

Analisis Daya Beda Tes Hasil Belajar Bahasa Arab Siswa SMA Mamba'ul Hikmah Paron Ngawi

Maziyyatul Muslimah¹, Amila Widiyanti²

¹ Institut Agama Islam Negeri (IAIN) Kediri

² Institut Agama Islam Negeri (IAIN) Kediri

Info Artikel :

Diterima 01 Juli 2023

Direvisi 12 September 2023

Dipublikasikan 13 Oktober 2023

Kata Kunci:

Analisis Daya Pembeda

Soal Pilihan Ganda

Tes Bahasa Arab

Keywords:

Discriminating Power Analysis

Multiple Choices

Arabic Language Test

Abstrak

Kualitas butir soal pilihan ganda memiliki dampak signifikan pada evaluasi dan pengukuran pemahaman siswa terhadap materi pelajaran bahasa Arab. Dengan mengevaluasi kualitas butir soal, guru dapat mengidentifikasi area yang perlu diperbaiki dalam penyusunan soal, serta meningkatkan efektivitas proses pembelajaran secara keseluruhan. Penelitian ini bertujuan untuk mendeskripsikan kualitas butir soal pilihan ganda pada pelajaran bahasa Arab di SMA Mamba'ul Hikmah Paron Ngawi. Metode penelitian ini menggunakan metode deskriptif kuantitatif, data yang diperoleh dalam bentuk angka dan dianalisis melalui statistik menggunakan aspek pengukuran, perhitungan rumus dan kepastian data numerik. Penelitian ini bertujuan untuk mengetahui kekuatan butir soal pilihan ganda pada pelajaran bahasa Arab. Hasil penelitian ini menunjukkan bahwa 10 butir soal pilihan ganda terdapat hasil sebagai berikut: 3 butir soal yang memiliki kriteria sangat baik, yang menunjukkan bahwa soal dapat membedakan antara siswa yang memiliki kemampuan tinggi dengan siswa yang berkemampuan rendah. Terdapat 1 soal yang memiliki kriteria baik. Ada 1 soal yang memiliki kriteria cukup. Dan terdapat 5 butir soal yang memiliki kriteria buruk.

Abstract

The quality of multiple-choice question items has a significant impact on the evaluation and measurement of students' understanding of Arabic subject matter. By evaluating the quality of the question items, teachers can identify areas that need improvement in the preparation of the questions, as well as improve the effectiveness of the overall learning process. This study aims to describe the quality of multiple-choice questions in Arabic lessons at SMA Mamba'ul Hikmah Paron Ngawi. This research method uses quantitative descriptive methods, data obtained in the form of numbers and analyzed through statistics using aspects of measurement, formula calculation and certainty of numerical data. This study aims to determine the strength of multiple-choice question items in Arabic lessons. The results of this study showed that 10 multiple-choice questions contained the following results: 3 questions that had very good criteria, which showed that the questions could distinguish between students who had high abilities and students who had low abilities. There is 1 question that has good criteria. There is 1 question that has sufficient criteria. And there are 5 questions that have bad criteria.



This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2019 by author.

Koresponden:

Maziyyatul Muslimah,

Email: maziyya@iainkediri.ac.id

Pendahuluan

Pendidikan memegang peranan penting dalam menumbuhkan suatu bangsa menjadi generasi yang berkarakter dan memiliki jiwa kompetitif.¹ Pendidikan juga memiliki peran dalam pencapaian negara maju dan mandiri dalam pembangunan negara. Guru harus memiliki pengetahuan tentang penyelenggaraan pendidikan yang bermutu. Faktor internal pelaksanaan tujuan pendidikan umum adalah kompetensi guru dalam mencetak lulusan yang berkualitas. Sebagai seorang guru, harus memiliki potensi dalam pengetahuan yang berkaitan dengan perencanaan, pelaksanaan dan evaluasi pendidikan.²

Dalam dunia pendidikan, penilaian disamakan dengan istilah evaluasi. Evaluasi merupakan upaya yang dapat dilakukan untuk mengetahui bagaimana pencapaian hasil belajar dan meningkatkan pembelajaran. Evaluasi mengacu pada proses yang sistematis dalam menentukan tujuan pendidikan yang dicapai siswa.³ Pencapaian tujuan pembelajaran dilaksanakan melalui pengukuran dan evaluasi. Dengan pengukuran dan penilaian, guru dapat mengetahui kemampuan siswa sehingga mereka dapat mengetahui tindakan yang akan dilakukan untuk meningkatkan kemampuan siswa.

