Psikosospen : Jurnal Psikososial dan Pendidikan
https://publisherqu.com/index.php/psikosospen/
Vol. 1 No. 2 Juni 2025
https://issn.brin.go.id/terbit/detail/20250326030437435

ANALISIS KUALITAS TES DAN BUTIR SOAL UJIAN MADRASAH MATA PELAJARAN BAHASA ARAB KELAS IX

[1]Azdiva Miftahur Rizkiyah, [2]Hikmah, [3]Masrun
Universitas Islam Negeri Sultan Syarif Kasim Riau
[1]azdivamiftahur@gmail.com,[2]hikmahananada2@gmail.com, [3]masrun@uin-suska.ac.id

Abstrak

Penelitian analisis butir soal ini bertujuan untuk mengetahui kualitas dari butir soal ujian madrasah mata pelajaran Bahasa Arab kelas IX. Dalam penelitian ini menggunakan metode deskriptif degan menggunakan pendekatan kuantitatif. Populasi pada penelitian ini adalah siswa kelas IX MTs Humairoh HNN Kampar tahun pelajaran 2024/2025, sampel pada penelitian ini adalah seluruh siswa kelas XI MTs yang terdiri dari 14 siswa. Soal ujian madrasah mata pelajaran Bahasa Arab terdiri dari 20 butir soal pilihan ganda. Hasil penelitian menunjukkan kulitas soal masih kurang baik karena tidak sesuai dengan soal standar, maka perlu untuk diperbaiki. Hasil uji validitas soal didapatkan 3 butir soal valid dan 17 butir soal tidak valid, reabilitas soal menyatakan tidak reliabel termasuk kategori rendah, untuk tingkat kesukaran terdapat 15 butir soal kategori sukar, 4 butir soal kategori sedang, dan 1 butir soal kategori mudah, sedangkan untuk daya pembeda terdapat 2 butir soal kategori baik, 2 butir soal kategori cukup, dan 13 butir soal kategori lemah.
Kata Kunci : Validitas, Reliabilitas, Tingkat Kesukaran, Daya Pembeda

PENDAHULUAN

Setiap program pembelajaran mengandung empat komponen utama, yaitu: tujuan, materi, metode dan evaluasi. Tujuan adalah sasaran dengan indikator kompetensi yang telah ditetapkan. Sedangkan Bahan ajar atau materi merupakan substansi ilmu pengetahuan yang ditransformasikan pendidik kepada peserta didik. Metode adalah langkah langkah atau cara maupun teknik penyajian materi kepada peserta didik, dan yang terakhir evaluasi adalah pengukuran atau penilaian komponen input, proses dan produk melalui instrumen tes yang ditetapkan. Pendidik harus mengetahui kualitas hasil belajar, yang dapat dikategorikan menjadi baik, tidak baik, bermanfaat, atau tidak bermanfaat. Hal ini penting untuk diketahui karena hasil belajar berfungsi sebagai alat untuk mengukur seberapa baik proses pembelajaran yang diterapkan. Proses pembelajaran dapat dikatakan berhasil jika hasil belajarnya baik, dan sebaliknya (Zainal, 2012).

Salah satu kompetensi yang harus dimiliki seorang pendidik adalah kemampuan mengadakan evaluasi, baik dalam proses pembelajaran maupun penilaian hasil belajar. Kemampuan melaksanakan evaluasi pembelajaran merupakan kemampuan dasar yang mesti dikuasai oleh seorang pendidik maupun calon pendidik sebagai salah satu kompetensi professionalnya. Dengan melihat hasil belajar, pendidik dapat mengetahui seberapa memahami pelajaran siswa dan dapat dijadikan balikan pendidik dalam memperbaiki dan menyempurnakan program dan kegiatan pembelajaran (Asrul et al., 2022; Rahman & Nasryah, 2019).

