

Analisis Butir Soal Tes Materi Pertidaksamaan Linier untuk Siswa Kelas XI

Annisah Kurniati^{1✉}, Depriwana Rahmi², Suci Yuniati³, Dhilla Rahmania⁴

^{1, 2, 3, 4} Program Studi Pendidikan Matematika, Fakultas Tarbiyah dan Keguruan, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
Jl. Subarntas No.155 Km. 15, Pekanbaru-Riau, Indonesia
depriwanarahmi@uin-suska.ac.id

Abstract

Analyzing the quality of test questions is important to ensure that the questions used can accurately measure students' abilities. This study aims to evaluate the quality of linear inequality test questions for Class XI based on aspects of validity, reliability, difficulty level, and discrimination power. The method used is quantitative, with the data consisting of 57 student answer sheets. The analysis results show that one of the six questions tested is invalid and therefore not suitable for use. The reliability of the questions is moderate, with the question format being essay. The discrimination index for question number 1 is categorized as poor and therefore needs to be removed, while questions 2 through 6 have acceptable discrimination indices. In terms of difficulty level, the remaining five questions are in the medium category. This finding indicates that most of the questions are of fairly good quality and can be used by teachers as a reference in evaluating students' abilities in linear inequalities. Thus, item analysis becomes an important step in improving the quality of learning evaluation.

Keywords: Test questions, validity, reliability, discriminatory power, level of difficulty

Abstrak

Analisis kualitas soal tes penting dilakukan untuk memastikan bahwa soal yang digunakan mampu mengukur kemampuan siswa secara tepat. Penelitian ini bertujuan untuk mengevaluasi kualitas soal tes materi pertidaksamaan linear kelas XI berdasarkan aspek validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Metode yang digunakan adalah kuantitatif dengan data berupa 57 lembar jawaban siswa. Hasil analisis menunjukkan bahwa dari enam soal yang diuji, satu soal tidak valid dan tidak layak digunakan. Reliabilitas soal tergolong sedang, dengan bentuk soal berupa uraian. Daya pembeda soal nomor 1 dikategorikan buruk sehingga perlu dihapus, sedangkan soal nomor 2 hingga 6 memiliki daya pembeda yang dapat diterima. Dari segi tingkat kesukaran, kelima soal yang tersisa berada dalam kategori sedang. Temuan ini mengindikasikan bahwa sebagian besar soal memiliki kualitas yang cukup baik dan dapat dimanfaatkan oleh guru sebagai acuan dalam mengevaluasi kemampuan siswa pada materi pertidaksamaan linear. Dengan demikian, analisis butir soal menjadi langkah penting dalam meningkatkan kualitas evaluasi pembelajaran.

Kata kunci: Soal tes, validitas, Reliabilitas, Daya pembeda, Tingkat kesukaran

Copyright (c) 2025 Annisah Kurniati, Depriwana Rahmi, Suci Yuniati, Dhilla Rahmania

✉ Corresponding author: Annisah Kurniati

Email Address: depriwanarahmi@uin-suska.ac.id (Jl. Subarntas No.155 Km. 15, Pekanbaru-Riau, Indonesia)

Received 27 April 2025, Accepted 22 August 2025, Published 28 November 2025

DoI: <https://doi.org/10.31004/cendekia.v9i3.4017>

PENDAHULUAN

Dalam dunia pendidikan, evaluasi merupakan bagian integral yang tidak dapat dipisahkan dari keseluruhan proses pembelajaran. Evaluasi tidak hanya berfungsi sebagai alat untuk mengukur pencapaian hasil belajar peserta didik, tetapi juga sebagai dasar bagi pendidik dalam mengambil keputusan yang tepat untuk meningkatkan mutu pembelajaran. Salah satu tanggung jawab pendidik adalah merencanakan sekaligus melaksanakan kegiatan evaluasi atau penilaian. Oleh karena itu kegiatan evaluasi merupakan proses penilaian terkait berbagai hal yang berhubungan dengan kegiatan pendidikan (Fitrianti, 2018). Proses evaluasi dilakukan dalam rangka untuk mengetahui ketercapaian hasil belajar peserta didik. Pendidik dituntut tidak hanya profesional dalam mengajar namun juga mampu menguasai teknik evaluasi hasil belajar (Balau, Pesik, dan Damai, 2021). Sudjana (1995)

menegaskan bahwa evaluasi atau penilaian hasil belajar berperan penting dalam menentukan mutu pendidikan, karena dari hasil evaluasi dapat diketahui keberhasilan pencapaian kompetensi, efektivitas metode pembelajaran, dan kualitas proses pembelajaran secara keseluruhan. Evaluasi juga sangat penting dilaksanakan untuk mengetahui apakah proses pembelajaran yang dilaksanakan sudah sesuai dengan tujuan yang diharapkan dan menjadi acuan bagi pendidik untuk melangkah lebih baik kedepannya. Evaluasi harus selalu menjadi perhatian penting dalam pembelajaran, karena dalam kegiatan evaluasi terdapat beberapa proses yang dilalui seperti tahap merencanakan, memperoleh dan menyediakan berbagai informasi yang berguna dalam membuat alternatif keputusan. Semua proses tersebut diarahkan untuk ketercapaian tujuan pembelajaran.

