

## **Analisis Butir Soal SAS Matematika Kelas X SMA Negeri 25 Makassar dengan Anates 4.0**

**Suryadi Ishak<sup>1</sup>, Irfawandi Samad<sup>2</sup>, Muh. Khairul Arfandi<sup>3</sup>**

*Universitas Negeri Makassar*

Email: [suryadi.ishak@unm.ac.id](mailto:suryadi.ishak@unm.ac.id), [ipbank05@gmail.com](mailto:ipbank05@gmail.com), [khairulfandi39411@gmail.com](mailto:khairulfandi39411@gmail.com)

### **Abstrak**

Penelitian ini bertujuan untuk menganalisis kualitas butir soal Sumatif Akhir Semester mata pelajaran Matematika kelas X di SMA Negeri 25 Makassar dengan menggunakan program Anates Versi 4.0 for Windows. Analisis dilakukan untuk mengetahui tingkat reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh pada setiap butir soal. Penelitian ini menggunakan jenis penelitian deskriptif dengan pendekatan kuantitatif. Data penelitian diperoleh dari lembar jawaban peserta didik kelas X yang mengikuti Sumatif Akhir Semester tahun ajaran 2025/2026. Data dianalisis menggunakan bantuan program Anates Versi 4.0 for Windows untuk memperoleh informasi mengenai karakteristik setiap butir soal. Hasil analisis menunjukkan bahwa instrumen tes memiliki koefisien reliabilitas sebesar 0,88 yang termasuk dalam kategori sangat tinggi, sehingga instrumen tes memiliki tingkat konsistensi yang baik. Analisis tingkat kesukaran menunjukkan bahwa dari 15 butir soal yang dianalisis terdapat 10 butir soal (66,67%) berkategori sedang, 4 butir soal (26,67%) berkategori mudah, dan 1 butir soal (6,67%) berkategori sangat mudah. Selanjutnya, hasil analisis daya pembeda menunjukkan bahwa sebagian besar butir soal memiliki daya pembeda yang cukup hingga sangat baik dengan rentang nilai 37,50% sampai 93,75%. Analisis kualitas pengecoh menunjukkan bahwa sebagian besar alternatif jawaban telah berfungsi dengan baik. Dengan demikian, secara umum instrumen tes yang digunakan memiliki kualitas yang cukup baik dan layak digunakan sebagai alat evaluasi pembelajaran.

**Kata Kunci:** *Analisis butir soal, Reliabilitas tes, Tingkat kesukaran, Daya pembeda, Anates.*

### **PENDAHULUAN**

Pendidikan merupakan proses sistematis yang bertujuan untuk mengembangkan potensi peserta didik secara optimal. Salah satu komponen penting dalam proses pendidikan adalah evaluasi pembelajaran. Evaluasi berfungsi untuk mengetahui tingkat ketercapaian tujuan pembelajaran, efektivitas proses pembelajaran, serta sebagai dasar pengambilan keputusan dalam perbaikan pembelajaran. Evaluasi pendidikan merupakan kegiatan untuk mengumpulkan informasi tentang pencapaian belajar peserta didik yang kemudian digunakan sebagai dasar dalam menentukan langkah selanjutnya (Suharsimi Arikunto, 2018). Dengan demikian, kualitas instrumen evaluasi menjadi faktor yang sangat menentukan ketepatan hasil penilaian.

Dalam konteks pembelajaran matematika, instrumen tes harus mampu mengukur kemampuan berpikir logis, analitis, sistematis, dan kritis peserta didik. Tes yang baik harus memenuhi persyaratan validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta efektivitas distraktor. Analisis butir soal bertujuan untuk mengetahui mutu setiap butir soal sehingga dapat ditentukan apakah soal tersebut layak digunakan, perlu direvisi, atau harus dibuang (Anas Sudijono, 2015). Tanpa analisis yang tepat, hasil tes berpotensi menghasilkan informasi yang bias dan kurang akurat dalam menggambarkan kemampuan siswa.

Secara teoretis, analisis butir soal merupakan bagian dari pendekatan evaluasi berbasis data. Analisis item membantu pendidik memahami karakteristik soal serta meningkatkan kualitas pengukuran hasil belajar (Anthony J. Nitko & Susan M. Brookhart, 2014). Dengan melakukan analisis tingkat kesukaran dan daya pembeda, guru dapat memastikan bahwa soal tidak terlalu mudah atau terlalu sulit, serta mampu membedakan siswa berkemampuan tinggi dan rendah secara proporsional.

