

## Analisis Butir Soal Ujian Akhir Semester Bahasa Arab Kelas X Menggunakan Software Anates

Fadhli<sup>1</sup>, Hikmah<sup>2</sup>, Masrun<sup>3</sup>

<sup>1,3</sup> Pascasarjana UIN Sultan Syarif Kasim Riau

<sup>2</sup> STAI Nurul Hidayah Riau

E-mail : [fadhlysyn91@gmail.com](mailto:fadhlysyn91@gmail.com)<sup>1</sup>, [hikmah@stai-nh.ac.id](mailto:hikmah@stai-nh.ac.id)<sup>2</sup>, [masrun@uin-suska.ac.id](mailto:masrun@uin-suska.ac.id)<sup>3</sup>

Volume 2 Nomor 3 Juni 2024: DOI: <https://doi.org/10.55102/hidmah.v2i3> Article History Submission: 10-5-2024 Revised: 10-6-2024 Accepted: 21-6-2024 Published: 26-6-2024

**Abstract:** *This study aims to see whether the question items used in the final exam for the even semester of the 2023-2024 school year which will be held on June 3-10, 2024 at MAN 1 Meranti Islands have good quality as a benchmark for achieving student competencies to move up to the next grade or whether they are not worthy of being quality questions. This study is qualitative descriptive research that describes the results of the analysis of question items using the Anates Ver 4.0.2 application. The questions in the form of multiple choice are 40 questions with 5 choice answers (a, b, c, d, e) and 32 students who are students of class X4. After the answer result data is input into the Anates application, the Validation Test of question items in the category is very good or the majority is valid. The validity of the question item has a reliable value of 0.85 (very high). In the difficulty level, it was found that out of a total of 40 questions, 60% were in the moderate category, 3 questions or 7.5% were in the easy category, and 13 questions, or a percentage of 32.5% were in the difficult category. The very easy level in this question was not found, aka 0%. The differentiation in question items was categorized as "Good" with the number of question items as many as 17 questions out of a total of 40 questions with a percentage of 42.5%. 12 questions (30%) with the category "Quite Good", 7 questions (17.5%) with the category "Poor" should not be used, and 4 questions (10%) with the category "Very Good". Meanwhile, an analysis of the effectiveness of deception found that there was an effectiveness of deception that functioned very well, namely in questions number 4, 6, 7, 13, 14, 33, 37, and 38.*

**Keywords:** *Question Item Analysis, UAS, Arabic Language, Anates*



## PENDAHULUAN

Analisis butir soal adalah kegiatan menganalisis untuk mengidentifikasi soal-soal pada suatu tes sehingga dapat dipergunakan dan dapat meningkatkan kualitas soal (Ida dan Musyarofah 2021). Sejalan dengan itu, Kegiatan menganalisis butir soal merupakan salah satu kewajiban bagi setiap pendidik karena setiap pendidik pada akhirnya harus dapat memberikan informasi kepada lembaga sekolah ataupun kepada peserta didik tentang kondisi penguasaan dan kemampuan yang telah dicapai peserta didik terhadap materi dan keterampilan-keterampilan dalam mata pelajaran yang telah diberikan (Wiguna 2021).

Analisis terhadap soal tes sangatlah penting dilakukan guna memperbaiki kualitas soal yang akan diujikan setiap tahunnya. Analisis butir soal yang dilakukan akan dapat meningkatkan kualitas soal melalui unsur validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan kualitas pengecoh. Lebih lanjut, analisis butir soal sangat membantu dalam menilai soal-soal yang berkualitas sehingga layak sebagai tolak ukur keberhasilan pembelajaran siswa (Salmina dan Adyansyah 2017). Dengan adanya analisis butir soal, dapat diidentifikasi soal yang baik dan soal yang kurang baik serta soal yang dapat masuk ke dalam bank soal dan direvisi atau dibuang.

Salah satu bentuk penilaian dalam evaluasi pembelajaran adalah ujian akhir semester. Sebagaimana dikutip Juandi dalam peraturan Menteri Pendidikan Nasional No. 20 tahun 2007 mengenai standar penilaian, Penilaian Akhir Semester didefinisikan sebagai kegiatan penilaian yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik di akhir semester meliputi seluruh indikator yang merepresentasikan semua kompetensi dasar pada semester tersebut (Juandi 2019).

Hasil dari ujian semester ini digunakan untuk berbagai kepentingan di antaranya adalah sebagai salah satu pertimbangan dalam pembinaan dan pemberian bantuan kepada satuan pendidikan dalam upaya meningkatkan mutu pendidikan. Selain itu, hasil pelaksanaan ujian semester ini digunakan sebagai salah satu penentu kelulusan peserta didik untuk jenjang semester dan kenaikan kelas berikutnya. Hasil dari ujian akhir semester ini juga digunakan sebagai standar kelulusan siswa, sehingga tim atau guru mapel pembuat soal ujian akhir semester harus mampu menyusun butir-butir soal agar memiliki tingkat validitas yang tinggi, memiliki daya beda yang baik, serta efektif dalam mengecoh pada opsi jawaban. Hal ini tentu suatu hal yang melekat pada penyusun soal ujian semester untuk membuat soal yang baik dan berkualitas sehingga tidak merugikan siswa dalam mencapai target belajarnya (Amalia dan Widayati 2012).



Setiap semesternya Madrasah Aliyah Negeri 1 Kepulauan Meranti menyelenggarakan ujian akhir semester yaitu ujian akhir semester ganjil dan ujian akhir semester genap (ujian kenaikan kelas). Ujian dilakukan pada mata pelajaran yang dipelajari pada semester itu dan soal-soal yang diujikan berbentuk soal pilihan ganda. Untuk jumlah butir soal sebanyak 35 soal pilihan ganda bagi mata pelajaran jenis eksakta dan 40 soal pilihan ganda bagi mata pelajaran non-eksakta. Soal-soal yang diujikan dalam ujian akhir semester ini disusun oleh guru pengampu mata pelajaran.