Secara harafiah evaluasi berasal dari bahasa Inggris evaluation yang berarti penilaian atau penaksiran. Secara garis besar dapat dikatakan bahwa evaluasi adalah suatu proses yang sistematis dan berkelanjutan untuk pemberian nilai terhadap kualitas sesuatu dan dapat dipandang juga sebagai proses merencanakan, memperoleh, dan menyediakan informasi yang sangat diperlukan untuk membuat alternatif keputusan. Evaluasi memiliki dua macam fungsi, yaitu fungsi formatif dan fungsi sumatif. Fungsi formatif dilakukan apabila hasil yang diperoleh dari kegiatan evaluasi diarahkan untuk memperbaiki bagian tertentu atau sebagian besar bagian kurikulum yang sedang dikembangkan. Sedangkan fungsi sumatif dihubungkan dengan penyimpulan mengenai kebaikan dari sistem secara keseluruhan (Zainal, 2012). Dengan demikian, Evaluasi menjadi suatu hal yang sangat penting untuk mengetahui apakah setiap kegiatan pembelajaran yang dilakukan dapat dikatakan berhasil atau tidak, maka dalam melakukan evaluasi harus terencana, terakurat, dan terprogram (Hasan, n.d.; Masullah et al., 2024; Rahman & Nasryah, 2019).

Teknik pengumpulan data dalam kegiatan evaluasi terbagi menjadi dua yaitu: teknik tes dan teknik non tes. Tes dapat didefiniskan sebagai suatu pertanyaan atau tugas atau seperangkat tugas yang direncanakan untuk memperoleh informasi tentang trait atau atribut pendidikan atau psikologik yang setiap butir pertanyaan atau tugas tersebut mempunyai jawaban atau ketentuan yang dianggap benar (Hasan, n.d.). Tes yang baik yang akan menghasilkan data yang tepat. Tingkat kebaikan suatu tes sekurang kurangnya dapat dilihat dari empat yaitu, validitas, reliabilitas, tingkat kesukaran, dan daya beda tes. Sehingga, tujuan dari pembelajaran dapat tercapai (Asrul et al., 2014; Fathiyah, 2019; Hasan, n.d.)

Validitas atau kesahihan berasal dari kata validity yang berarti sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Reliabilitas berasal dari kata reliability berarti sejauh mana hasil suatu pengukuran dapat dipercaya. Reliabilitas

dibedakan menjadi dua, yaitu: Reliabilitas konsistensi tanggapan dan Reliabilitas konsistensi gabungan item. Suatu tes dapat dikatakan reliabel jika selalu memberikan hasil yang sama bila diteskan pada kelompok yang sama pada waktu atau kesempatan yang berbeda (Hikmah, 2024; Ummah, 2019; Zainal, 2012).

Analisis butir soal merupakan proses yang dilakukan untuk menyelidiki, menelititi dan mengkaji pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai. Dalam analisis butir soal, ada empat hal yang akan dianalisis yaitu tingkat kesukaran soal, daya pembeda soal, analisis pengecoh soal, dan analisis homogenitas soal. Analisis butis soal bertujuan untuk meningkatkan derajat validitas dan reliabilitas soal secara menyeluruh (Zainal, 2012).

Berdasarkan penelitian sebelumnya Penelitian (Masullah et al., 2024) terkait analisis butir soal ujian akhir semester mata pelajaran matematika SMP Negeri 6 Praya Timur yang terdiri dari validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektifitas pengecoh, dan terdapat 25 soal kategori tidak valid dan 5 soal kategori valid, reliabilitasnya termasuk kategori rendah, tingkat kesukaran terdapat 2 butir soal kategori sukar, 11 butir soal kategori sedang, dan 17 butir soal kategori mudah, sedangkan untuk daya pembeda terdapat 2 butir soal kategori dan 25 butir soal kategori lemah.

Selanjutnya (Fahmi et al., 2022) tentang analisis butir soal mata pelajaran bahasa arab pada ujian semester bagi siswa MAS Pondok Pesantren Asaalam Kab. Kampar, Riau Tahun Ajaran 2020/2021, dimana hasil penelitiannya menunjukkan bahwa hasil uji daya pembeda menunjukkan terdapat 7 soal pada kategori sangat tinggi, 9 soal pada kategori tinggi, 10 soal soal pada kategori sedang, 13 soal soal pada kategori rendah, dan 11 soal soal pada kategori sangat rendah. Hasil Analisis Tingkat Kesukaran diketahui bahwa mayoritas soal-soal berada dalam kategori mudah dengan perbandingan statistiknya 43:7, 43 soal dalam kategori mudah dan 7 soal dalam kategori sedang. hasil uji validitas menunjukkan bahwa terdapat 24 soal yang berstatus valid dan 26 soal yang berstatus tidak valid. hasil uji reliabilitas menunjukkan bahwa soal dinyatakan sangat reliabel dengan nilai Cronbach's Alpha 0,868.