Pengukuran dan penilaian menjadi kegiatan utama dalam evaluasi pembelajaran. Terdapat dua jenis alat evaluasi untuk menilai hasil belajar yaitu tes dan non tes. Teknik tes berguna dalam menilai kemampuan kognitif dan psikomotor, sedangkan teknik non tes berguna dalam menilai tentang karakteristik afektif dari suatu objek yang diteliti (Syaripudin, 2012). Kedua teknik tersebut harus menggunakan prinsip penilaian yang telah ditetapkan. Pada proses penilaian kemampuan kognitif seperti tes dikatakan baik apabila hasil dari tes tersebut dapat mengungkapkan apa yang diukur sebenarnya (Syaifudin, 2023). Beberapa prinsip dalam penilaian diantaranya dengan melihat tingkat kevalidan soal, daya beda dan reliabilitas soal. Jadi agar prinsip tersebut dapat terlaksana maka dibutuhkan kegiatan analisis butir dari suatu soal. Analisis butir tes merupakan prosedur mengkonstruksi tes guna mendapatkan gambaran tentang kehandalan atau mutu dari suatu soal tes, baik secara keseluruhan maupun pada setiap butir tes (Magdalena, Anggraini, dan Khoiriah, 2021).

Menganalisis setiap butir soal merupakan suatu kegiatan yang sangat penting dan bermanfaat dalam memperoleh berbagai informasi yang digunakan sebagai umpan balik dengan tujuan agar soal tes yang telah disusun dapat mengukur kemampuan secara berkelanjutan, sistematis dan transparan. Tujuan menganalisis butir soal untuk mengetahui kevalidan, reliabilitas, indeks kesukaran dan daya pembeda soal yang diujikan. Jika tes dari soal yang diuji masih kurang baik, maka hasil yang diperoleh tidak akan baik. Agar mutu dari soal tes harus baik dan berkualitas dengan melakukan analisis kualitas butir soal pada suatu tes (Iskandar dan Rizal, 2018). Analisis butir soal dapat mengevaluasi daya pembeda butir soal yang sangat dibutuhkan dalam mengukur sejauh mana sebuah soal mampu membedakan berbagai kemampuan peserta didik yang kategori kemampuan tinggi, sedang dan rendah. Selain itu, analisis butir soal ini juga berguna dalam melakukan evaluasi terhadap ketercapaian hasil pembelajaran yang sesuai harapan dan dapat dijadikan sebagai usaha perbaikan terhadap instrumen penilaian baik tes maupun non tes yang digunakan. Jadi, kegiatan menganalisis setiap butir soal sebaiknya dilakukan secara berkala dan berkesinambungan sehingga menghasilkan soal tes yang efektif digunakan sebagai alat ukur (Pamujo dan Romadhoni, 2022)

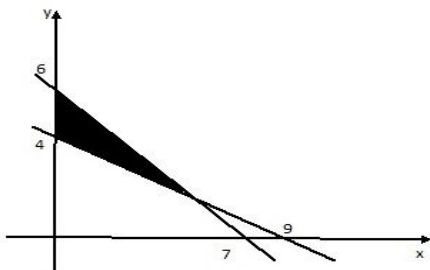
Ada dua jenis tes untuk mengetahui tingkat kemampuan yaitu tes bentuk uraian dan pilihan ganda. Tes bentuk uraian biasanya pada soal ulangan harian, tugas rumah, sedangkan tes bentuk pilihan ganda biasanya pada soal ujian akhir semester atau ujian sekolah (Halik, Mania, dan Nur, 2019).

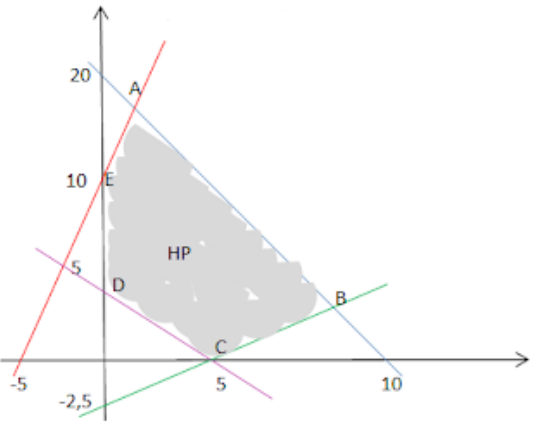
Penelitian yang dilakukan Fadillah dkk. (2024) menunjukkan bahwa tes bentuk uraian lebih efektif untuk mengukur kemampuan penalaran matematis, sedangkan tes bentuk pilihan ganda lebih sesuai untuk mengukur penguasaan konsep secara luas. Temuan ini menguatkan bahwa pemilihan bentuk tes perlu disesuaikan dengan tujuan pembelajaran dan karakteristik materi yang diujikan. Penelitian ini berfokus pada menganalisis validitas, reliabilitas, daya pembeda, dan indeks kesukaran soal matematika kelas XI tahun ajaran 2023/2024 pada materi pertidaksamaan linear dalam bentuk soal uraian.

