



ANALISIS BUTIR SOAL BAHASA ARAB SISWA MAS PONDOK PESANTREN ASAALAM KAMPAR RIAU

Bahrudin Fahmi¹, Syahrul Rizqi², Nurul Elmira H³

^{1,2,3} UIN Maulana Malik Ibrahim Malang, Indonesia

Corresponding E-mail: syahrulrizqi@uin-malang.ac.id

ABSTRACT

This study aims to analyze the questions of Arabic subjects in the semester exam for students of MAS Pondok Pesantren Asaalam Kab. Kampar, Riau Academic Year 2020/2021. After doing research using descriptive quantitative methods, the researchers obtained research results that (1) the results of the distinguishing power test showed that there were 7 questions in the very high category, 9 questions in the high category, 10 questions in the medium category, 13 questions in the low category, and 11 questions in the very low category. (2) the results of the difficulty level analysis show that the majority of the questions are in the easy category with a statistical comparison of 43:7, 43 questions are in the easy category and 7 questions are in the medium category. (3) the results of the validity test show that there are 24 questions with valid status and 26 questions with invalid status. (3) the results of the reliability test showed that the questions were very reliable with a Cronbach's Alpha value of 0.868.

Keywords: Analysis Question Item, Arabic Language, MAS PP Asaalam.

ABSTRAK

Penelitian ini bertujuan untuk menganalisis butir soal mata pelajaran bahasa arab pada ujian semester bagi siswa MAS Pondok Pesantren Asaalam Kab. Kampar, Riau Tahun Ajaran 2020/2021. Setelah dilakukan penelitian dengan menggunakan metode kuantitatif deskriptif, peneliti memperoleh hasil penelitian bahwa (1) hasil uji daya pembeda menunjukkan terdapat 7 soal pada kategori sangat tinggi, 9 soal pada kategori tinggi, 10 soal pada kategori sedang, 13 soal pada kategori rendah, dan 11 soal pada kategori sangat rendah. (2) 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. (3) hasil uji validitas menunjukkan bahwa terdapat 24 soal yang berstatus valid dan 26 soal yang berstatus tidak valid. (3) hasil uji reliabilitas menunjukkan bahwa soal dinyatakan sangat reliabel dengan nilai *Cronbach's Alpha* 0,868.

Kata Kunci: Analisis Butir Soal, Bahasa Arab, MAS PP Asaalam.

PENDAHULUAN

Pelaksanaan pembelajaran merupakan suatu proses yang memiliki tujuan yang hendak dicapai oleh guru dan murid. Dengan adanya tujuan, maka memudahkan guru untuk menetapkan arah proses pembelajaran agar dapat membantu dalam pencapaian tujuan pembelajaran. Dengan berbagai usaha yang dilakukan oleh guru, diharapkan mampu mengarahkan siswa untuk mencapai tujuan pembelajaran, baik penggunaan strategi atau metode yang bermacam-macam, media pembelajaran yang bermacam-macam, bahan ajar yang bervariasi, dan berbagai cara lainnya.

Berbagai usaha guru yang telah dilakukan dalam pembelajaran, maka kita perlu melihat apakah pelaksanaan pembelajaran yang telah dilakukan sudah mencapai tujuan atau belum. Oleh karena itu, perlu diadakan evaluasi yang dapat melihat sejauh mana kemampuan siswa dalam belajar dan sejauh mana tujuan pembelajaran telah tercapai. (Maulana, 2020, hal. 13) Maka, pada tahap ini butuh penilaian hasil belajar terhadap kemampuan siswa yang biasanya dilaksanakan pada ujian-ujian sekolah, seperti ujian tengah semester, ujian semester, ujian akhir madrasah berstandar nasional (UAMBN), dan ujian-ujian lainnya. Penilaian hasil belajar itu dengan pemberian tes kepada siswa berupa soal-soal yang berkaitan dengan materi yang telah mereka pelajari. (Fathimah, 2019, hal. 143)

Di MAS Pondok Pesantren As-salam, soal-soal ujian yang digunakan dalam evaluasi pembelajaran bahasa arab dibuat oleh pemerintah dengan melibatkan KKG/MGMP yang ada di daerah tersebut. Soal-soal ujian tidak dibuat sendiri oleh guru-guru yang ada di MAS Pondok Pesantren As-salam atau oleh guru-guru dari sekolah lain yang juga terlibat dalam KKG/MGMP. Maka soal-soal ujian terlebih dahulu dilakukan dengan diskusi guru-guru yang terlibat dalam KKG/MGMP kemudian menghasilkan satu bentuk tes yang sudah siap pakai untuk penilaian hasil belajar siswa. cara seperti ini merupakan, cara yang sudah ditempuh oleh beberapa sekolah di daerah tersebut yang terlibat dalam KKG/MGMP.