Berdasarkan informasi yang diperoleh dari Bidang Kurikulum Madrasah Aliyah Negeri 1 Kepulauan Meranti, selama ini guru pembuat soal ujian akhir semester belum pernah melakukan analisis butir soal yang telah disusun. Hal ini disebabkan kurang maksimalnya kemampuan penyusun soal untuk melakukan analisis butir soal ujian semester. Baik atau tidaknya kualitas soal selama ini hanya diketahui berdasarkan hasil ujian siswa. Dengan adanya temuan fakta di lapangan sebagaimana disebut, soal ujian akhir semester yang disusun oleh guru pengampu mata pelajaran belum diketahui kehandalan dan keterpercayaannya. Maka dari itu, Untuk mengetahui butir soal yang berkualitas dan baik sebagai alat ukur hendaklah dilakukan suatu analisis butir soal sehingga dapat diketahui butir soal mana yang harus direvisi atau bahkan mungkin harus dihilangkan sama sekali.

Melihat adanya keadaan tersebut, maka kajian terhadap analisis butir soal ujian akhir semester di MAN 1 Kepulauan Meranti bagi mata pelajaran Bahasa Arab kelas X menarik untuk dilakukan. Penelitian ini bertujuan untuk melihat apakah butir-butir soal yang digunakan dalam ujian akhir semester genap tahun ajaran 2023-2024 yang dilaksanakan pada 3-10 Juni 2024 memiliki kualitas yang baik sebagai tolak ukur pencapaian kompetensi siswa untuk naik kelas selanjutnya ataukah belum layak menjadi soal yang berkualitas.

## **METODE PENELITIAN**

Penelitian ini adalah penelitian kualitatif yang dilakukan untuk mengetahui kualitas butir soal Bahasa Arab dalam Ujian Akhir Semester (UAS) semester Genap tahun ajaran 2023-2024 kelas X4 yang berjumlah 32 Siswa di Madrasah Aliyah Negeri 1 Kepulauan Meranti Riau. Tujuan penelitian ini adalah untuk mengetahui kriteria soal Ujian Akhir Semester Genap. Teknik pengumpulan data yang digunakan adalah dokumentasi hasil ujian siswa dalam mengerjakan soal. Data yang didokumentasikan dianalisis menggunakan aplikasi *Anates V4*. Soal dalam bentuk pilihan ganda dengan jumlah soal sebanyak 40 nomor soal dan pilihan jawaban ada 5 butir (a,b,c,d,e). Data diinput ke dalam software Anates sesuai pilihan jawaban



yang dijawab siswa. Kemudian diproses untuk semua analisis yang ingin diperoleh. Hasil akan muncul secara otomatis sesuai dengan kebutuhan. Hasil analisis dalam bentuk persentase kemudian dideskriptifkan sehingga dapat diketahui soal yang baik dan soal yang butuh perbaikan atau soal yang tidak perlu digunakan kembali. Hasil analisis soal kemudian dideskriptifkan dengan beberapa ketentuan sebagai berikut.

#### A. Validitas dan Reliabilitas Butir Soal

Validitas butir soal digunakan untuk mengetahui dukungan suatu butir soal terhadap skor total. Untuk menguji validitas setiap butir soal, skor-skor yang ada pada butir soal yang dimaksud dikorelasikan dengan skor total. Sebuah soal akan memiliki validitas yang tinggi jika skor soal tersebut memiliki dukungan yang besar terhadap skor total. Dukungan setiap butir soal dinyatakan dalam bentuk korelasi sehingga untuk mendapatkan validitas suatu butir soal digunakan rumus korelasi. Interpretasi besarnya koefisien korelasi dapat dilihat pada Tabel 1 berikut.

Tabel 1. Kategori Validitas Butir Soal

Koefisien	Kategori
$0.80 < r_{xy} \leq 1.00$	Sangat Tinggi
$0.60 < r_{xy} \leq 0.80$	Tinggi
$0.40 < r_{xy} \leq 0.60$	Cukup
$0.20 < r_{xy} \leq 0.40$	Rendah
$0.00 < r_{xy} \leq 0.20$	Sangat Rendah

Menurut Masri Singarimbun dalam Ancok, realibilitas adalah indeks yang menunjukkan sejauh mana suatu alat ukur dapat dipercaya atau dapat diandalkan. Bila suatu alat pengukur dipakai dua kali untuk mengukur gejala yang sama dan hasil pengukuran yang diperoleh relative konsisten, maka alat pengukur tersebut reliable. Dengan kata lain, realibitas menunjukkan konsistensi suatu alat pengukur di dalam pengukur gejala yang sama (Ancok 1989). Kategori dalam mengolah hasil analisis soal menggunakan software anates adalah sebagai berikut

Tabel 2. Kategori Reliabelitas Soal

Reliabilitas Soal	Nilai Koefisien
Sangat Rendah	0,00 sampai 1,19
Rendah	0,20 sampai 0,39
Cukup	0,40 sampai 0,59
Tinggi	0,60 sampai 0,79
Sangat Tinggi	0,80 sampai 1,00



## B. Daya Pembeda

Daya pembeda soal dapat ditentukan dengan melihat selisih antara banyaknya siswa tergolong kelompok unggul dengan banyaknya siswa tergolong kelompok asor pada masing-masing butir soal. Beberapa kategori klasifikasi daya pembeda sebagaimana dijelaskan dalam tabel berikut:

Tabel 3. Klasifikasi Daya Pembeda

Besarnya Daya Pembeda	Keterangan
0,00 - 0,20	Buruk, yaitu tidak adanya daya pembeda yang bernilai baik
0,21 - 0,40	Cukup, yaitu daya pembeda cukup
0,41 - 0,70	Baik, yaitu daya pembeda baik
0,71 - 1,00	Baik sekali, yaitu adanya daya pembeda yang baik sekali

Jika terdapat koefisien daya beda yang bernilai negative, maka soal dapat dikategorikan tidak baik. Maka sebaiknya tidak digunakan (Anggreini dan Darmawan 2016).

