



Analisis Butir Soal Ujian Akhir Bentuk Pilihan Ganda Mata Pelajaran Bahasa Arab Kelas Xi Madrasah Aliyah Negeri 2 Cilacap Tahun Pelajaran 2023/2024

Audi Ismatul Mukarromah,¹ Zamakhsari,² Muhajir³

123UIN Sunan Kalijaga Yogyakarta, Indonesia

Email: ¹21104020044@student.uin-suka.ac.id, ²zamakhsari@uin-suka.ac.id,

³muhajir@gmail.com

Article Info

Article History

Received:

Revised:

Accepted:

Published:

Keyword:
End of Semester
Sumatif,
Evaluation,
Analysis,
Madrasah Aliyah
Negeri 2 Cilacap و
Quality of
Questions

Abstract

Assessment is a critical component of the learning process, serving to determine the achievement of learning objectives. This study aims to evaluate the quality of Arabic language multiple-choice final exam items for eleventh-grade students of class "L" at Madrasah Aliyah Negeri 2 Cilacap in the 2023/2024 academic year, focusing on validity, reliability, difficulty level, discrimination index, and distractor effectiveness. This quantitative descriptive study used data sources including question sheets, student answer sheets, question grids, and answer keys. Data were collected through interviews and documentation and analyzed using the Anates V4 application. The findings revealed that the overall quality of the items was low. In terms of validity, only 7 items (23%) were categorized as high and none as very high; 14 items (47%) were classified as very low. Reliability was high with a coefficient of 0.72. Regarding difficulty level, most items were too easy, with 15 items (50%) in the very easy category. The discrimination index showed 11 items (37%) were not good, while none reached the very good category. Distractor effectiveness showed poor results, with 12 items (40%) in the very bad category. These findings suggest that the exam items require significant revision to meet quality standards.

الملخص

يُعد التقييم جزءاً أساسياً في عملية التعلم، ويهدف إلى قياس مدى تحقيق الأهداف التعليمية. تهدف هذه الدراسة إلى تقويم جودة أسئلة الاختبار النهائي لمادة اللغة العربية بصيغة الاختيار من متعدد لطلاب الصف الحادي عشر "L" في المدرسة الثانوية الإسلامية الحكومية الثانية بـ "جلاكاب" للعام الدراسي 2023/2024، من حيث الصدق والثبات ومستوى الصعوبة ومعامل التمييز وفعالية المشتقات. تعتمد هذه الدراسة على المنهج الوصفي الكمي، ومصادر البيانات تشمل أوراق الأسئلة، وأوراق إجابات الطلاب، وجدول المواصفات، ومفتاح الإجابة. جُمعت البيانات من خلال المقابلات والتوثيق، وتم تحليلها باستخدام برنامج Anates V4. أظهرت النتائج

أن جودة الأسئلة لا تزال منخفضة بشكل عام. من حيث الصدق، تبيّن أن 7 أسئلة (23٪) فقط كانت ضمن الفئة العالية، ولم يكن هناك أي سؤال ضمن الفئة العالية جداً، بينما صنفت 14 سؤالاً (47٪) ضمن الفئة المنخفضة جداً. أما الثبات فكان مرتفعاً بقيمة 0.72. من حيث مستوى الصعوبة، كانت أغلب الأسئلة سهلة جداً (15 سؤالاً بنسبة 50٪). أظهر معامل التمييز أن 11 سؤالاً (37٪) لم تكن جيدة، ولم يصل أي سؤال إلى فئة التمييز العالى جداً. أما فعالية المشتتات، فقد كانت ضعيفة، حيث تبيّن أن 12 سؤالاً (40٪) ضمن الفئة السيئة جداً. وتشير هذه النتائج إلى ضرورة إعادة صياغة الأسئلة بشكل كبير لتوافق مع معايير الجودة.