Walaupun soal-soal ujian tersebut merupakan tanggung jawab KKG/MGMP dalam pembuatannya, tetap saja harus memperhatikan syarat-syarat tes yang baik. Artinya, soal-soal tersebut harus memiliki kualitas yang baik dan memenuhi kriteria dalam mengukur kemampuan siswa. Tes yang baik, juga akan mempengaruhi bagaimana kualitas siswa dalam pembelajaran. Oleh karena itu, soal-soal yang telah dibuat perlu adanya analisis untuk menentukan kualitas dari soal tersebut dan apakah sudah memenuhi kriteria soal-soal yang baik.

Menurut Daryanto dan Ishak, analisis butir soal merupakan kegiatan mengkaji kualitas pada setiap butir soal. Hal bertujuan untuk mengidentifikasi soal yang baik, kurang baik, dan jelek. Dengan mengetahui kualitas tiap soal, maka memberikan isyarat untuk dilakukannya perbaikan pada soal-soal yang belum memenuhi kriteria agar soal-soal tersebut memiliki mutu yang baik. (Rahmi, 2020,

hal. 46) Dengan analisis butir soal, dapat diketahui soal-soal yang mana saja yang banyak dijawab benar oleh siswa dan begitu pula sebaliknya.

Dalam analisis butir soal, ada empat hal yang akan dianalisis yaitu, *pertama*, analisis tingkat kesukaran soal. Dalam hal ini, soal akan dianalisis berdasarkan tingkat kesulitan suatu soal. (Bagiyono, 2017, hal. 3) Sehingga akan diperoleh soal-soal yang masuk pada kelompok mudah, sedang, dan sulit. *Kedua*, Analisis daya pembeda soal. Analisis ini merupakan analisis terhadap soal-soal dengan melihat kemampuan soal dalam mengelompokkan siswa yang memiliki kemampuan tinggi dan kemampuan rendah.

Dengan analisis ini, dapat diketahui dan dikelompokkan siswa-siswa yang pintar dan siswa yang kurang mampu dalam pembelajaran. Jika soal tidak mampu membedakan kedua kemampuan siswa, maka ada kemungkinan soal tersebut tidak memenuhi kriteria soal yang baik. (Solichin, 2019) *Ketiga*, analisis validitas soal. Tes yang baik adalah tes yang mampu mengukur apa yang seharusnya diukur. Validitas suatu soal dilihat sejauh mana keakuratan suatu tes dapat berfungsi sebagai pengukuran. (Otomo, 2018, hal. 147) *Keempat*, Analisis reliabilitas soal. Reliabilitas soal berkaitan dengan masalah kepercayaan. Reliabilitas menggambarkan apakah suatu tes dapat mengukur secara konsisten. Artinya, ketika soal-soal tersebut diujikan kembali pada subjek dan kondisi yang sama akan memberikan hasil yang sama (konsisten).

METODE

Penelitian ini menggunakan metode kuantitatif deskriptif. Menurut (Arikunto S. , 2010, hal. 143) tujuan penelitian deskriptif yaitu mencari informasi tentang peristiwa yang ada, yaitu keadaan yang bersifat apa adanya pada saat penelitian dilakukan. Sedangkan penelitian kuantitatif yaitu penelitian untuk mencari informasi menggunakan data berupa angka sebagai suatu alat untuk mengetahui atau menemukan sebuah keterangan (Margono, 2010, hal. 106).

Data dari penelitian ini berupa skor mentah siswa yang kemudian datanya diolah dengan mencari tingkat kesukaran soal, daya beda soal, validitas soal, dan reliabilitas soal untuk mengetahui kualitas soal di MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau.