METODE

Jenis penelitian ini merupakan penelitian kuantitatif yaitu menganalisis setiap butir soal yang berbentuk uraian. Analisis yang dilakukan berupa analisis kelayakan soal matematika kelas XI tahun ajaran 2023/2024 yang memenuhi aspek validitas, reliabilitas, daya pembeda, dan indeks kesukaran. Adapun butir soal yang dianalisis yaitu:

Tabel 1. Butir Soal

Indikator Soal	Soal	Nomor Soal
Peserta didik mampu menyajikan masalah dalam bentuk grafik maupun tabel guna memberikan gambaran yang lebih jelas serta mendukung proses pencarian solusi.	Dari pertidaksamaan berikut: a. $0 \leq y \leq 8$ b. $2 \leq x \leq 6$ Tentukan daerah himpunan penyelesaian	1
	Tentukan daerah himpunan penyelesaian dari pertidaksamaan berikut: $5x + y \geq 10$	2
Peserta didik dapat mengidentifikasi bentuk pertidaksamaan linear dua variabel berdasarkan wilayah himpunan penyelesaiannya yang telah ditentukan.	Tentukanlah pertidaksamaan dari gambar berikut: 	3

	<p>Tentukanlah pertidaksamaan dari gambar berikut:</p> 	4
<p>Peserta didik mampu memberikan jawaban terhadap soal disertai uraian dalam bentuk narasi atau penjelasan tertulis.</p>	<p>Susi berencana membeli dua varian jeruk: jenis A seharga Rp6.000 per kilogram dan jenis B dengan harga Rp4.000 per kilogram. Ia memiliki dana terbatas, yaitu sebesar Rp50.000, dan keranjangnya hanya mampu menampung maksimal 10 kilogram. Jika jumlah jeruk jenis A dinyatakan dengan variabel x dan jeruk jenis B dengan y, tentukan bentuk model matematis dari situasi tersebut.</p>	5
	<p>Sebuah perusahaan yang memproduksi tas dan sepatu membutuhkan total 4 unit bahan A dan 6 unit bahan B setiap minggunya untuk mendukung kegiatan produksinya. Untuk membuat satu tas, diperlukan 1 unit bahan A dan 2 unit bahan B, sedangkan memproduksi satu sepatu membutuhkan 2 unit bahan A dan 2 unit bahan B. Keuntungan yang diperoleh dari setiap tas adalah Rp3.000, dan dari setiap sepatu adalah Rp2.000. Tentukan berapa banyak tas dan sepatu yang sebaiknya diproduksi agar keuntungan maksimal dapat dicapai.</p>	6

Pengujian Validitas

Untuk menghitung validitas adalah rumus korelasi *produt moment pearson* (Supriadi, 2021). Tujuan dari uji validitas adalah untuk menilai apakah suatu kuesioner benar-benar mengukur apa yang seharusnya diukur. Jika butir pertanyaan dalam kuesioner mampu merepresentasikan aspek yang ingin digali, maka kuesioner tersebut dianggap valid.

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum Y^2 - (\sum Y)^2\}}} \quad (1)$$

Dimana r_{xy} Koefisien validitas menunjukkan tingkat keterkaitan antara skor suatu butir soal dengan skor total. Dalam hal ini, X merepresentasikan nilai atau jawaban benar dari masing-masing item soal, Y merupakan skor keseluruhan, dan N adalah jumlah responden atau peserta yang mengikuti tes. Dengan kriteria sebagai berikut (Iskandar dan Rizal, 2018).

Tabel 2. Kriteria Validitas Butir Soal

Besarnya Nilai r	Kriteria
$0,80 \leq r_{xy} \leq 1,00$	Sangat Tinggi
$0,60 \leq r_{xy} \leq 0,79$	Tinggi
$0,40 \leq r_{xy} \leq 0,59$	Sedang
$0,20 \leq r_{xy} \leq 0,39$	Rendah
$0,00 \leq r_{xy} \leq 0,19$	Sangat Rendah

Pengujian Reliabilitas

Pengujian reabilitas tes butir soal menggunakan rumus *Alpha Cronbach* dengan rumus berikut.

$$r_{11} = \left(\frac{k}{k-1} \right) \left(1 - \frac{\sum S_i^2}{S_t^2} \right) \quad (2)$$

