KITABA: Journal of Interdisciplinary Arabic Learning, 01 (02): 69-74 (2023)
DOI: 10.18860/kitaba.v1i2.21796

Magister Pendidikan Bahasa Arab
Pascasarjana UIN Maulana Malik Ibrahim
Malang, Indonesia

KITABA
JOURNAL OF INTERDISCIPLINARY ARABIC LEARNING

ANALISIS BUTIR SOAL UJIAN AKHIR SEMESTER BAHASA ARAB: PENDEKATAN STATISTIK DAN PENDIDIKAN DI LEMBAGA BAHASA ASING (LPBA)

Durrotus Tsaminah [1], Mohammad Sofi Anwar [2*], Hilman Sholehuddin [3], Nur Qomari [4]
[1, 2, 3, 4] Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia

Article History:
Received: June 06, 2023
Revised: June 27, 2023
Accepted: July 05, 2023
Published: July 30, 2023

Keywords:
Analysis, Arabic Letters, Final Exam Semester

*** Correspondence Address:**
sofianwar1997@gmail.com

Abstract: This study analyzed the grains of Arabic for the final assessment of the semester in the Pantoktan Pesantren Al Yasini Pasuruan. The grain of analysis is called a double option amounts to 50 items. This analysis aims to know the characteristics of good tests based on validity, differentness, casting rate, and difficulty level. This research is quantitative research whose data analysis is presented in descriptive. The result of the item grains is said to be 100% valid and divided into three easy, moderate, and difficult categories. In addition, the letters of Arabic are also analyzed by the power to know the difference in leaf ability, so obtained five categories are ugly, enough, well, very good, and inappropriate. In addition to analyzing the matter, this study also analyzed the pattern of answers as a castled rate that the result of 144 options may be loaded and the remaining six may not.

PENDAHULUAN

Pembelajaran merupakan suatu kegiatan yang terdiri atas beberapa komponen yang saling terkait satu sama lain. Komponen tersebut terdiri atas perencanaan, materi, strategi dan metode, serta penilaian pembelajaran. Penilaian merupakan salah satu komponen yang harus ada dalam kegiatan pembelajaran. Menurut Gronlund dan Linn penilaian adalah proses yang sistematis untuk mengumpulkan, menganalisis, dan menafsirkan informasi tentang ketercapaian peserta didik dalam kegiatan pembelajaran (Ainin, 2020). Penilaian tidak hanya bertujuan mengetahui pencapaian peserta didik (Ridho, 2018). Penilaian juga bertujuan memastikan seluruh ragkaian kegiatan pembelajaran sesuai dengan tujuan pembelajaran yang diharapkan, baik dari segi perencanaan, pelaksanaan, metode, strategi, dan media yang digunakan (Arifianto et al., 2021). Dalam konteks pembelajaran bahasa Arab, tujuannya adalah peserta didik mampu menguasai unsur-unsur bahasa dan menerapkannya dalam bentuk empat keterampilan bahasa.

Penilaian berkaitan dengan alat ukur yang digunakan. Jika alat ukurnya baik dan benar, maka penilaian juga akan akurat. Sebaliknya jika alat ukurnya tidak baik, maka penilaian tidak akan akurat artinya tidak tepat sasaran. Alat ukur yang digunakan dalam kegiatan penilaian juga bermacam-macam. Salah satu yang sering digunakan adalah tes. Menurut Sudjiono dan Matsna kata tes berasal dari bahasa Prancis *testum*. Dalam konteks pembelajaran bahasa tes diartikan sebagai alat yang digunakan untuk mengetahui seberapa jauh seseorang mempelajari suatu bahasa. Untuk itu Muhbib mendefinisikan tes bahasa Arab dengan sejumlah pertanyaan atau pernyataan yang didalamnya mengandung materi bahasa Arab dan dirancang sedemikian rupa agar dijawab oleh peserta didik (Wahab, 2013).

Tes berisi butir-butir soal yang menjadi instrumen utama dalam kegiatan penilaian

pembelajaran, dalam hal ini bahasa Arab. Menurut *Al Khuli* ada enam karakteristik atau ciri dari tes bahasa yang baik. Jika salah satu dari kelimanya tidak ada maka tes bahasa belum bisa dikatakan memenuhi standar sebagaimana yang diungkapkan *Al Khuli* (Khuli, 2000). Adapun enam karakteristik tersebut adalah *al shidq* (valid), *al tsabat* (reliabel), *al maudhu'iyyah* (obyektifitas), *al 'amaliyah* (praktis), *al iqtishodiyah* (ekonomis), dan *al tamyiz* (memiliki daya beda). Sedangkan Ainin dalam bukunya menyebutkan empat karakteristik saja yang harus dipenuhi tes bahasa yang baik yaitu validitas, reliabel, memiliki daya beda yang baik dan tingkat kesulitan yang sesuai (Ainin, 2021). Adapun dua lainnya yang tidak disebutkan adalah tambahan yang membuat tes bahasa semakin sempurna. Ainin lebih merinci validitas soal dengan membaginya menjadi validitas isi, tampak luar, dan konstruk.