Pendahuluan

Pelaksanaan pembelajaran adalah proses yang mencakup tujuan yang dimaksudkan untuk dicapai oleh guru dan siswa. Dengan adanya tujuan ini, guru dapat lebih mudah menetapkan alur pembelajaran yang akan membantu mereka mencapai tujuan tersebut. Dengan berbagai upaya yang dilakukan oleh guru, diharapkan mereka dapat membantu mengarahkan siswa mereka untuk mencapai tujuan tersebut. Dengan banyaknya upaya yang dilakukan guru dalam pelaksanaan pembelajaran, maka perlu dinilai apakah pelaksanaan pembelajaran telah mencapai tujuan atau tidak. Oleh karena itu, evaluasi perlu dilakukan untuk mengetahui kemampuan siswa dalam belajar dan apakah tujuan pembelajaran telah tercapai.¹

Dalam pembahasan evaluasi pembelajaran, terdapat juga istilah pengukuran dan penilaian. Ketiganya memiliki makna yang berbeda namun sering disamakan. Evaluasi sendiri sering disamaartikan dengan ujian, akan tetapi arti tersebut belum mencakup keseluruhan makna evaluasi yang sebenarnya. Karena pada dasarnya, evaluasi pembelajaran bukan hanya menilai hasil belajar, akan tetapi juga menilai proses tindakan guru dan siswa selama proses pembelajaran. Sedangkan pengukuran adalah metode untuk menghitung kuantitas sesuatu. Dan penilaian adalah suatu prosedur atau kegiatan yang teratur dan berkelanjutan yang dimaksudkan untuk mengumpulkan informasi tentang proses dan hasil belajar siswa dalam proses pengambilan keputusan dengan mempertimbangkan pertimbangan dan kriteria tertentu.²

Evaluasi adalah bagian dari proses dan tidak dapat dipisahkan dari kegiatan pembelajaran. Menurut BAB XVI Pasal 58 Ayat 1 dari Undang-Undang Republik Indonesia No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional, "Evaluasi hasil belajar peserta didik dilakukan oleh pendidik untuk memantau proses, kemajuan dan perbaikan hasil belajar peserta didik secara berkesinambungan". Oleh karena itu, evaluasi belajar bertujuan untuk mengevaluasi pencapaian kompetensi, meningkatkan proses pembelajaran dan menetapkan standar untuk penyusunan laporan kemajuan hasil

¹ Bahrudin Fahmi, Syahrul Rizqi dan Nurul Elmira, "Analisis Butir Soal Bahasa Arab Siswa MAS Pondok Pesantren Asaalam Kampar Riau", Jurnal Pendidikan Bahasa Arab & Kebahasaaraban, Vol.6 No.1 (2022), 96.

² Asrul, Abdul Hasan Saragih dan Mukhtar, *Evaluasi Pembelajaran*, (Medan: Perdana Publishing, 2022), hal.18.

belajar siswa.³ Secara umum, instrumen evaluasi dapat dibagi menjadi dua kategori yaitu tes dan non tes. Tes dapat dilakukan secara tertulis maupun tidak tertulis, sedangkan non tes dilakukan dengan menilai sikap, keterampilan, tingkah laku dan sebagainya.⁴

Tes merupakan kumpulan tugas yang diberikan kepada siswa untuk menilai atau mengevaluasi tingkat pembelajaran mereka. Langkah-langkah dalam pelaksanaan tes meliputi perencanaan, persiapan, uji coba dan penilaian hasil ukuran.⁵ Tes yang diujikan kepada siswa harus berkualitas. Tes yang berkualitas adalah tes yang dapat memberikan informasi yang akurat tentang hasil belajar siswa, sehingga dapat digunakan untuk mengetahui seberapa jauh program pengajaran yang telah ditetapkan.⁶ Dengan menggunakan tes, seorang pendidik dapat mengetahui apakah tujuannya dalam proses pembelajaran yang sudah ditetapkan dalam kurikulum seperti kompetensi dasar, standar kompetensi dan indikator sudah tercapai atau belum.⁷ Tes yang baik harus memenuhi validitas rasional, yang mencakup validitas empirik, validitas konstruk dan validitas isi. Kegunaan tes itu sendiri selain digunakan untuk mengetahui tujuan pembelajaran sudah tercapai atau belum, digunakan juga untuk seleksi, penempatan, diagnosa, remedial, umpan balik, motivasi dan bimbingan, perbaikan kurikulum, program pendidikan dan pengembangan ilmu.⁸

