

## ANALISIS BUTIR SOAL PENGEMBANGAN TEKNOLOGI DI BIDANG INDUSTRI DAN DUNIA KERJA

Dian Intan Khairiah \*<sup>1</sup>

Siti Nabila Indah <sup>2</sup>

Luqman Hakim <sup>3</sup>

Vivi Pratiwi <sup>4</sup>

<sup>1,2,3,4</sup> Universitas Negeri Surabaya

\*e-mail: [dian.23005@mhs.unesa.ac.id](mailto:dian.23005@mhs.unesa.ac.id) <sup>1</sup>, [siti.23082@mhs.unesa.ac.id](mailto:siti.23082@mhs.unesa.ac.id) <sup>2</sup> [luqmanhakim@unesa.ac.id](mailto:luqmanhakim@unesa.ac.id) <sup>3</sup>  
[vivipratiwi@unesa.ac.id](mailto:vivipratiwi@unesa.ac.id) <sup>4</sup>

### Abstrak

Penelitian ini bertujuan untuk menganalisis kualitas butir soal dalam mata pelajaran Pengembangan Teknologi di Industri dan Dunia Kerja menggunakan aplikasi Anates. Melalui analisis yang melibatkan 50 siswa SMKN 2 Mataram, penelitian ini mengevaluasi aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh dari 10 butir soal pilihan ganda. Hasil analisis menunjukkan bahwa sebagian besar butir soal memiliki daya beda yang baik, meskipun terdapat beberapa soal yang perlu diperbaiki. Selain itu, efektivitas pengecoh juga bervariasi, dengan beberapa soal menunjukkan kategori sangat baik, sementara yang lain tergolong kurang baik. Temuan ini diharapkan dapat memberikan rekomendasi untuk perbaikan kualitas soal dan mendukung peningkatan proses pembelajaran di sekolah.

**Kata kunci :** Teknologi, Aplikasi Anates, efektivitas pengecoh, pendidikan.

### Abstract

This study aims to analyze the quality of test items in the subject of Technology Development in Industry and the World of Work using the Anates application. Through an analysis involving 50 students of SMKN 2 Mataram, this study covers aspects of validity, reliability, level of difficulty, discriminating power, and distractor effectiveness of 10 multiple-choice test items. The results of the analysis show that most test items have good discriminating power, although there are some questions that need to be improved. In addition, the effectiveness of distractors also varies, with some questions showing a very good category, while others are classified as less good. These findings are expected to provide recommendations for improving the quality of questions and supporting the improvement of the learning process in schools.

**Keywords:** Technology, Anates Application, distractor effectiveness, education.

### PENDAHULUAN

Salah satu aspek yang sangat penting dalam hidup adalah pendidikan. Pendidikan dapat berlangsung efektif dan optimal jika semua elemennya berfungsi dengan baik. Salah satu elemen yang mendukung proses pendidikan adalah evaluasi, yang juga merupakan hal yang sangat signifikan di dalamnya. Pendidikan memerlukan evaluasi sebagai cara atau aktivitas untuk mengendalikan, menjamin, dan menetapkan kualitas pendidikan. Evaluasi dilaksanakan secara menyeluruh di setiap jalur atau tingkat pendidikan. Dengan kata lain, saat membahas pendidikan, kita tidak bisa terlepas dari evaluasi pendidikan. Pendidikan adalah proses interaksi antara pengajar dan siswa untuk mencapai tujuan yang telah ditentukan. Pengajar, siswa, dan tujuan pendidikan adalah elemen utama dalam pendidikan. Jika salah satu dari ketiganya tidak ada, maka gagasan pendidikan itu sendiri akan hilang.

Evaluasi di bidang pendidikan di Indonesia sering dianggap sebagai fase akhir dalam proses pembelajaran. Namun, pandangan itu tidak sepenuhnya akurat karena ada sejumlah evaluasi yang dilakukan secara bertahap. Dari sini, kita dapat melihat bahwa penggunaan evaluasi yang sesuai dapat secara efektif mengukur hasil dari proses belajar. Selain itu, ini juga bisa menjadi alat untuk mengembangkan tipe soal yang digunakan (Muhson dkk., 2016). Dan juga untuk membandingkan

mutu soal yang diterapkan (Subali dkk., 2021). Tujuan dari evaluasi yang dilakukan adalah untuk mengevaluasi seberapa jauh pencapaian tujuan pendidikan (Hamzah, 2014). Dalam melakukan evaluasi di dunia pendidikan peran pendidik juga sangat diperlukan untuk memberikan penilaian hasil belajar terhadap peserta didiknya. Dengan kata lain, penilaian berfungsi sebagai alat untuk mengetahui keberhasilan proses dan hasil belajar siswa.