Populasi penelitian ini berupa seluruh siswa kelas 11 MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau. Peneliti mengambil sampel berupa 1 kelas dari keseluruhan kelas 11 di sekolah MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau. Teknik pengumpulan data peneliti menggunakan teknik dokumentasi berupa lembar soal dan lembar jawaban siswa kemudian di analisis menggunakan aplikasi Microsoft Excel dan IBM SPSS STATISTIK 25 untuk mencari hasil statistik dari penelitian tersebut.

DATA DAN PEMBAHASAN

Analisis butir soal dilakukan untuk mengidentifikasi mana butir-butir soal yang baik dan yang buruk. Butir soal yang diambil oleh peneliti adalah soal ujian akhir semester kelas 11 MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau. Soal tersebut berupa pilihan ganda berjumlah 50 butir soal dan 5 (lima) opsi jawaban pada soal tersebut. Dari butir soal tersebut akan diperoleh informasi yang dapat digunakan untuk menghasilkan soal-soal yang lebih berkualitas, sehingga dapat diperoleh informasi yang baik terhadap peserta tes.

Penelitian ini menggunakan analisis kuantitatif dengan bantuan 2 aplikasi komputer yaitu aplikasi IBM SPSS STATISTIK 25 dan Microsoft Excel. Populasi dalam penelitian ini merupakan seluruh siswa kelas 11 MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau dan sampel yang diambil berupa 1 kelas yaitu kelas 11 sebanyak 32 siswa dari seluruh siswa kelas 11 MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau. Data pada penelitian ini berupa soal-soal Ujian Akhir semester Mata Pelajaran bahasa arab siswa dan lembar jawaban ujian Akhir semester Mata Pelajaran bahasa arab siswa kelas 11 MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau.

Setelah mengumpulkan data-data tersebut peneliti melakukan analisis butir soal Ujian Akhir Semester Mata Pelajaran bahasa arab siswa kelas 11 MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau peneliti mendapatkan hasil pendiskripsian data dari validitas soal, reliabilitas soal, daya pembeda soal dan tingkat kesukaran soal Ujian Akhir Semester Bahasa Arab siswa kelas 11 MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau. Berikut merupakan hasil statistik yang dihasilkan oleh peneliti terkait analisis butir soal menggunakan bantuan 2 aplikasi komputer yaitu aplikasi Microsoft Excel dan aplikasi IBM SPSS STATISTIK 25:

1. VALIDITAS

Validitas merupakan kecermatan tes dalam menjalankan fungsi pengukurannya. Suatu tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur. Menurut Arikunto validitas sebuah tes dapat diketahui dari hasil pengalaman. Validitas tes tersebut dibagi menjadi dua kelompok, yaitu validitas logis dan validitas empiris. Validitas logis meliputi validitas isi (*content validity*) dan validitas konstruk (*construct validity*), sedangkan validitas empiris meliputi validitas “ada sekarang” atau konkruen (*concurrent validity*) dan validitas prediksi (*predictive validity*). (Arikunto S. , 2008, hal. 68-69)

Menurut Sukardi bahwa suatu tes valid apabila koefisien 0,5 dapat diterima, jika hanya satu-satunya, sebaliknya jika ternyata ada tes prediksi lain yang sejenis dan mempunyai koefisien lebih tinggi maka koefisien 0,5 tidak diterima. (Sukardi, 2008, hal. 38) Maka suatu tes valid minimal koefiennya 0,5. Menurut Azwar butir soal dikatakan valid apabila suatu koefisien validitas dianggap

memuaskan apabila koefisien diperoleh berkisar antara 0,30 sampai 0,50. (Azwar, 2008, hal. 158)

Pada penelitian ini berdasarkan analisis data menggunakan aplikasi *Microsoft Excel*, dengan menggunakan R Tabel sebagai alat uji validitas dan berfungsi untuk menguji validitas suatu instrumen yang digunakan, pada penelitian ini lembar jawaban dan butir soal ujian akhir semester bahasa Arab siswa kelas 11 MAS PONDOK PESANTREN ASAALAM menjadi instrumen untuk pengumpulan data. Peneliti mendapatkan hasil bahwa terdapat 24 jumlah butir soal yang berstatus “valid” dan 26 jumlah butir soal lainnya berstatus “tidak valid” yang diambil dari nilai probabilitas 0.05, dan ditarik dari R Tabel yakni mendapatkan nilai 0,361. Berikut data validitas butir soal berbentuk tabel:

Tabel 5

Hasil Validitas Butir Soal Bahasa Arab

VALIDITAS		
Status	Nomor Soal	Jumlah
Valid	2,3,4,5,15,23,26,27,28,31,32,33,34,35,36,38,39,40,44,45,46,47,48,50	24
Tidak Valid	1,6,7,8,9,10,11,12,13,14,16,17,18,19,20,21,22,24,25,29,30,37,41,42,43,49	26

Terdapat hasil statistik yang menunjukkan bahwa terdapat 24 butir soal yang berstatus Valid yakni: No.2, No.3, No.4, No.5, No.10, No.23, No.26, No.27, No.28, No.31, No.32, No.33, No.34, No.35, No.36, No.38, No.39, No.40, No.44, No.45, No.46, No.47, No.48, No.50.

Kemudian terdapat 26 butir soal yang berstatus Tidak Valid yakni : No.1, No.6, No.7, No.8, No.9, No.10, No.11, No.12, No.13, No.14, No.16, No.17, No.18, No.19, No.20, No.21, No.22, No.24, No.25, No.29, No.30, No.37, No.41, No.42, No.43, No.49.

Suatu tes mempunyai validitas tinggi apabila tes tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dan tujuan diadakannya tes tersebut. Sebaliknya, suatu tes yang menghasilkan data tidak relevan dengan tujuan pengukuran dikatakan sebagai tes yang memiliki validitas rendah.

2. RELIABILITAS

Kata reliabilitas diambil dari bahasa Inggris yaitu *reliability*, dari kata *reliable* yang artinya dapat dipercaya. Apabila suatu alat pengukur itu menghasilkan data yang ajeg, maka alat pengukur tersebut bisa disebut *reliable* atau dapat dipercaya. (Ainin, 2006, hal. 31)

Reliabilitas suatu tes merupakan konsistensi dari suatu tes dalam mengukur apa yang seharusnya diukur sehingga pengukuran itu memberikan

informasi yang dapat dipercaya. Hasil pengukuran dikatakan mempunyai reliabilitas yang tinggi apabila pengukuran pertama dan kedua menunjukkan hasil yang hampir sama, sebaliknya apabila pengukuran pertama dan kedua jauh berbeda maka reliabilitas suatu tes mempunyai reliabilitas rendah. Tujuan utama mengestimasi reliabilitas adalah untuk menentukan seberapa besar variabilitas yang terjadi karena adanya kesalahan pengukuran dan seberapa besar variabilitas tes yang sebenarnya.

Untuk mengukur apakah butir soal bahasa Arab kelas 11 MAS Pondok Pesantren Asaalam reliabel atau tidak maka peneliti menggunakan aplikasi IBM SPSS STATISTIK 25 untuk menganalisis reliabilitas soal tersebut. Peneliti menggunakan metode *Cronbach's Alpha* dengan dasar apabila nilai *Cronbach's Alpha* > 0,60 maka butir soal konsisten dan reliabel dan apabila nilainya < 0,60 maka butir soal bahasa Arab siswa kelas 11 tidak reliabel. Berikut merupakan hasil analisis data yang diperoleh peneliti menggunakan aplikasi IBM SPSS STATISTIK 25:

Tabel 6
Hasil Analisis Reliabilitas Soal Bahasa Arab

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.868	.828	50

Diketahui dari hasil analisis data di atas bahwa dalam mencari realibilitas soal bahasa Arab kelas 11 MAS Pondok Pesantren Asaalam menggunakan metode *Cronbach's Alpha*, kemudian peneliti mendapatkan hasil bahwa nilai *Cronbach's Alpha* sebesar 0,868 dan > 0,60 maka butir soal bahasa Arab siswa kelas 11 MAS Pondok Pesantren Asaalam reliabel dan nilai item pada butir soal ini bernilai 0,828 dengan jumlah butir soal (*N of item*) berjumlah 50.