Dengan r adalah koefisien reabilitas, n adalah banyak butir soal, S_i^2 adalah variansi skor butir soal ke-I dan S_t^2 adalah variansi skor total (Supriadi, 2021). Untuk mengetahui apakah suatu tes memiliki reliabilitas tinggi, sedang, atau rendah dapat dilihat dalam tabel berikut ini (Iskandar dkk, 2022):

Tabel 3. Standar Penafsiran Tingkat Reliabilitas Butir Soal

Rentang Nilai r	Kategori Interpretasi
0,80 – 1,00	Sangat Tinggi
0,60 – 0,79	Tinggi
0,40 – 0,59	Sedang
0,20 – 0,39	Rendah
0,00 – 0,19	Sangat Rendah

Daya Pembeda Soal

Penghitungan daya pembeda bertujuan untuk menilai kemampuan suatu butir soal dalam membedakan antara peserta didik yang telah memahami materi pelajaran dengan mereka yang belum menguasainya. Terdapat rumus tertentu yang digunakan untuk menentukan nilai daya pembeda tersebut. (Lestari Eka dan Yudhanegara, 2018) adalah $DP = \frac{\bar{X}_A - \bar{X}_B}{SMI}$ dimana DP adalah koefisien daya pembeda, \bar{X}_A merepresentasikan nilai rata-rata jawaban dari kelompok siswa berkemampuan tinggi, sedangkan \bar{X}_B adalah rata-rata jawaban dari kelompok siswa berkemampuan rendah. Sementara itu, SMI (Skor Maksimum Ideal) menunjukkan skor tertinggi yang mungkin dicapai. Untuk menilai kualitas daya pembeda suatu butir soal, digunakan kategori tertentu berdasarkan nilai koefisien daya bedanya. (Iskandar dkk, 2022).

Tabel 4. Standar Penilaian Daya Pembeda untuk Setiap Butir Soal

Rentang Nilai	Penilaian Daya Pembeda
$0,70 < DP \leq 1,00$	Sangat Baik
$0,40 < DP \leq 0,70$	Baik
$0,20 < DP \leq 0,40$	Cukup
$0,00 < DP \leq 0,20$	Buruk
$DP \leq 0,00$	Sangat Buruk

Indeks Kesukaran Soal

Indeks kesukaran soal menggambarkan probabilitas siswa dapat memberikan jawaban yang benar pada suatu butir soal, berdasarkan tingkat kemampuan mereka. Berikut adalah rumus yang digunakan untuk menghitung tingkat kesukaran sebuah butir soal: $IK = \frac{\bar{X}}{SMI}$

Dimana IK adalah indeks kesukaran, \bar{X} adalah rata-rata skor jawaban siswa pada butir soal dan SMI adalah skor maksimum ideal yaitu skor maksimum yang akan diperoleh siswa jika menjawab butir soal tersebut dengan sempurna. Untuk mengukur indeks kesukaran soal ialah sebagai berikut (Lestari Eka dan Yudhanegara, 2018).

Tabel 5. Kriteria Penafsiran Tingkat Kesukaran Soal

Indeks Kesukaran	Kategori
0,0 – 0,3	Sukar
0,3 – 0,7	Sedang
0,7 – 1,0	Mudah

HASIL DAN DISKUSI

Pada penelitian ini diawali dengan pembuatan soal yang terdiri dari 6 soal uraian dengan materi pertidaksamaan linear. Bagian ini menyajikan informasi terkait aktivitas penelitian yang dilakukan. Data yang ditampilkan berupa jawaban siswa dan butir soal dari tes uraian pada materi pertidaksamaan linear. Data tersebut diolah menggunakan Microsoft Excel untuk menganalisis validitas soal, tingkat kesukaran, daya pembeda, serta reliabilitasnya. Berdasarkan uji analisis uji validitas menggunakan koefisien korelasi *product moment* Pearson, diperoleh hasil seperti pada Tabel 6 berikut.

Tabel 6. Hasil Uji Validitas Butir Soal

No	Koefisien Korelasi	Interpretasi	Keputusan
1	0.198494144	Sangat Rendah	Tidak Valid
2	0.841829405	Tinggi	Valid
3	0.611527502	Tinggi	Valid
4	0.723753541	Tinggi	Valid
5	0.680018664	Sangat Tinggi	Valid
6	0.870971616	Sangat Tinggi	Valid

Dari tabel diketahui hasil dari enam soal terdapat lima soal yang valid dan satu soal yang tidak valid sehingga tidak dapat digunakan dengan hasil 0,19. Menurut Hazlita, Zulkardi, dan Darmawijoyo (2015) Sebuah soal dikatakan valid apabila butir pertanyaannya dapat secara tepat merepresentasikan aspek yang hendak diukur. Secara teoretis, validitas butir soal mengacu pada sejauh mana suatu instrumen mampu mengukur apa yang seharusnya diukur (Azwar, 2016). Menurut Arikunto (2002) nilai koefisien korelasi yang tinggi menunjukkan bahwa butir tersebut memiliki hubungan yang kuat dengan konstruk yang diukur, sehingga dianggap valid. Sebaliknya, nilai yang rendah menunjukkan bahwa butir tidak cukup representatif untuk mengukur konstruk yang dimaksud. Hasil pengujian pada Tabel 6 menunjukkan bahwa lima butir soal nomor 2 sampai 6 termasuk kategori “tinggi” hingga

“sangat tinggi,” sehingga dinyatakan valid. Sementara itu, satu butir soal nomor 1 memiliki koefisien korelasi 0,198 yang termasuk kategori “sangat rendah” sehingga dinyatakan tidak valid.