Dalam proses evaluasi, diperlukan data terkait objek yang sedang dinilai. Dalam konteks kegiatan belajar dan mengajar, data yang relevan mencakup perilaku atau penampilan siswa selama proses pembelajaran, hasil ujian, dan nilai akhir semester. Untuk mengevaluasi hasil belajar, penting untuk mengikuti langkah-langkah evaluasi sesuai dengan jenis tes atau alat yang akan digunakan untuk menilai hasil pengajaran, serta memperhatikan materi dan metode pengajaran yang diterapkan serta tujuan pengajaran yang telah ditetapkan. Purwanto (2011) menyatakan bahwa evaluasi merupakan usaha yang sistematis dan teliti untuk memahami kemampuan serta perkembangan siswa, baik sebelum, selama, maupun setelah proses belajar, melalui pengumpulan data, serta melakukan perbandingan dengan norma atau kriteria tertentu. Tes merupakan aktivitas atau proses yang terstruktur untuk menilai kemampuan individu. Aktivitas tes selalu memanfaatkan instrumen yang juga dikenal sebagai tes. Sesuai dengan pendapat Sudjana dan Ibrahim, terdapat dua jenis tes, yaitu tes pencapaian belajar dan tes psikologi. Tes pencapaian belajar terdiri dari dua jenis soal, yaitu soal berbentuk uraian dan soal pilihan ganda. Tes yang berkualitas disusun mengikuti prosedur dan prinsip tertentu dalam pembuatannya. Tes yang efektif dapat digunakan kembali dengan beberapa penyesuaian. Sementara itu, tes yang berkualitas rendah sebaiknya dihapus, dan jika perlu, tidak digunakan untuk memberikan penilaian kepada siswa.

Analisis tes merupakan salah satu aktivitas yang bertujuan untuk merancang tes guna mendapatkan pemahaman mengenai kualitas tes, baik secara keseluruhan maupun pada setiap pertanyaan. Hubungan antara tes, pengukuran, dan penilaian menunjukkan bahwa evaluasi hasil belajar dapat dilakukan dengan efektif dan akurat hanya jika didasarkan pada informasi yang diperoleh dari pengukuran hasil belajar yang memanfaatkan tes sebagai instrumen. Fungsi tes, pengukuran, dan penilaian di bidang pendidikan meliputi seleksi, penempatan, diagnosis, perbaikan, umpan balik, pendorong motivasi, bimbingan, serta peningkatan kurikulum, program pendidikan, dan pengembangan ilmu pengetahuan. Untuk mempermudah analisis sebuah tes, penting bagi guru dan calon guru memiliki alat yang dapat membantu dalam menilai hasil dari soal-soal tersebut. Salah satu alat yang dapat digunakan adalah aplikasi ANATES, yang berguna untuk menganalisis kualitas dari setiap butir soal. Untuk mengumpulkan data yang relevan, diperlukan tes yang berkualitas baik. Umumnya, analisis tes memenuhi standar tinggi untuk validitas, reliabilitas yang kuat, kemampuan membedakan yang baik, serta tingkat kesulitan yang sesuai. Untuk mengevaluasi kriteria tes yang berkaitan dengan Pengembangan Teknologi di Industri dan Dunia Kerja, telah dilakukan pengujian instrumen dan analisisnya untuk memperoleh informasi tentang validitas, kemampuan membedakan, dan tingkat kesulitan dari soal-soal yang disusun.

## **METODE**

Penelitian ini menggunakan pendekatan kuantitatif dengan desain deskriptif untuk menganalisis kualitas butir soal dalam pengembangan teknologi di industri dan dunia kerja. Penelitian ini melibatkan siswa kelas X di SMKN 2 Mataram, dengan sampel sebanyak 50 siswa yang dipilih secara acak. Instrumen yang digunakan adalah seperangkat soal pilihan ganda yang terdiri dari 10 butir soal dengan 5 alternatif jawaban (a, b, c, d, e). Pengumpulan data dilakukan melalui Google Form (g-form), yang memungkinkan pengumpulan jawaban siswa secara otomatis dan efisien.

Setelah data terkumpul, data dianalisis menggunakan program ANATES untuk mengevaluasi kualitas butir soal berdasarkan kriteria validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Hasil analisis ini kemudian diinterpretasikan untuk memberikan gambaran

tentang kualitas butir soal serta rekomendasi perbaikan yang diperlukan. Dengan demikian, penelitian ini diharapkan dapat berkontribusi pada peningkatan kualitas pendidikan dan proses pembelajaran di sekolah.