Daya pembeda (DB) adalah kemampuan butir soal THB (Tes Hasil Belajar) untuk membedakan kemampuan siswa yang memiliki kemampuan tinggi dan rendah.⁴ Analisis daya pembeda ini bertujuan untuk mengetahui kemampuan siswa yang tergolong mampu dan siswa yang Performanya relatif buruk.⁵ Jika soal diajukan kepada anak yang memiliki kemampuan unggul akan menunjukkan prestasi yang tinggi, dan jika diberikan kepada anak dengan tingkat kognisi rendah, maka hasilnya buruk. Semakin tinggi daya pembeda butir soal, semakin baik soal yang bisa membedakan antara siswa yang sudah faham dan tidak memahami materi. Jika daya pembeda negatif (-) atau memiliki indeks daya pembeda kurang dari 0 maka lebih banyak Kelompok bawah yang menjawab benar dari pada kelompok atas, atau daya pembeda tersebut tidak berfungsi. Daya pembeda memiliki beberapa tahapan. Langkah pertama adalah menentukan kekuatan daya pembeda, melalui kelompok kecil (kurang dari 100) dan kelompok besar (100 lebih).

Salah satu alat untuk menilai hasil belajar adalah tes. Tes adalah alat ukur yang digunakan terutama untuk menentukan kemampuan kognitif siswa. Tes biasanya menggunakan soal-soal yang objektif, terutama pada soal-soal pilihan ganda. Soal pilihan ganda merupakan bentuk tes

¹ Rahmasari, D., & Ismiyati. (2016). Analisis Butir Soal Mata Pelajaran Pengantar Administrasi perkantoran. *Economic Education Analysis Journal*, 5(1), 317–330

² Fitriatun, A., & Sukanti. (2016). Analisis Validitas, Reliabilitas dan Butir Soal Latihan Ujian Nasional Ekonomi Akuntansi di MAN Maguwaharjo. *Jurnal Kajian Pendidikan Akuntansi Indonesia*, 8(3), 1–11

³ Kusumawati, M., & Hadi, S. (2018). An analysis of Multiple Choice Questions (MCQs): Item and Test Statistics From Mathematics Assessments in Senior High School. *Research and Evaluation in Education*, 4(1), 70–78. <https://doi.org/10.21831/reid.v4i1.20202>

⁴ Purwanto. 2010. *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Belajar

⁵ Sudjana, N. 2012. *Penilaian Hasil Proses Belajar Mengajar*. Bandung: Remaja Rosdakarya Offset

⁶ Yusrizal. (2016). Analysis of Difficulty Level of Physics National Examination's Questions. *Jurnal Pendidikan IPA Indonesia*, 5(1), 140–149. <https://doi.org/10.15294/jpii.v5i1.5803>

dengan alternatif jawaban sehingga siswa dapat memilih jawaban yang menurutnya benar.⁶ Pertanyaan Pilihan ganda dapat mencakup lebih banyak materi, efisien dan memudahkan guru dalam penilaian. Tes dibagi menjadi dua bagian, yaitu tes standar dan tes yang disiapkan oleh guru. Tes standar adalah tes yang telah dibakukan melalui proses validitas dan reliabilitas. Sehingga tes dinyatakan valid dan reliabel untuk menguji siswa. Sebaliknya Tes yang diberikan oleh seorang guru adalah tes yang bertujuan untuk mengetahui hasil belajar peserta didik. Tes yang dilakukan oleh seorang guru terbatas pada guru yang sama atau satu sekolah.⁷