Temuan ini sejalan dengan penelitian Fitrianty, Yunita, dan Juwita (2022) yang menunjukkan bahwa butir soal dengan korelasi rendah cenderung tidak mengukur kemampuan yang sama dengan tes secara keseluruhan, sehingga perlu direvisi atau dihapus. Selain itu, penelitian yang dilakukan Wulandari dan Ayda (2021) juga menegaskan bahwa validitas tinggi sangat penting agar tes mampu memberikan gambaran akurat tentang kemampuan peserta, khususnya dalam konteks evaluasi pembelajaran matematika berbasis pendekatan tertentu. Dengan demikian, berdasarkan hasil uji validitas ini, instrumen yang digunakan dalam penelitian layak dipakai setelah mengeluarkan butir yang tidak valid nomor 1. Hal ini memastikan bahwa instrumen yang digunakan memiliki tingkat kesesuaian yang tinggi dengan konstruk yang ingin diukur, sehingga data yang diperoleh akan lebih akurat dan dapat dipertanggungjawabkan secara ilmiah.

Tabel 7. Temuan Reliabilitas Butir Soal

<i>r_{hitung}</i>	Kesimpulan	Interpretasi
0,585289477	Reliabel	Sedang

Selanjutnya Reliabilitas soal diperoleh dari hasil analisis reliabilitas soal yang berbentuk uraian Pada analisis data ini memperoleh hasil bahwasanya indeks reliabilitas adalah 0,585 hal itu menandakan hasil analisis butir soal memiliki reliabilitas sedang. Reliabilitas instrumen menunjukkan tingkat konsistensi hasil pengukuran apabila tes yang sama diberikan pada kelompok yang sama dalam kondisi yang relatif sama (Azwar, 2016). Dalam konteks penelitian pendidikan, reliabilitas yang tinggi berarti instrumen mampu memberikan hasil yang stabil dan tidak dipengaruhi oleh faktor kebetulan. Namun, nilai reliabilitas kategori sedang, seperti pada penelitian ini, menunjukkan bahwa instrumen masih memiliki konsistensi yang cukup baik meskipun terdapat ruang untuk perbaikan.

Nilai 0,585 yang diperoleh mengindikasikan bahwa soal uraian yang digunakan dalam penelitian ini relatif konsisten dalam mengukur konstruk yang dimaksud. Hal ini sejalan dengan temuan Susanto, Rinaldi, dan Novalia (2015) yang menunjukkan bahwa instrumen tes dengan reliabilitas kategori cukup masih dapat digunakan untuk tujuan penelitian, selama validitas butir soal berada pada kategori baik. Temuan serupa juga dilaporkan oleh Saputri dkk (2023) yang menyatakan bahwa reliabilitas tidak harus tinggi secara absolut untuk dapat diterapkan, asalkan validitas setiap butir soal telah terverifikasi dengan baik. Dengan demikian, meskipun reliabilitas instrumen tergolong sedang, instrumen ini tetap dapat digunakan dalam penelitian dengan catatan hasil interpretasi disertai pertimbangan terhadap keterbatasan konsistensi pengukuran. Selain itu, validitas tinggi yang dimiliki sebagian besar butir soal akan membantu menjaga kualitas keseluruhan instrumen dalam mengukur kemampuan.

Tabel 8. Rekapitulasi Hasil Analisis Daya Pembeda Butir Soal

No	DP	Indeks Daya Beda	Kriteria
1	0.0825	$0,00 < DP \leq 0,20$	Buruk
2	0.4048	$0,40 < DP \leq 0,70$	Baik
3	0.2315	$0,20 < DP \leq 0,40$	Cukup
4	0.2254	$0,20 < DP \leq 0,40$	Cukup
5	0.2098	$0,20 < DP \leq 0,40$	Cukup
6	0.4068	$0,40 < DP \leq 0,70$	Baik

