

**ANALISIS BUTIR SOAL PELAJARAN BAHASA ARAB SUMATIF AKHIR SEMESTER
GANJIL TAHUN AJARAN 2022/2023 KELAS XI SEKOLAH MENENGAH ATAS
AL-FATTAH SIDOARJO**

Muhammad Taufiq Ismail¹, Farikh Marzuki Ammar²
1,2PBA FAI Universitas Muhammadiyah Sidoarjo
1tnsimail18@gmail.com, 2farikh@umsida.ac.id

ABSTRACT

The Arabic language learning implemented at Al-Fattah High School uses the Durusullughah textbook which does not refer to the state curriculum, but rather the book is specifically for students who are active in daily Arabic conversations. The aim of this research is to find out what the quality is like and how it is developed so that the questions that have been created by the teacher are able to meet the criteria based on the level of validity, level of reliability and level of difficulty of the Arabic language questions in the summative mid-semester odd class XI held at SMA Al-Fattah Sidoarjo. This type of research uses descriptive quantitative methods. The research location is Al-Fattah Sidoarjo High School with the data to be analyzed coming from the responses of 10 students and 10 female students in class Final semester exam questions Data analysis techniques use validity tests, reliability tests and difficulty level tests. The research results obtained contained 11 questions (44%) in the medium difficulty level, 10 questions (40%) in the easy difficulty level and there were 4 questions (16%) in the very easy difficulty level and the research results of the difficulty level were included in the group good at adjusting the difficulty of the questions to the students' abilities. Keywords: Blended Learning, Alef Education Platform, Arabic Language Learning

Keywords – Question Items, Arabic, Summative

ABSTRAK

Pembelajaran bahasa arab yang diterapkan di SMA Al-Fattah menggunakan buku ajar durusullughah yang mana tidak beracuan pada kurikulum negara, melainkan buku tersebut dikhususkan bagi pelajar yang aktif dalam percakapan bahasa arab dalam keseharian. Tujuan dari penelitian ini untuk mengetahui seperti apakah kualitas dan bagaimana pengembangannya agar butir soal yang telah dibuat oleh guru mampu memenuhi kriteria berdasarkan tingkat validitas, tingkat reliabelitas, dan tingkat kesukaran soal bahasa Arab pada sumatif tengah semester ganjil kelas XI yang diadakan di SMA Al-Fattah Sidoarjo. Jenis penelitian ini menggunakan metode kuantitatif deskriptif. Lokasi penelitian di SMA Al-Fattah Sidoarjo dengan data yang akan dianalisis bersumber dari respon jawaban siswa kelas XI yang berjumlah 10 siswa dan 10 siswi pada sumatif tes akhir semester ganjil tahun ajaran 2022/2023 dengan butir soal berjumlah 25. Teknik pengumpulan data menggunakan angket berisikan soal ujian akhir semester Teknik analisis data menggunakan uji validitas, uji reliabilitas dan uji tingkat kesukaran. Hasil penelitian yang diperoleh terdapat 11 butir soal (44%) dalam tingkat kesukaran sedang, 10 butir soal (40%) dalam tingkat kesukaran mudah dan terdapat 4 butir soal (16%) dalam tingkat kesukaran sangat mudah dan hasil penelitian

tingkat kesukaran sudah termasuk kedalam golongan baik dalam penyesuaian kesukaran soal dengan kemampuan siswa.