## C. Tingkat Kesukaran

Kriteria untuk tingkat kesukaran adalah jika hasil analisis diperoleh nilai rentang 0,71-1,00 termasuk dalam kategori mudah. Untuk kategori sedang dengan rentang 0,31-0,70 dan untuk kategori sulit adalah dengan kriteria 0-0,30 (Yoshita Cahyaningrum, Fuady, dan Sunismi 2023).

## D. Efektivitas Pengecoh

Pengecoh dikategorikan baik bila jumlah siswa memilih pengecoh sama atau mendekati ideal, yaitu 
$$\frac{P}{(N-B)/(n-1)} \times 100\%$$

Keterangan :

Ip = Indeks Pengecoh

P = Jumlah siswa yang memilih pengecoh

N = Jumlah siswa yang mengikuti tes

B = Jumlah siswa yang menjawab benar

n = Jumlah alternatif jawaban

1 = Bilangan tetap

Adapun kualitas pengecoh berdasarkan indeks pengecoh adalah sebagai berikut.

Indeks Pengecoh kategori Sangat Baik = 76% hingga 125%

Indeks Pengecoh kategori Baik = 51% hingga 75% atau 126% hingga 150%

Indeks Pengecoh kategori Kurang Baik = 26% hingga 50% atau 151% hingga 175%

Indeks Pengecoh kategori Buruk = 0% hingga 25% atau 176% hingga 200%

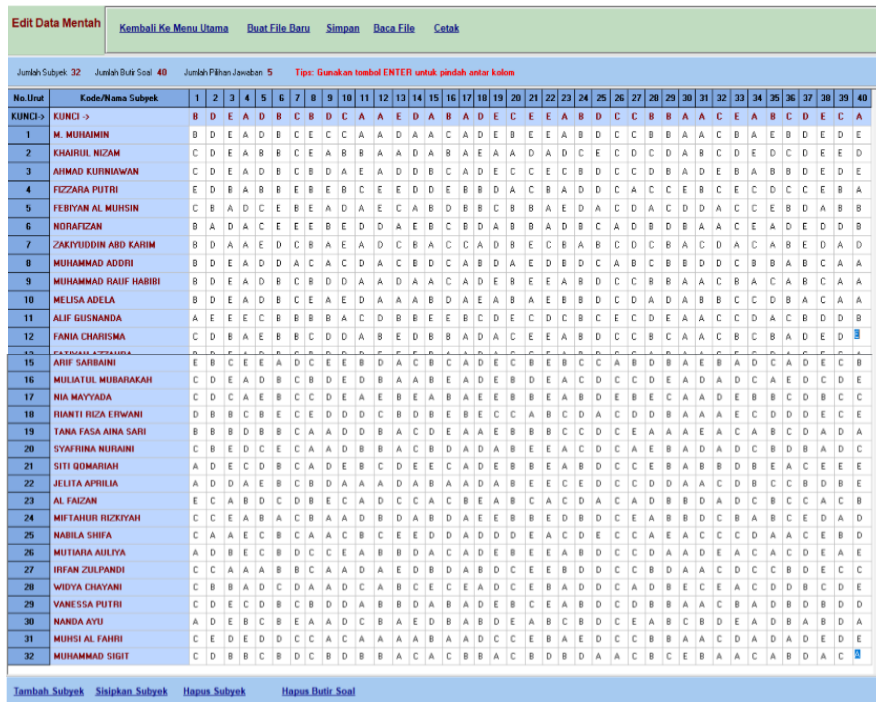
Indeks Pengecoh kategori Sangat Buruk = di atas 200%

Catatan : Bila peserta menjawab soal benar semua sehingga IP = 0 maka efektivitas pengecoh tidak ada fungsi dan soal jelek (Wiguna 2021).



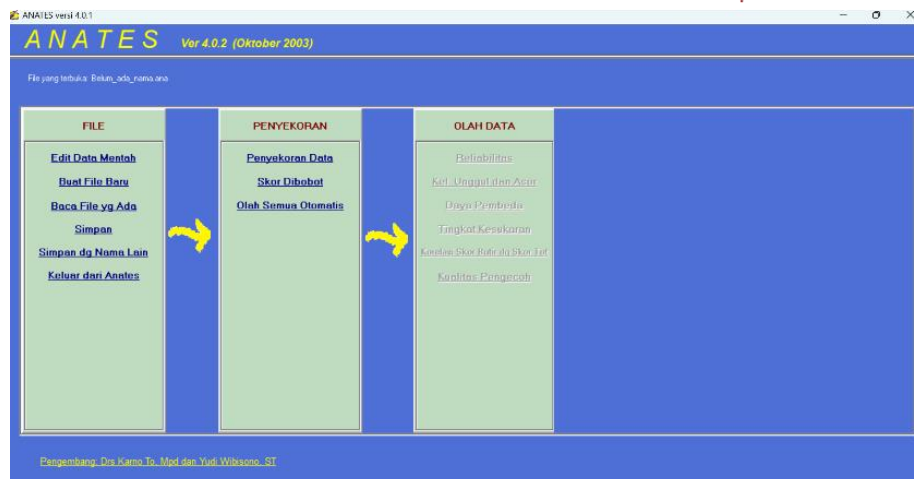
### HASIL DAN PEMBAHASAN

Penelitian yang dilakukan pada Ujian Akhir Semester genap bidang studi bahasa Arab kelas X tahun ajaran 2023-2024 memiliki kriteria soal yaitu pilihan ganda dengan opsi 5 (lima) jawaban. Jawaban siswa dikumpulkan dan diinput pada software Anates dan menghasilkan tampilan gambar sebagai berikut:



Gambar 1 Input Manual Hasil Ujian Siswa Kelas X4

Hasil jawaban siswa yang sudah dikumpulkan diawali dengan menginput satu per satu jawaban siswa. Jumlah subyek sebanyak 32 siswa dengan jumlah soal sebanyak 40 butir soal pilihan ganda. Kolom ke bawah adalah urutan nama siswa sesuai abjad pada absensi kelas, sedangkan kolom ke kanan menunjukkan nomor soal dan kunci jawaban di atasnya. Setelah data diinput semua, selanjutnya kembali pada menu utama maka akan muncul tampilan sesuai gambar berikut:



Gambar 2 Menu Utama Anates

Selanjutnya, setelah memasukkan data, pilih salah satu menu diantaranya penyekoran skor, skor dibobot, dan olah semua otomatis. Jika memilih semua olah data otomatis sebagaimana pada gambar dua, maka akan secara otomatis menampilkan semua hasil analisis yaitu, validitas dan reliabilitas soal, kelompok unggul dan asor, daya pembeda, hubungan skor bobot dengan skor total, kualitas pengecoh, dan rangkuman analisis butir soal.

