

Nurul Fikriyah
Universitas Muhammadiyah
Yogyakarta

nurulfikriyah3008@gmail.com

Analisis Butir Soal Ulangan Tengah Semester Mata Pelajaran Bahasa Arab Kelas VII Semester Ganjil SMP Muhammadiyah 1 Yogyakarta Tahun Ajaran 2019/2020

DOI: 10.18196/mht.v3i2.10501

Abstract

This study aims to determine the quality and feasibility of the Middle Semester Test items in Arabic for Class VII Even Semester SMP Muhammadiyah 1 Yogyakarta in the 2019/2020 academic year. The approach used in this research is a quantitative approach because all information obtained will be presented in numerical form. The data obtained is by using the documentation method in the form of midterm test questions, answer keys, and student responses. In this study, using the entire population consisting of all mid-semester test questions in Arabic class VII. The data obtained were analyzed using the SPSS version 25 and Microsoft Excel 2010 program. The results showed that: (1) validity of valid questions is 31 items (89%) and invalid is 4 items (11%), (2) difficulty level of easy questions is 3 items (10%) while 25 items (80%) are difficult 3 items (10%), (3) the distinguishing power of feasible questions was 25 items (81%) and not feasible (19%). Thus, it can be concluded that the quality and appropriateness of the midterm test items in Arabic class VII SMP Muhammadiyah 1 Yogyakarta in the 2019/2020 academic year are in the sufficient category.

Keywords: *analysis, evaluation, questions, midterm tests*

Abstrak

Penelitian ini bertujuan agar mengetahui kelayakan serta kualitas butir soal Ulangan Tengah Semester pada Mata Pelajaran Bahasa Arab Kelas VII Semester Genap SMP Muhammadiyah 1 Yogyakarta Tahun Ajaran 2019/2020. Penelitian ini menggunakan pendekatan kuantitatif karena semua informasi yang didapat akan disajikan dalam bentuk angka. Data yang didapatkan dengan menggunakan metode dokumentasi berupa soal ulangan tengah semester, kunci jawaban, dan respon jawaban siswa. Penelitian ini menggunakan semua populasi berupa seluruh soal ulangan tengah semester mata pelajaran Bahasa Arab kelas VII. Setelah data diperoleh selanjutnya akan dianalisis menggunakan aplikasi spss versi 25 dan Microsoft Excel 2010. Hasil dari penelitian ini menghasilkan: (1) validitas soal valid berjumlah 31 butir (89%) dan tidak valid berjumlah 4 butir (11%), (2) tingkat kesukaran soal mudah berjumlah 3

butir (10%) sedang berjumlah 25 butir (80%) sukar 3 butir (10%), (3) daya pembeda soal layak berjumlah 25 butir (81%) dan tidak layak berjumlah (19%). Oleh karena itu dapat disimpulkan bahwa kelayakan dan kualitas butir soal ulangan tengah semester mata pelajaran bahasa arab kelas VII SMP Muhammadiyah 1 Yogyakarta tahun ajaran 2019/2020 sudah termasuk ke dalam kategori cukup.

Kata kunci: analisis, evaluasi, soal, ulangan tengah semester.

PENDAHULUAN

Menurut (Herlambang, 2015:1), pendidik seharusnya lebih dulu melaksanakan pengukuran dan penilaian sebelum akan melaksanakan evaluasi. Evaluasi yang dilakukan secara berkelanjutan akan mampu membantu evaluator dalam memperkirakan apakah tujuan dari evaluasi yang telah disusun tersebut akan berhasil tercapai dalam waktu yang telah ditentukan maupun yang tidak.