Kata Kunci - Butir Soal, Bahasa Arab, Sumatif

A. Pendahuluan

Beberapa hal dalam lingkup pendidikan yang saling berhubungan dan mempunyai pengaruh satu sama lain diantaranya adalah pemaparan materi pengujian materi dan evaluasi hasil belajar tiga hal tersebut adalah urutan yang harus ada pada proses pembelajaran secara formal maupun pembelajaran yang bersifat tambahan sesuai minat juga bakat, pemaparan materi yang baik adalah dimana pengajar bisa mengetahui kemampuan siswa dalam memahami penyampaian dengan jelas dan benar-benar faham,(Minsih dan D 2018) proses selanjutnya adalah pengujian materi yang bertujuan mengetahui seberapa paham siswa tentang materi yang dipaparkan oleh seorang guru, Sedangkan evaluasi menurut Eveline Siregar dan Hartini Nara merupakan proses menentukan hasil prestasi belajar dengan menggunakan ketetapan tertentu agar mencapai hasil pembelajaran yang sesuai dengan tujuan awal.(Siregar dan Widyaningrum 2015) Hal ini juga mempengaruhi praktik dan sikap guru karena membantu mereka menemukan efektivitas kurikulum, buku teks, metode pengajaran, metode dan kegiatan yang dilakukan sehingga menentukan tindakan apa yang akan dilakukan dalam proses pembelajaran selanjutnya.(2023 د.فرج المبروك)

Salah satu bentuk evaluasi pembelajaran berupa sumatif tes yang dilakukan agar guru dapat mengukur proses pencapaian siswa dalam proses pembelajaran serta merencanakan perencanaan proses pembelajaran

selanjutnya(Adi 2023) dengan waktu yang telah di tentukan yaitu selama beberapa pertemuan yang mana biasanya dilakukan dalam satu semester atau 6 bualan KBM (kegiatan belajar mengajar), juga dilakukan pada akhir tahun pelajaran pada setiap tahunnya yang bertujuan menentukan kelayakan siswa untuk menempuh jenjang selanjutnya dalam tingkatan pembelajaran, namun sumatif juga mengharuskan guru untuk memperhatikan secara detail apa yang akan di ujikan kepada siswa.(Dr. Sumardi 2020)

Selain evaluasi pembelajaran yang diberikan kepada siswa guru juga harus mengevaluasi proses pengajaran yang telah dilaksanakan yang dapat ditinjau dari sisi ketercapaian siswa(Dahlillah 2015) juga dari hasil ujian siswa yang di mana guru bisa menilai dari data lembar jawaban yang telah diisi oleh siswa siswa dalam ujian dan disini guru bisa menganalisis dari setiap butir soal yang dipaparkan dalam ujian.

Analisis butir soal adalah penelitian atau penyelidikan suatu bagian dari serangkaian yang harus ditanggapi oleh setiap peserta didik. Analisis soal berfungsi sebagai alat penilaian yang telah ditetapkan baik oleh guru maupun oleh tim.(SAWALUDDIN 2018) dalam pendapat tersebut bisa dijadikan acuan bahwasanya guru juga harus memberikan soal yang sesuai dengan apa yang diajarkan dalam pembelajaran di kelas. Selain itu analisis butir soal juga memiliki tujuan untuk menghubungkan skor dari setiap butir soal serta jawaban kepada keseluruhan soal serta jawaban.

Dengan teori pengukuran klasik yang sesuai dengan jumlah obyek yang ada dapat memudahkan penganalisa untuk mengerjakan dengan perhitungan kalkulator ataupun dengan aplikasi pengolah angka yang di gunakan pada umumnya(Fatayah Fatayah, Ika Farida Yuliana 2022)

Berdasarkan penelitian terdahulu yang dilakukan oleh Siti Fatimah 2019 dengan judul ANALISIS BUTIR SOAL PELAJARAN BAHASA ARAB DI MA ROUDLOTUL ULUM PAGAK MALANG yang mana dalam penelitian tersebut menggunakan metode kualitatif dengan sumber dokumen dari lembar jawaban siswa pada ujian dengan penjabaran berdasarkan tingkat validitas, tingkat reliabilitas, tingkat kesukaran, daya beda dan tingkat pengecoh pada setiap butir pilihan ganda, jika ditinjau dari segi validitas soal, pada penelitian tersebut didapat dari 15 soal hanya 3 soal yang termasuk kategori valid sedangkan 12 soal lainnya tidak valid dengan hasil dari analisis validasi soal berdasarkan daya pembeda. Dan dalam penelitian tersebut disimpulkan bahwa soal dalam penelitian terbilang cukup bagus untuk digunakan dalam evaluasi pembuatan soal.(Al Fathiyah 2019)