Analisis butir soal jarang dilakukan oleh guru dikarenakan beberapa alasan, termasuk fakta bahwa soal-soal yang diberikan tidak sepenuhnya mencerminkan tingkat pencapaian kompetensi siswa yang sebenarnya. Oleh karena itu tujuan dari penelitian ini adalah menganalisis Kualitas Tes dan Butir Soal ujian madrasah mata pelajaran Bahasa Arab kelas

IX MTs Humairoh HNN Kampar terkait dengan 4 kriteria instrumen butir soal yaitu: validitas butir soal, reliabilitas, tingkat kesukaran butir dan daya pembeda.

METODE PENELITIAN

Dalam penelitian ini merupakan penelitian deskriptif yang menggunakan pendekatan kuantitatif dikarenakan penelitian ini bertujuan untuk mengetahui kelayakan soal ujian madrasah mata pelajaran Bahasa Arab kelas IX, yang memenuhi kriteria validitas, reliabilitas, tingkat kesukaran butir, dan daya pembeda. Teknik dalam pengumpulan data pada penelitian ini adalah dokumentasi, dilakukan agar mendapatkan data yang dibutuhkan berupa daftar nama siswa, soal ujian madrasah beserta kunci jawaban, dan kertas ujian hasil jawaban siswa kelas IX MTs Humairah HNN. Analisis data ini dilakukan dengan mendata nilai yang di dapatkan siswa dalam bentuk tabulasi data, dan selanjutnya peneliti menentukan validitas, reliabilitas, tingkat kesukaran butir, dan daya beda.

HASIL DAN PEMBAHASAN

Analisis butir soal merupakan salah satu cara yang bisa digunakan untuk mengetahui tingkat kesalahan dalam membuar butir soal pada soal ujian, jadi dengan melakukan analisis butir soal ini akan diperoleh tes yang baik, penelitian ini bertujuan untuk menganalisis butir soal ujian madrasah mata pelajaran Bahasa Arab kelas IX, yang dimana terdapat 14 siswa, analisis pada penelitian ini meliputi validitas butir soal, reliabilitas, tingkat kesukaran, dan daya pembeda dan perhitungannya dilakukan dengan bantuan software SPSS dan microsoft excel. Kemudian hasil jawaban siswa tersebut dijadikan pertimbangan untuk menentukan kualitas soal ujian madrasah mata pelajaran Bahasa Arab kelas IX.

Validitas Soal

Validitas atau kesahihan berasal dari kata validity yang berarti sejauh mana ketepatan dan kecermatan suatu tes dalam menghasilkan data atau informasi yang relevan dengan tujuan atau keputusan yang akan dibuat (Hasan, n.d.). Suatu tes atau intrumen pengukuran dikatakan memiliki validitas tinggi apabila alat tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut.

Validitas memiliki beberapa macam yaitu: ***Pertama***, validitas permukaan, merupakan validitas yang menggunakan kriteria yang sangat sederhana, karena hanya melihat dari sisi muka atau tampang dari instrumen itu sendiri. Artinya, jika suatu tes secara sepintas telah

dianggap baik untuk mengungkap fenomena yang akan diukur, maka tes tersebut sudah dapat dikatakan memenuhi syarat validitas permukaan, sehingga tidak perlu lagi adanya judgement yang mendalam. ***Kedua***, validitas isi (content validity) yaitu suatu tes untuk melihat seberapa jauh tingkat penguasaan terhadap isi suatu materi tertentu yang seharusnya dikuasi sesuai dengan tujuan pengajaran. Validitas isi sering digunakan dalam pengukuran hasil belajar. Jika dilihat dari segi kegunaannya dalam penilaian hasil belajar, validitas isi ini sering disebut juga validitas kurikuler dan validitas perumusan. ***Ketiga***, validitas konstruk (contruk validity) yaitu validitas yang mempertimbangkan seberapa jauh item tes mampu mengukur apa yang benar-benar dimaksudkan untuk diukur berdasarkan konsep, struktur, atau definisi konseptual yang telah ditetapkan. Validitas konstruk banyak dikenal dan digunakan dalam tes-tes psikologis untuk mengukur gejala perilaku yang abstrak, seperti kesetiakawanan, kematangan emosi, sikap, motivasi, minat, dan sebagainya. ***Keempat***, validitas empiris atau validitas kriteria, yaitu validitas yang ditentukan berdasarkan kriteria, baik kriteria internal maupun kriteria eksternal. Kriteria internal adalah tes atau instrument yang menggunakan validitas butir yang terdiri atas indeks pembeda dan tingkat kesukaran, sedangkan kriteria eksternal adalah hasil ukur instrument atau tes dilakukn dengan cara mengkorelasikan alat pengukur baru dengan alat pengukur yang sudah valid Reliabilitas alat ukur menunjuk kepada sejauh mana perbedaan-perbedaan skor perolehan itu mencerminkan perbedaan perbedaan atribut yang sebenarnya. Validitas ini biasanya menggunakan teknik statistik, yaitu analisis korelasi. Hal ini disebabkan validitas empiris mencari hubungan antara skor tes dengan suatu kriteria tertentu yang merupakan suatu tolok ukur di luar tes yang bersangkutan. ***Kelima*** validitas faktor. validitas faktor ini dapat diketahui dengan menghitung homoginitas skor setiap faktor dengan total skor, dan antara skor dari faktor yang satu dengan skor dari faktor yang lain. Validitas ini diperoleh berdasarkan dimensi atau indikator dari variabel yang diukur sesuai dengan apa yang terungkap dalam konstruksi teoritisnya (Hikmah, 2024; Ummah, 2019; Zainal, 2012)**.**