Dengan menggunakan alat bantu berupa tes, maka guru akan dengan mudah mengetahui hasil belajar dari para siswa. Tes dalam dunia pendidikan biasanya terdiri dari 2 jenis, yaitu tes standar dan tes non standar (Arikunto, 2013:159). Tes standar yaitu tes yang dilakukan untuk meningkatkan kemampuan yang patut dimiliki oleh peserta didik. Tes standar juga wajib untuk mengamati berbagai aspek dari sebuah tes yaitu seperti validitas, reliabilitas, kelayakan butir soal serta kisi-kisi. Sedangkan tes non standar yaitu tes yang telah dirangkai oleh seorang guru pengampu mata pelajaran tersebut, contohnya seperti tes yang disusun untuk ulangan tengah semester pada mata pelajaran bahasa Arab. Tes non standar tersebut dapat berupa pilihan ganda dan tes uraian.

Dengan menggunakan tes tersebut, maka seorang pendidik akan mengetahui informasi mengenai berhasil atau tidaknya tujuannya agar proses pembelajaran yang sudah ditetapkan pada kurikulum seperti kompetensi dasar, standar kompetensi, dan indikator.

Pada umumnya tes yang dilaksanakan di SMP Muhammadiyah 1 Yogyakarta sama dengan sekolah-sekolah lain, yaitu berupa tes formatif yang umumnya dilaksanakan pada pertengahan semester/ulangan tengah semester (UTS), dan tes sumatif yang umumnya dilaksanakan pada akhir semester/ulangan akhir semester (UAS). Butir soal ujian di SMP Muhammadiyah 1 Yogyakarta yang akan diujikan pada ulangan tengah

semester tersebut dirancang oleh guru mata pelajaran bahasa Arab sehingga tes itu disebut dengan tes nonstandar.

Untuk menganalisis butir soal, hanya butir soal yang berkualitas yang akan diujikan kepada siswa dan siswi untuk dapat mengetahui sejauh mana kemampuan yang dimiliki oleh siswa dan siswi sekolah tersebut. Teori pengukuran untuk menganalisis butir soal yang dilakukan pada penelitian ini menggunakan teori pengukuran klasik, yaitu untuk menghitung daya pembeda dan tingkat kesukaran. Sedangkan validitas dan reliabilitas dipakai untuk mengetahui keajegan pada soal. Biasanya dalam menyusun butir soal, guru hanya ingin mengelompokkan siswa yang pandai dan kurang pandai sehingga kualitas dari soal tersebut belum diketahui kelayakannya. Dari uraian kriteria-kriteria di atas seharusnya menjadi acuan untuk guru dalam menyusun butir soal.

Berdasarkan latar belakang dari permasalahan di atas, maka penelitian ini akan membahas tentang "Analisis Butir Soal Ulangan Tengah Semester Mata Pelajaran Bahasa Arab Kelas VII Semester Genap SMP Muhammadiyah 1 Yogyakarta Tahun Ajaran 2019/2020". Peneliti mengangkat judul ini untuk mengetahui kualitas butir soal ulangan tengah semester mata pelajaran Bahasa Arab kelas VII di SMP Muhammadiyah 1 Yogyakarta tahun ajaran 2019/2020 yang ditinjau dari segi validitas, reliabilitasnya, tingkat kesukaran dan daya pembedanya.

Landasan Teori

1. Analisis Butir Soal

Analisis butir soal adalah proses yang dilakukan secara sistematis untuk memberikan informasi mengenai butir soal yang telah disusun, Arikunto (2006, 205). Menganalisis butir soal ditujukan untuk menghubungkan antara skor dari masing-masing butir soal dengan skor keseluruhan serta membandingkan jawaban dari masing-masing butir soal dengan jawaban dari skor keseluruhan tes.

Berdasarkan Antara (2020, 7), terdapat 2 teori pengukuran mengenai analisis butir soal, yaitu teori pengukuran klasik dan teori respon butir. Teori pengukuran yang biasa digunakan untuk menganalisis butir soal adalah teori pengukuran klasik karena lebih sedikit dalam jumlah peserta didik dan dalam menganalisis butir soal akan lebih

mudah, bisa dilakukan dengan perhitungan kalkulator maupun dengan menggunakan program komputer.