Pada pembelajaran Bahasa Arab, perancangan soal tes yang baik memiliki peranan krusial dalam mengevaluasi pemahaman siswa terhadap materi pembelajaran serta mengukur sejauh mana siswa telah berhasil mencapai tujuan pembelajaran yang telah ditetapkan⁸. Tes biasanya memfokuskan pada evaluasi kemampuan kognitif siswa, memastikan bahwa penilaian dilakukan secara obyektif dan terukur⁹. Selain itu, tes juga menjadi alat bagi guru untuk mengukur efektivitas metode pengajaran yang telah diterapkan, mengidentifikasi area di mana siswa mungkin menghadapi kesulitan, dan mengadaptasi pendekatan pembelajaran untuk memenuhi kebutuhan individu siswa¹⁰. Dengan demikian, tes yang diberikan oleh guru berfungsi sebagai alat evaluasi dan alat bantu pengajaran yang membantu menciptakan lingkungan pembelajaran yang lebih responsif dan efisien¹¹.

Sebagai opsi alternatif, soal pilihan ganda memiliki peranan yang sangat penting dalam tes bahasa Arab¹². Dengan adanya pilihan jawaban yang tersedia, siswa dapat diuji dalam kemampuan mereka untuk menganalisis informasi, mengenali pola, dan membuat keputusan berdasarkan pemahaman terhadap materi¹³. Keunggulan lain dari soal pilihan ganda adalah dapat mencakup berbagai aspek materi secara lebih komprehensif¹⁴. Dalam beberapa pertanyaan, berbagai elemen materi dapat diuji secara efisien, memungkinkan guru untuk menilai sejumlah kompetensi sekaligus dalam satu sesi tes.

⁶ Amelia, M. A. (2017). Analisis Soal Tes Hasil Belajar High Order Thinking Skills (HOTs) Matematika Materi Pecahan Untuk Kelas 5 Sekolah Dasar. *Jurnal Penelitian*, 20(2), 123–131. Retrieved from <http://ejournal.usd.ac.id/index.php/JP/article/viewFile/869/685>

⁷ Septiana, N. (2016). Analisis Butir Soal Ulangan Akhir Semester (Uas) Biologi Tahun Pelajaran 2015/2016 Kelas X Dan Xi Pada Man Sampit. *Jurnal Pendidikan Sains Dan Matematika*, 4(2), 115–121

⁸ Rohman, F. (2014). Strategi Pengelolaan Komponen Pembelajaran Bahasa Arab. *Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban*, 1(1), 63-78.

⁹ Kamila, I. Z. (2022). *EVALUASI PEMBELAJARAN BAHASA ARAB DENGAN PENILAIAN AUTENTIK DI KELAS XI IPS MADRASAH ALIYAH AL AMIRIYAH BLOKAGUNG BANYUWANGI TAHUN AJARAN 2021-2022* (Doctoral dissertation, INSTITUT AGAMA ISLAM DARUSSALAM BLOKAGUNG BANYUWANGI).

¹⁰ Tanduklangi, A., & Amri, C. (2019). *Manajemen Sumber Daya Pembelajaran Bahasa Berbantuan Komputer: Computer Assisted Language Learning*. Deepublish.

¹¹ Amrullah, A. F., & Hum, S. (2021). *Manajemen Kurikulum Pembelajaran Bahasa Arab*. Prenada Media.

¹² Nashrullah, M. (2021). Penggunaan Media Google Form Dalam Evaluasi Pembelajaran Bahasa Arab (Pilihan Ganda). *Naskhi: Jurnal Kajian Pendidikan dan Bahasa Arab*, 3(1), 30-40.

¹³ Munip, A. (2017). Penilaian Pembelajaran Bahasa Arab. *Yogyakarta: FITK UIN Sunan Kalijaga*.