## Validitas

Validitas adalah sebagai cerminan sejauh mana ketepatan dan kecermatan suatu instrumen tes berfungsi sebagai alat ukur hasil belajar. Kelayakan soal dapat diukur melalui validitas. Ayunita menyatakan bahwa validitas adalah kecermatan suatu instrumen dalam pengukuran (NN 2018). Sejalan dengan itu, Sugiyono menyatakan bahwa validitas memiliki tujuan untuk mengukur instrumen yang akan digunakan (Sugiyono 2013). Validitas butir soal yang dianalisis menggunakan aplikasi Anates sebagaimana dipaparkan berikut ini.

KORELASI SKOR BUTIR DG SKOR TOTAL  
-----

Jumlah Subyek= 32  
Butir Soal= 40  
Nama berkas: D:\DOCUMENTS\PASCA\S2 PENDIDIKAN B. ARAB\TUGAS\SEMESTER 2\PENGEMBANGAN EVALUASI BAHASA ARAB\X4 UJIAN SEP

No Butir	Korelasi	Signifikansi
1	0.164	-
2	0.562	Sangat Signifikan
3	0.387	Signifikan
4	0.380	Signifikan
5	0.471	Sangat Signifikan
6	0.560	Sangat Signifikan
7	0.506	Sangat Signifikan
8	0.282	-
9	0.619	Sangat Signifikan
10	0.061	-
11	0.380	Signifikan
12	0.358	Signifikan
13	0.228	-
14	0.253	-
15	0.284	-
16	0.267	-
17	0.618	Sangat Signifikan
18	0.499	Sangat Signifikan
19	0.422	Sangat Signifikan
20	0.096	-
21	0.259	-
22	0.641	Sangat Signifikan
23	0.589	Sangat Signifikan
24	0.509	Sangat Signifikan
25	0.666	Sangat Signifikan
26	0.335	Signifikan
27	0.537	Sangat Signifikan
28	0.401	Sangat Signifikan
29	0.228	-
30	0.545	Sangat Signifikan
31	0.456	Sangat Signifikan
32	0.522	Sangat Signifikan
33	0.087	-
34	0.398	Sangat Signifikan
35	0.067	-
36	-0.068	-
37	0.265	-
38	0.389	Signifikan
39	-0.033	-
40	-0.006	-

Gambar 3 Hasil Uji Validasi

Berdasarkan pada gambar di atas, ditemukan bahwa hasil olahan Anates dilihat dari signifikansi menyatakan sangat signifikan, signifikan, dan keterangan -. Valid atau tidak validnya soal dapat digunakan dengan teknik korelasi sebagai teknik analisisnya. Butir soal

yang valid bisa digunakan untuk tes hasil belajar, sedangkan butir soal yang tidak valid maka soal tidak bisa digunakan atau dibuang (Alpusari 2014).

Gambar di atas menunjukkan bahwa dari total 40 butir soal, terdapat 24 butir soal yang valid dan 16 soal lainnya tidak valid. Butir soal yang memiliki signifikansi yang tinggi adalah butir soal nomor 2, 5, 6, 7, 9, 17, 18, 19, 22, 23, 24, 25, 27, 28, 30, 31, 32, dan 34. Butir soal dengan kategori signifikan adalah 3, 4, 11, 12, 26, dan 38. Sedangkan butir soal yang tidak valid adalah soal nomor 1, 8, 10, 13, 14, 15, 16, 20, 21, 29, 35, 36, 37, 39, dan 40.

Salah satu butir soal yang valid adalah terdapat pada soal nomor 9 dengan korelasi yang diperoleh yaitu 0.619. Korelasi ini merupakan koefisien yang sangat tinggi. Bentuk soal sebagaimana berikut.

ليلا تتعلم في البيت بعد العشاء في الساعة 8.15.....

أ. الثامنة وخمسة عشرة د. الثامنة والرابع

ب. السابعة والرابع هـ. الثامنة والثالث

ج. الثامنة والنصف

Adapun soal yang tidak valid dan tidak layak untuk digunakan adalah seperti soal pada nomor 39 dengan korelasi yang diperoleh yaitu 0.033. korelasi ini merupakan koefisien yang tidak signifikan atau sangat rendah. Bentuk soal sebagaimana berikut.

ذهب الأب للعمل في الساعة صباحا ويستريح ساعة ويرجع في الخامسة مساء. كم ساعة يعمل الأب؟

أ. سبع ساعات د. عشر ساعات

ب. ثمان ساعات هـ. نصف اليوم

ج. تسع ساعات

## Reliabilitas

Hasil analisis pada soal Ujian Akhir Semester (UAS) bidang studi bahasa Arab kelas X4, diperoleh hasil koefisien korelasi xy sebesar 0,74. Lebih lanjut, dapat diketahui nilai reliabelnya berada di angka 0,85. Hal ini dapat dilihat pada gambar berikut.