2. Validitas

Arikunto (2011, 58) mengatakan bahwa validitas suatu data akan dinilai valid jika sesuai dengan keadaan yang nyata atau sebenarnya. Arikunto (2011, 65) juga mengemukakan secara umum bahwa validitas dibagi menjadi dua, yaitu validitas logis (*Logical Validity*) dan validitas empiris (*Empirical Validity*).

Pada penelitian dalam menganalisis validitas akan difokuskan menggunakan validitas logis yang mencakup pada validitas isi dan validitas konstruk, karena pada validitas logis akan menjadikan tingkat validitas butir soal sesuai dengan silabus dan RPP.

Rentang koefisien yang digunakan untuk menganalisis validitas adalah:

Antara 0,8 sampai 1	: sangat tinggi
Antara 0,6 sampai 0,8	: tinggi
Antara 0,4 sampai 0,6	: cukup
Antara 0,2 sampai 0,4	: rendah
Antara 0 sampai 0,2	: sangat rendah

3. 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, 31).

Menurut Iskandarwassid dan Sunendar (2009, 188) nilai untuk interpretasi reliabilitas adalah:

0,8 – 1	= tinggi
0,6 – 0,8	= cukup
0,4 – 0,6	= agak rendah
0,2 – 0,4	= rendah
0 – 0,2	= sangat rendah

4. Tingkat Kesukaran

Sudijono (2011, 370) mengatakan butir item tes dikatakan baik jika butir item tidak terlalu sulit dan tidak terlalu mudah yang berarti derajat item tersebut berada pada

tingkat kesukaran yang sedang atau cukup. Butir soal yang sangat mudah tidak bisa mengasah siswa untuk berusaha dalam menjawab soal tersebut. Sedangkan butir soal yang sangat sulit akan menyebabkan siswa mudah putus asa karena soal tersebut berada di luar jangkauannya. Kesanggupan siswa merupakan hal yang harus diperhatikan dalam menganalisis tingkat kesukaran, bukan dari bentuk soal yang dibuat oleh guru.

Menurut Arikunto (2011, 210) indeks kesukaran diklasifikasikan nilainya sebagai berikut:

P 0,00 – 0,30 = sukar

P 0,30 – 0,70 = sedang

P 0,70 – 1,00 = mudah

5. Daya Pembeda

Daryanto (2007, 183) menjelaskan bahwa daya pembeda yaitu kemampuan membedakan mana siswa yang berkemampuan tinggi dan mana siswa yang berkemampuan rendah. Langkah yang harus dilakukan untuk menghitung indeks daya pembeda adalah:

- a. Mengurutkan skor peserta tes dari yang tertinggi sampai terendah
- b. Menghitung proporsi 27% untuk siswa yang menjawab benar pada soal kelompok atas dan 27% untuk siswa yang menjawab pada soal kelompok bawah
- c. Membandingkan proporsi jawaban benar untuk kedua kelompok

Menurut Arikunto (2011, 118), klasifikasi daya pembeda adalah sebagai berikut:

D : 0,00 – 0,20 = jelek

D : 0,20 – 0,40 = cukup

D : 0,40 – 0,70 = baik

D : 0,70 – 1,00 = baik sekali

D : negatif, semua butir soal yang memiliki nilai D negatif sebaiknya tidak digunakan atau dibuang.

METODE PENELITIAN

Metode yang digunakan pada penelitian ini adalah deskriptif kuantitatif. Menurut (Arikunto 2010, 234) 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, 106).

Data yang digunakan pada penelitian ini adalah soal-soal ulangan tengah semester mata pelajaran bahasa Arab kelas VII, kunci jawaban, dan jawaban siswa kelas VII SMP Muhammadiyah 1 Yogyakarta.

Populasi pada penelitian ini yaitu seluruh respon jawaban siswa soal ulangan tengah semester mata pelajaran bahasa Arab kelas VII SMP Muhammadiyah 1 Yogyakarta. Sedangkan pengambilan sampel yaitu dengan menggunakan sampel acak sehingga diperoleh 115 respon jawaban siswa yang dijadikan contoh/sampel.