Uji validitas berdasarkan validitas isi, validitas ini dilakukan dengan melihat kesesuaian pada soal yang diujiankan. Soal dikatakan sesuai apabila setiap soal memenuhi kaidah penulisan yang benar. Analisis ini dilakukan dengan bantuan software SPSS dan microsoft excel. Selanjutnya validitas butir soal ini dikorelasikan dengan nilai $r\ tabel$. Jika $r\ hitung > r\ tabel$, maka butir soal valid, sebaliknya jika $r\ hitung <$ r table maka butir soal tidak valid. Nilai probibalitas 0,05 dan ditarik dari R table yakni mendapatkan nilai untuk 14

adalah 0.4973. hasil dari analisis validitas butir soal pilihan ganda didapatkan butir-butir yang termasuk dalam kategori valid dan tidak valid sebagai berikut:

Tabel 1. Hasil Validitas Butir Soal Bahasa Arab

Interpretasi	Jumlah Soal	Nomor Soal
Valid	3	8,10,19
Tidak Valid	17	1,2,3,4,5,6,7,9,11,12,13,14
		15,16,17,18,19,20

Pada tabel tersebut didapatkan 3 butir soal termasuk kategori valid sedangkan 17 butir soal lainnya termasuk kategori tidak valid dan menunjukkan bahwa tingkat kevalidan soal ujian madrasah siswa kelas IX MTs Humairah HNN Mata Pelajaran Bahasa Arab siswa kelas MTs Humairoh HNN Kampar dalam bentuk pilihan ganda ini termasuk kategori tidak valid.

Reabilitas Soal

Reliabilitas berasal dari kata reliability berarti sejauh mana hasil suatu pengukuran dapat dipercaya. Reliabilitas merupakan pengukuran suatu alat ukur sama akan memberikan hasil yang sama, digunakan oleh orang yang sama dalam waktu yang berlainan atau digunakan oleh orang yang berlainan dalam waktu yang bersamaan atau dalam waktu yang berlainan (Hikmah, 2024; Ummah, 2019).

Reliabilitas dibedakan menjadi dua, yaitu: Reliabilitas konsistensi tanggapan, yaitu responden mempersoalkan apakah tanggapan responden atau obyek ukur terhadap instrument atau tes tersebut sudah baik atau konsisten. Reliabilitas konsistensi gabungan item, yaitu kemantapan atau konsitensi item-item suatu tes (Ummah, 2019). Suatu tes dapat dikatakan reliabel jika selalu memberikan hasil yang sama bila diteskan pada kelom pok yang sama pada waktu atau kesempatan yang berbeda (Zainal, 2012).

Uji reliabilitas bertujuan untuk mengetahui tingkat konsitensi dari sebuah tes yang diujikan, untuk menentukan nilai reliabilitas suatu tes bisa menggunakan formula *Kuder-Richatdson KR-21 dan Alpa.* Hasil uji ini ada dua jenis yaitu, koefisien reliabilitas > 0,60, maka reliable atau nilai reliabilitas yang tinggi dan apabila koefisien reliabilitas < 0,60, maka tidak reliable atau nilai reliabilitas rendah (Asrul et al., 2014).