Adapun penelitian yang dilakukan oleh Fikriyah 2021 dengan judul ANALISIS BUTIR SOAL ULANGAN TENGAH SEMESTER MATA PELAJARAN BAHASA ARAB KELAS VII SEMESTER GENAP SMP MUHAMMADIYAH 1 YOGYAKARTA TAHUN AJARAN 2019/2020, dengan metode penelitian kuantitatif dengan seluruh sumber data dari jawaban siswa pada ulangan tengah semester, dengan analisis aplikasi olah data spss dan Microsoft excel dari segi kesukaran soal didapati dari 35 butir soal, 3 soal tergolong mudah, 25 soal tergolong sedang, dan 3 soal tergolong sukar, dan pada penelitian tersebut berkesimpulan bahwa butir

soal yang ada sudah baik dan layak untuk di masukkan dalam bank soal yang dapat digunakan pada evaluasi pembelajaran selanjutnya.(Fikriyah 2021)

Mengamati dari pelajaran kebahasaan dalam bahasa inggris penelitian yang di lakukan oleh Rusmawan 2018 yaitu ANALISIS BUTIR SOAL UJIAN AKHIR SEMESTER MATA PELAJARAN BAHASA INGGRIS KELAS VII SMPN 2 TEGAL SIWALAN PUTU menggunakan metode penelitian kualitatif deskriptif dalam penilaian reliabilitas yang dianalisa dengan rumus KR-20 adalah 0,88% tingkat kesulitan yang memenuhi persyaratan pengukuran sebesar 55%. Daya pembeda sebesar 0,25% yang telah memenuhi syarat pembeda sedangkan daya pengecoh sudah lebih dari 5% dan telah memenuhi kriteria pengukuran. Makin reliabilitas butir soal pada soal pilihan ganda mempunyai hasil yang cukup tinggi dari beberapa analisis yang di terapkan.(Putu Ngurah Rusmawan 2018)

Mengacu pada para peneliti sebelumnya yang telah menggunakan berbagai validasi yang di terapkan dalam penelitian mereka mulai pada peneliti pertama dengan menggunakan uji validitas isi mendapatkan tingkat validasi yang tinggi, pada peneliti kedua menggunakan rumus KR – 20 dalam menentukan reliabilitas, dan peneliti ketiga memiliki keunggulan pada validitas kesukaran butir soal.

Pembelajaran bahasa Arab di SMA Al-Fattah yang menggunakan buku ajar Durusullughah menimbulkan beberapa pertanyaan penting. pertama, penggunaan buku ajar yang tidak mengikuti kurikulum negara dapat mempengaruhi motivasi belajar siswa, baik secara positif maupun negatif. kedua, efektivitas metode pembelajaran yang diterapkan dalam meningkatkan-

kemampuan berbahasa Arab siswa menjadi pertanyaan, terutama dalam konteks pembelajaran didalam kelas dan aspek penerapan bahasa. Ketiga, penyusunan soal yang dapat menentukan perbedaan antara siswa

yang kurang mampun dan siswa yang mudah memahami pelajaran bahasa Arab memerlukan pendekatan khusus.

Terakhir, sistem evaluasi dan penilaian harus dirancang sedemikian rupa sehingga dapat mengukur kemampuan bahasa Arab siswa secara adil dan akurat,(Ahmad 2024) mempertimbangkan perbedaan dalam keterlibatan mereka dengan bahasa tersebut. Pertanyaan- pertanyaan ini harus dijawab untuk mengevaluasi dan mengembangkan strategi pembelajaran bahasa Arab yang efektif di sekolah tersebut.