Teknik yang dilakukan untuk mengumpulkan data pada penelitian ini adalah dengan menggunakan teknik dokumentasi. Teknik dokumentasi yaitu suatu proses pengumpulan data dengan cara melihat secara langsung sumber-sumber dokumen yang terkait dengan penelitian. Dengan arti lain metode dokumentasi ini yaitu melakukan pengambilan data melalui dokumen-dokumen tertulis maupun elektronik.

Secara garis besar, terdapat 3 langkah dalam menganalisis sebuah data yaitu (Arikunto 2010, 278):

1. Persiapan
2. Tabulasi
3. Penerapan data sesuai dengan pendekatan penelitian

Rumus yang akan digunakan dalam menganalisis butir soal ulangan tengah semester mata pelajaran bahasa Arab kelas VII SMP Muhammadiyah 1 Yogyakarta dari segi validitas, reliabilitas, daya pembeda, dan tingkat kesukaran adalah:

- 1) Analisis Validitas

Dalam menganalisis validitas butir soal rumus yang digunakan dengan rumus korelasi product moment dengan angka kasar (Arikunto 2011, 72):

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N \sum X^2 - (\sum X)^2\} \{N \sum y^2 - (\sum y)^2\}}}$$

Langkah-langkah untuk mencari validitas adalah sebagai berikut:

a. Tabel untuk mencari validitas

No.	Nama	X	Y	X ²	Y ²	XY
1.						
2.						
	Jumlah					

b. Masukkan kedalam rumus dengan rentang koefisien sebagai berikut:

- Antara 0,800 sampai 1,00 : sangat tinggi
 Antara 0,600 sampai 0,800 : tinggi
 Antara 0,400 sampai 0,600 : cukup
 Antara 0,200 sampai 0,400 : rendah
 Antara 0,00 sampai 0,200 : sangat rendah

2) Analisis Reliabilitas

Pada analisis reliabilitas ini memakai rumus K-R 20 pada soal yang berbentuk pilihan ganda. Rumus tersebut adalah sebagai berikut:

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

Langkah-langkah untuk mencari reliabilitas adalah sebagai berikut:

a. Membuat tabel analisis

No.	Nama	Nomor Item					Skor Total	Kuadrat Skor Total
		1	2	3	4	Dst		
1.								
2.								
3.								
	Jumlah							
	Jumlah Kuadrat							

b. Mencari varian dari tiap-tiap item dengan interpretasi reliabilitas adalah sebagai berikut:

0,80 – 1,00 = tinggi

0,60 – 0,80 = cukup

0,40 – 0,60 = agak rendah

0,20 – 0,40 = rendah

0,00 – 0,20 = sangat rendah

3) Analisis Tingkat Kesukaran

Rumus yang digunakan untuk menganalisis tingkat kesukaran menurut Arikunto (2011, 208):

$$P = \frac{B}{JS}$$

Langkah-langkah untuk mencari indeks tingkat kesukaran adalah sebagai berikut:

a. Membuat tabel analisis

Siswa	Nomor Soal											Skor Siswa	
	1	2	3	4	5	6	7	8	9	10	Dst		
A													
B													
C													
Dst													
Jumlah													

b. Masukkan ke dalam rumus dengan indeks kesukaran sebagai berikut :

P 1,00 – 0,30 = sukar

P 0,30 – 0,70 = sedang

P 0,70 – 1,00 = mudah

4) Analisis Daya Pembeda

Berikut adalah rumus yang digunakan untuk menganalisis daya pembeda menurut Arikunto (2011, 213):

$$D = \frac{BA}{JB} - \frac{BB}{JB} = PA - PB$$

Langkah-langkah untuk mencari daya pembeda adalah sebagai berikut:

a. Membuat tabel analisis

Siswa	Kelompok	Nilai Soal					Skor Siswa
		1	2	3	4	Dst	
A							
B							
C							
Jumlah							

b. Mengurutkan data dari skor yang paling tinggi ke skor yang paling rendah dengan mengambil 27% skor teratas sebagai kelompok atas (Ja) dan 27% skor terbawah sebagai kelompok bawah (Jb).