Tabel 2. Hasil Reliabilitas Butir Soal

Reliability Statistics

Cronbach's Alpha	Part 1	Value	.417
		N of Items	10[a]
	Part 2	Value	-.082[b]
		N of Items	10[c]
	Total N of Items		20
Correlation Between Forms			-.120[d]
Spearman-Brown Coefficient	Equal Length		-.273[d]
	Unequal Length		-.214[d]
Guttman Split-Half Coefficient			-.261

Diketahui dari hasil analisis data di atas bahwa dalam mencari realibilitas soal ujian madrasah mata pelajaran Bahasa Arab kelas IX MTs Humairoh HNN menggunakan metode Cronbach's Alpha, kemudian peneliti mendapatkan hasil bahwa nilai Cronbach's Alpha sebesar -,120 dan < 0,60 maka butir soal ujian madrasah mata pelajaran Bahasa Arab kelas IX MTs Humairoh HNN tidak reliable dengan jumlah butir soal (N of item) berjumlah 20.

Tingkat Kesukaran Soal

Tingkat kesukaran Tingkat kesukaran yakni ukuran yang menunjukkan derajat kesulitan soal untuk diselesaikan oleh siswa. Secara sederhana dapat kita katakan bahwa sebuah soal dikatakan sukar jika sebagian besar testi gagal menyelesaikannya begitupun sebaliknya (Hasan, n.d.). Secara sederhana menjelaskan bahwa apabila peluang siswa menjawab benar butir soal, dan jika semakin banyak siswa yang menjawab soal dengan benar maka soal tergolong mudah, sebaliknya jika semakin sedikit siswa yang menjawab soal dengan benar maka soal tergolong sukar, soal dikatakan baik adalah soal yang tidak terlalu sukar atau mudah, dikarenakan dengan soal yang terlalu mudah tidak mampu meningkatkan kemampuan siswa, sedangkan soal yang terlalu sukar menjadikan siswa putus asa mengerjakannya.

Tabel 3. Interpretasi Kesukaran Butir Soal

Nilai P	Kategori
0,00 < 0,30	Sukar
0,31 < 0,70	Sedang
0,71 < 1,00	Mudah

Tabel 4. Hasil Tingkat Kesukaran Butir Soal

Interpretasi	Jumlah Soal	Nomor Soal
Sukar	15	1,2,3,4,5,9,10,11,12,13,14,15,17,18,20
Sedang	4	6,8,16,19
Mudah	1	7

Pada tabel tersebut didapatkan 1 butir soal termasuk kategori mudah sedangkan 4 butir soal termasuk kategori sedang dan 15 butir soal termasuk kategori sukar. Perlu dikaji kembali untuk mengetahui penyebab dari soal-soal tersebut dianggap sukar oleh siswa.

Daya Pembeda Soal

Daya pembeda soal adalah kemampuan sesuatu soal untuk membedakan antara siswa bertujuan untuk mengetahui tingkat tes dalam membedakan kemampuan siswa yang memiliki kemampuan tinggi dan siswa yang berkemampuan rendah. Angka yang menunjukkan besarnya beda pembeda disebut indeks Diskriminasi, disingkat D. Seperti halnya indeks kesukaraan, indeks diskriminasi (daya pembeda) ini berkisar antara 0,00 sampai 1,00 hanya bedanya indeks kesukaraan tidak mengenal tanda negatif (Rahman & Nasryah, 2019).

Tabel 5. Daya Pembeda Butir Soal

Nilai P	Kategori
0,70 < 1,00	Sangat Baik
0,40 < 0,69	Baik
0,20 < 0,39	Cukup
0,00 < 0,19	Lemah

Tabel 6. Hasil Daya Pembeda Butir Soal

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
soal1	3.50	2.577	-.090	-.055[a]
soal2	3.50	2.577	-.090	-.055[a]
soal3	3.43	1.956	.476	-.378[a]
soal4	3.36	2.863	-.328	.098
soal5	3.43	1.956	.476	-.378[a]
soal9	3.29	3.143	-.476	.198
soal10	3.29	2.066	.212	-.255[a]
soal11	3.36	2.401	-.008	-.088[a]
soal12	3.29	3.143	-.476	.198
soal13	3.29	2.220	.094	-.164[a]
soal14	3.36	1.940	.380	-.363[a]
soal15	3.43	2.418	.019	-.102[a]
soal16	3.14	2.440	-.082	-.034[a]
soal17	3.29	2.220	.094	-.164[a]
soal18	3.50	2.731	-.261	.009
soal19	3.14	2.132	.117	-.192[a]
soal20	3.57	2.571	.000	-.087[a]