Pertama, Validitas disebut juga kesahihan yang artinya ketepatan atau kebenaran. Suatu tes dikatakan valid jika ada kesesuaian antara hasil dan tujuan (Ainin et al., 2006). *Kedua*, Reliabel mengacu pada istilah kestabilan dan kekonsistenan. Artinya suatu tes dikatakan reliabel jika tes tersebut memiliki konsitensi yang tinggi (Arifin, 2017). Ketiga, Daya beda merupakan kemampuan butir soal sebagai alat ukur untuk membedakan kemampuan peserta didik, apakah kemampuan peserta didik masuk dalam kategori tinggi, sedang, maupun rendah (Munip, 2017). Keempat, fungsi pengecoh adalah jawaban-jawaban salah yang disediakan untuk mengecoh siswa (Wachidun, 2021). Pengecoh setidaknya dipilih dari 5% dari peserta tes. Selain itu jangan sampai pengecoh dipilih melebihi jumlah peserta tes yang memilih kunci jawaban. Kelima, tingkat kesulitan yaitu kemampuan tes yang tidak terlalu mudah dan tidak terlalu sulit dijawab (Ainin, 2020). Semakin sedikit peserta tes yang menjawab dengan benar, berarti soal tersebut soal kategori sulit. Sebaliknya, semakin banyak peserta tes yang menjawab dengan benar berarti soal tersebut kategori soal mudah.

Pada praktiknya di lapangan masih banyak guru maupun pembuat soal yang kurang memerhatikan lima karakteristik diatas. Hal ini sebagaimana yang diungkapkan Ainin (2019) bahwa kesalahan umum yang sering terjadi dalam butir-butir soal adalah ketidaksesuaian antara Indikator Pencapaian Kompetensi (IPK) dan soal yang diberikan (Ainin, 2019) . Misalnya IPK untuk pembelajaran maharah al kalam adalah siswa mampu berbicara dengan baik dan benar. Tetapi soal yang diberikan siswa diminta menjawab pertanyaan dari teks yang telah disajikan. Hal serupa juga ditemukan dalam penilaian maharah al istima' yang lebih menekankan siswa untuk memahami wacana lisan yang telah diperdengarkan, bukan menceritakan kembali wacana lisan dengan bahasanya sendiri. Ainin juga menemukan cuplikan teks bacaan dalam buku bahasa Arab MA/SMA/SMK yang masuk dalam kategori sedang. Beberapa soal diambil dari teks bacaan tersebut bisa dijawab hanya dengan membaca tanpa harus memahami (Ainin, 2018). Hal-hal semacam ini menunjukkan bahwa butir-butir soal tersebut belum memenuhi karakteristik tes yang baik yaitu valid. Oleh karenanya perlu adanya kajian analisis terhadap butir-butir soal bahasa Arab untuk mengetahui kesesuaian butir-butir soal tersebut dengan empat karakteristik tes yang baik yaitu validitas, daya beda, fungsi pengecoh, dan tingkat kesulitan.

METODE

Penelitian ini merupakan penelitian kuantitatif yang menggunakan analisis deskriptif untuk mendeskripsikan hasilnya. Menurut Arikunto analisis deskriptif bertujuan menemukan informasi dari suatu peristiwa yang sedang berlangsung (Arikunto, 2010). Sedangkan penelitian kuantitaif menurut Margono bertujuan menemukan informasi dengan menggunakan data berupa angka sebagai alatnya (Margono, 2010).

Adapun data yang digunakan dalam penelitian ini adalah soal bahasa Arab untuk ujian akhir semester di Pondok Pesantren Al Yasini Pasuruan. Hasil penelitian ini bertujuan mengetahui kualitas butir-butir soal bahasa Arab di pondok pesantren Al Yasini Pasuruan berdasarkan karakteristik validitias, reliabelitas, daya beda, dan tingkat kesulitannya.