RELIABILITAS TES  
\*\*\*\*\*

Rata2= 15.84  
Simpang Baku= 6.81  
KorelasiXY= 0.74  
Reliabilitas Tes= 0.85  
Nama berkas: D:\DOCUMENTS\PASCA\S2 PENDIDIKAN B. ARAB\TUGAS\SEMESTER 2\PENGEBAANGAN EVALUASI BAHASA ARAB\X4 UJIAN SEMESTER.ANA

No.Ucut	Kode>Nama	Subyek	Skor	Ganjil	Skor	Genap	Skor	Total
1	M. MUHAMMID		14	15				29
2	KHAIRUL WIZAH		11	5				16
3	AHMAD KURNIAWAN		14	11				25
4	FIZZARA PUTRI		9	4				13
5	FEBRYAN AL MU...		1	2				3
6	NORAFIZAH		4	2				6
7	ZAKIYUDDIN AB...		5	6				11
8	MUHAMMAD ADDRI		4	7				11
9	MUHAMMAD RAUF...		13	15				28
10	MELISA ADELA		8	6				14
11	ALIF GUSNANDA		6	3				9
12	FANIA CHARTISMA		14	11				25
13	FATIYAH AZZAHRA		12	14				26
14	FATIMAH ZAHRA...		7	3				10
15	ARIF SARBAINI		5	5				10
16	MULIATUL MUBA...		8	10				18
17	NIA MAYYADA		8	13				21
18	RIANTI RIZA E...		5	7				12
19	TANA FASA AIN...		5	7				12
20	SYAFRIHA NURA...		4	8				12
21	SITI QOMARIHA		8	10				18
22	JELITA APRILIA		11	8				19
23	AL FAIZAN		4	6				10
24	MIFTAHUR RIZK...		8	7				15
25	NABILA SHIFA		7	5				12
26	MUTIARA AULIYA		10	9				19
27	IRFAN ZULPANDI		11	8				19
28	WIDYA CHAYANI		4	5				9
29	VANESSA PUTRI		13	13				26
30	HANIDA AYU		6	5				11
31	MUHSI AL FAHRI		10	11				21
32	MUHAMMAD SIGIT		4	4				8

Gambar 4 Reliabilitas Tes pada Aplikasi Anates

Hasil analisis pada soal menunjukkan bahwa dari gambar di atas menunjukkan reliabilitas soal yang dikategorikan sangat tinggi berdasarkan pada kategori reliabel soal yang dipaparkan pada tabel 2. Hal ini menunjukkan bahwa keajegan soal sudah sangat layak untuk diberikan dan diujikan kepada siswa.

## Daya Pembeda

Anates selanjutnya menunjukkan adanya daya pembeda pada soal yang dianalisis. Adapun hasil analisis soal pada software Anates untuk mengetahui daya pembeda soal dapat dilihat pada gambar berikut.

DAYA PEMBEDA  
\*\*\*\*\*

Jumlah Subyek= 32  
Klp atas/bawah(n)= 9  
Butir Soal= 40  
Nama berkas: D:\DOCUMENTS\PASCA\S2 PENDIDIKAN B. ARAB\TUGAS\SEMESTER 2\PENGEBAANGAN EVALUASI BAHASA ARAB\X4 UJIAN SEMESTER.ANA

No Butir	Kel. Atas	Kel. Bawah	Beda	Indeks DP (%)
1	3	2	1	11.11
2	8	2	6	66.67
3	5	1	4	44.44
4	7	3	4	44.44
5	6	2	4	44.44
6	8	2	6	66.67
7	8	1	7	77.78
8	5	3	2	22.22
9	7	0	7	77.78
10	2	1	1	11.11
11	7	3	4	44.44
12	5	1	4	44.44
13	2	0	2	22.22
14	3	1	2	22.22
15	4	3	1	11.11
16	3	0	3	33.33
17	9	2	7	77.78
18	8	3	5	55.56
19	5	2	3	33.33
20	4	4	0	0.00
21	5	2	3	33.33
22	8	2	6	66.67
23	7	1	6	66.67
24	7	2	5	55.56
25	9	1	8	88.89
26	8	5	3	33.33
27	7	2	5	55.56
28	5	2	3	33.33
29	6	4	2	22.22
30	9	4	5	55.56
31	8	3	5	55.56
32	7	1	6	66.67
33	2	0	2	22.22
34	5	1	4	44.44
35	3	1	2	22.22
36	2	2	0	0.00
37	6	3	3	33.33
38	5	1	4	44.44
39	2	3	-1	-11.11
40	2	2	0	0.00

Gambar 5 Daya Pembeda



Berdasarkan gambar di atas, data yang muncul diolah kemudian didistribusikan berdasarkan tabel 3 untuk setiap soal. Beberapa klasifikasi yang dapat dimunculkan sesuai dengan anates yang keluar. Pendistribusian soal sebagaimana pada tabel berikut.

Tabel 4. Distribusi Klasifikasi Daya Pembeda

Nomor Soal	Keterangan	Persentase
1,10,15,20,36,39,40	Soal buruk, sebaiknya tidak digunakan	$\frac{7}{40} \times 100 = 17,5\%$
8,13,14,16,19,21,26,28,29,33,35,37	Cukup	$\frac{12}{40} \times 100 = 30\%$
2,3,4,5,6,11,12,18,22,23,24,27,30,31,32,34,38	Baik	$\frac{17}{40} \times 100 = 42,5\%$
7,9,17,25	Baik Sekali	$\frac{4}{40} \times 100 = 10\%$

Berdasarkan tabel di atas, dapat dilihat bahwa 7 dari 40 atau 17,5% soal berada pada kategori tidak baik atau tidak layak digunakan. Dapat dilihat juga bahwa 12 dari 40 soal atau 30% soal berada pada kategori cukup baik. Selanjutnya, 17 dari 40 soal atau 42,5% berada pada kategori soal yang baik, sedangkan soal pada kategori baik sekali terdapat 4 dari 40 soal atau 10%.

### Tingkat Kesukaran

Selanjutnya adalah analisis tingkat kesukaran soal. Hasil dari olahan data Anates sebagaimana dipaparkan dalam gambar berikut.