Berdasarkan data di atas, soal nomor 1 memiliki daya pembeda dengan kategori buruk, sehingga soal tersebut perlu dieliminasi. Sementara itu, soal nomor 2 hingga soal nomor 6 memenuhi kriteria layak dan dapat digunakan. Namun soal nomor 1 jika ingin juga dipergunakan maka masih bisa diperbaiki, sesuai pendapat Firmansyah dan Sridadi (2018) butir soal masih bisa diperbaiki yang mempunyai tingkat daya pembeda rendah. Temuan ini konsisten dengan penelitian Rizqa dkk (2025) yang menegaskan bahwa butir dengan daya pembeda rendah perlu dihapus atau diperbaiki karena berpotensi mengurangi akurasi pengukuran kemampuan siswa. Selain itu, studi Nurhayati, Halini, dan Indriani (2024) menunjukkan bahwa peningkatan kualitas redaksi soal, kejelasan konteks, dan kesesuaian tingkat kesulitan dapat memperbaiki daya pembeda suatu butir. Dengan demikian, berdasarkan hasil ini, butir soal kategori buruk perlu dieliminasi atau direvisi, butir kategori cukup perlu perbaikan untuk meningkatkan efektivitasnya, dan butir kategori baik dapat dipertahankan sebagai komponen instrumen yang andal untuk mengukur kemampuan representasi matematis mahasiswa.

Tabel 9. Hasil Tingkat Kesukaran Butir Soal

No	TK	Indeks Kesukaran	Kriteria
1	0.65497	0,3-0,7	Sedang
2	0.80117	0,7-1,0	Mudah
3	0.73684	0,7-1,0	Mudah
4	0.57895	0,3-0,7	Sedang
5	0.69005	0,3-0,7	Sedang
6	0.561403	0,3-0,7	Sedang

Dari hasil analisis mengenai tingkat kesukaran butir soal diperoleh hasil sebagian besar soal berada pada tingkat kesukaran sedang yaitu sebanyak empat soal. Ini berarti keempat soal tersebut termasuk kategori soal yang baik. Suatu soal termasuk kategori baik jika pertanyaan yang disajikan soal tersebut memiliki tingkat kesukaran yang tidak terlalu susah dan tidak juga terlalu mudah (Hanna dan Retnawati, 2022). Ini dikarenakan jika soal terlalu mudah maka semua peserta didik dapat menjawab soal tersebut sehingga tidak dapat mengukur kemampuan peserta didik secara keseluruhan, sedangkan sebaliknya jika soal terlalu susah maka banyak peserta didik yang tidak dapat menjawab soal tersebut karena tingkat kesukaran soalnya di atas kemampuan rata-rata. Soal yang terlalu sukar menjadikan peserta didik putus asa mengerjakannya (Halik, Mania, dan Nur, 2019). Hal ini menunjukkan bahwa soal-soal perlu disesuaikan agar tingkat kesukaran lebih merata dan dapat

mengevaluasi kemampuan peserta didik secara lebih menyeluruh (Surbakti, 2025).

Penelitian ini memiliki beberapa kelebihan yang patut dicatat. Pertama, penyusunan instrumen tes dilakukan melalui tahapan analisis yang komprehensif, meliputi validitas, reliabilitas, daya pembeda, dan tingkat kesukaran. Hal ini sesuai dengan rekomendasi Arikunto (2018) bahwa uji instrumen secara menyeluruh akan meningkatkan kualitas pengukuran. Kedua, penggunaan Microsoft Excel sebagai alat analisis mempermudah perhitungan dan visualisasi data, sehingga proses analisis menjadi lebih efisien dan transparan (Junaidi dkk, 2024). Ketiga, penelitian ini menggunakan data empiris langsung dari peserta didik melalui soal uraian, yang memungkinkan peneliti menilai kemampuan berpikir matematis secara mendalam dan bukan sekadar hasil akhir jawaban (Nitko dan Brookhart, 2011).

Di sisi lain, penelitian ini juga memiliki beberapa keterbatasan. Pertama, jumlah butir soal yang relatif sedikit (hanya 6 soal) dapat mempengaruhi cakupan materi dan estimasi reliabilitas, karena instrumen dengan jumlah butir terbatas cenderung memiliki koefisien reliabilitas yang lebih rendah (Azwar, 2016). Kedua, hasil uji reliabilitas berada pada kategori sedang ($r = 0,585$), yang menunjukkan bahwa konsistensi pengukuran belum maksimal. Ketiga, penelitian ini tidak menyertakan uji coba instrumen pada sampel yang berbeda sebelum digunakan, padahal pengujian di dua kelompok berbeda dapat meningkatkan generalisasi dan akurasi instrument (Creswell, 2018). Keempat, beberapa butir soal memiliki daya pembeda yang rendah atau kategori buruk, yang menunjukkan perlunya revisi redaksi atau perubahan tingkat kesukaran (Lestari, 2018).

KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan terkait analisis butir soal, ditemukan bahwa lima dari enam soal termasuk dalam kategori valid dan memiliki reliabilitas sedang. Hanya satu soal yang memiliki daya pembeda yang buruk sehingga perlu dihapus atau diperbaiki. Adapun tingkat kesukaran soal menunjukkan bahwa dua soal tergolong mudah dan empat soal tergolong sedang. Dengan demikian, kelima soal yang memenuhi kriteria tersebut dapat digunakan oleh guru dalam proses evaluasi pembelajaran khususnya pada materi pertidaksamaan linear kelas XI. Dalam pelaksanaan penelitian ini, terdapat beberapa kendala, antara lain keterbatasan jumlah soal yang dianalisis dan keterbatasan jumlah responden. Sehingga hasil belum dapat digeneralisasikan secara luas. Oleh karena itu, untuk penelitian selanjutnya disarankan agar menggunakan jumlah soal yang lebih banyak dan melibatkan lebih banyak responden, guna memperoleh gambaran yang lebih menyeluruh tentang kualitas soal. Selain itu, penggunaan perangkat lunak analisis butir soal yang lebih lengkap dapat membantu meningkatkan ketepatan analisis.

UCAPAN TERIMA KASIH

Kami ucapkan terimakasih kepada semua pihak yang telah membantu dalam penelitian ini.

REFERENSI

- Arikunto, Suharsimi. 2002. *Prosedur penelitian: suatu pendekatan praktik/Suharsimi Arikunto*. Rineka Cipta.
- Arikunto, Suharsimi; 2018. *Dasar-dasar evaluasi pendidikan / Suharsimi Arikunto; editor: Restu Damayanti*. Jakarta: Bumi Aksara.
- Azwar, Saifuddin. 2016. "Reliabilitas Dan Validitas Aitem." *Buletin Psikologi* 3(1):19–26. doi:10.22146/bpsi.13381.
- Balau, Metriks, Anekke Pesik, dan I. Wayan Damai. 2021. "Analisis Kualitas Butir Soal Buatan Guru Mata Pelajaran Matematika Kelas VIII SMP Negeri Satap Matabulu Kabupaten Bolaang Mongondow Timur." *MARISEKOLA: Jurnal Matematika Riset Edukasi Dan Kolaborasi* 2(1):13–18. doi:10.53682/marisekola.v2i1.1112.
- Creswell, John W., dan J. David Creswell. 2018. "Research Design: Qualitative, Quantitative, and Mixed Methods Approaches."
- Fadillah, Ila, Ririn Arini, Erwin Salpa Riansi, dan Muhammad Khuluqin Hasan. 2024. "Peran Tes Objektif Dan Subjektif Dalam Mengukur Keterampilan Berfikir Matematis Siswa Sekolah Dasar." *Wilangan: Jurnal Inovasi dan Riset Pendidikan Matematika* 5(3):236–44. doi:10.62870/wjirpm.v5i3.26099.
- Firmansyah, Hasim, dan Sridadi. 2018. "Analisis Butir Soal Tes Pilihan Ganda Mata Pelajaran Pendidikan Jasmani Olahraga Dan Kesehatan Kelas Viii SMP N 1 Wonosari Tahun Ajaran 2017/2018." *Pendidikan Jasmani Kesehatan Dan Rekreasi* 7(12). <https://journal.student.uny.ac.id/ojs/pjkr/article/view/14605>.
- Fitrianti, Leni. 2018. "Prinsip Kontinuitas Dalam Evaluasi Proses Pembelajaran." *AL-ISHLAH: Jurnal Pendidikan* 10(1):89–102. doi:10.35445/alishlah.v10i1.68.
- Fitrianty, Fenny, Alfi Yunita, dan Ratulani Juwita. 2022. "Pengembangan Instrumen Tes Kemampuan Pemecahan Masalah Matematika di SMP Negeri 12 Padang." *Lattice Journal : Journal of Mathematics Education and Applied* 2(1):91. doi:10.30983/lattice.v2i1.5337.
- Halik, Andi Surahma, Sitti Mania, dan Fitriani Nur. 2019b. "Analisis Butir Soal Ujian Akhir Sekolah (UAS) Mata Pelajaran Matematika Pada Tahun Ajaran 2015/2016 Smp Negeri 36 Makassar." *Al asma : Journal of Islamic Education* 1(1):11–17. doi:10.24252/asma.v1i1.11249.
- Hanna, Wiwin Fajriah, dan Heri Retnawati. 2022. "Analisis Kualitas Butir Soal Matematika Menggunakan Model Rasch Dengan Bantuan Software Quest." *AKSIOMA: Jurnal Program Studi Pendidikan Matematika* 11(4):3695–3704. doi:10.24127/ajpm.v11i4.5908.
- Hazlita, Hazlita, Zulkardi Zulkardi, dan Darmawijoyo Darmawijoyo. 2015. "Pengembangan Soal Penalaran Model TIMSS Konteks Sumatera Selatan di Kelas IX SMP." *Kreano, Jurnal Matematika Kreatif-Inovatif* 5(2):170. doi:10.15294/kreano.v5i2.3326.
- Iskandar, Akbar, dan Muhammad Rizal. 2018b. "Analisis kualitas soal di perguruan tinggi berbasis aplikasi TAP." *Jurnal Penelitian dan Evaluasi Pendidikan* 22(1):12–23.

doi:10.21831/pep.v22i1.15609.