TEMUAN

Berdasarkan analisis butir soal yang dilakukan oleh peneliti di Lembaga Pengembangan Bahasa Asing (LPBA) Pondok Pesantren Terpadu Al-Yasini Pasuruan terhadap soal yang diujikan sebagai alat evaluasi kelulusan dengan jumlah 50 butir soal pilihan ganda dan masing-masing nomor memiliki 3 pilihan jawaban ditemukan beberapa hasil sebagai berikut:

Pertama, analisis validitas. Analisis validitas ini mengacu pada validnya butir soal terhadap materi yang telah diujikan dalam kurun waktu tertentu sesuai kalender pendidikan yang dibuat oleh lembaga tersebut. Setelah dilakukannya analisis validitas ini diketahui bahwa 50 butir soal pilihan ganda yang diujikan terhadap muta'allimah 100% valid. Karena sudah sesuai dengan materi-materi yang sudah dipelajari selama kegiatan belajar mengajar di LPBA Al-Yasini Pasuruan.

Kedua, analisis tingkat kesulitan. Analisis tingkat kesulitan ini bertujuan untuk mengetahui seberapa sulit butir soal tersebut ketika diujikan kepada muta'allimah LPBA Al-Yasini Pasuruan. Dalam analisis tingkat kesulitan ini terdapat tiga kriteria soal yaitu tergolong mudah, tergolong sedang, dan tergolong sulit. Dari 50 butir soal pilihan ganda tersebut ditemukan 23 butir soal mudah, 24 butir soal sedang dan 3 butir soal sulit.

Analisis tingkat kesulitan ini diperoleh dari perhitungan jumlah jawaban benar (dari peserta tes) dibagi jumlah keseluruh peserta tes kemudian dikali 100%. Kegiatan tersebut dilakukan satu persatu tiap nomor soal. Sehingga dapat diketahui butir soal mana yang tergolong mudah, sedang, ataupun sulit. Perhitungan ini bisa dilakukan dengan bantuan Microsoft excel yang sudah disesuaikan dengan rumus yang berlaku. Agar lebih jelas, perhatikan tabel di bawah ini:

Tabel 1: Analisis Tingkat Kesulitan

Kriteria Soal	Prosentase	Jumlah Soal
Mudah	0,71 – 1,00	23
Sedang	0,31 – 0,70	24
Sulit	0,00 – 0,30	3

Tabel 2: Rincian Hasil Analisis Tingkat Kesulitan

Kriteria Soal	Prosentase	Jumlah Peserta Tes	Jumlah Jawaban Benar	Nomor Butir Soal
Sulit	0,11	64	7	15
Sulit	0,25	64	16	34
Sulit	0,28	64	18	38

Ketiga, analisis daya beda. Analisis daya beda bertujuan untuk mengetahui jenis butir soal yang bisa membedakan antara muta'allimah yang menguasai materi ataupun sebaliknya. Analisis ini juga diperlukan dalam pembuatan butir-butir soal yang akan diujikan selanjutnya, sebab jika butir soal tersebut tidak bisa membedakan kemampuan peserta didik maka butir soal tersebut dianggap kurang baik untuk digunakan dalam alat evaluasi peserta didik. Analisis daya beda ini memiliki lima kategori soal yaitu tergolong jelek, cukup, baik, baik sekali, dan tidak pantas atau disimbolkan dengan hasil minus (-) nantinya.

Hasil analisis daya beda terhadap butir soal alat evaluasi kelulusan LPBA Al-Yasini Pasuruan adalah sebagai berikut: 1) butir soal yang tergolong jelek ada 10 nomor, 2) butir soal yang tergolong cukup ada 16 nomor, 3) butir soal yang tergolong baik ada 20 nomor, 4) butir soal yang tergolong baik sekali ada 2 nomor, dan 5) butir soal yang tergolong tidak pantas ada 2 nomor. Perhitungan analisis daya beda ini juga sudah sesuai dengan rumus yang berlaku. Perhatikan tabel di bawah ini:

Tabel 1: Analisis Daya Beda

Kriteria Soal	Prosentase	Jumlah Soal
Jelek	0,00 – 0,20	10
Cukup	0,21 – 0,40	16
Baik	0,41 – 0,70	20
Baik Sekali	0,71 – 1,00	2
Tidak Pantas	(-)	2

Keempat, analisis pola jawaban. Analisis pola jawaban ini bertujuan untuk mengetahui fungsi pengecoh yang disajikan sebagai pilihan ganda dari masing-masing butir soal. Analisis pola jawaban ini juga berfungsi untuk mengetahui jenis-jenis pilihan ganda yang baik dan benar dan sesuai untuk dimuat sebagai pilihan ganda dari masing-masing butir soal. Terkadang ada beberapa pilihan ganda yang tidak berfungsi sebagai pengecoh sehingga peserta didik tidak ada yang menjawab pilihan ganda tersebut. Jika terjadi seperti itu, maka alangkah baiknya bila pilihan ganda tersebut tidak dimuat kembali sebagai alat evaluasi.