Namun demikian, dalam praktik di lapangan, analisis butir soal belum sepenuhnya dilakukan secara sistematis. Banyak pendidik masih melakukan penilaian tanpa analisis kuantitatif yang mendalam, terutama karena keterbatasan waktu dan kompleksitas perhitungan manual. Padahal, standar penilaian pendidikan menegaskan bahwa penilaian hasil belajar harus dilakukan secara objektif, terpadu, dan berkesinambungan (Kementerian Pendidikan dan Kebudayaan Republik Indonesia, 2016). Hal ini menunjukkan bahwa penggunaan instrumen yang valid dan reliabel merupakan suatu keharusan dalam proses evaluasi pembelajaran.

Perkembangan teknologi memberikan solusi terhadap kendala tersebut melalui penggunaan perangkat lunak analisis tes. Salah satu program yang dapat digunakan adalah Anates Versi 4.0 for Windows yang dirancang untuk membantu guru menganalisis kualitas butir soal secara cepat dan akurat, meliputi perhitungan validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta efektivitas pengecoh. Pemanfaatan perangkat lunak ini dapat meningkatkan efisiensi, akurasi, serta objektivitas dalam proses analisis soal.

Berdasarkan pembahasan diatas, peneliti tertarik untuk menganalisis kualitas butir soal Sumatif Akhir Semester Matematika kelas X di SMA Negeri 25 Makassar menggunakan program Anates Versi 4.0 for Windows. Analisis ini bertujuan untuk mengetahui tingkat validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta efektivitas distraktor pada setiap butir soal. Hasil penelitian diharapkan dapat menjadi dasar perbaikan instrumen evaluasi serta meningkatkan mutu penilaian pembelajaran matematika di sekolah.

## **METODE**

### **Jenis dan Pendekatan Penelitian**

Penelitian ini menggunakan jenis penelitian deskriptif dengan pendekatan kuantitatif. Penelitian deskriptif bertujuan untuk menggambarkan secara sistematis dan akurat mengenai fakta serta karakteristik objek yang diteliti tanpa memberikan perlakuan tertentu terhadap variabel penelitian (Suharsimi Arikunto, 2018). Pendekatan kuantitatif digunakan karena data yang dianalisis berupa skor hasil tes siswa yang kemudian diolah secara statistik untuk mengetahui kualitas butir soal berdasarkan indikator validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas distraktor. Pendekatan kuantitatif menekankan pada analisis data numerik yang diolah menggunakan prosedur statistik untuk memperoleh kesimpulan yang objektif (Sugiyono, 2019).

Pemilihan jenis dan pendekatan ini didasarkan pada tujuan penelitian, yaitu mendeskripsikan kualitas butir soal Sumatif Akhir Semester Matematika kelas X secara empiris berdasarkan hasil analisis statistik. Data penelitian diperoleh dari lembar jawaban siswa yang kemudian dianalisis menggunakan bantuan program Anates Versi 4.0 for

Windows. Melalui analisis tersebut, diperoleh informasi kuantitatif mengenai karakteristik setiap butir soal sehingga dapat ditentukan kategori kelayakannya. Penelitian deskriptif kuantitatif dinilai tepat karena mampu memberikan gambaran objektif mengenai kondisi instrumen tes berdasarkan data yang terukur (Anas Sudijono, 2015).

### **Subjek Penelitian**

Subjek dalam penelitian ini adalah peserta didik kelas X di SMA Negeri 25 Makassar yang mengikuti Sumatif Akhir Semester 1 mata pelajaran Matematika pada tahun ajaran 2025/2026. Pemilihan subjek dilakukan secara purposive, yaitu berdasarkan pertimbangan bahwa siswa kelas X merupakan peserta yang secara langsung mengikuti dan mengerjakan instrumen tes yang dianalisis. Subjek penelitian ini berperan sebagai sumber data utama berupa lembar jawaban siswa yang digunakan untuk menganalisis kualitas butir soal.