TINGKAT KESUKARAN

Jumlah Subyek= 32  
Butir Soal= 40  
Nama berkas: D:\DOCUMENTS\PASCA\S2 PENDIDIKAN B. ARAB\TUGAS\SEMESTER 2\PENGEMBANGAN EVALUASI BAHASA ARAB\X4 UJIAN SEMESTER.ANA

No Butir	Jml Betul	Tkt. Kesukaran(%)	Tafsiran
1	8	25.00	Sukac
2	18	56.25	Sedang
3	14	43.75	Sedang
4	17	53.13	Sedang
5	12	37.50	Sedang
6	20	62.50	Sedang
7	18	56.25	Sedang
8	11	34.38	Sedang
9	10	31.25	Sedang
10	5	15.63	Sukac
11	11	34.38	Sedang
12	10	31.25	Sedang
13	5	15.63	Sukac
14	9	28.13	Sukac
15	9	28.13	Sukac
16	5	15.63	Sukac
17	23	71.88	Mudah
18	16	50.00	Sedang
19	12	37.50	Sedang
20	10	31.25	Sedang
21	12	37.50	Sedang
22	17	53.13	Sedang
23	13	40.63	Sedang
24	14	43.75	Sedang
25	20	62.50	Sedang
26	26	81.25	Mudah
27	14	43.75	Sedang
28	8	25.00	Sukac
29	16	50.00	Sedang
30	23	71.88	Mudah
31	13	40.63	Sedang
32	13	40.63	Sedang
33	5	15.63	Sukac
34	9	28.13	Sukac
35	8	25.00	Sukac
36	9	28.13	Sukac
37	15	46.88	Sedang
38	13	40.63	Sedang
39	7	21.88	Sukac

Gambar 6 Tingkat Kesukaran



Data diolah dan dikelompokkan berdasarkan kategori yang sudah dijelaskan di bagian metode penelitian. Kategori tingkat kesukaran soal yaitu, sangat mudah, mudah, sedang, dan sukar. Berikut adalah hasil klasifikasi kategori tingkat kesukaran yang diambil dari data Anates.

Tabel 5. Data Klasifikasi Tingkat Kesukaran Soal

Level Kesukaran	Nomor Soal	Jumlah	Persentase
Sangat Mudah	-	0	0%
Mudah	17,26,30	3	7,5%
Sedang	2,3,4,5,6,7,8,9,11,12,18,19,20,21,22,23,24,25, 27,29,31,32,37,38	24	60%
Sukar	1,10,13,14,15,16,28,33,34,35,36,39,40	13	32,5%

Berdasarkan tabel di atas, dapat dilihat bahwa dari total 40 soal, diantaranya 60% berada pada kategori sedang, 3 soal atau 7.5% nya berada pada kategori mudah, dan 13 soal atau dengan persentase 32.5% berada pada kategori sukar. Adapun level sangat mudah pada soal ini tidak ditemukan alias 0% berdasarkan analisis anates.

### Kualitas Pengecoh

Analisis berikutnya adalah analisis pada kualitas pengecoh. Setelah menggunakan software Anates, maka muncul hasil sebagaimana dipaparkan berikut.

```

KUALITAS PENGECOH
-----
Jumlah Subyek- 32
Butir Soal- 40
Nama berkas: D:\DOCUMENTS\PASCA\S2 PENDIDIKAN B. ARAB\TUGAS\SEMESTER 2\PENGEMBANGAN EVALUASI BAHASA ARAB\X4 UJIAN SEMESTER.ANA

No Butir  a    b    c    d    e  *
1  5++  8**  15--  1--  3-  0
2  3++  6-   3++  18**  2+  0
3  6+   7-   2-   3+   14** 0
4  17** 4++  3++  3+   5+   0
5  1--  5++  8-   12**  6++  0
6  2+  20** 2+   4+   4+   0
7  2+  4++  18** 5+   3++  0
8  7+  11** 8-   0--  6++  0
9  14--- 3+  2-   10** 3+  0
10 5+  2-   5**  12-- 8++  0
11 11** 7+  4++  9-   1--  0
12 10** 10-- 3+  5++  4+  0
13 18+  7++  4+  6++  5**  0
14 8+  3+  6++  9**  6++  0
15 9**  14--- 0--  5++  4+  0
16 3-  5**  12-- 7++  5+  0
17 23** 7--- 1-   0--  1-  0
18 3+  6+  1--  16** 6+  0
19 9--  0--  3+  8-   12** 0
20 3+  16--- 10** 1--  2-  0
21 4++  8-   6++  2-   12** 0
22 5+  6-   1-   3++  17** 0
23 13** 6+  9--  3+  1--  0
24 1--  14** 6+  9--  2-  0
25 5-   1-   4+  20** 2+  0
26 4--- 0--  26** 0--  2+  0
27 4++  3+  14** 8--  3+  0
28 8+  8**  4+  9+  3-  0
29 2-  16** 5++  6+  3+  0
30 23** 4--  1-  1-  3+  0
31 13** 6+  3+  8-  2-  0
32 9--  3+  13** 4++  3+  0
33 5+  6++  8++  8++  5**  0
34 9**  4+  14--- 3+  2-  0
35 7++  8**  5++  9+  3-  0
36 8+  10-  9**  4+  1--  0
37 3+  6+  5++  15** 3+  0
38 5++  4++  5++  5++  13** 0
39 6++  4+  7**  12-- 3-  0
40 9**  5++  3+  5++  10-  0

```

Gambar 7 Kualitas Pengecoh



Gambar di atas adalah hasil data yang diolah melalui software Anates untuk mengetahui kualitas pengecoh soal. Dapat dilihat bahwa terdapat daftar kebawah yaitu nomor soal yang berjumlah 40 soal dengan daftar kesamping yaitu pilihan ganda a sampai e. Untuk membacanya dapat dilihat melalui keterangan yang dipaparkan melalui gambar

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

Gambar 8 Keterangan

Setelah menganalisis dengan Anates terkait dengan kualitas pengecoh, maka data diolah untuk dapat menemukan gambaran yang jelas. Analisis efektivitas pengecoh butir soal dapat dilihat dari hasil penyebaran jawaban siswa di setiap butir soal.

Untuk mengetahui efektivitas pengecoh pada setiap butir soal, digunakan kriteria yang diadaptasi dari skala likert. Skala likert dikembangkan oleh seorang ilmuwan asal Amerika Serikat yang bernama Rensis Likert pada tahun 1932. Skala Likert adalah skala atau pengukuran psikologis yang digunakan dalam penelitian untuk mengukur sikap, pengetahuan, pendapat, maupun persepsi responden terhadap suatu objek melalui alat pengumpulan data yang berupa angket atau kuesioner (Arikunto 1998).