- Iskandar, Askar Jaya, Rini Wartu, dan Zaini. 2022. *Statistik Pendidikan: Teori Dan Aplikasi Spss*. Penerbit NEM.
- Junaidi, Juliandry Kurniawan, Liza Husnita, Meldawati Meldawati, dan Kaksim Kaksim. 2024. "Peningkatan Aktifitas Belajar Mahasiswa Dengan Memanfaatkan Microsoft Excel Dalam Analisis Kualitas Butir Soal." *Journal of Education Research* 5(1):875–81. doi:10.37985/jer.v5i1.921.
- Lestari Eka, Karunia, dan Mokhammad Ridwan Yudhanegara. 2018. *Penelitian Pendidikan Matematika*. Cet 3. Bandung: Refika Aditama.
- Magdalena, Ina, Indah Ayu Anggraini, dan Siti Khoiriah. 2021. "Analisis Daya Pembeda Dan Taraf Kesukaran Pada Soal Bilangan Romawi Kelas 4 SDN Tobat 1 Balaraja." *NUSANTARA* 3(1):151–58. doi:10.36088/nusantara.v3i1.1284.
- Nitko, Anthony J., dan Susan M. Brookhart. 2011. *Educational Assessment of Students*. Pearson/Allyn & Bacon.
- Nurhayati, Siti, Halini Halini, dan Tari Indriani. 2024. "Analisis Daya Pembeda Butir Pada Materi Persamaan dan Pertidaksamaan Nilai Mutlak dengan Metode Teori Tes Klasik." *Jurnal MathEdu (Mathematic Education Journal)* 7(1):57–63.
- Pamujo, Pamujo, dan Dynar Rizqi Mara Romadhoni. 2022b. "Analisis Butir Soal Evaluasi Pembelajaran PKn Kelas IV Sekolah Dasar Dengan ITEMAN DI Kecamatan Banjarnegara." *PROCEEDING UMSURABAYA* 1(1). <https://journal.um-surabaya.ac.id/Pro/article/view/14939>.
- Rizqa, Miftahir, Melly Andrina, Putri Puja Lestari, Kurniawati Fitri Indriyani, dan Ulmi Hidayanti. 2025. "Analisis Butir Soal Tes Kemampuan Siswa Pada Mata Pelajaran Manajemen Perkantoran." *DIAJAR: Jurnal Pendidikan Dan Pembelajaran* 4(3):337–46. doi:10.54259/diajar.v4i3.4029.
- Saputri, Hera Apriliana Saputri, Zulhijrah, Nabila Joti Larasati, dan Shaleh. 2023. "Analisis Instrumen Assesmen: Validitas, Reliabilitas, Tingkat Kesukaran Dan Daya Beda Butir Soal." *Didaktik : Jurnal Ilmiah PGSD STKIP Subang* 9(5):2986–95. doi:10.36989/didaktik.v9i5.2268.
- Sudjana, Nana. 1995. *Penilaian hasil proses belajar mengajar*. Bandung: PT Remaja Rosdakarya.
- Supriadi, Gito. 2021. *Statistik Penelitian Pendidikan*. 1 ed. Yogyakarta: UNY Press.
- Surbakti, Icca Widya. 2025. "Analisis Kualitas Butir Soal Pada Uji Coba Evaluasi Pembelajaran Matematika." *CENDEKIA: Jurnal Ilmu Pengetahuan* 5(1):116–25. doi:10.51878/cendekia.v5i1.4097.
- Susanto, Hery, Achi Rinaldi, dan Novalia Novalia. 2015. "Analisis Validitas Reliabilitas Tingkat Kesukaran Dan Daya Beda Pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika Kelas Xii Ips Di Sma Negeri 12 Bandar Lampung Tahun Ajaran 2014/2015." *Al-Jabar : Jurnal Pendidikan Matematika* 6(2):203–18. doi:10.24042/ajpm.v6i2.50.

Syaifudin, Syaifudin. 2023. "Analisis Butir Soal Penilaian Akhir Tahun Mata Pelajaran Matematika Dengan Item And Test Analysis." *JURNAL JENDELA MATEMATIKA* 1(2):34–42. doi:10.57008/jjm.v1i2.559.

Syaripudin, Tatang. 2012. *Ilmu Pendidikan*. 1 ed. Bandung: Pustaka Setia.

Wulandari, Enika, dan Erni Ayda. 2021. "Analysis of Open-Ended Based Mathematics Learning Tools Oriented to Higher Order Thinking Skills Viewed from the Aspect of Validity." *Edumatica : Jurnal Pendidikan Matematika* 11(03):32–39. doi:10.22437/edumatica.v11i03.14404.