Berdasarkan fenomena yang terdapat pada pembelajaran tersebut dapat dirumuskan masalah, diantaranya: 1). Seberapa valid soal yang di berikan ke siswa dalam ujian sumatif tes ?. 2). Seberapa tingkat reliabilitas yang terdapat dalam butir soal yang di ujikan ?. 3). Seberapa tinggi tingkat kesukaran butir soal yang ada. Maka disini peneliti bertujuan menerapkan dari setiap alat yang di gunakan dari para peneliti tersebut untuk menerapkannya di SMA AL-Fattah Sidorajo guna mengetahui seperti apakah kualitas dan bagaimana pengembangannya agar butir soal yang telah dibuat oleh guru mampu memenuhi kriteria berdasarkan tingkat validitas, tingkat reliabelitas, dan tingkat kesukaran soal bahasa Arab pada sumatif tengah semester genap kelas XI yang diadakan di SMA Al-Fattah Sidoarjo.

B. Metode Penelitian

Penelitian ini dilakukan menggunakan metode kuantitatif deskriptif. Lokasi penelitian di SMA Al-Fattah Sidoarjo dengan data yang akan dianalisis bersumber dari respon jawaban siswa kelas XI yang berjumlah 10 siswa dan 10 siswi pada sumatif tes akhir semester genap tahunajaran 2022/2023 dengan butir soal berjumlah 25 dengan alternatif lima pilihan jawaban yaitu a, b, c, d, dan e, dan 5 soal berbentuk esai. Yang kemudian dilakukan analisis dengan mengukur tingkat validitas soal, reliabilitas, dan tingkat kesukaran dari setiap butir soal yang tersusun.

Pengujian validasi butir soal peneliti menggunakan validitas isi yang mana tujuannya adalah untuk memahami apa yang terjadi pada siswa setelah mereka menyelesaikan proses pembelajaran saat ini dan perubahan psikologis apa saja yang terjadi pada siswa setelah mereka selesai.(Febrianawati Yusup 2018) validitas isi juga memiliki posisi yang cukup penting dalam menentukan pengujian konstruk dengan begitu jika didapati dalam pengujian ini terbukti valid dan sesuai maka tidak perlu lagi diadakan pengujian selanjutnya yaitu dengan uji konstruk atau uji validitas yang lainnya,(Zaenal 2016) karenanya uji ini bertumpu pada validitas rasional dan logis sementara untuk pengujian yang lain berfokus pada validitas empiris.

Untuk pengujian reliabelitas peneliti menggunakan rumus KR-20 yang mana rumus tersebut banyak di gunakan untuk mengukur konsistensi suatu keadaan yang mempunyai pengaruh pada uji sebelumnya dengan hasil yang relatif sama.(Musfiqon 2012)

KR-20

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{s^2 - \sum pq}{s^2} \right)$$

Ketrangan:

r11 = Reliabilitas tes secara keseluruhan.

p = Proporsi subyek yang menjawab item dengan benar.

q = Proporsi subyek yang menjawab item dengan salah.

n = Banyak item.

S2 = Standar deviasi dari tes (standar deviasi adalah akar varians).

Σpq = Jumlah hasil perkalian antara p dan q.

Sedangkan untuk mengukur kesukaran soal peneliti menggunakan teknik yang meninjau dari kesukaran butir tes dan kesukaran alat tes dengan menggunakan rumus sebagai berikut: Tingkat kesukaran butir tes (IK) Tingkat kesukaran perangkat tes (IKP)

