika Peserta Didik Sekolah Dasar Berorientasi pada Berpikir Tingkat Tinggi. *Premiere Educandum: Jurnal Pendidikan Dasar dan Pembelajaran*, 10(1), 94–111. <https://doi.org/10.25273/pe.v10i1.6274>.
- Qomariyah, L. (2022). Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal TOAFL Universitas Hasyim Asy'ari Tebuireng Jombang. *Lisanan Arabiya: Jurnal Pendidikan Bahasa Arab*, 6(1), 1–18. <https://doi.org/10.32699/liar.v6i1.2549>.
- Rachmawati, D., & Bayu Pradana, A. (2025). Analisis Butir Soal Mata Pelajaran Ekonomi: Validitas, Reliabilitas, Tingkat Kesukaran, dan Daya Pembeda. *Jurnal Pendidikan Ekonomi (JUPE)*, 13(3), 273–284. <https://doi.org/10.26740/jupe.v13n3.p273-284>.

- Ramadhan, M. F., Siroj, R. A., & Afgani, M. W. (2024). Validitas dan Reliabilitas Instrumen Pengukuran. *Journal on Education*, 6(2), 10967-10975. <https://doi.org/10.31004/joe.v6i2.4885>
- Riana, J. D., & Hasanudin, C. (2024). Uji Validitas dan Reliabilitas Instrumen Penelitian Tingkat Pemahaman Materi Teks Berita Siswa. *Seminar Nasional dan Gelar Karya Produk Hasil Pembelajaran*, 2(2), 1411-1426. Retrieved from <https://prosiding.ikipgribojonegoro.ac.id/index.php/SNGK/article/view/3071>.
- Rofida, A., & Anggraini, D. (2026). Pengembangan Instrumen Penilaian Bahasa Indonesia Berbantuan GenAI untuk Mengukur Kemampuan Berpikir Kritis Peserta Didik SD Kelas IV. *Jurnal Ilmiah Profesi Pendidikan*, 11(2), 1524-1531. <https://doi.org/10.29303/jipp.v11i2.4905>.
- Sabrina, A. B., Hery, A., & Irianti, S. (2025). Evaluasi Kualitas Butir Soal Try Out Usaha Layanan Pariwisata. *Jurnal Ilmiah Ilmu Pendidikan*, 8(10), 11717-11722. <https://doi.org/10.54371/jiip.v8i10.9625>.
- Saputra, H., Saputra, R. A., Julianto, J., & Ramadhani, A. (2025). Penilaian Hasil Belajar Siswa pada Kurikulum Merdeka di Sekolah Dasar. *Jurnal Inovasi Pendidikan dan Teknologi Informasi*, 6(1), 128-138. <https://doi.org/10.52060/jipti.v6i1.2743>.
- Sobiri, M., & Nashrullah. (2026). Analisis Validitas dan Reliabilitas Instrumen Penelitian dalam Pendekatan Kuantitatif. *Jurnal Psikososial dan Pendidikan*, 2(1), 1048-1057. Retrieved from <https://publisherqu.com/index.php/psikosospen/article/view/4042>.
- Subekti, E. (2019). Peningkatan Keterampilan Menulis Teks Narasi dengan Menggunakan Media Gambar pada Siswa Kelas VII SMP N 1 Baturaden. *Metafora: Jurnal Pembelajaran Bahasa dan Sastra*, 5(1), 10-23. <https://doi.org/10.30595/mtf.v5i1.5023>.
- 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>
- Taneo, N., & Henukh, S. (2026). Uji Validitas Isi dan Konstruk pada Pengembangan Instrumen Evaluasi Pembelajaran. *Educational Journal*, 1(4), 1679-1689. <https://doi.org/10.63822/kdaefy18>.
- Tavakol, M., & Dennick, R. (2011). Making Sense of Cronbach's Alpha. *International Journal of Medical Education*, 2, 53-55. <https://doi.org/10.5116/ijme.4dfb.8d>.
- 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), 59-67. [https://doi.org/10.25299/jrec.2024.vol6\(1\).17430](https://doi.org/10.25299/jrec.2024.vol6(1).17430).

- Wiranata, D. A., Cahyawati, R. S., & Hasanudin, C. (2024). Uji Validitas dan Reliabilitas Instrumen Penelitian Tingkat Pemahaman Teks Prosedur. *Seminar Nasional dan Gelar Karya Produk Hasil Pembelajaran*, 6(2), 1-12. <https://doi.org/10.70118/prosidingukmpr.v2i2.2936>.
- Yusrizal. (2016). Analysis of Difficulty Level of Physics National Examination's Questions. *Jurnal Pendidikan IPA Indonesia*, 5(1), 140-149. <https://doi.org/10.15294/jpii.v5i1.5803>.
- Yusup, F. (2018). Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif. *Jurnal Tarbiyah: Jurnal Ilmiah Kependidikan*, 7(1), 17-23. <https://doi.org/10.18592/tarbiyah.v7i1.2100>.
- Zayrin, A. A., Nupus, H., Maizia, K. K., Marsela, S., Hidayatullah, R., & Harmonedi, H. (2025). Analisis Instrumen Penelitian Pendidikan (Uji Validitas dan Reliabilitas Instrumen Penelitian). *QOSIM: Jurnal Pendidikan Sosial & Humaniora*, 3(2), 780-789. <https://doi.org/10.61104/jq.v3i2.1070>.