### **Instrumen**

Instrumen yang digunakan dalam penelitian ini adalah tes hasil belajar berupa soal Sumatif Akhir Semester mata pelajaran Matematika kelas X di SMA Negeri 25 Makassar. Bentuk tes yang digunakan adalah pilihan ganda dengan empat atau lima opsi jawaban. Instrumen tes merupakan alat yang digunakan untuk mengukur kemampuan peserta didik secara sistematis dan objektif (Suharsimi Arikunto, 2018). Data yang diperoleh dari lembar jawaban siswa selanjutnya dianalisis untuk mengetahui kualitas setiap butir soal secara kuantitatif.

Analisis butir soal dilakukan untuk mengetahui tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Analisis ini bertujuan untuk menentukan mutu setiap butir soal sehingga dapat diputuskan apakah soal tersebut layak digunakan, direvisi, atau dibuang (Anas Sudijono, 2015). Pengolahan data dilakukan menggunakan bantuan program Anates Versi 4.0 for Windows untuk memperoleh hasil perhitungan yang akurat.

### **Validitas Butir Soal**

Validitas butir soal bertujuan untuk mengetahui sejauh mana suatu item memberikan kontribusi terhadap skor total tes. Pengujian validitas dilakukan dengan cara mengorelasikan skor pada setiap butir soal dengan skor total yang diperoleh peserta didik. Suatu butir soal dinyatakan memiliki validitas tinggi apabila skor pada butir tersebut menunjukkan hubungan yang kuat dengan skor total tes. Besarnya kontribusi masing-masing butir soal dinyatakan dalam bentuk koefisien korelasi, sehingga perhitungan validitas item dilakukan menggunakan rumus korelasi yang sesuai. Perhitungan dilakukan dengan menggunakan rumus korelasi product moment Pearson (Suharsimi Arikunto, 2005). Interpretasi besarnya koefisien korelasi dapat dilihat pada Tabel 1.

Tabel 1. Kategori Validitas Soal

Tingkat	Kategori
0,81 – 1,00	Sangat Tinggi
0,61 – 0,80	Tinggi
0,41 – 0,60	Cukup
0,21 – 0,40	Rendah
0,00 – 0,20	Sangat Rendah

Tingkat kesukaran menunjukkan proporsi siswa yang menjawab benar suatu butir soal. Rumus yang digunakan adalah  $P = B/N$  ( $P$  = indeks tingkat kesukaran;  $B$  = jumlah siswa yang menjawab benar;  $N$  = jumlah seluruh siswa). Kriteria interpretasi:

Tabel 2. Kategori Tingkat Kesukaran

Tingkat	Kategori
0,00 – 0,30	sukar
0,31 – 0,70	sedang
0,71 – 1,00	mudah

(Suharsimi Arikunto, 2018)

Daya pembeda menunjukkan kemampuan suatu butir soal dalam membedakan siswa berkemampuan tinggi dan rendah. Rumus yang digunakan adalah  $D = (B_A - B_B) / (N - 2)$  ( $D$  = indeks daya pembeda;  $B_A$  = jumlah siswa kelompok atas yang menjawab benar;  $B_B$  = jumlah siswa kelompok bawah yang menjawab benar;  $N$  = jumlah seluruh siswa). Kriteria interpretasi:

Tabel 3. Kategori Daya Pembeda

Tingkat	Kategori
0,00 – 0,20	jelek
0,21 – 0,40	cukup
0,41 – 0,70	baik
0,71 – 1,00	sangat baik

(Anas Sudijono, 2015)

Efektivitas pengecoh (distractor effectiveness) merupakan salah satu indikator penting dalam analisis butir soal pilihan ganda. Pengecoh adalah alternatif jawaban salah yang disusun untuk mengecoh peserta didik yang belum menguasai materi. Pengecoh yang baik harus mampu menarik perhatian siswa yang berkemampuan rendah, namun tidak dipilih oleh siswa yang berkemampuan tinggi. Analisis efektivitas pengecoh bertujuan untuk mengetahui apakah setiap opsi jawaban salah telah berfungsi sebagaimana mestinya dalam membedakan kemampuan peserta didik (Suharsimi Arikunto, 2018). Secara kuantitatif, efektivitas pengecoh dapat dihitung dengan melihat persentase siswa yang memilih setiap opsi jawaban. Rumus yang digunakan adalah  $IP = (P/N) \times 100\%$  ( $IP$  = indeks pengecoh;  $P$  = jumlah siswa yang memilih pengecoh;  $N$  = jumlah seluruh peserta tes).