Interpretasi	**Jumlah Soal**	**Nomor Soal**
Sangat Baik	-	-
Baik	2	3,5
Cukup	2	10,14
Lemah	13	1,2,4,9,11,12,13,15,16,17,18,19,20

Pada tabel tersebut didapatkan 2 butir soal termasuk kategori baik sedangkan 2 butir soal termasuk kategori cukup dan 13 butir soal termasuk kategori lemah. soal yang mempunyai tingkat daya beda lemah disarankan untuk diganti namun masih bias digunakan dengan melakukan revisi, sedangkan untuk ketegori baik dan cukup bias digunakan lansung.

KESIMPULAN

Meninjau dari hasil analisis butir soal, peneliti menyimpulkan bahwa butir soal ujian madrasah mata pelajaran Bahasa Arab kelas IX memiliki validitas dan reliabilitas, untuk hasil soal yang valid berjumlah 3 dan yang kategori tidak valid berjumlah 17. Sedangkan untuk reliabilitas pada soal ujian madrasah bahasa Arab siswa kelas IX ini menyatakan tidak reliabel sebagai soal soal ujian madrasah mata pelajaran Bahasa Arab kelas IX. Selanjutnya adalah daya pembeda dan tingkat kesukaran. Daya beda soal bahasa Arab berkriteria sangat tinggi tidak memiliki butir soal, kategori baik memiliki 2 butir soal, kategori cukup memiliki 2 butir soal , kategori lemah memiliki 13 butir soal. Sedangkan tingkat kesukaran pada soal madrasah mata pelajaran bahasa Arab siswa kelas IX ini berkategori sukar memiliki 15 butir soal, kategori sedang memiliki 4 butir soal, dan kategori mudah memiliki 1 butir soal.

DAFTAR PUSTAKA

Asrul, Ananda, R., & Rosinta. (2014). Evaluasi Pembajalaran. In *Ciptapustaka Media*.

Asrul, Sarigih, A. H., & Mukhtar. (2022). Evaluasi Pembelajaran. In *Perdana Publishing*. Http://Repo.Iain-Tulungagung.Ac.Id/5510/5/Bab 2.Pdf

Fahmi, B., Rizqi, S., & Harmeilinda, N. E. (2022). Analisis Butir Soal Bahasa Arab Mas Pondok Pesantren Assalam Kampar Riau. *Ta'lim Al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab & Kebahasaaraban*, *6*(1), 95–105. Https://Doi.Org/10.15575/Jpba.V6i1.16193

Fathiyah, S. F. Al. (2019). Analisis Butir Soal Pelajaran Bahasa Arab Di Ma Roudlotul Ulum Pagak Malang. *Tarbiyatuna: Jurnal Pendidikan Ilmiah*, *4*(1), 76–100.

Hasan, M. (N.D.). *Evaluasi Pembelajaran : Teori Dan Praktek Tahta Media Group*.

Hikmah, A. Z. (2024). *Evaluasi Pembelajaran Bahasa Arab Berkelanjutan*. Budi Utama.

Masullah, B. D., Zuhry, L. H., & Usman, L. H. (2024). *Analisis Butir Soal Ujian Akhir Semester Mata Pelajaran Matematika Smp Negeri 6 Praya Timur*. *5*(September), 152–161.

Rahman, A. A., & Nasryah, C. E. (2019). Evaluasi Pembelajaran. In *Uwais Inspirasi Indonesia*.

Ummah, M. S. (2019). Evaluasi Pembelajaran. In *Sustainability (Switzerland)* (Vol. 11, Issue

1). Http://Scioteca.Caf.Com/Bitstream/Handle/123456789/1091/Red2017-Eng-8ene.Pdf?Sequence=12&Isallowed=Y%0ahttp://Dx.Doi.Org/10.1016/J.Regsciurbeco.2008.06.005%0ahttps://Www.Researchgate.Net/Publication/305320484_Sistem_Pembetungan_Terpusat_Strategi_Melestari

Zainal, A. (2012). *Evaluasi Pembelajaran: Prinsip, Teknik, Prosedur*.