Tes tertulis tidak cukup untuk menentukan seberapa jauh kemampuan siswa dalam memahami topik yang telah dipelajari. Guru juga dapat menggunakan non tes, seperti melihat teman sekelasnya atau lingkungan masyarakat tentang hal-hal yang berkaitan dengan materi pelajaran. Tujuan dari tes tulis atau lisan adalah untuk mengevaluasi pengetahuan kognitif siswa, sedangkan untuk non tes untuk mengevaluasi kemampuan siswa dalam menerapkan pengetahuan pada aspek psikomotorik dan afektif.⁹

Dengan melakukan evaluasi, sekolah dapat mendorong siswanya untuk belajar lebih baik lagi dan dapat mendorong guru untuk meningkatkan proses pembelajaran. Dengan demikian, optimalisasi sistem evaluasi memiliki dua arti: pertama, sistem evaluasi yang memberikan evaluasi yang optimal, dan kedua, manfaat yang dicapai dari evaluasi. Manfaat yang paling signifikan dari evaluasi adalah meningkatkan kualitas pembelajaran dan memperbaiki kualitas pendidikan.¹⁰ Fakta di lapangan menunjukkan

³ Tutut Kurniawan, "Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran IPS Sekolah Dasar", *Jurnal of Elementary Education*, (Januari, 2015), 2.

⁴ Farida dan Anna Musyarofah, "Validitas dan Reliabilitas dalam Analisis Butir Soal", *Jurnal Al-Mu'arrib*, Vol.1 No.1 (Juni, 2021), 37.

⁵ Eka Mayasari, "Instrumen Tes Sebagai Alat Evaluasi", *Jurnal Pendidikan dan Pengajaran*, Vol.2 No.1 (Januari, 2023), 59.

⁶ Rahmatika Rahayu, M. Djazari, "Analisis Kualitas Soal Pra Ujian Nasional Mata Pelajaran Ekonomi Akuntansi", *Jurnal Pendidikan Akuntansi Indonesia*, Vol.14 No.1 (2016), 86.

⁷ Nurul Fikriyah, "Analisis Butir Soal Ulangan Tengah Semester Mata Pelajaran Bahasa Arab Kelas VII Semester Ganjil SMP Muhammadiyah 1 Yogyakarta Tahun Ajaran 2019/2020", *Jurnal Maharaat*, Vol.3 No.2 (April, 2021), 129.

⁸ Deni Maulana dan Anwar Sanusi, "Analisis Butir Soal Bahasa Arab Ujian Akhir Madrasah Bersama Daerah (UAMBD) Madrasah Ibtidaiyah Tahun 2017-2018", *Jurnal Ta'lim al-'Arabiyyah*, Vol.4 No.1 (2020), 13.

⁹ Angga Widiyarto dan Nurul Latifatul Inayati, "Penerapan Evaluasi Pembelajaran Tes dan Non-Tes Pendidikan Agama Islam di Sekolah Menengah Kejuruan", *Jurnal Munaddhomah*, Vol.4 No.2 (2023), 308.

¹⁰ Annisa Rakhmaningrum, Skripsi: *Analisis Butir Soal Penilaian Tengah Semester Mata Pelajaran Bahasa Arab Wajib Kelas XII Keagamaan Madrasah Aliyah Negeri 1 Banyumas Kecamatan Purwokerto Timur Kabupaten Banyumas Tahun Ajaran 2021/2022*, (Purwokerto: UIN Saizu, 2022), hal.3.



bahwa analisis butir soal sering diabaikan. Oleh karena itu, informasi tentang validitas, reliabilitas, tingkat kesukaran, daya pembeda dan efektivitas pengecoh sering kali tidak dapat diukur. Sangat penting bagi seorang guru untuk melakukan analisis butir soal untuk meningkatkan kualitas soal yang mereka buat. Kegiatan ini melibatkan pengumpulan, ringkasan dan pemanfaatan informasi dari jawaban siswa untuk membantu membuat keputusan terkait semua aspek penilaian.¹¹