Suatu pengecoh dinyatakan efektif apabila memenuhi kriteria berikut:

- Dipilih oleh sekurang-kurangnya 5% dari jumlah peserta tes.
- Lebih banyak dipilih oleh kelompok bawah dibandingkan kelompok atas.
- Tidak ada pengecoh yang sama sekali tidak dipilih (0%).

Apabila suatu pengecoh tidak memenuhi kriteria tersebut, maka pengecoh dianggap tidak berfungsi dan perlu direvisi. Pengecoh yang tidak berfungsi dapat mengurangi kualitas butir soal karena tidak mampu menjalankan perannya sebagai pembeda kemampuan siswa (Anas Sudijono, 2015).

Reliabilitas merupakan salah satu syarat penting dalam penyusunan instrumen evaluasi karena berkaitan dengan tingkat konsistensi dan keajegan hasil pengukuran. Suatu

tes dikatakan reliabel apabila memberikan hasil yang relatif sama ketika digunakan berulang kali pada kelompok yang setara dalam kondisi yang sama (Suharsimi Arikunto, 2018). Dengan kata lain, reliabilitas menunjukkan sejauh mana suatu instrumen dapat dipercaya sebagai alat ukur hasil belajar peserta didik.

Dalam analisis butir soal, reliabilitas tes objektif (seperti pilihan ganda) umumnya dihitung menggunakan rumus Kuder-Richardson (KR-20) atau Alpha Cronbach. Rumus KR-20 yang sering digunakan untuk soal dengan skor dikotomi (benar = 1, salah = 0). Menurut Anas Sudijono (2015), nilai koefisien reliabilitas berkisar antara 0 sampai 1. Semakin mendekati angka 1, maka tingkat reliabilitas tes semakin tinggi. Secara umum, kriteria reliabilitas dapat diinterpretasikan sebagai berikut:

Tabel 4. Kategori Reliabilitas Instrumen

Tingkat	Kategori
0,00 – 0,20	Sangat rendah
0,21 – 0,40	Rendah
0,41 – 0,60	Sedang
0,61 – 0,80	Tinggi
0,81 – 1,00	Sangat tinggi

## HASIL DAN PEMBAHASAN

### Reliabilitas tes

Reliabilitas merupakan tingkat konsistensi suatu alat ukur dalam menghasilkan data ketika digunakan secara berulang pada subjek yang sama. Reliabilitas adalah konsistensi pengukuran dan stabilitas skor artinya pengukuran yang sama menghasilkan hasil yang sama (Defrijon dkk., 2025). Suatu instrumen dikatakan memiliki reliabilitas tinggi apabila mampu memberikan hasil yang stabil dan relatif tetap, meskipun pengukuran dilakukan pada waktu yang berbeda. Dengan demikian, reliabilitas menjadi salah satu syarat penting yang harus dipenuhi oleh instrumen penelitian sebelum digunakan dalam proses pengumpulan data, karena menunjukkan ketepatan dan konsistensi hasil pengukuran. Adapun hasil analisis reliabilitas terhadap instrumen tes yang diteliti disajikan sebagai berikut:

```

RELIABILITAS TES
=====

Rata2= 10,52
Simpang Baku= 4,20
KorelasiXY= 0,79
Reliabilitas Tes= 0,88
Nama berkas: C:\USERS\SUDAR\ONEDRIVE\DOKUMEN\ARTIKEL.ANA
    
```

Gambar 1. Hasil Perhitungan Reliabilitas Tes

Berdasarkan hasil analisis reliabilitas tes, diperoleh nilai rata-rata skor sebesar 10,52 dengan simpangan baku sebesar 4,20. Hasil perhitungan juga menunjukkan bahwa korelasi antara dua bagian tes (XY) sebesar 0,79, sehingga diperoleh koefisien reliabilitas tes sebesar 0,88. Nilai tersebut menunjukkan bahwa soal yang diberikan memiliki tingkat reliabilitas yang sangat tinggi, sehingga soal yang digunakan dalam penelitian ini dapat dinyatakan memiliki konsistensi yang baik dan layak digunakan sebagai alat pengumpulan data

penelitian. Hal ini sejalan dengan pendapat Anas Sudijono (2015) yang menyatakan bahwa semakin tinggi koefisien reliabilitas suatu tes, maka semakin tinggi pula tingkat keajegan atau konsistensi tes tersebut dalam mengukur kemampuan peserta didik.