$$IK = \frac{nB}{n}$$

IK = indeks kesukaran butir tes

nB = banyaknya testee yang menjawab benar

n = banyaknya testee seluruhnya

$$IKP = \frac{\sum IK}{N}$$

IKP = Indeks kesukaran perangkat tes

IK = Indeks kesukaran tiap butir

N = Banyaknya butir tes

pengukuran daya tingkat kesukaran butir soal dituliskan dengan angka indeks kesukaran (difficulty indexs). Indeks kesukaran berada pada rentang nilai 0,00 sampai dengan 1,00. Butir soal dengan indeks 0,00 dengan artian soal tersebut terlalu sukar, sebaliknya jika indeks menunjukkan 1,00 berarti soal sangat mudah. (I Wayan Eka Mahendra 2019) Pengujian kesukaran dengan rumus tersebut ditinjau lebih jelas dengan jumlah soal dan obyek yang ada dalam penentuan indeks kesukaran setiap butir soal, juga menentukan indeks kesukaran perangkat tes yang di ujikan. (Sujiono 2016)

C. Hasil Dan Pembahasan

Hasil data yang diperoleh peneliti berupa soal dan jawaban Akhir Semester ganjil Pelajaran Bahasa Arab Sumatif kelas XI Sekolah Menengah Atas Al-Fattah Sidoarjo tahun 2022/2023.

Tabel 1. Data nilai peserta didik dikelas XI

Jawaban Benar	Nilai	Banyak Siswa
7	28	1
9	36	1
10	40	1
14	56	1
16	64	2
17	68	1
18	72	2
19	76	1
20	80	2
21	84	3
22	88	2
23	92	2
24	96	1
Total		20

Berdasarkan tabel diatas menjelaskan nilai minimum atau paling rendah diperoleh oleh 1 siswa dengan nilai 28, sementara nilai maksimum atau tertinggi diperoleh 1 orang siswa dengan nilai 96. Nilai rata-rata dari keseluruhan nilai yang diperoleh siswa dalam Pelajaran Bahasa Arab Sumatif sebesar 44. Median atau nilai Tengah dari hasil ujian akhir semester adalah 74. Modus atau nilai yang sering muncul diperoleh oleh 3 siswa dengan nilai 84, sementara jumlah siswa yang mendapatkan nilai lebih besar rata-rata terdapat 17 siswa.

1. Uji Validitas

validitas tujuannya adalah untuk memahami apa yang terjadi pada siswa setelah mereka menyelesaikan proses pembelajaran saat ini dan perubahan

Butir Soal	r _{hitung}		r _{tabel}	Keterangan
1	0,286	<	0,3783	Tidak Valid
2	0,242	<	0,3783	Tidak Valid
3	0,138	<	0,3783	Tidak Valid

psikologis apa saja yang terjadi pada siswa setelah mereka selesai. Dalam pengujian ini data dapat dinyatakan valid apabila nilai $r_{hitung} > r_{tabel}$ dan dinyatakan tidak valid jika $r_{hitung} < r_{tabel}$. Nilai r_{tabel} pada penelitian ini sebesar 0,3783 dengan taraf signifikansi 0,05. (Widiyanto n.d.) Pengujian Validitas ini dibantu dengan software SPSS 22. Berikut hasil pengujian validitas nilai Tengah dari hasil ujian akhir semester Pelajaran Bahasa Arab Sumatif :

4	0,747	>	0,3783	Valid
5	0,542	>	0,3783	Valid
6	0,329	<	0,3783	Tidak Valid
7	0,498	>	0,3783	Valid
8	0,521	>	0,3783	Valid
9	0,545	>	0,3783	Valid
10	0,371	<	0,3783	Tidak Valid
11	0,377	<	0,3783	Tidak Valid
12	0,608	<	0,3783	Tidak Valid
13	0,398	>	0,3783	Valid
14	0,679	>	0,3783	Valid
15	0,440	>	0,3783	Valid
16	0,817	>	0,3783	Valid
17	0,618	>	0,3783	Valid
18	0,030	<	0,3783	Tidak Valid
19	0,267	<	0,3783	Tidak Valid
20	0,520	>	0,3783	Valid
21	0,694	>	0,3783	Valid
22	0	<	0,3783	Tidak Valid
23	0,687	>	0,3783	Valid
24	0,537	>	0,3783	Valid
25	0,251	<	0,3783	Tidak Valid