Dalam bukunya Joko Widiyanto (2010, hal. 108) menjelaskan bahwa dasar pengambilan keputusan uji reliabilitas adalah jika nilai *Cronbach's Alpha* > r tabel maka reliabel jika nilai *Cronbach's Alpha* < r tabel maka tidak reliabel. Berdasarkan *output "reliability statistic"* di atas di ketahui nilai *Cronbach's Alpha* sebesar 0,868. Nilai tersebut di bandingkan dengan nilai r tabel dengan nilai N=30 dicari nilai r tabel di tingkat signifikansi 0,05 maka diperoleh nilai r tabel 0,3494.

Karena nilai *Cronbach's Alpha* $0,868 > 0,3494$ maka sebagaimana dasar pengambilan keputusan di atas, dapat peneliti simpulkan bahwa butir soal bahasa Arab siswa kelas 11 MAS Pondok Pesantren Asaalam dinyatakan reliabel atau konsisten sebagai soal ujian akhir semester bahasa Arab siswa kelas 11 MAS Pondok Pesantren Asaalam Kab. Kampar, Riau.

3. DAYA PEMBEDA

Daya beda butir soal merupakan suatu indeks yang menunjukkan tingkat kemampuan butir soal membedakan kelompok atas (peserta tes yang berprestasi tinggi) dari kelompok bawah (peserta tes yang berprestasi rendah). Pernyataan ini sesuai dengan pendapat Sundayana bahwa daya pembeda merupakan kemampuan soal untuk dapat membedakan siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah.

Daya pembeda soal atau daya beda butir soal merupakan indeks yang menunjukkan tingkat kemampuan butir soal membedakan kelompok atas (unggul) dari kelompok bawah (asor). Daya pembeda suatu butir berfungsi untuk menentukan apakah suatu butir soal dapat membedakan kelompok dalam aspek yang diukur sesuai dengan perbedaan pada kelompok tersebut.

Sehingga butir soal tersebut dapat membedakan siswa yang mempunyai kemampuan tinggi dan siswa yang mempunyai kemampuan rendah. Diambil dari teori Sundayana bahwa terdapat indeks yang menunjukkan standar nilai yang harus dicapai untuk mencapai kriteria tersebut. Kriteria *Sangat Rendah* " <0 ", *Rendah* " $0.00 < 0.20$ ", *Sedang* " $0.21 < 0.40$ ", *Baik* " $0.41 < 0.70$ ", *Sangat Baik* " $0.71 < 1.00$ ", kriteria pada soal sangat rendah sangat direkomendasikan untuk dibuang dan untuk kriteria rendah sebaiknya untuk dibuang. (Sundayana, 2016, hal. 76) Berikut tabel signifikansi indeks dan kriteria Daya Pembeda:

Tabel 1

Indeks Daya Pembeda

Indeks Daya Pembeda	Kriteria
$0.71 < 1.00$	Sangat baik
$0.41 < 0.70$	Baik
$0.21 < 0.40$	Sedang
$0.00 < 0.20$	Rendah
< 0	Sangat Rendah

Hasil analisis daya pembeda pada butir soal ujian akhir semester bahasa Arab siswa kelas 11 MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau yang diperoleh peneliti menggunakan *Microsoft Excel* sebagai berikut:

Tabel 2
Hasil Statistik Daya Pembeda

DAYA PEMBEDA		
Status	Nomor Soal	Jumlah
Sangat Tinggi	No.4, No.26, No.32, No.34, No.35, No.36, No.40	7
Tinggi	No.2, No.3, 27, No.28, No.33, No.38, No.39, No.42, No.43	9
Sedang	No.5, No.15, No.23, No.30, No.31, No.44, No.45, No.48, No.49, No.50	10
Rendah	No.7, No.10, No.11, No.13, No.16, No.17, No.20, No.21, No.29, No.25, No.41, No.46, No.47.	13
Sangat Rendah	No.1, No.6, No.8, No.9, No.12, No.14, No.18, No.19, No.22, No.24, No.37.	11
Total		50