### Tingkat Kesukaran

Tingkat kesukaran soal merupakan salah satu indikator penting dalam analisis butir soal yang digunakan untuk mengetahui seberapa mudah atau sulit suatu soal bagi peserta didik. Tingkat kesukaran dihitung berdasarkan proporsi jumlah peserta didik yang dapat menjawab soal dengan benar dibandingkan dengan seluruh peserta tes. Semakin besar nilai indeks kesukaran, maka soal tersebut semakin mudah, sedangkan semakin kecil nilainya maka soal tersebut semakin sulit. Menurut Anas Sudijono (2015), indeks kesukaran menunjukkan derajat kesulitan suatu butir soal yang ditentukan dari banyaknya peserta tes yang mampu menjawab soal tersebut dengan benar. Oleh karena itu, analisis tingkat kesukaran penting dilakukan untuk mengetahui kualitas butir soal yang digunakan dalam penelitian. Hasil analisis tingkat kesukaran soal dalam penelitian ini dapat dilihat pada tabel berikut.

TINGKAT KESUKARAN  
 \*\*\*\*\*

Jumlah Subyek= 60  
 Butir Soal= 15  
 Nama berkas: C:\USERS\SUDAR\ONEDRIVE\DOKUMEN\ARTIKEL.ANA

No Butir	Jml Betul	Tkt. Kesukaran(%)	Tafsiran
1	39	65,00	Sedang
2	41	68,33	Sedang
3	42	70,00	Sedang
4	49	81,67	Mudah
5	46	76,67	Mudah
6	39	65,00	Sedang
7	34	56,67	Sedang
8	50	83,33	Mudah
9	52	86,67	Sangat Mudah
10	35	58,33	Sedang
11	44	73,33	Mudah
12	39	65,00	Sedang
13	41	68,33	Sedang
14	41	68,33	Sedang
15	39	65,00	Sedang

Gambar 2. Hasil Perhitungan Tingkat Kesukaran

Berdasarkan Tabel 4 hasil analisis tingkat kesukaran terhadap 15 butir soal, diketahui bahwa mayoritas soal tergolong dalam kategori sedang. Sebanyak 10 butir soal atau sebesar 66,67% berada pada kategori sedang, yaitu pada butir 1, 2, 3, 6, 7, 10, 12, 13, 14, dan 15. Selain itu, terdapat 4 butir soal atau sebesar 26,67% yang termasuk dalam kategori mudah, yaitu butir 4, 5, 8, dan 11. Sementara itu, terdapat 1 butir soal atau sebesar 6,67% yang tergolong sangat mudah, yaitu butir 9.

Berdasarkan hasil tersebut dapat disimpulkan bahwa sebagian besar soal berada pada tingkat kesukaran sedang, sehingga instrumen soal yang digunakan dalam penelitian sudah cukup baik karena tidak terlalu mudah maupun terlalu sulit. Proporsi tingkat kesukaran yang didominasi oleh kategori sedang dapat membantu tes dalam mengukur kemampuan peserta didik secara lebih tepat serta memberikan peluang yang lebih baik untuk membedakan tingkat kemampuan antar peserta tes.

## Daya Pembeda

Daya pembeda merupakan salah satu indikator penting dalam analisis butir soal yang digunakan untuk mengetahui kemampuan suatu soal dalam membedakan peserta didik yang memiliki kemampuan tinggi dan kemampuan rendah. Suatu butir soal dikatakan memiliki daya pembeda yang baik apabila peserta didik pada kelompok atas lebih banyak menjawab benar dibandingkan dengan peserta didik pada kelompok bawah. Analisis daya pembeda dilakukan dengan membandingkan jumlah jawaban benar antara kelompok atas dan kelompok bawah. Menurut Siti Nurjanah (2020) dalam penelitiannya tentang analisis butir soal, daya pembeda berfungsi untuk mengetahui sejauh mana suatu butir soal mampu membedakan kemampuan peserta didik sehingga dapat digunakan untuk menentukan kualitas soal dalam suatu tes. Hasil analisis daya pembeda pada penelitian ini dapat dilihat pada tabel berikut.

```
DAYA PEMBEDA
=====

Jumlah Subyek= 60
Klp atas/bawah(n)= 16
Butir Soal= 15
Nama berkas: C:\USERS\SUDAR\ONEDRIVE\DOKUMEN\ARTIKEL.ANA
```