Ja : jumlah kelompok atas

Jb : jumlah kelompok bawah

Kelompok Atas

Kelompok Bawah

Nama	Nilai	Nama	Nilai
A		B	
C		D	
F		E	
Dst		Dst	

c. Masukkan kedalam rumus dengan klasifikasi daya pembeda adalah sebagai berikut:

D : 0,00 – 0,20 = jelek

D : 0,20 – 0,40 = cukup

D : 0,40 – 0,70 = baik

D : 0,70 – 1,00 = baik sekali

D : negatif, semua butir soal yang memiliki nilai D negatif sebaiknya tidak digunakan atau dibuang.

Hasil dan Pembahasan Penelitian

Hasil penelitian ini dipaparkan oleh peneliti yaitu meliputi validitas, reliabilitas, tingkat kesulitan, dan daya pembeda.

1. Validitas

Berikut adalah tabel hasil analisis validitas soal pilihan ganda:

Tabel 1. Distribusi butir soal berdasarkan validitas

Indeks validitas (r_{xy})	Nomor butir soal	Jumlah	Presentase
1. Valid $0,182 \leq r_{xy} \leq 1$, 2, 3, 4, 6, 7, 8, 9, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 35	31	89%
2. Tidak valid $r_{xy} \leq 0,182$	5, 10, 15, 28	4	11%

Tabel 1 memaparkan hasil presentase validitas butir soal yang terdiri dari dua kategori yaitu valid dan tidak valid. Butir soal yang masuk ke dalam kategori valid berjumlah 31 butir dengan hasil persentase sebesar 89%. Sedangkan butir soal yang masuk ke dalam kategori tidak valid berjumlah 4 butir dengan hasil persentase sebesar 11%.

2. Reliabilitas

Hasil r_{11} yang didapat akan dikonsultasikan ke dalam tabel product moment dengan taraf signifikansi 5%. Soal dikatakan reliabel apabila $r_{11} > r$ tabel.

Tabel 2. Reliabilitas Soal Pilihan Ganda

Reliability Statistics	
Cronbach's Alpha	N of Items
.854	31

Tabel 2 menunjukkan bahwa butir soal pada penilaian tengah semester tahun ajaran 2019/2020 mata pelajaran bahasa Arab kelas VII SMP Muhammadiyah 1 Yogyakarta yang terdiri dari 31 butir soal memiliki koefisien reliabel dengan kategori tinggi yaitu sebesar 0,854. Hasil analisis reliabilitas menunjukkan bahwa soal tersebut memiliki butir-butir soal yang ajeg atas pengukuran yang telah dilakukan.

3. Tingkat Kesukaran

Hasil pengelompokan kategori tingkat kesukaran dapat dilihat pada presentase sebagai berikut:

Tabel 3. Tingkat Kesukaran Soal Pilihan Ganda

No.	Kategori	Nomor soal	Jumlah	Presentase
1.	Mudah	12, 14, 20	3	10%
2.	Sedang	, 2, 3, 4, 7, 8, 9, 11, 13, 16, 17, 18, 19, 21, 22, 23, 24, 26, 29, 30, 31, 32, 33, 34, 35.	25	80%
3.	Sukar	6, 25, 27	3	10%
Jumlah			31	

Tabel 3 menunjukkan bahwa butir soal yang termasuk ke dalam kategori mudah dengan hasil persentase sebesar 10%. Butir soal yang termasuk ke dalam kategori sedang dengan hasil persentase sebesar 80%. Sedangkan butir soal nomor 6, 25, dan 27 termasuk ke dalam kategori sukar dengan persentase sebesar 10%.

4. Daya Pembeda

Distribusi soal berdasarkan kategori daya pembeda bisa dilihat pada:

Tabel 4. Daya Pembeda Soal Pilihan Ganda

Kategori	Nomor Soal	Jumlah	Presentase %
Baik	2, 7, 12, 16, 17, 19, 21, 30, 31, 32, 33, 35	12	39%
Cukup	1, 3, 4,8, 9, 13, 14, 18, 20, 22, 23, 24, 34	13	42%
Jelek	6, 11, 25, 26, 27, 29	6	19%

Berdasarkan tabel 4, dapat dilihat bahwa butir soal pada kategori baik sebesar 39%, kategori cukup sebesar 42% dan kategori jelek sebesar 19%.