Berdasarkan tabel diatas menunjukkan perhitungan Software SPSS 22 menghasilkan 14 Butir soal atau 56% dinyatakan valid, sedangkan 11 butir soal atau 44% dinyatakan tidak valid. Berikut butir soal yang dinyatakan valid dan tidak valid dalam software SPSS 22 meliputi soal valid pada butir soal 4,5,7,8,9,13,14,15,16,17,20,21,23 dan 24, sementara soal tidak valid pada butir soal 1,2,3,6,10,11,12,18,19,22 dan 25

2. Uji Reliabilitas

Uji reliabilitas merupakan indeks yang digunakan dalam menunjukkan sejauh mana suatu alat pengukuran dapat dipercaya atau diandalkan, sehingga uji ini dapat digunakan dalam mengetahui konsistensi alat ukur apakah alat ukur tetap konsisten jika dilakukan

Reliability Statistics	
Cronbach's Alpha	N of Items
.861	14

pengukuran ulang, dan dinyatakan reliabel apabila tidak ada perubahan hasil meskipun diukur berkali-kali. (PURNOMO 2016) Terdapat banyak metode uji reabilitas

namun pada penelitian ini menggunakan rumus KR-20 atau metode *Cronbach's Alpha* menggunakan SPSS 22. Dalam pengukuran pengujian ini di nyatakan reliabel apabila $r_{hitung} > r_{tabel}$ pada taraf signifikansi 5% atau 0,05 (Pengaruh et al. 2024). Berikut hasil pengujian reliabilitas dari hasil ujian akhir semester Pelajaran Bahasa Arab Sumatif:

Tabel 3. Hasil Uji Reliabilitas

Pada uji reliabilitas ini nilai r_{tabel} dengan 20 sampel sebesar 0,444, pada tabel diatas *N of Items* menunjukkan nilai 14 yang artinya data yang dimasukkan sebanyak 14 yang merupakan soal valid meliputi butir soal 4,5,7,8,9,13,14,15,16,17,20,21,23 dan 24. Sementara nilai r_{hitung} terletak pada tabel *Cronbach's Alpha* dengan nilai sebesar 0,861, sehingga dapat disimpulkan $r_{hitung} > r_{tabel}$ ($0,861 > 0,444$)

Butir Soal	Indeks	Tingkat Kesukaran
1	0,90	Sangat Mudah
2	0,85	Mudah
3	0,80	Mudah
4	0,80	Mudah
5	0,55	Sedang
6	0,55	Sedang
7	0,50	Sedang
8	0,55	Sedang
9	0,75	Mudah
10	0,50	Sedang
11	0,70	Sedang
12	0,65	Sedang
13	0,75	Mudah
14	0,60	Sedang
15	0,95	Sangat Mudah
16	0,70	Sedang
17	0,75	Mudah
18	0,65	Sedang
19	0,45	Sedang
20	0,75	Mudah
21	0,80	Mudah
22	1,00	Sangat Mudah
23	0,85	Mudah
24	0,95	Sangat Mudah
25	0,75	Mudah

maka data dinyatakan reliabel atau dapat dipercaya dan konsisten.

3. Analisi Tingkat Kesukaran

Tingkat kesukaran soal ialah ukuran guna menentukan seberapa sulit atau mudah suatu soal bagi peserta ujian. Kesukaran soal dapat diukur dengan menggunakan indeks kesukaran soal (Difficulty index) dengan membagi banyaknya sampel yang menjawab butir soal dengan benar dan jumlah seluruh sampel. Pada pengujian menggunakan Software SPSS dapat membandingkan mean pada output data pada SPSS dengan indek tingkat kesukaran.