No Butir	Kel. Atas	Kel. Bawah	Beda	Indeks DP (%)
1	16	5	11	68,75
2	16	3	13	81,25
3	16	2	14	87,50
4	16	9	7	43,75
5	16	7	9	56,25
6	16	5	11	68,75
7	16	1	15	93,75
8	16	10	6	37,50
9	16	10	6	37,50
10	15	5	10	62,50
11	16	4	12	75,00
12	16	1	15	93,75
13	16	4	12	75,00
14	16	6	10	62,50
15	16	6	10	62,50

Gambar 3. Hasil Perhitungan Daya Pembeda

Berdasarkan hasil analisis daya pembeda terhadap 15 butir soal, diketahui bahwa nilai indeks daya pembeda berkisar antara 37,50% hingga 93,75%. Nilai tertinggi terdapat pada butir soal nomor 7 dan 12 dengan indeks daya pembeda sebesar 93,75%, yang menunjukkan bahwa soal tersebut memiliki kemampuan yang sangat baik dalam membedakan peserta didik kelompok atas dan kelompok bawah. Sementara itu, nilai indeks daya pembeda terendah terdapat pada butir soal nomor 8 dan 9 dengan nilai 37,50%, namun masih termasuk kategori cukup sehingga masih dapat digunakan dalam instrumen tes.

Secara keseluruhan, hasil analisis menunjukkan bahwa sebagian besar butir soal memiliki daya pembeda yang baik, sehingga instrumen tes yang digunakan dalam penelitian ini mampu membedakan kemampuan peserta didik secara efektif.

## Kualitas Pengecoh

Kualitas pengecoh (distractor) merupakan salah satu komponen penting dalam analisis butir soal pilihan ganda. Pengecoh adalah alternatif jawaban selain kunci jawaban yang berfungsi untuk mengecoh peserta didik yang belum memahami materi dengan baik.

Pengecoh yang baik seharusnya dipilih oleh sebagian peserta didik, khususnya dari kelompok berkemampuan rendah, sehingga dapat membantu meningkatkan kualitas butir soal. Apabila suatu pengecoh tidak dipilih oleh peserta didik, maka pengecoh tersebut dianggap tidak berfungsi dengan baik dan perlu diperbaiki. Menurut Djemari Mardapi (2012), pengecoh yang baik adalah pengecoh yang mampu menarik perhatian peserta tes dan dipilih oleh sejumlah peserta didik sehingga dapat meningkatkan kualitas soal dalam suatu tes. Oleh karena itu, analisis kualitas pengecoh perlu dilakukan untuk mengetahui apakah setiap alternatif jawaban telah berfungsi dengan baik atau belum. Hasil analisis kualitas pengecoh pada penelitian ini dapat dilihat pada tabel berikut.

KUALITAS PENGECOH  
 \*\*\*\*\*

Jumlah Subyek= 60  
 Butir Soal= 15  
 Nama berkas: C:\USERS\SUDAR\ONEDRIVE\DOKUMEN\ARTIKEL.ANA

No Butir	a	b	c	d	e	*
1	39**	6++	9-	6++	0--	0
2	6+	41**	1--	9--	3+	0
3	3+	4++	3+	8--	42**	0
4	49**	2+	8---	1-	0--	0
5	4++	46**	4++	1-	5+	0
6	1--	39**	18---	1--	1--	0
7	3-	10-	34**	7++	6++	0
8	1-	4-	50**	3++	2++	0
9	0--	2++	52**	5---	1-	0
10	4+	17---	35**	4+	0--	0
11	44**	8--	6+	1--	1--	0
12	5++	10--	3+	39**	3+	0
13	7+	41**	8-	3+	1--	0
14	5++	41**	1--	6+	7+	0
15	12---	4++	4++	39**	1--	0

Keterangan:  
 \*\* : Kunci Jawaban  
 ++ : Sangat Baik  
 + : Baik  
 - : Kurang Baik  
 -- : Buruk  
 --- : Sangat Buruk

Gambar 4. Hasil Perhitungan Kualitas Pengecoh

Berdasarkan hasil analisis kualitas pengecoh terhadap 15 butir soal, diketahui bahwa sebagian besar alternatif jawaban pengecoh telah berfungsi dengan baik. Hal ini terlihat dari adanya peserta didik yang memilih setiap alternatif jawaban selain kunci jawaban pada beberapa butir soal. Beberapa pengecoh bahkan termasuk dalam kategori sangat baik dan baik, yang menunjukkan bahwa pengecoh tersebut mampu menarik perhatian peserta didik yang belum menguasai materi dengan baik.