Dalam menggunakan skala likert, soal dikatakan memiliki efektivitas pengecoh yang sangat baik jika terdapat empat jawaban pengecoh. Selanjutnya, soal dapat dikatakan memiliki efektivitas pengecoh yang baik jika terdapat tiga jawaban pengecoh. Soal memiliki efektivitas pengecoh yang cukup baik jika terdapat dua jawaban pengecoh dan jika hanya terdapat satu jawaban pengecoh, maka soal dikatakan tidak baik dan jika semua jawaban pengecoh tidak berfungsi, maka soal dikatakan memiliki efektivitas pengecoh yang sangat tidak baik. Dari penyebaran jawaban yang diolah menggunakan Anates diperoleh hasil efektivitas pengecoh pada setiap butir soal sebagaimana dipaparkan pada tabel berikut

Tabel 6. Analisis Efektivitas Pengecoh

No	Kualitas Pengecoh					Efektifitas Pengecoh
	A	B	C	D	E	
1	Sgt Baik	Kunci	Sgt Buruk	Buruk	Krg Baik	Tidak Baik
2	Sgt Baik	Krg Baik	Sgt Baik	Kunci	Baik	Baik
3	Baik	Krg Baik	Krg Baik	Baik	Kunci	Baik
4	Kunci	Sgt Baik	Sgt Baik	Sgt Baik	Baik	Sangat Baik
5	Sgt Buruk	Sgt Baik	Krg Baik	Kunci	Sgt Baik	Cukup Baik
6	Baik	Kunci	Baik	Baik	Baik	Sangat Baik
7	Baik	Sgt Baik	Kunci	Baik	Sgt Baik	Sangat Baik
8	Baik	Kunci	Krg Baik	Buruk	Sgt Baik	Cukup Baik
9	Sgt Buruk	Baik	Krg Baik	Kunci	Baik	Cukup Baik
10	Baik	Krg Baik	Kunci	Buruk	Sgt Baik	Cukup Baik
11	Kunci	Baik	Sgt Baik	Krg Baik	Buruk	Cukup Baik
12	Kunci	Buruk	Baik	Sgt Baik	Baik	Baik
13	Baik	Sgt Baik	Baik	Sgt Baik	Kunci	Sangat Baik
14	Baik	Baik	Sgt Baik	Kunci	Sgt Baik	Sangat Baik
15	Kunci	Sgt Buruk	Buruk	Sgt Baik	Baik	Cukup Baik
16	Krg Baik	Kunci	Buruk	Sgt Baik	Baik	Cukup Baik
17	Kunci	Sgt Buruk	Krg Baik	Buruk	Krg Baik	Sangat Tidak Baik
18	Baik	Baik	Buruk	Kunci	Baik	Baik
19	Buruk	Buruk	Baik	Krg Baik	Kunci	Tidak Baik
20	Baik	Sgt Buruk	Kunci	Buruk	Krg Baik	Tidak Baik
21	Sgt Baik	Krg Baik	Sgt Baik	Krg Baik	Kunci	Cukup Baik
22	Baik	Krg Baik	Krg Baik	Sgt Baik	Kunci	Cukup Baik
23	Kunci	Baik	Buruk	Baik	Buruk	Cukup Baik
24	Buruk	Krg Baik	Sgt Baik	Krg Baik	Kunci	Tidak Baik
25	Krg Baik	Krg Baik	Baik	Kunci	Baik	Cukup Baik
26	Sgt Buruk	Buruk	Kunci	Buruk	Baik	Tidak Baik
27	Sgt Baik	Baik	Kunci	Buruk	Baik	Baik
28	Baik	Kunci	Baik	Baik	Krg Baik	Baik
29	Krg Baik	Kunci	Sgt Baik	Baik	Baik	Baik
30	Kunci	Buruk	Krg Baik	Krg Baik	Baik	Tidak Baik
31	Kunci	Baik	Baik	Krg Baik	Krg Baik	Cukup Baik
32	Buruk	Baik	Kunci	Sgt Baik	Baik	Baik
33	Baik	Sgt Baik	Sgt Baik	Sgt Baik	Kunci	Sangat Baik
34	Kunci	Baik	Sgt Buruk	Baik	Krg Baik	Cukup Baik
35	Sgt Baik	Kunci	Sgt Baik	Baik	Krg Baik	Baik
36	Baik	Krg Baik	Kunci	Baik	Buruk	Cukup Baik
37	Baik	Baik	Sgt Baik	Kunci	Baik	Sangat Baik
38	Sgt Baik	Sgt Baik	Sgt Baik	Sgt Baik	Kunci	Sangat Baik
39	Sgt Baik	Baik	Kunci	Buruk	Krg Baik	Cukup Baik
40	Kunci	Sgt Baik	Baik	Sgt Baik	Krg Baik	Baik

Berdasarkan tabel 6 di atas, ditemukan adanya efektivitas pengecoh yang berfungsi dengan sangat baik yaitu terdapat pada soal nomor 4, 6, 7, 13, 14, 33, 37, dan 38. Pengecoh yang berfungsi dengan baik yaitu terdapat pada soal nomor 2, 3, 12, 18, 27, 28, 29, 32, 35, dan 40. Selanjutnya, pengecoh yang berfungsi cukup baik terdapat pada soal nomor 5, 8, 9, 10, 11, 15, 16, 21, 22, 23, 25, 31, 34, 36, dan 39. Adapun pengecoh yang berfungsi dengan tidak baik terdapat pada soal nomor 1, 19, 20, 24, 26, dan 30. Sedangkan hanya satu nomor soal saja yang memiliki pengecoh yang tidak berfungsi sama sekali yaitu pada soal nomor 17.