#### 1. Validitas

Validitas diartikan sebagai interpretasi skor dari instrumen yang digunakan untuk mengukur atribut yang hendak diukur (Azwar, 2015). Suatu instrumen dianggap memiliki tingkat validitas yang tinggi jika hasil yang diperoleh sesuai dengan tujuan pengukurannya. Untuk validitas yang terkait dengan Pengembangan Teknologi di Dunia Kerja dan Industri, digunakan validitas isi, yang dapat dianalisis secara rasional melalui g-form sebelum soal diberikan kepada 50 partisipan. Selain itu, validitas isi menjamin bahwa soal yang disusun sudah mencerminkan sekumpulan item yang dapat mengeksplorasi area ukur atau konsep yang menjadi fokus pengukuran.

#### 2. Realibilitas

Reliabilitas adalah karakteristik yang terdapat pada nilai yang dihasilkan oleh suatu alat ukur, atau bisa juga diartikan sebagai pertanyaan yang diberikan kepada peserta. Reliabilitas berkaitan dengan aspek kepercayaan dan konsistensi dari hasil tes. Instrumen yang baik adalah yang dapat secara konsisten memberikan data yang mencerminkan kenyataan. Tingkat reliabilitas sebuah instrumen dapat bervariasi dari sangat tinggi hingga sangat rendah (Djali, 2021). Untuk menganalisis reliabilitas soal Pengembangan Teknologi di Industri dan Dunia Kerja, digunakan anates, dan hasilnya dibandingkan dengan kriteria tingkat reliabilitas soal. Berikut ini adalah kriteria tingkat reliabilitas (Hairun, 2020):

Reliabilitas Sangat Tinggi : 0.80 - 1.00

Reliabilitas Tinggi : 0.61 - 0.80

Reliabilitas Sedang : 0.41 - 0.60

Reliabilitas Rendah : 0.21 - 0.40

Reliabilitas Sangat Rendah : 0.00 - 0.20

#### 3. Daya Pembeda Soal

Setelah melakukan pengujian terhadap validitas dan reabilitas, langkah berikutnya adalah menganalisis daya pembeda. Menghitung hasil daya beda soal Pengembangan Teknologi di Industri dan Dunia Kerja menunjukkan seberapa efektif soal-soal tersebut dapat membedakan antara siswa yang telah menguasai materi dengan kriteria tertentu. Apabila soal memiliki nilai daya beda  $<0,19$  maka dinilai jelek,  $0,20-0,29$  berarti cukup,  $0,30-0,39$  dinyatakan baik dan  $>0,40$  berarti baik sekali (Sukmela, 2018).

#### 4. Tingkat Kesukaran

Cara menghitung tingkat kesukaran soal Pengembangan Teknologi di Industri dan Dunia Kerja menggunakan software Anates. Penggunaan perangkat lunak anates diperoleh karena bisa diunduh tanpa biaya dan lebih praktis saat mengaksesnya. Beberapa manfaat dari anates adalah ia menggunakan bahasa Indonesia. Selain itu, software ini dapat dipakai langsung untuk menganalisis soal pilihan ganda tanpa kesulitan dalam membuat rumus untuk perhitungan atau menghitung secara manual dalam sebuah program (Ariany & Al-Ghifari, 2018). Soal yang memiliki indeks kesulitan 0,00 dianggap sebagai soal yang sangat sulit, sementara indeks 1,0 menunjukkan soal tersebut terlalu mudah. Indeks kesulitan dilambangkan dengan P yang berarti "proporsi", sehingga soal dengan  $P=0,70$  dianggap lebih mudah dibanding  $P=0,20$ . Sebaliknya, soal dengan  $P=0,30$  lebih sulit daripada soal dengan  $P=0,80$ . Meskipun angka indeks ini lebih cocok sebagai indeks kemudahan atau indeks fasilitas, telah disetujui bahwa meskipun indeks yang lebih tinggi menunjukkan lebih mudah, tetap disebut sebagai indeks kesukaran.

#### 5. Efektivitas Pengecoh

Cara menghitung efektivitas pengecoh soal Pengembangan Teknologi di Industri dan Dunia

Kerja adalah dengan menggunakan software Anates. Dalam menyimpulkan seberapa efektif Pengecoh pada setiap butir soal, peneliti menerapkan kriteria yang berasal dari Skala Likert seperti berikut (Sugiyono, 2012:134).

3 Baik

2 Cukup baik

1 Kurang baik

0 Tidak baik

Keterangan mengenai kriteria penilaian efektivitas pengecoh adalah:

1. Jika keempat jawaban pengecoh berfungsi, maka soal dianggap memiliki efektivitas pengecoh yang sangat baik.
2. Jika tiga jawaban pengecoh berfungsi, maka soal tersebut dinilai memiliki efektivitas pengecoh yang baik.
3. Jika hanya dua jawaban pengecoh yang berfungsi, maka soal dikatakan memiliki efektivitas pengecoh yang cukup baik.
4. Jika adanya satu jawaban pengecoh yang berfungsi, maka soal disebut memiliki efektivitas pengecoh yang kurang baik.
5. Jika tidak ada jawaban pengecoh yang berfungsi sama sekali, maka soal dianggap tidak baik/jelek.