KESIMPULAN DAN SARAN

Berdasarkan dari hasil penelitian dan juga pembahasan tentang analisis butir soal Ujian Penilaian Tengah Semester Mata Pelajaran Bahasa Arab Kelas VII SMP Muhammadiyah 1 Yogyakarta Tahun Ajaran 2019/2020 yang dilihat dari segi validitas menunjukkan bahwa soal dalam kategori valid berjumlah 31 butir soal (89%) dan soal yang termasuk ke dalam kategori tidak valid berjumlah 4 butir soal (11%). Berdasarkan dari segi reliabilitas menunjukkan nilai sebesar 0,854. Oleh karena itu, butir soal tersebut memiliki reliabilitas yang tinggi. Sedangkan pada segi tingkat kesulitan menunjukkan bahwa 3 butir soal (10%) masuk dalam kategori mudah, 25 butir soal (80%) termasuk ke dalam kategori sedang, dan 3 butir soal (10%) termasuk ke dalam kategori sulit. dari segi daya pembeda menunjukkan bahwa 12 butir soal (39%) termasuk ke dalam kategori baik, 13 butir soal (42%) termasuk ke dalam kategori cukup, dan 6 butir soal (19%) masuk dalam kategori jelek.

Butir soal dengan kualitas baik dapat dimasukkan ke dalam bank soal dan dapat diujikan kembali pada ujian yang akan dilaksanakan kembali. Butir soal yang berkualitas kurang dapat dilakukan pemeriksaan kembali untuk proses perbaikan butir soal. Butir soal yang berkualitas jelek sebaiknya dikeluarkan atau dibuang dari bank soal dan tidak diujikan kembali untuk ujian yang akan dilangsungkan lagi karena butir soal tersebut memiliki kualitas yang kurang memadai.

Sebaiknya sekolah dapat memperhatikan dan lebih peduli lagi terhadap keterampilan guru dalam pembuatan dan penyusunan instrumen tes/ujian. Salah satu cara yang dapat dilakukan yaitu diadakannya pelatihan menyusun instrumen evaluasi yang baik dan benar. Sehingga pengajar/guru bisa melaksanakan analisis secara keseluruhan terhadap soal yang dibuatnya sendiri.

REFERENSI

- Ainin, dkk. 2006. *Evaluasi Dalam Pembelajaran Bahasa Arab*. Malang: Misykat.
- Antara, Anak Agung Purwa. 2020. *Penyetaraan Vertikal dengan Pendekatan Klasik dan Item Response Theory (Teori dan Aplikasi)*. Yogyakarta: Deepublish.
- Arikunto, Suharsimi. 2006. *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Arikunto, Suharsimi. 2010. *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Arikunto, Suharsimi. 2011. *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Arikunto, Suharsimi. 2013. *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Daryanto, Daryanto. 2007. *Evaluasi Pendidikan*. Jakarta: Rineka Cipta.
- Herlambang, Bima Kartika. 2015. *Analisis Butir Soal Ulangan Tengah Semester Mata Pelajaran Pendidikan Jasmani Olahraga Kesehatan Kelas VII Semester Genap SMP N 2 Wonosari Tahun Ajaran 2014/2015*. Universitas Negeri Yogyakarta.
- Iskandarwassid, dan Dadang Sunendar. 2009. *Strategi Pembelajaran Bahasa*. Bandung: PT Remaja Rosdakarya.
- Margono, Margono. 2010. *Metodologi Penelitian Pendidikan*. Jakarta: Rineka Cipta.
- Sudijono, Anas. 2011. *Pengantar Evaluasi Pendidikan*. Jakarta: PT. Raja Grafindo Persada.