Dalam analisis pola jawaban ini hanya ada 2 kategori, yaitu jika masing-masing pilihan ganda tersebut dipilih oleh peserta didik dalam prosentase 5% ke atas maka baik untuk dimuat kembali dalam tes berikutnya. Jika pilihan ganda tersebut tidak terpilih sama sekali dari peserta didik yang mengikuti tes, maka harus dibuang dan tidak boleh dimuat kembali dalam tes-tes selanjutnya. Berdasarkan analisis pola jawaban soal terhadap alat evaluasi LPBA Al-Yasini Pasuruan yang terdiri dari 50 butir soal, ada beberapa soal yang memiliki jawaban pengecoh. Diantaranya adalah sebagai berikut:

a. Soal nomor 4, siswa diminta menentukan isim muannats. Dari tiga pilihan jawaban semuanya diakhiri *ta' marbuthoh*. Padahal tidak semua yang diakhiri *ta' marbuthoh* dikatakan *isim muannats*.

b. Soal nomor 13, siswa diminta menentukan terjemahan dari "14 kitab". 2 dari 3 pilihan jawaban *ma'dud*-nya berupa isim mufrod dan 1 jawaban sisanya jama'. Jika *adad* berupa *isim jama'* maka *ma'dud* belum tentu harus *jama'.* ada beberapa kaidah yang membuat ma'dud harus tetap *mufrod*.

c. Soal nomor 14, siswa diminta menentukan terjemahan dari kata "11 murid perempuan". Seperti soal nomor 13, 2 dari 3 pilihan jawaban ma'dudnya berupa isim jama'. Padahal *ma'dud* tidak selalu jama' walaupun *adad*-nya berupa *isim jama'*.

d. Soal nomor 17, siswa diminta menentukan *'alamat rafa'* dari *asma'ul khomsah*. *'alamat rofa'* ada beberapa macam, yang paling identik adalah dhommah. Jika siswa tidak jeli mereka akan terkecoh dengan jawaban *dhommah*. Padahal tidak semua isim memiliki *'alamat rofa' dhommah*.

e. Soal nomor 22, siswa diminta menerjemahkan kalimat jumlah fi'liyah yang memiliki dua fa'il (mutsana). Aturan dalam jumlah fi'liyah adalah fi'il didahulukan sehingga berapapun

jumlah fa'il, fi'ilnya tetap satu. Jika tidak jeli siswa akan memilih jawaban A yang mana fi'ilnya berupa mutsana.

f. Soal nomor 27, siswa diminta menentukan *kaana* dan saudara-saudaranya. Setiap jawaban menyediakan 3 saudara-saudara kaana. Untuk menjawab soal ini siswa diharuskan menghafal kaana dan saudara-saudaranya. Jika siswa tidak benar-benar hafal maka mereka akan terjebak dengan salah satu jawaban yang salah.

g. Soal nomor 30, siswa diminta menentukan pengamalan *kaana* dalam kalimat. Siswa yang tidak memahami pengalaman *kaana*, mereka akan terjebak pada jawaban A. Begitupun dengan jawaban C, jika tidak jeli siswa akan beranggapan *khobar* dari pilihan jawaban C adalah *isim tasniyah*. Padahal tidak demikian.

DISKUSI

Berdasarkan hasil analisis yang dilakukan terhadap butir-butir soal bahasa Arab untuk penilaian akhir semester di LPBA Al Yasini Pasuruan telah memenuhi 4 karakteristik ciri soal yang baik. Akan tetapi hal ini tidak berlaku secara mutlak, selain 1 karakterisik lagi yang belum dibahas dalam penelitian ini. *Pertama,* dari karakteristik validitas soal. Dari 50 soal yang disajikan semuanya dinyatakan valid karena sesuai dengan materi yang diajarkan di LPBA Al Yasini Pasuruan. Meski demikian butir-butir soal ini memiliki tingkat kesulitan yang baik. Hal ini dapat dilihat dari hasil analisisnya bahwa peserta yang berhasil menjawab dengan benar tidak mencapai 40%.

Kemudian berdasarkan analisis daya bedanya soal ini belum bisa digunakan untuk membedakan kemampuan siswa, apakah mereka mempunyai kemampuan tingkat tinggi, sedang maupun rendah. Hal ini dikarenakan masih ada 10 soal yang memiliki daya beda jelek dan 16 soal yang daya bedanya masih pada kategori cukup.