## HASIL DAN PEMBAHASAN

Hasil analisis ini bertujuan untuk mengevaluasi mutu soal pada aspek Pengembangan Teknologi di Industri dan Dunia Kerja untuk kelas X, dengan mengamati faktor-faktor seperti validitas, reliabilitas, tingkat kesulitan, daya pembedaan, dan efektivitas pilihan jawaban yang menyesatkan. Instrumen yang digunakan adalah seperangkat pertanyaan yang disajikan dalam bentuk *g-form*, beserta kunci jawaban dari 10 soal pilihan ganda yang memiliki 5 alternatif jawaban (a, b, c, d, e), yang telah diuji kepada 50 siswa SMKN 2 Mataram dari kelas X Akuntansi. Data diperoleh dengan menggunakan metode pengisian *g-form* yang telah disediakan, sehingga jawaban dari para siswa telah terekam secara otomatis.

Selanjutnya, data tersebut dianalisis dengan pendekatan kuantitatif. Untuk soal pilihan ganda, digunakan perangkat lunak ANATES untuk menilai kualitas butir soal berdasarkan aspek validitas, reliabilitas, tingkat kesulitan, kemampuan membedakan, serta efektivitas pengalih perhatian. Berikut adalah hasil analisis untuk soal pilihan ganda:

### 1. Validitas

Suatu nilai koefisien validitas dianggap baik jika nilainya berada di antara 0,30 hingga 0,50 (Saifuddin Azwar, 2008: 158). Validitas atau kesahihan dari suatu tes dapat dijelaskan sebagai ketepatan dan akurasi tes dalam melaksanakan perannya sebagai alat ukur. Koefisien yang lebih tinggi menunjukkan bahwa tes tersebut lebih akurat. Sebuah tes dikatakan memiliki validitas yang tinggi jika tes tersebut efektif dalam melaksanakan fungsi ukurnya, atau memberikan hasil yang selaras dengan tujuan dari pelaksanaannya. Di sisi lain, sebuah tes yang menghasilkan informasi yang tidak berkaitan dengan maksud pengukurannya dinilai memiliki validitas yang rendah (Azwar, 1987:146). Validitas dihitung menggunakan rumus korelasi point biserial. Total siswa di SMKN 2 Mataram kelas X Akuntansi mencapai 50 siswa, sehingga diketahui  $n=50$ , dan nilai  $r$  tabel tercatat pada angka 0,215. Dari analisis terhadap 10 soal pilihan ganda, ditemukan bahwa 5 soal (50%) dinyatakan valid dan 5 soal (50%) tidak valid. Berikut adalah distribusi 10 soal tersebut menurut indeks validitasnya:

No	Indeks Validitas	Butir Soal	Jumlah	Presentase
1	$\geq 0,215$ (Soal Dinyatakan Valid)	1, 2, 3, 4, 6	5	50%
2	$< 0,215$ (Soal Dinyatakan Tidak Valid)	5, 7, 8, 9, 10	5	50%

	Tidak Valid)			
--	--------------	--	--	--

Dari tabel 1 dapat dilihat bahwa soal yang valid ada pada soal nomor 1, 2, 3, 4, dan 6. Sedangkan soal yang tidak valid ada pada soal nomor 5, 7, 8, 9, dan 10. Dengan demikian terdapat 5 soal pilihan ganda dapat dikatakan valid dan 5 soal pilihan ganda yang dikatakan tidak valid.

**2. Realibilitas**

Reliabilitas berasal dari kata reliability yaitu terdiri dari kata rely dan ability. Pengukuran yang memiliki reliabilitas tinggi disebut sebagai pengukuran yang reliabel (Saifuddin Azwar, 2008: 4). Reliabilitas suatu tes adalah konsistensi dari suatu tes dalam mengukur apa yang seharusnya diukur sehingga pengukuran itu memberikan informasi yang dapat dipercaya. Pengertian itu disimpulkan dari pendapat dua ahli pengukuran berikut ini. Arikunto (2013) menyatakan bahwa reliabilitas dalam pengukuran berhubungan dengan masalah kepercayaan. Menurut Suryabrata (2004), reliabilitas menunjukkan sejauh mana hasil pengukuran dengan alat tersebut dapat dipercaya.

```

RELIABILITAS TES
=====
Rata2= 4,98
Simpang Baku= 2,03
KorelasiXY= 0,42
Reliabilitas Tes= 0,59
Nama berkas: C:\USERS\THINKPAD\ONEDRIVE\DOCUMENTS\BUTIR SOAL 1-10.ANA
    
```