Metode Penelitian

Dalam penelitian ini, peneliti menggunakan jenis penelitian kuantitatif deskriptif. Metode penelitian kuantitatif deskriptif merupakan metode dengan suatu rumusan masalah yang memadukan penelitian untuk mengeksplorasi atau memotret situasi sosial yang akan diteliti secara menyeluruh, luas dan mendalam. Tujuan dari penelitian kuantitatif deskriptif adalah mengungkapkan fakta atau karakteristik populasi atau bidang tertentu secara akurat dan sistematis.¹²

Data dalam penelitian ini adalah berupa jawaban-jawaban dari 30 soal ujian pilihan ganda siswa dalam mata pelajaran Bahasa Arab Kelas XI L Madrasah Aliyah Negeri 2 Cilacap Tahun pelajaran 2023/2024. Data akan diolah sampai diketahui berapa nilai validitas, reliabilitas, tingkat kesukaran, daya pembeda dan efektivitas pengecohnya untuk diketahui kualitas soal yang telah diujikan kepada siswa.

Hasil dan Pembahasan

Validitas

No	Butir	Korelasi	Signifikansi
1		0.045	-
2		0.130	-
3		-0.125	-
4		-0.120	-
5		-0.349	-
6		0.118	-
7		-0.191	-
8		0.146	-
9		0.296	-
10		0.216	-
11		0.129	-
12		-0.007	-
13		0.625	Sangat Signifikan
14		0.064	-
15		0.660	Sangat Signifikan
16		NAN	NAN
17		-0.058	-
18		0.533	Sangat Signifikan
19		0.417	Signifikan
20		0.558	Sangat Signifikan
21		0.454	Sangat Signifikan
22		0.596	Sangat Signifikan
23		0.625	Sangat Signifikan
24		0.752	Sangat Signifikan
25		-0.166	-
26		0.306	-
27		0.716	Sangat Signifikan
28		0.673	Sangat Signifikan
29		-0.107	-
30		0.766	Sangat Signifikan

Gambar 1. Hasil Uji Validitas Instrumen

¹¹ Leony Eka Pratiwi, H. Rufi'i, "Analisis Soal Sumatif IPA Mengenai Gaya pada Peserta Didik kelas IV di SD YP Nasional", *Jurnal Didaktika*, Vol.12 No. 4 (November, 2023), 600-601.

¹² Agus Rustamana, Putri Wahyuning, Muhammad Fikri Azka dan Pipit Wahyu, "Penelitian Metode Kuantitatif", *Jurnal Sindoro Cendikia Pendidikan*, Vol.5 No.6 (2024), 4.

Dari 30 butir soal yang dianalisis diperoleh 11 soal sangat signifikan, 1 soal signifikan, 17 soal tidak signifikan dan 1 soal NAN (*Not a Number*) yang artinya soal tersebut tidak bisa dihitung atau dianalisis karena korelasinya 0,000 sehingga menampilkan kata NAN. Jika diklasifikasikan berdasarkan kategori validitasnya terdapat 7 soal yaitu soal nomor 13, 15, 23, 24, 27, 28 dan 30 dengan kategori tinggi dengan persentase 23%, tidak terdapat soal dengan kategori sangat tinggi (0%), 5 soal yaitu soal nomor 18, 19, 20, 21, dan 22 dengan kategori cukup dengan persentase 17%, 4 soal yaitu soal nomor 5, 9, 10 dan 26 dengan kategori rendah dengan persentase 13%, 14 soal yaitu soal nomor 1, 2, 3, 4, 6, 7, 8, 11, 12, 14, 16, 17, 25 dan 29 dengan kategori sangat rendah dengan persentase 47%.