Terdapat hasil statistik yang menunjukkan bahwa terdapat 7 soal yang berkriteria *sangat tinggi* daya pembedanya terdapat pada soal: No.4, No.26, No.32, No.34, No.35, No.36, No.40. Kemudian 9 kriteria *tinggi* pada soal: No.2, No.3, No.27, No.28, No.33, No.38, No.39, No.42, No.43. Kemudian 10 soal kriteria *sedang*: No.5, No.15, No.23, No.30, No.31, No.44, No.45, No.48, No.49, No.50. Kemudian 13 soal *rendah*: No.7, No.10, No.11, No.13, No.16, No.17, No.20, No.21, No.29, No.25, No.41, No.46, No.47 (sebaiknya dibuang). Kemudian 11 soal *sangat rendah*: No.1, No.6, No.8, No.9, No.12, No.14, No.18, No.19, No.22, No.24, No.37(direkomendasikan untuk dibuang).

Dapat disimpulkan bahwa daya pembedanya pada 50 butir soal ujian akhir semester bahasa Arab siswa kelas 11 MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau memiliki daya beda yang signifikan seperti yang tertera pada tabel bahwa kriteria *Sangat Tinggi* (7), *Tinggi* (9), *Sedang* (10), *Rendah* (13), *Sangat Rendah* (11).

4. TINGKAT KESUKARAN

Tingkat kesukaran (*difficulty level*) disebut juga sebagai indeks kesukaran butir soal yang dilambangkan dengan huruf p yaitu *proportion*. Tingkat kesukaran adalah peluang untuk menjawab benar pada suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Tingkat kesukaran menunjukkan proporsi peserta didik yang dapat mengerjakan soal secara benar dari suatu tes/ujian. (Sumintono, 2015, hal. 13)

Dapat dikatakan bahwa taraf kesukaran merupakan proporsi peserta menjawab benar terhadap butir soal. Suatu tes dikatakan baik apabila butir soal pada tes tersebut tidak terlalu sukar dan juga tidak terlalu mudah. Soal yang terlalu mudah tidak merangsang mahasiswa untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan siswa

menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya. (Arikunto S. , 2008, hal. 207)

Terdapat indeks angka yang menunjukkan sukar mudahnya suatu butir soal. Terdapat beberapa indeks tingkat kesukaran yaitu pada kriteria *sulit* memiliki nilai “ $0.00 < P < 0.03$ ”. pada kriteria *sedang* “ $0.00 < P < 0.70$ ”. pada kriteria *mudah* “ $0.70 < P < 1.00$ ”. Berikut table indeks tingkat kesukaran beserta kriteria tingkat kesukaran:

Tabel 3
Indeks Tingkat Kesukaran

Indeks kesukaran	Kriteria
$0.00 < P < 0.03$	Sulit
$0.00 < P < 0.70$	Sedang
$0.70 < P < 1.00$	Mudah

Hasil analisis tingkat kesukaran pada butir soal ujian akhir semester bahasa Arab siswa kelas 11 MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau yang diperoleh peneliti menggunakan *Microsoft Excel* sebagai berikut:

Tabel 4
Hasil Statistik Tingkat Kesukaran

TINGKAT KESUKARAN		
Status	Nomor Soal	Jumlah
Mudah	No.1, No.2, No.3, No.5, No.6, No.7, No.8, No.9, No.10, No.11, No.12, No.13, No.14, No.15, No.16, No.17, No.18, No.19, No.20, No.21, No.22, No.23, No.24, No.25, No.27, No.28, No.29, No.30, No.31, No.33, No.37, No.38, No.39, No.41, No.42, No.43, No.44, No.45, No.46, No.47, No.48, No.49, No.50	43
Sedang	No.4, No.26, No.32, No.34, No.35, No.36, No.40	7
Sulit		0
Total		50

Terdapat hasil statistik yang menunjukkan bahwa terdapat 43 soal yang mempunyai kriteria *mudah* yaitu : No.1, No.2, No.3, No.5, No.6, No.7, No.8, No.9, No.10, No.11, No.12, No.13, No.14, No.15, No.16, No.17, No.18, No.19, No.20, No.21, No.22, No.23, No.24, No.25, No.27, No.28, No.29, No.30, No.31, No.33, No.37, No.38, No.39, No.41, No.42, No.43, No.44, No.45, No.46, No.47, No.48, No.49, No.50. kemudian yang mempunyai kriteria *sedang* terdapat 7 butir soal yaitu : No.4, No.26, No.32, No.34, No.35, No.36, No.40. Kemudian peneliti tidak menemukan soal yang berkriteria *sulit* pada soal ini dan hasil untuk kriteria Sulit.