Nilai reliabilitas yang diperoleh dari hasil analisis butir soal adalah sebesar 0,59 yang termasuk pada kategori cukup. Berdasarkan hasil tersebut, butir soal pengembangan teknologi di industri dan dunia kerja menunjukkan tingkat reabilitas atau keandalan yang cukup baik dalam menilai kemampuan siswa. Butir soal tersebut layak digunakan sebagai alat penilaian untuk mengukur kemampuan siswa. Tingkat keandalan dipengaruhi oleh beberapa elemen seperti pelaksanaan tes, jumlah peserta, dan tingkat kesulitan butir soal (Rahmasari dan Ismiyati, 2016). Pelaksanaan tes, yang mencakup petunjuk pengerjaan, kesiapan siswa untuk menjawab soal, dan lingkungan tempat pelaksanaan, dapat berpengaruh terhadap tingkat keandalan suatu item soal. Semakin banyak peserta tes akan menyebabkan semakin banyak variasi jawaban dan hal tersebut akan mempengaruhi nilai reliabilitas dari butir soal. Selanjutnya, semakin sukar tes, maka akan mengakibatkan semakin rendahnya nilai reliabilitas karena soal yang sukar akan mengakibatkan peserta didik menebak jawaban karena putus asa ketika sedang mengerjakan soal sukar (Anita et al., 2018).

**3. Tingkat Kesukaran**

Perhitungan tingkat kesulitan soal ditentukan oleh seberapa sulit suatu soal ketika dihadapi oleh siswa, dan soal dinyatakan sulit jika hanya sedikit siswa yang mampu menjawabnya. Tingkat kesulitan yang didapat dari hasil pengujian soal dapat dilihat pada tabel yang ada di bawah ini:

No	Kategori Soal	Jumlah	Nomor Soal
1	Sangat Sukar	0	0
2	Sukar	0	0
3	Sedang	9	1, 2, 3, 4, 5, 6, 8, 9, 10
4	Mudah	1	7

Berdasarkan hasil analisis tingkat kesukaran maka diketahui bahwa dari 10 butir soal Pengembangan Teknologi di Industri dan Dunia Kerja, tidak ada soal yang masuk dalam kategori sangat sukar. Jumlah soal dalam kategori sulit juga tidak ada. Terdapat sembilan soal yang termasuk dalam kategori sedang, yakni pada nomor 1, 2, 3, 4, 5, 6, 8, 9, dan 10. Butir soal dengan kategori mudah berjumlah 1 butir soal nomer 7. Melihat analisis tersebut, soal-soal yang dianggap sedang adalah yang tidak terlalu mudah maupun terlalu sulit, sehingga dapat memotivasi peserta didik untuk mencoba menyelesaikannya. Sebaliknya, jika soal terlalu sulit, hal ini dapat membuat peserta didik merasa putus asa karena merasa tidak mampu untuk menyelesaikannya. Hal ini sesuai dengan

pendapat Fatimah & Alfath (2019) yang menyatakan bahwa soal yang baik merupakan soal yang tidak terlalu mudah dan tidak terlalu sukar.

#### 4. Daya Pembeda

Daya pembeda soal mengukur seberapa efektif sebuah soal dalam membedakan antara siswa yang telah menguasai kompetensi dan mereka yang belum berdasarkan kriteria yang ditetapkan. Hasil dari pengujian soal tersebut mengenai daya pembeda dapat ditemukan dalam tabel di bawah ini:

##### DAYA PEMBEDA =====

Jumlah Subyek= 50  
Klp atas/bawah(n)= 14  
Butir Soal= 10  
Nama berkas: C:\USERS\THINKPAD\ONEDRIVE\DOCUMENTS\BUTIR SOAL 1-10.ANA

No Butir Baru	No Butir Asli	Kel. Atas	Kel. Bawah	Beda	Indeks DP (%)
1	1	10	1	9	64,29
2	2	13	1	12	85,71
3	3	14	3	11	78,57
4	4	11	1	10	71,43
5	5	13	7	6	42,86
6	6	13	5	8	57,14
7	7	12	6	6	42,86
8	8	8	5	3	21,43
9	9	5	3	2	14,29
10	10	5	4	1	7,14

Nilai daya beda yang diperoleh pada soal Pengembangan Teknologi di Industri dan Dunia Kerja yang dianalisis menggunakan program Anates menunjukkan bahwa terdapat 3 soal yang termasuk dalam kategori sangat baik, yaitu soal nomor 2, 3, dan 4. Sementara itu, 4 soal lainnya berada dalam kategori baik, yang terdiri dari soal nomor 1, 5, 6, dan 7. Terdapat 1 soal yang masuk dalam kategori cukup baik, yaitu soal nomor 8, dan 2 soal lainnya berada dalam kategori tidak baik (jelek), yaitu soal nomor 9 dan 10. Berdasarkan hasil tersebut dapat disimpulkan bahwa daya beda soal Pengembangan Teknologi di Industri dan Dunia Kerja materi dapat dikategorikan sangat baik, namun masih ada butir soal yang perlu diperbaiki.