¹⁴ Ainin, M. (2016). Kesahihan Dalam Penyusunan Tes Bahasa Arab Di Madrasah/Sekolah. *Prosiding Konferensi Nasional Bahasa Arab*, 1(2).

Dalam mengembangkan soal pilihan ganda Bahasa Arab, penting bagi guru untuk memperhatikan kedalaman pemahaman yang diuji dan berusaha menciptakan variasi opsi jawaban yang mewakili berbagai alternatif yang mungkin dipertimbangkan oleh siswa¹⁵. Daya pembeda yang baik akan memastikan bahwa tes bahasa Arab dengan soal pilihan ganda dapat menjadi alat evaluasi yang akurat dan berharga dalam mengukur pemahaman siswa terhadap materi serta kemampuan siswa dalam menerapkan konsep-konsep bahasa Arab dalam konteks yang berbeda¹⁶.

Pada penyusunan soal pilihan ganda untuk tes Bahasa Arab, daya beda menjadi faktor utama yang perlu dipertimbangkan¹⁷. Daya beda merujuk pada tingkat kesulitan dan perbedaan antara opsi jawaban yang benar dan opsi jawaban yang salah. Dalam konteks pembelajaran bahasa Arab, daya beda yang baik dapat mengukur tingkat pemahaman dan penguasaan siswa terhadap berbagai aspek bahasa, seperti tata bahasa, kosakata, konteks komunikatif, dan nuansa budaya¹⁸.

Dengan mengatur daya beda yang tepat pada soal tes bahasa Arab pilihan ganda, guru dapat menyesuaikan tingkat kesulitan soal sesuai dengan tingkat kemampuan siswa¹⁹. Soal-soal dengan daya beda yang beragam akan memungkinkan siswa yang berbeda-beda kemampuannya untuk diuji secara adil²⁰. Para siswa yang telah menguasai materi dengan baik akan mampu menjawab soal-soal dengan tingkat kesulitan yang lebih tinggi, sementara siswa yang masih dalam proses pembelajaran akan memiliki kesempatan untuk menjawab soal-soal dengan tingkat kesulitan yang lebih rendah.

Selain itu, penerapan daya beda yang baik juga mendorong siswa untuk benar-benar memahami konsep-konsep bahasa Arab yang diajarkan²¹. Ketika opsi jawaban yang salah dihasilkan dari pemahaman yang keliru atau tipuan yang halus, siswa harus dapat menganalisis dan memilah informasi dengan cermat untuk menemukan jawaban yang benar. Hal ini tidak hanya mengukur pengetahuan, tetapi juga mengembangkan keterampilan kritis dan analitis siswa dalam memecahkan masalah bahasa²².

Mempertimbangkan dampak yang sangat penting dari kualitas butir soal pilihan ganda dalam pelaksanaan evaluasi serta pengukuran pemahaman siswa terhadap materi pelajaran

¹⁵ Rukajat, A. (2018). *Teknik evaluasi pembelajaran*. Deepublish.

¹⁶ Widodo, H. (2021). *Evaluasi Pendidikan*. UAD PRESS.

¹⁷ Ida, F. F., & Musyarofah, A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *AL-MU'ARRIB: JOURNAL OF ARABIC EDUCATION*, 1(1), 34-44.

¹⁸ Arifianto, M. L., Ainin, M., Ahsanuddin, M., Irhamni, I., Fitria, N., Nikmah, K., & Anwar, M. S. (2021). Evaluasi Pembelajaran dan Pengembangan Tes Interaktif Bahasa Arab.

¹⁹ Hajaroh, S. (2016). Perbandingan Reliabilitas Butir Soal Bahasa Arab Antara Pilihan Ganda dengan Menjodohkan pada Tes Buatan Guru. *EL-HIKMAH: Jurnal Kajian dan Penelitian Pendidikan Islam*, 10(2), 157-178.