Reliabilitas

```
RELIABILITAS TES
=====
Rata2= 23.53
Simpang Baku= 3.08
KorelasiXY= 0.56
Reliabilitas Tes= 0.72
Nama berkas: C:\USERS\ACER\ONEDRIVE\DOCUMENTS\SKRIPSI BGTTTTTT\DATA FIX.ANA
```

Gambar 2. Hasil Uji Reliabilitas

Berdasarkan hasil analisis 30 butir soal pilihan ganda Mata Pelajaran Bahasa Arab Kelas XI L MAN 2 Cilacap Tahun Ajaran 2023/2024 melalui aplikasi Anates V4 maka dapat diperoleh hasil bahwa reliabilitas tes sebesar 0,72. Dengan hasil ini maka dapat dikategorikan bahwa reliabilitas tes memiliki interpretasi yang tinggi karena termasuk ke dalam reliabilitas tes 0,60 – 0,79 berdasarkan rumus kriteria reliabilitas.

Tingkat Kesukaran

```
TINGKAT KESUKARAN
=====
Jumlah Subjek= 32
Butir Soal= 30
Nama berkas: C:\USERS\ACER\ONEDRIVE\DOCUMENTS\SKRIPSI BGTTTTTT\DATA FIX.ANA
```

No Butir	Jml Betul	Tkt. Kesukaran(%)	Tafsiran
1	30	93.75	Sangat Mudah
2	30	93.75	Sangat Mudah
3	30	93.75	Sangat Mudah
4	29	90.63	Sangat Mudah
5	23	71.88	Mudah
6	22	68.75	Sedang
7	29	90.63	Sangat Mudah
8	1	3.13	Sangat Sukar
9	26	81.25	Mudah
10	30	93.75	Sangat Mudah
11	28	87.50	Sangat Mudah
12	25	78.13	Mudah
13	24	75.00	Mudah
14	23	71.88	Mudah
15	23	71.88	Mudah
16	32	100.00	Sangat Mudah
17	28	87.50	Sangat Mudah
18	26	81.25	Mudah
19	25	78.13	Mudah
20	27	84.38	Mudah
21	26	81.25	Mudah
22	28	87.50	Sangat Mudah
23	24	75.00	Mudah
24	28	87.50	Sangat Mudah
25	23	71.88	Mudah
26	20	62.50	Sedang
27	25	78.13	Mudah
28	24	75.00	Mudah
29	19	59.38	Sedang
30	25	78.13	Mudah

Gambar 3. Hasil Uji Tingkat Kesukaran Butir Soal

Dari 30 butir soal yang dianalisis diperoleh 11 soal dengan kategori sangat mudah, 15 soal dengan kategori mudah, 3 soal dengan kategori sedang dan 1 soal dengan kategori sangat sukar. Jika diklasifikasikan berdasarkan kategori tingkat kesukarannya terdapat 1 soal nomor 8 dengan kategori sangat sukar (3%), tidak terdapat soal dengan kategori sukar (0%), 1 soal nomor 29 dengan kategori sedang (3%), 13 soal nomor 5, 6, 12, 13, 14, 15, 19, 23, 25, 26, 27, 28 dan 30 dengan kategori mudah (44%) dan 15 soal nomor 1, 2, 3, 4, 7, 9, 10, 11, 16, 17, 18, 20, 21, 22 dan 24 dengan kategori sangat mudah (50%).

Daya Pembeda

DAYA PEMBEDA						
=====						
<u>Jumlah Subyek= 32</u>						
<u>Klp atas/bawah(n)= 9</u>						
<u>Butir Soal= 30</u>						
Nama berkas: C:\USERS\ACER\ONEDRIVE\DOCUMENTS\SKRIPSI BGTTTTTT\DATA FIX.ANA						
No	Butir	Kel. Atas	Kel. Bawah	Beda	Indeks	DP (%)
1		9	8	1	11.11	
2		9	7	2	22.22	
3		8	9	-1	-11.11	
4		8	9	-1	-11.11	
5		5	9	-4	-44.44	
6		6	7	-1	-11.11	
7		7	9	-2	-22.22	
8		1	0	1	11.11	
9		6	7	-1	-11.11	
10		9	7	2	22.22	
11		8	7	1	11.11	
12		9	7	2	22.22	
13		8	3	5	55.56	
14		9	5	4	44.44	
15		8	2	6	66.67	
16		9	9	0	0.00	
17		7	7	0	0.00	
18		7	5	2	22.22	
19		7	6	1	11.11	
20		8	6	2	22.22	
21		8	6	2	22.22	
22		9	6	3	33.33	
23		8	2	6	66.67	
24		9	5	4	44.44	
25		8	7	1	11.11	
26		7	5	2	22.22	
27		8	3	5	55.56	
28		8	3	5	55.56	
29		9	7	2	22.22	
30		9	3	6	66.67	