#### 5. Efektivitas Pengecoh

Soal tes bentuk pilihan ganda terdiri atas soal yang mencerminkan suatu masalah dan dilengkapi dengan beberapa pilihan yang masing-masing menawarkan kemungkinan jawaban untuk pertanyaan tersebut. Pilihan jawaban tersebut mencakup satu jawaban yang benar (kunci jawaban) dan beberapa jawaban yang tidak tepat, yang dikenal sebagai pengecoh. Jumlah pilihan jawaban biasanya berkisar antara tiga hingga lima. Tujuan dari pengecoh adalah untuk mengalihkan perhatian siswa yang tidak yakin mengenai jawaban yang benar (Bhakti, 2015).

Hasil analisis butir soal ulangan Pengembangan Teknologi di Industri dan Dunia Kerja dengan menggunakan program Anates pada aspek efektivitas pengecoh terdiri atas 5 kategori, yaitu sangat baik, baik, kurang baik, buruk, dan sangat buruk. Berikut merupakan hasil analisis butir soal pengecoh pada gambar sebagai berikut:

## KUALITAS PENGECOH

=====

Jumlah Subyek= 50

Butir Soal= 10

Nama berkas: C:\USERS\THINKPAD\ONEDRIVE\DOCUMENTS\BUTIR SOAL 1-10.ANA

No Butir Baru	No Butir Asli	a	b	c	d	e	*
1	1	31---	0--	18**	0--	1--	0
2	2	4++	29**	3+	3+	11---	0
3	3	5++	24**	4+	6++	11-	0
4	4	17**	9++	17---	3-	4-	0
5	5	4++	1--	33**	7-	5++	0
6	6	27**	6++	5++	6++	6++	0
7	7	2+	38**	4+	3++	3++	0
8	8	7++	25**	7++	5++	6++	0
9	9	4-	21---	4-	17**	4-	0
10	10	11-	8++	9++	21**	1--	0

## Keterangan:

\*\* : Kunci Jawaban

++ : Sangat Baik

+ : Baik

- : Kurang Baik

-- : Buruk

--- : Sangat Buruk

Gambar tersebut menunjukkan hasil analisis butir soal Pengembangan Teknologi di Industri dan Dunia Kerja menggunakan program Anates pada aspek efektivitas pengecoh. Berdasarkan hasil analisis diperoleh bahwa terdapat 2 butir soal dengan pengecoh yang berkategori sangat baik, yaitu pada nomor soal 6 dan 8. Terdapat 2 soal lainnya yang termasuk dalam kategori baik untuk efektivitas pengecoh, yaitu soal nomor 7 dan 3. Satu soal dikategorikan kurang baik dalam hal efektivitas pengecoh, yaitu soal nomor 4. Dua soal lainnya dianggap buruk dalam efektivitas pengecoh, yaitu soal nomor 5 dan 10. Tiga soal dikategorikan sangat buruk dalam efektivitas pengecoh, yaitu soal nomor 1, 2, dan 9. Hasil penelitian sesuai dengan penelitian yang telah dilakukan oleh Toksöz & Ertunç (2017) yang menyatakan bahwa setengah dari total butir soal memiliki efektivitas pengecoh dengan kategori kurang baik dan sangat tidak baik karena pengecoh tersebut kurang memiliki daya tarik bagi peserta didik. Pengecoh dikatakan sangat baik jika dipilih secara merata oleh para responden. Membuat jawaban pengecoh untuk soal evaluasi bukanlah hal yang mudah, karena hal ini akan mempengaruhi daya pembeda dan tingkat kesulitan dari butir soal itu sendiri. Dengan kata lain, apabila kualitas pengecoh buruk, maka indikator kemampuan membedakan akan rendah dan tingkat kesulitan soal juga cenderung rendah. Sebaliknya, jika kualitas pengecoh dari suatu butir soal baik, maka indikator daya pembeda akan baik dan tingkat kesulitan soal akan meningkat.