Indek kesukaran sendiri meliputi sangat sukar (0,00-0,15), sukar (0,16-0,30), sedang (0,31-0,70), mudah (0,71-0,85) dan sangat mudah (0,86-1,00). (Fatayah Fatayah, Ika Farida Yuliana 2022) Berikut hasil pengujian dari hasil ujian akhir semester Pelajaran Bahasa Arab Sumatif menggunakan SPSS 22

Tabel 3. Tingkat Kesukaran

Berdasarkan tabel diatas hasil perhitungan terhadap 25 butir soal ujian akhir semester terdapat 11 butir soal (44%) dalam tingkat kesukaran sedang, 10 butir soal (40%) dalam tingkat kesukaran mudah dan terdapat 4 butir soal (16%) dalam tingkat kesukaran sangat mudah. Sementara butir soal yang dalam kategori sedang terdapat pada nomor 5,6,7,8,10,11,12,14,16,18 dan 19. Butir soal yang dalam kategori mudah terdapat pada nomor 2,3,4,9,13,20,21,23 dan 25. Butir soal yang dalam kategori sangat mudah terdapat pada nomor 1,15,22 dan 24. Maka dari hasil perhitungan rumus ini memiliki pendekatan dengan mencari letak salah pada jawaban reponden, sehingga semakin responden menjawab salah maka soal semakin terlihat tingkat kesukaran butir soal semakin sukar.

Dapat dinyatakan dari pengujian data diatas bahwa butir soal ujian akhir semester Pelajaran Bahasa Arab Sumatif yang berbentuk pilihan ganda mempunyai kualitas soal yang baik dimana terdapat keseimbangan dari kesulitan soal tersebut. Pada dasarnya tingkat kesukaran soal bisa dilihat melalalui pemahaman siswa bukan dari guru, tingkat kesukaran soal dapat menggunakan pola 3-5-2 yang artinya 30% soal berkategori mudah, 50% soal berkategori sedang dan 20% berkategori sukar. Jika tingkat kesukaran tersebut dibandingkan dengan hasil penelitian tingkat kesukaran sudah termasuk kedalam golongan baik dalam penyesuaian kesukaran soal dengan kemampuan siswa.

D. Kesimpulan

Berdasarkan hasil penelitian dan pembahasan dapat diperoleh

kesimpulan yaitu terdapat 14 (56%) butir soal yang valid dan 11 (44%) butir soal yang tidak valid. Pada pengujian reliabilitas termasuk kedalam golongan baik dilihat dari hasil perhitungan menggunakan SPSS. terdapat 11 butir soal (44%) dalam tingkat kesukaran sedang, 10 butir soal (40%) dalam tingkat kesukaran mudah dan terdapat 4 butir soal (16%) dalam tingkat kesukaran sangat mudah dan hasil penelitian tingkat kesukaran sudah termasuk kedalam golongan baik dalam penyesuaian kesukaran soal dengan kemampuan siswa.