Gambar 4. Hasil Uji Daya Pembeda Butir Soal

Dari 30 butir soal yang dianalisis diperoleh 11 soal dengan kategori kurang baik, 10 soal dengan kategori cukup, 9 soal dengan kategori baik dan tidak ada soal dengan kategori baik sekali. Jika diklasifikasikan berdasarkan kategori daya pembedanya terdapat 11 soal yaitu soal nomor 1, 3, 4, 6, 8, 9, 11, 16, 17, 19 dan 25 dengan persentase 37% masuk ke dalam kategori kurang baik. 10 soal yaitu soal nomor 2, 7, 10, 12, 18, 20, 21, 22, 26 dan 29 dengan persentase 33% masuk ke dalam kategori cukup, 9 soal yaitu soal nomor 5, 13, 14, 15, 23, 24, 27, 28 dan 30 dengan persentase 30% masuk ke dalam kategori baik dan tidak ada soal dengan kategori baik sekali (0%).

Efektivitas Pengecoh



KUALITAS PENGECHOH						
No	Butir	a	b	c	d	e
1	2---	30**	0--	0--	0--	0
2	30**	2---	0--	0--	0--	0
3	2---	30**	0--	0--	0--	0
4	3---	0--	29**	0--	0--	0
5	4--	0--	3+	23**	2++	0
6	4-	1-	3++	22**	2++	0
7	1+	0--	2---	0--	29**	0
8	7++	24---	0--	1**	0--	0
9	6---	26**	0--	0--	0--	0
10	30**	0--	0--	2---	0--	0
11	2--	0--	0--	2--	28**	0
12	0--	2++	0--	25**	5---	0
13	2++	24**	4--	0--	2++	0
14	23**	0--	2++	2++	5---	0
15	2++	23**	4--	1-	2++	0
16	32**	0	0	0	0	0
17	2--	0--	28**	2--	0--	0
18	2+	26**	2+	2+	0--	0
19	4---	3-	0--	0--	25**	0
20	2-	27**	1++	2-	0--	0
21	5---	1+	0--	26**	0--	0
22	28**	1++	2--	0--	1++	0
23	3+	0--	24**	3+	2++	0
24	4---	0--	0--	0--	28**	0
25	23**	0--	0--	7---	2++	0
26	20**	0--	0--	12---	0--	0
27	2++	0--	0--	25**	5---	0
28	1-	1-	4--	2++	24**	0
29	2+	0--	19**	5-	6--	0
30	2++	3-	25**	0--	2++	0

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

Gambar 5. Kualitas Pengecoh Butir Soal

Berdasarkan hasil analisis butir soal ujian akhir Mata Pelajaran Bahasa Arab Kelas XI L MAN 2 Cilacap Tahun Ajaran 2023/2024 yang berjumlah 30 soal dalam segi efektivitas pengecoh terdapat 6 soal yaitu soal nomor 6, 13, 14, 15, 22 dan 30 dengan persentase 20% masuk dalam kategori sangat baik, 4 soal yaitu soal nomor 18, 20, 23 dan 28 dengan persentase 13% masuk dalam kategori baik, 2 soal nomor 5 dan 29 dengan persentase 7% masuk dalam kategori kurang baik, 6 soal yaitu soal nomor 8, 12, 19, 21, 25 dan 27 dengan persentase 20% masuk dalam kategori buruk dan 12 soal yaitu soal nomor 1, 2, 3, 4, 7, 9, 10, 11, 16, 17, 24 dan 26 dengan persentase 40% masuk dalam kategori sangat buruk.

Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa soal ujian akhir bentuk pilihan ganda Mata Pelajaran Bahasa Arab Kelas XI L MAN 2 Cilacap Tahun Ajaran 2023/2024 yang berjumlah 30 soal belum bisa dikatakan berkualitas baik karena masih terdapat beberapa soal dengan kategori rendah.