Hasil analisis yang diperoleh bahwasannya soal Pengembangan Teknologi di Industri dan Dunia Kerja yang ditujukan kepada siswa kelas X SMKN 2 Mataram untuk mata pelajaran spreadsheet memiliki standar yang memuaskan. Hal ini terlihat dari berbagai jenis soal yang diberikan kepada siswa yang menunjukkan analisis yang baik. Validitas daari setiap jenis soal tersebut menunjukkan nilai yang bagus. Daya beda setiap butir soal memiliki hasil yang beragam. Tingkat kesulitan soal cenderung seimbang. Efektivitas pengecoh soal memiliki nilai yang kurang baik hingga ada yang baik. Oleh karena itu, dapat disimpulkan bahwasanya soal Pengembangan Teknologi di Industri dan Dunia Kerja yang ditujukan kepada siswa kelas X SMKN 2 Mataram memiliki kualitas yang baik.

**KESIMPULAN**

Berdasarkan hasil penelitian, dapat disimpulkan bahwa kualitas butir soal berdasarkan tingkat kesukaran menunjukkan sebanyak 9 (90%) butir soal berada dalam kategori sedang, sementara hanya 1 (10%) butir soal termasuk dalam kategori mudah, tidak ada butir soal yang dikategorikan sebagai sukar. Dari segi daya pembeda, terdapat 4 (40%) butir soal memiliki daya pembeda baik, 1(10%) butir soal cukup baik, dan 50% butir soal tidak baik, menunjukkan kemampuan soal dalam membedakan antara siswa yang menguasai materi dan yang tidak. Kualitas butir soal dari efektivitas

pengecoh menunjukkan 2(20%) butir soal memiliki efektivitas pengecoh sangat baik, 2 (20%) butir soal baik, 1 (10%) butir soal kurang baik, dan 5 (50%) butir soal buruk, mengindikasikan banyak pengecoh yang tidak efektif. Validitas butir soal menunjukkan 8 (80%) soal valid dan 2 (20%) soal tidak valid, dengan reliabilitas 0,59 yang termasuk cukup baik namun masih di bawah standar ideal 0,70, menunjukkan tes memiliki reliabilitas rendah.

Secara keseluruhan, hasil analisis ini menunjukkan bahwa meskipun sebagian besar butir soal memiliki kualitas yang baik dalam hal tingkat kesukaran dan validitas, masih terdapat beberapa aspek yang perlu diperbaiki, terutama dalam daya pembeda dan efektivitas pengecoh. Oleh karena itu, revisi terhadap butir soal yang kurang efektif sangat diperlukan untuk meningkatkan kualitas evaluasi dan hasil belajar siswa.

#### DAFTAR PUSTAKA

- Akhmadi, M. N. (2021). Analisis Butir Soal Evaluasi Tema 1 Kelas 4 SDN Plumbungan Menggunakan Program Anates. *Ed-Humanistics*, 6(1), 799-806.
- Aliman, M., Budijanto, Sumarmi, Astina, I. K., Putri, R. E., & Arif, M. (2019). The Effect of Earthcomm Learning Model and Spatial Thinking Ability on Geography Learning Outcomes. *Journal of Baltic Science Education*, 18(3), 323-334
- Aliman, M., Mutia, T., & Yustesia, A. (2018). Integritas Kebangsaan dalam Tes Berpikir Spasial. *Prosiding Seminar Nasional Pendidikan Geografi FKIP UMP 2018*, 82-89. Purwokerto: UM Purwokerto Press
- Anas Sudijono. (2011). *Pengantar Evaluasi Pendidikan*, Jakarta: Raja Grafindo Persada.
- Ani Interdiana Candra Sari dan Mirna Herawati. (2014). Aplikasi Anates Versi 4 Dalam Menganalisis Butir Soal, *Faktor Jurnal Ilmiah Kependidikan*, Volume 1, no.2.
- Anica., Idi, A., & Ismail, F. 2023. Evaluasi Hasil Belajar Siswa Pada Sistem Pembelajaran Daring Selama Masa Peandemi. *TAFANI: Jurnal Pengabdian Kepada Masyarakat*, 1(1), 21-33.
- Arifin, Z. 2019. *Evaluasi Pembelajaran*. PT. Remaja Rosdakarya: Bandung
- Bagiyono. 2017. Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal Ujian Pelatihan Radiografi Tingkat 1. *Widyanuklida*, 16(1), 1-12.
- Elementary School of Social Studies Subjects). *Journal of Elementary Education*, 4(1), 1-6. Magdalena, Ina. (20220). Menjadi evaluator pembelajaran yang baik dan benar. CV. Jejak. Sukabumi
- Fiska, J. M., Hidayati, Y., Qomaria, N., & Hadi, W. P. (2021). Analisis butir soal ulangan harian IPA menggunakan software Anates pada pendekatan teori tes klasik. *Natural Science Education Research*, 4(1), 65-76.
- Gierl, Mark J.; Lai, Hollis; Pugh, Debra; Touchie, Claire; Boulais, André; De Champlain, André (2016). Evaluating the Psychometric Characteristics of Generated Multiple-Choice Test Items. *Applied Measurement in Education*, (), 08957347.2016.1171768-. doi:10.1080/08957347.2016.1171768.
- Gimo, Farida Nugrahani. (2019). Analisis Butir Soal Ulangan Akhir Semester (UAS) Mata Pelajaran Bahasa Indonesia Kelas XI Sekolah Menengah Kejuruan Negeri 1 Wonogiri Tahun Pelajaran 2015/2016. *Stilistika*, Vol. 5, No. 1 2019 : 35 - 46.
- Gumay, A., Nurika, Y., Herningsih, S. W., & Magdalena, S. (2022). Validitas dan Reliabilitas Ujian Keahlian Pelaut Bahasa Inggris Maritim. *JOEAI (Journal of Education and Instruction)*, 5(2), 382-394.
- Halik, Andi Surahma; Sitti Mania; Fitriani Nur. (2019). Analisis Butir Soal Ujian Akhir Sekolah (UAS) Mata Pelajaran Matematika pada Tahun Ajaran 2015/2016 SMP Negeri 36 Makassar.
- Handani, H. A., & Prayitno, H. J. (2015). Validitas dan Reliabilitas Soal Tengah Semester Genap Kaitannya dengan Ketercapaian Tujuan Pembelajaran Bahasa Indonesia Kelas VIII A SMP Negeri 2 Banyudono Tahun Ajaran 2013/2014. *University Research Colloquium*, 2(1), 193-206.