CONCLUSIONS | خاتمة | SIMPULAN

Berdasarkan analisis yang telah dilakukan peneliti terhadap butir soal bahasa Arab untuk ujian akhir semester di pondok pesantren Al Yasini Pasuruan, dapat disimpulkan bahwa butir-butir soal tersebut dinyatakan 100% valid yang terbagi atas 3 kategori soal mudah, sedang, dan sulit. Selain itu butir-butir soal bahasa Arab tersebut juga dianalisis berdasarkan daya beda untuk mengetahui perbedaan kemampuan pembelajar sehingga diperoleh 5 kategori yaitu jelek, cukup, baik, baik sekali, dan tidak pantas. Selain menganalisis soal, penelitian ini juga menganalisis pola jawaban sebagi tingkat pengecoh yang hasilnya 144 pilihan ganda boleh dimuat kembali dan 6 sisanya tidak boleh dimuat kembali.

Penelitian ini belum mengungkapkan reliabilitas butir soal bahasa Arab untuk ujian akhir semester di pondok pesantren Al Yasini Pasuruan. Oleh karena itu, diharapkan penelitian selanjutnya bisa mengungkapkan reliabilitas atau analisis kejelekan butir soal tersebut berdasarkan data-data yang akurat. Selain itu pada ujian akhir semester selanjutnya diharapakan LPBA Miftahul Ulum Al Yasini bisa menambah jumlah butir soal kategori mudah agar seimbang dengan soal kategori sulit baik dari segi jumlah dan persentasenya.

ACKNOWLEDGEMENTS | شكر وتقدير | TERIMA KASIH

Penulis mengucapkan terima kasih kepada keluarga yang memberi dukungan dan

semangat baik secara lahir dan batin. Selanjutnya penulis juga mengucapkan terima kasih sebesar-besarnya kepada dosen pembimbing yang telah memberi saran dan bimbingan terhadap penulisan artikel ini. Ucapan terima kasih juga tidak lupa penulis ucapkan kepada pengajar di LPBA Al Yasini Pasuruan yang membantu penulis mengumpulkan soal-soal ujian akhir bahasa Arab di lembaga tersebut. Semoga artikel ini dapat menjadi salah satu referensi pedoman pembuatan soal ujian akhir bahasa Arab yang baik, khususnya yang memenuhi kriteria yang telah disebutkan.

DAFTAR PUSTAKA

Ainin, M. (2018). Penilaian Dalam Pembelajaran Bahasa Arab Di Madrasah Atau Sekolah: Hots, Mots Atau Lots? In JSA UM (Ed.), *Prosiding Konferensi Nasional Bahasa Arab* (Vol. 4, Issue 4, pp. 155–165).

Ainin, M. (2019). Kesalahan Dalam Penyusunan Rencana Pembelajaran Bahasa Arab di Madrasah/Sekolah. *Prosiding Konferensi Nasional Bahasa Arab V*, 199–209.

Ainin, M. (2021). *Kesahihan Instrumen Penialaian Dalam Pba: Sebuah Prasyarat Penilaian Berbasis Ict* [Slide]. Webinar Penilaian Pembelajaran Bahasa Arab, Malang.

Ainin, Moh. (2020). *Penilaian Berfikir Tingkat Tinggi (HOTS) Dalam Pembelajaran Bahasa Arab* (T. CV. B. Sejahtera, Ed.; 1st ed.). CV. Bintang Sejahtera.

Arifianto, M. L., Ainin, Moh., Irhamni, Ahasanuddin, M., Nikmah, K., Anwar, M. S., & Fitria, N. (2021). *Evaluasi Pembelajaran dan PengembanganTes Interaktif Bahasa Arab* (L. A. R. Muhammad Lukman Arifianto, Nurul Fitria, Ed.; 1st ed., Issue February, p. 202).

Arikunto, S. (2010). *Prosedur Penelitian: Suatu Pendekatan Praktik*. Rineka Cipta.

Khuli, M. ’Ali Al. (2000). *Ikhtibaaraat Al Lughawiyah*. Dar Al Falah.

Margono. (2010). *Metodologi Penelitian Pendidikan*. Rineka Cipta.

Ridho, U. (2018). Evaluasi Dalam Pembelajaran Bahasa Arab. *An Nabighoh Jurnal Pendidikan Dan Pembelajaran Bahasa Arab*, 19.

Wahab, M. A. (2013). Pengembangan Tes Bahasa Arab Standar di Indonesia. *Journal of Chemical Information and Modeling*, *53*(9), 1689–1699.