Soal yang dibuat oleh guru pengampu mata pelajaran bahasa Arab ini merupakan kumpulan dari sekian banyak butir soal, dengan harapan penyusun dapat mengetahui sejauh mana peserta didik memahami apa yang telah dipelajari. Hal ini sesuai dengan



pendapat Sudijono analisis butir soal yang dilakukan para tenaga pendidik di sekolah adalah mengungkap hasil belajar peserta didik setelah mengikuti proses pembelajaran dalam jangka waktu tertentu. Setiap butir soal yang terdapat pada tes hasil belajar tersebut merupakan bagian yang tidak bisa terpisahkan dari tes hasil belajar sebagai suatu totalitas (Sudijono 2001).

## KESIMPULAN

Analisis butir soal adalah kegiatan menganalisis untuk mengidentifikasi soal-soal pada suatu tes sehingga dapat dipergunakan dan dapat meningkatkan kualitas soal. Sejalan dengan itu, Kegiatan menganalisis butir soal merupakan salah satu kewajiban bagi setiap pendidik karena setiap pendidik pada akhirnya harus dapat memberikan informasi kepada lembaga sekolah ataupun kepada peserta didik tentang kondisi penguasaan dan kemampuan yang telah dicapai peserta didik terhadap materi dan keterampilan-keterampilan dalam mata pelajaran yang telah diberikan.

Berdasarkan hasil penelitian yang telah dilakukan dengan aplikasi Anates terkait analisis butir soal Ujian Akhir Semester Genap kelas X Madrasah Aliyah Negeri 1 Kepulauan Meranti TA. 2023-2023. Didapatkan Uji Validasi butir soal pada kategori sangat baik atau mayoritas valid. Validitas butir soal yang berjumlah 40 bernilai reliabel sebesar 0.85 (sangat tinggi).

Hasil analisis dengan anates terhadap 40 butir soal Bahasa Arab pada Ujian Akhir Semester Genap yang bisa digunakan adalah sebanyak 24 soal. Sedangkan 16 soal lainnya lebih baik diganti, direvisi kembali, atau bahkan dibuang. Pada Tingkat kesukaran ditemukan bahwa bahwa dari total 40 soal, diantaranya 60% berada pada kategori sedang, 3 soal atau 7.5% nya berada pada kategori mudah, dan 13 soal atau dengan persentase 32.5% berada pada kategori sukar. Adapun level sangat mudah pada soal ini tidak ditemukan alias 0%.

Daya pembeda pada butir soal dikategorikan "Baik" dengan jumlah butir soal sebanyak 17 soal dari total 40 soal dengan persentase 42,5%. 12 Soal (30%) dengan kategori "Cukup Baik", 7 soal atau (17,5%) dengan kategori "Buruk" sebaiknya tidak digunakan dan 4 soal (10%) dengan kategori "Baik Sekali". Sementara itu, analisis tentang efektivitas pengecoh ditemukan adanya efektivitas pengecoh yang berfungsi dengan sangat baik yaitu terdapat pada soal nomor 4, 6, 7, 13, 14, 33, 37, dan 38. Pengecoh yang berfungsi dengan baik yaitu terdapat pada soal nomor 2, 3, 12, 18, 27, 28, 29, 32, 35, dan 40. Selanjutnya, pengecoh



yang berfungsi cukup baik terdapat pada soal nomor 5, 8, 9, 10, 11, 15, 16, 21, 22, 23, 25, 31, 34, 36, dan 39. Adapun pengecoh yang berfungsi dengan tidak baik terdapat pada soal nomor 1, 19, 20, 24, 26, dan 30. Sedangkan hanya satu nomor soal saja yang memiliki pengecoh yang tidak berfungsi sama sekali yaitu pada soal nomor 17.

## DAFTAR PUSTAKA

- Alpusari, Mahmud. 2014. "Analisis Butir Soal, Program Komputer Anates Versi 4.0 For Windows Mahmud Alpusari." *Jurnal Primary Program Studi Pendidikan Guru Sekolah Dasar Universitas Riau* 3(2):107.
- Amalia, Ata Nayla, dan Ani Widayati. 2012. "Analisis Butir Soal Tes Kendali Mutu Kelas Xii Sma Mata Pelajaran Ekonomi Akuntansi Di Kota Yogyakarta Tahun 2012." *Jurnal Pendidikan Akuntansi Indonesia* 10(1). doi: 10.21831/jpai.v10i1.919.
- Ancok, Djameludin. 1989. "Validitas dan Reliabilitas Instrumen-Penelitian, dalam Masri Singarimbun dan Sofian Effendi (Ed.) Metode Penelitian Survey." *Jakarta: LP3ES*.
- Anggreini, Dewi, dan Caesar Angga Darmawan. 2016. "Analisis kualitas soal try out ujian nasional dengan menggunakan aplikasi program ANATES." *JP2M (Jurnal Pendidikan dan Pembelajaran Matematika)* 2(1):20-34.
- Arikunto, Suharsimi. 1998. "Pendekatan Penelitian." *Jakarta: Rineka Cipta*.
- Ida, Farida Far, dan Anna Musyarofah. 2021. "Validitas dan Reliabilitas dalam Analisis Butir Soal." *Al-Mu'arrib: Jurnal Pendidikan Bahasa Arab* 1(1):34-44.
- Juandi, Aghi. 2019. "Standar penilaian pendidikan."
- NN, Dewi Dian Ayunita. 2018. "Modul Uji Validitas dan Reliabilitas." *Semarang: Universitas Diponegoro*.
- Salmina, Mik, dan Fadlillah Adyansyah. 2017. "Analisis kualitas soal ujian matematika semester genap kelas XI SMA Inshafuddin Kota Banda Aceh." *Numeracy* 4(1):37-47.
- Sudijono, Anas. 2001. *Pengantar evaluasi pendidikan*. PT. Raja Grafindo Persada.
- Sugiyono, Dr. 2013. "Metode penelitian pendidikan pendekatan kuantitatif, kualitatif dan R&D."
- Wiguna, Satria. 2021. "Aplikasi anates dalam evaluasi pembelajaran."
- Yoshita Cahyaningrum, Ira, Anies Fuady, dan Universitas Islam Malang. 2023. "Analisis Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran Matematika Kelas VII dengan Berbantuan Aplikasi Software Anates." *Mathema Journal* 5(2):67-81.