- Ismail, Ilyas. (2020). Asesmen Dan Evaluasi Pembelajaran. Makassar: Cendekia Publisher
- Ismail, Muhammad Ilyas. (2020). Evaluasi pembelajaran: Konsep dasar, prinsip, teknik, dan prosedur. PT. Rajagrafindo Persada. Depok
- Kurniawan, T. (2015). Analisis Butir soal ulangan akhir semester gasal mata pelajaran IPS Sekolah Dasar (Analysis of Odd Semester Final Test Items in
- Mawardi, M.S., Fuady A., & Sunismi 2023. Analisis Butir Soal Pilihan Ganda Menggunakan Anates pada Penilaian Tengah Semester Kelas VII D SMP Negeri 1 Ngajum Kabupaten Malang. Wahana: Tridarma Perguruan Tinggi, 75(1), 31-41
- Muchlizani, N. A., Mania, S., & Rasyid, M. N. A. 2023. Analisis Kualitas Butir Soal Ujian Akhir Semester Mata Pelajaran Akidah Akhlak Kelas V Mi Radhiatul Adawiyah Makassar. Jurnal Inspiratif Pendidikan, 12(1), 224–240.
- Pamilu, Ahmad Fikri Aji. 2014. "Analisis Butir Soal pad Ulangan Akhir Semester Gasal Mata Pelajaran Aqidah Akhlak Kelas X MAN Yogyakarta III Tahun Pelajaran 2013/2014". Skripsi. Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
- Purwanto. (2011) Prinsip-Prinsip dan Teknik Evaluasi Pengajaran. Bandung: Remaja Rosdakarya
- Rahmasari, D., & Ismiyati. (2016). Analisis Butir Soal Mata Pelajaran Pengantar Administrasi Perkantoran. Economic Education Analysis Journal, 5(1), 317–330.
- Riduwan, Dr. 2012. Skala Pengukuran Variabel-Variabel Penelitian. Jakarta:Alfabeta
- Silalahi, Tauada. (2020). Evaluasi pembelajaran. Yayasan Kita Menulis.
- Solichin, Mujiyanto. (2017). Analisis daya beda soal taraf kesukaran, butir tes, validitas butir tes, interpretasi hasil tes valliditas ramalan dalam evaluasi pendidikan. Jurnal Manajemen Dan Pendidikan Islam, 2(2), 192–213.
- Sugiyono. 2010. Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif, dan R&D. Bandung: CV Alfabeta.
- Suharsimi, Arikunto. 2013. Dasar-Dasar Evaluasi Pendidikan Edisi 2. Jakarta: PT Bumi Aksara.
- Supriyadi. (2021). Evaluasi pendidikan. PT. Nasya Expanding Management.
- Unpad. (2021). Dosen bisa terapkan metode asesmen untuk pembelajaran daring.(2021, 1 Februari). Diakses tanggal 16 Mei 2021 ddari artikel
- Widyoko, Eko Putra. 2014. Penilaian Hasil Pembelajaran di Sekolah. Yogyakarta: Pustaka Pelajar.
- Yektiana, N., & Nursikin, M. 2023. Konsep Dasar Pengukuran, Penilaian, Dan Evaluasi Hasil Belajar. J-CEKI: Jurnal Cendekia Ilmiah, 2(2), 263–266.