Referensi

Asrul, Abdul Hasan Saragih, Mukhtar. *Evaluasi Pembelajaran*. Medan: Perdana Publishing, 2022.

- Fahmi, Bahrudin, Syahrul Rizqi, dan Nurul Elmira Harmeilinda. "Analisis Butir Soal Bahasa Arab MAS Pondok Pesantren Assalam Kampar Riau." *Ta'lim al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab & Kebahasaaraban* 6, no. 1 (1 Juni 2022): 95–105. <https://doi.org/10.15575/jpba.v6i1.16193>.
- Fikriyah, Nurul. "Analisis Butir Soal Ulangan Tengah Semester Mata Pelajaran Bahasa Arab Kelas VII Semester Ganjil SMP Muhammadiyah 1 Yogyakarta Tahun Ajaran 2019/2020." *Maharaat: Jurnal Pendidikan Bahasa Arab* 3, no. 2 (22 September 2021). <https://doi.org/10.18196/mht.v3i2.10501>.
- Ida, Farida Far, dan Anna Musyarofah. "Validitas dan Reliabilitas dalam Analisis Butir Soal." *AL-MU'ARRIB: JOURNAL OF ARABIC EDUCATION* 1, no. 1 (6 Desember 2021): 34–44. <https://doi.org/10.32923/al-muarrib.v1i1.2100>.
- Kurniawan, Tutut. "ANALISIS BUTIR SOAL ULANGAN AKHIR SEMESTER GASAL MATA PELAJARAN IPS SEKOLAH DASAR," 2015.
- Maulana, Deni, dan Anwar Sanusi. "ANALISIS BUTIR SOAL BAHASA ARAB UJIAN AKHIR MADRASAH BERSAMA DAERAH (UAMBD) MADRASAH IBTIDAIYAH TAHUN 2017-2018." *Ta'lim al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab & Kebahasaaraban* 4, no. 1 (15 Juni 2020): 12–24. <https://doi.org/10.15575/jpba.v4i1.8054>.
- Mayasari, Eka. "INSTRUMEN TES SEBAGAI ALAT EVALUASI (Analisis Soal, Indek Kesukaran, Daya Pembeda dan Fungsi Distraktor)," t.t.
- Pratiwi, Leony Eka, H. Rufi'i. "Analisis Soal Sumatif IPA Mengenai Gaya pada Peserta Didik Kelas IV di SD YP Nasional." *Jurnal Didaktika* Vol.12 No.4 (November 2023).
- Rahayu, Rahmatika, M. Djazari. "ANALISIS KUALITAS SOAL PRA UJIAN NASIONAL MATA PELAJARAN EKONOMI AKUNTANSI." *Jurnal Pendidikan Akuntansi Indonesia* Vol. 14 No.1 (2016).
- Rakhmaningrum, Annisa. "ANALISIS BUTIR SOAL PENILAIAN TENGAH SEMESTER MATA PELAJARAN BAHASA ARAB WAJIB KELAS XII KEAGAMAAN MADRASAH ALIYAH NEGERI 1 BANYUMAS KECAMATAN PURWOKERTO TIMUR KABUPATEN BANYUMAS TAHUN AJARAN 2021/2022." UIN Saifuddin Zuhri Purwokerto, t.t.
- Rustamana, Agus, Putri Wahyuningsih, Muhammad Fikri Azka, Pipit Wahyu, Sultan Ageng Tirtayasa, dan Ciwaru Raya No. "PENELITIAN METODE KUANTITATIF," no. 6 (2024).
- Widiyarto, Angga, dan Nurul Latifatul Inayati. "Penerapan Evaluasi Pembelajaran Tes Dan Non-Tes Pendidikan Agama Islam di Sekolah Menengah Kejuruan." *Munaddhomah: Jurnal Manajemen Pendidikan Islam* 4, no. 2 (28 April 2023): 307–16. <https://doi.org/10.31538/munaddhomah.v4i2.439>.