DAFTAR PUSTAKA

- Adi, Guritno. 2023. "Penilaian Sumatif: Pengertian, Fungsi, Tujuan, Bentuk dan Contohnya - Esai Edukasi." Diambil 18 Desember 2023 (<https://www.esaiedukasi.com/2023/06/pengertian-fungsi-contoh-dan-bentuk-penilaian-sumatif.html>).
- Ahmad, Muhlis. 2024. "ANALISIS KESESUAIAN MATERI BUKU TEKS UTAMA BAHASA INDONESIA SMA/SMK KELAS X DENGAN CAPAIAN PEMBELAJARAN (CP) KURIKULUM MERDEKA." Universitas Jambi.
- Dahlillah, DR. Ayu. 2015. *التعلم ونظرياته*. Lebanon: Dar Al Kotob Al Alamiyah.
- Dr. Sumardi, M. Hum. 2020. *Teknik Pengukuran Dan Penilaian Hasil Belajar*. 1 ed. Sleman Jogja: CV Budi Utama.
- Fatayah Fatayah, Ika Farida Yuliana, Layla Mufidah. 2022. "Validity and Reliability Analysis in Supporting Mastery Learning STEM Model." *Jurna Buana Pendidikan* 18(1):60.
- Al Fathiyah, S. F. 2019. "ANALISIS BUTIR SOAL PELAJARAN BAHASA ARAB DI MA ROUDLOTUL ULUM PAGAK MALANG." *Jurnal Pendidikan Ilmiah* 4(1):76–100.
- Febrianawati Yusup. 2018. "Uji Validitas dan Reliabilitas Instrumen Penelitian Kuantitatif." *Tarbiyah : Jurnal Ilmiah Kependidikan* 7(1):17–23. doi: 10.18592/TARBIYAH.V7I1.2100.
- Fikriyah, Nurul. 2021. "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(2):128–40. doi: 10.18196/MHT.V3I2.10501.
- I Wayan Eka Mahendra. 2019. "Analisis Butir Soal." *Journal of Chemical Information and Modeling* 53(9):1689–99.
- Minsih, Minsih, dan Aninda Galih D. 2018. "PERAN GURU DALAM PENGELOLAAN KELAS." *Profesi Pendidikan Dasar* 5(1):20–27. doi: 10.23917/PPD.V1I1.6144.
- Musfiqon, M. 2012. *Metode Penelitian Pendidikan*.
- Pengaruh, Analisis, Mata Kuliah, Statistik Untuk, Penelitian Tugas, Akhir Mahasiswa, Rini Ramadhani, Rizka Fehriza Siregar, dan Hetty Elfina. 2024. "Analisis Pengaruh Mata Kuliah Statistik Untuk Penelitian Tugas Akhir Mahasiswa." *Journal on Education* 6(2):14790–96. doi: 10.31004/JOE.V6I2.5293.
- PURNOMO, ROCHMAT ALDY. 2016. "ANALISIS STATISTIK EKONOMI DAN BISNIS DENGAN SPSS (UNTUK MAHASISWA, DOSEN DAN PRAKTISI)" diedit oleh P. C. Ambarwati. *Cv. Wade Group* 1–177.
- Putu Ngurah Rusmawan. 2018. "Analisis Butir Soal Ujian Akhir Semester Mata Pelajaran Bahasa Inggris Kelas VII SMPN 2 Tegalsiwalan." *Linguista: Jurnal Ilmiah Bahasa,*

- Sastra, dan Pembelajarannya
2(1):39–46. doi:
10.25273/LINGUISTA.V2I1.2686.
- SAWALUDDIN. 2018. "Konsep Evaluasi
Dalam Pembelajaran Pendidikan
Islam." *Jurnal Pendidikan Agama
Islam Al-Thariqah* 3(1):39–52. doi:
10.25299/ALTHARIQAH.2018.VO
L3(1).1775.
- Siregar, Eveline, dan Reto
Widyaningrum. 2015. *Belajar Dan
Pembelajaran*. Vol. 09. 3 ed. Bogor:
Penerbit Ghalia Indonesia.
- Sujiono, Anas. 2016. *Pengantar
Evaluasi Pendidikan*. Jakarta:
Rajawali press.
- Widiyanto. n.d. "SPSS For Windows
Untuk Analisis Data Statistik... -
Google Scholar." Diambil 7
Desember 2023
(https://scholar.google.co.id/scholar?hl=id&as_sdt=0,5&cluster=54530954216254599).
- Zaenal, Arifin. 2016. *Evaluasi
pembelajaran*. Vol. 2. 2 ed. diedit
oleh W. B. Shidiq. Jakarta pusat:
Direktorat Jendral Pendidikan
Kementrian Agama 2012.
- د. فرج المبروك. 2023. *التقويم و القياس التربوي الحديث*.
The Anglo Egyptian Bookshop.