Dapat disimpulkan bahwa tingkat kesukaran pada butir soal ujian akhir semester bahasa Arab siswa kelas 11 MAS PONDOK PESANTREN ASAALAM Kab. Kampar, Riau mayoritas soalnya adalah mudah karena perbandingan statistiknya adalah 43:7, yang telah dipaparkan sebelumnya bahwa jumlah 50 butir soal diantaranya 43 (mudah) dan 7 (sedang).

KESIMPULAN

Meninjau dari hasil analisis butir soal, peneliti menyimpulkan bahwa butir soal ujian akhir semester bahasa Arab siswa kelas 11 MAS Pondok Pesantren Asaalam Kab. Kampar, Riau memiliki validitas dan reliabilitas, untuk hasil soal yang valid berjumlah 24 dan yang kategori tidak valid berjumlah 26. Sedangkan untuk reliabilitas pada soal UAS bahasa Arab siswa kelas 11 ini menyatakan sangat reliabel atau konsisten sebagai soal ujian akhir semester bahasa Arab siswa kelas 11 MAS Pondok Pesantren Asaalam Kab. Kampar, Riau. Selanjutnya adalah daya pembeda dan tingkat kesukaran. Daya beda soal bahasa Arab berkriteria *sangat tinggi* (7), *tinggi* (9), *sedang* (10), *rendah* (13), *sangat rendah* (11). Sedangkan tingkat kesukaran pada soal UAS bahasa Arab siswa kelas 11 ini mayoritas berkriteria mudah = 43 dan sedang = 7.

DAFTAR PUSTAKA

- Ainin, d. (2006). *Evaluasi Dalam Pembelajaran Bahasa Arab*. Malang: Misykat.
- Arikunto, S. (2008). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, S. (2010). *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Azwar, S. (2008). *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Bagiyono. (2017). Analisis Tingkat Kesukaran Soal dan Daya Pembeda Butir Soal Ujian Pelatihan Radiografi Tingkat 1, *Widyanyuklida*, Hal. 3.
- Fathimah, S. (2019). Analisis Butir Soal Pelajaran Bahasa Arab Di MA Roudlotul Ulum Pagak Malang. *Tarbiyatuna: Jurnal Pendidikan*, 143.
- Margono. (2010). *Metodologi Penelitian Pendidikan*. Jakarta: PT. Raja Grafindo Persada.
- Maulana, d. (2020). Analisis Butir Soal Bahasa Arab Ujian Akhir Madrasah Bersama Daerah (UAMBD) Madrasah Ibtidaiyah Tahun 2017-2018. *Ta'lim al-'Arabiyyah*, 13.

- Otomo, B. (2018). Analisis Validitas Isi Butir Soal Sebagai Salah Satu Upaya Peningkatan Kualitas Pembelajaran Di Madrasah Berbasis Nilai-Nilai Islam. *Jurnal Pendidikan Matematika*, hal. 147.
- Rahmi, I. (2020). Analisis Kualitas Tes Bahasa Arab Berbasis Higher Order Thinking Skill (HOTS). *Lisanuna*, hal. 46.
- Solichin, M. (2019). Analisis Daya Beda Soal, Taraf Kesukaran, Validitas, Butir Tes, Interpretasi Hasil Tes Dan Validitas Ramalan Dalam Evaluasi Pendidikan. *Dirasat* , hal. 196-198.
- Sukardi. (2008). *Evaluasi Pendidikan Prinsip & Operasionalnya*. Jakarta: Bumi Aksara.
- Sumintono. (2015). *Bambang dan Widhiarso, Wahyu. Aplikasi Pemodelan Rasch pada Assessment Pendidikan*. Bandung:: Trim Komunikata.
- Sundayana, R. (2016). *Statistika Penelitian Pendidikan*. Bandung: Alfabeta.
- Widiyanto, J. (2010). *SPSS for Windows Untuk Analisis Data Statistik dan Penelitian*. Surakarta: BP-FKIP UMS.