Namun demikian, pada beberapa butir soal masih terdapat pengecoh yang termasuk dalam kategori kurang baik hingga sangat buruk, yang ditunjukkan dengan sedikitnya atau bahkan tidak adanya peserta didik yang memilih alternatif jawaban tersebut. Kondisi ini menunjukkan bahwa pengecoh pada beberapa soal perlu diperbaiki agar dapat berfungsi lebih efektif dalam mengecoh peserta didik yang memiliki pemahaman materi yang rendah.

Secara keseluruhan, hasil analisis menunjukkan bahwa kualitas pengecoh pada instrumen tes dalam penelitian ini tergolong cukup baik, sehingga sebagian besar butir soal sudah memiliki alternatif jawaban yang berfungsi dengan baik dan dapat digunakan dalam proses pengukuran kemampuan peserta didik.

## **KESIMPULAN**

Berdasarkan hasil analisis instrumen tes yang meliputi reliabilitas, tingkat kesukaran, daya pembeda, dan kualitas pengecoh, dapat disimpulkan bahwa secara umum butir soal yang digunakan dalam penelitian ini memiliki kualitas yang cukup baik. Hasil analisis reliabilitas menunjukkan bahwa instrumen tes memiliki koefisien reliabilitas sebesar 0,88, yang termasuk dalam kategori sangat tinggi, sehingga instrumen tersebut dapat dikatakan konsisten dan layak digunakan sebagai alat pengumpulan data.

Hasil analisis tingkat kesukaran menunjukkan bahwa dari 15 butir soal yang dianalisis, terdapat 10 butir soal (66,67%) berkategori sedang, 4 butir soal (26,67%) berkategori mudah, dan 1 butir soal (6,67%) berkategori sangat mudah. Hal ini menunjukkan bahwa sebagian besar soal memiliki tingkat kesukaran yang baik karena berada pada kategori sedang. Selanjutnya, hasil analisis daya pembeda menunjukkan bahwa nilai indeks daya pembeda berkisar antara 37,50% hingga 93,75%, yang menunjukkan bahwa sebagian besar butir soal memiliki daya pembeda yang cukup baik hingga sangat baik dalam membedakan kemampuan peserta didik kelompok atas dan kelompok bawah.

Berdasarkan analisis kualitas pengecoh, sebagian besar alternatif jawaban pengecoh telah berfungsi dengan baik, meskipun masih terdapat beberapa pengecoh yang kurang optimal karena dipilih oleh sedikit peserta didik. Oleh karena itu, beberapa butir soal masih memerlukan perbaikan pada alternatif jawaban pengecoh agar dapat berfungsi lebih efektif.

Secara keseluruhan, instrumen tes dalam penelitian ini layak digunakan sebagai alat pengukuran kemampuan peserta didik, dengan beberapa perbaikan pada bagian tertentu agar kualitas butir soal menjadi lebih optimal.

## **DAFTAR PUSTAKA**

- Arikunto, S. (2018). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Defrijon dkk. (2025). Reliabilitas alat ukur, jenis-jenis dan faktor-faktor yang mempengaruhinya. *Jurnal Studi Multidisipliner*, 9(1).
- Kementerian Pendidikan dan Kebudayaan Republik Indonesia. (2016). *Permendikbud Nomor 23 Tahun 2016 tentang Standar Penilaian Pendidikan*. Jakarta: Kemendikbud.
- Mardapi, D. (2012). *Pengukuran, Penilaian, dan Evaluasi Pendidikan*. Yogyakarta: Nuha Medika
- Nitko, A. J., & Brookhart, S. M. (2014). *Educational Assessment of Students*. Pearson Education.
- Nurjanah, S. (2020). Analisis Butir Soal Pilihan Ganda pada Ujian Sekolah. *Jurnal Pendidikan dan Evaluasi Pendidikan*, 4(2), 120–128
- Sudijono, A. (2015). *Pengantar Evaluasi Pendidikan*. Jakarta: PT RajaGrafindo Persada.
- Sugiyono. (2019). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Alfabeta.