²⁰ Pribadi, W. S. (2014). *Mobile computer based test menggunakan metode item response theory untuk evaluasi pembelajaran Bahasa Arab* (Doctoral dissertation, Universitas Islam Negeri Maulana Malik Ibrahim).

²¹ Munip, A. (2017). Penilaian Pembelajaran Bahasa Arab. *Yogyakarta: FITK UIN Sunan Kalijaga*.

²² Ambiya, I. Z., Mulyawan, S., & Saefuloh, H. (2022). Analisis Soal Ujian Mata Pelajaran Bahasa Arab Kelas 12 di Madrasah Aliyyah. *EL-IBTIKAR: Jurnal Pendidikan Bahasa Arab*, 11(1), 70-87.

bahasa Arab, penelitian ini dilaksanakan dengan tujuan untuk memberikan deskripsi yang komprehensif mengenai analisis daya beda soal pilihan ganda pada mata pelajaran bahasa Arab di lingkungan SMA Mamba'ul Hikmah Paron Ngawi. Dengan melakukan evaluasi terhadap kualitas butir soal, para guru diharapkan dapat mengidentifikasi area-area yang memerlukan perbaikan dalam penyusunan soal, sehingga mampu meningkatkan efektivitas keseluruhan proses pembelajaran.

Metode

Penelitian ini merupakan penelitian deskriptif kuantitatif. Pendekatan yang digunakan dalam penelitian ini adalah pendekatan kuantitatif karena data yang diperoleh dalam bentuk angka dan dianalisis melalui statistik menggunakan aspek pengukuran, perhitungan rumus dan kepastian data numerik. Dalam penelitian ini, semua siswa kelas X SMA Mamba'ul Hikmah berpartisipasi dalam percobaan tersebut pada semester genap. Teknik pengumpulan data dilakukan pada pembelajaran siswa kelas X antara lain:

- 1) Tes: dengan memberikan soal-soal latihan untuk mengukur pengetahuan dan penelitian ini menggunakan kemampuan individu berupa 10 soal pilihan ganda untuk siswa kelas X dan lembar jawaban akan dikumpulkan untuk dievaluasi kembali.
- 2) Wawancara: dengan melaksanakan tanya jawab secara lisan untuk memperoleh informasi yang bertujuan untuk mendapatkan informasi yang tepat dari narasumber terpercaya. Dalam hal ini, penulis melakukan komunikasi langsung dengan guru bahasa Arab kelas X melalui panduan wawancara yang sudah disiapkan. Tujuannya agar penulis mengetahui tantangan yang muncul saat proses pemberian tes kepada siswa.

Untuk menentukan hasil analisis daya pembeda, penulis menghitung daya pembeda menggunakan rumus:

$$DP = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

Keterangan:

- DP : Jumlah peserta tes
 J_A : Banyaknya peserta kelompok atas
 J_B : Banyaknya peserta kelompok bawah
 B_A : Banyaknya peserta kelompok atas yang menjawab soal dengan benar
 B_B : Banyaknya peserta kelompok bawah yang menjawab soal dengan benar
 P_A : Proporsi peserta kelompok atas yang menjawab soal dengan benar
 P_B : Proporsi peserta kelompok atas yang menjawab soal dengan benar

Dalam hal ini penulis menganalisis butir soal pilihan ganda berdasarkan data lembaran jawaban untuk memperoleh data daya pembeda. Hasil perhitungan daya pembeda butir soal dihubungkan dengan kriteria daya pembeda butir soal yaitu:

Tabel 1. Kriteria Daya Pembeda

Daya Pembeda	Kriteria Daya Pembeda
$DP \geq 0,70$	Baik sekali (digunakan)

$0,40 \leq DP < 0,70$	Baik digunakan (digunakan)
$0,20 \leq DP < 0,40$	
$DP < 0,20$	
	Cukup
	Buruk

Setelah data skor hasil uji coba diperoleh, diurutkan dari yang terbesar sampai terkecil. Kemudian dari mulai urutan teratas diambil 50% sebagai kelompok atas dan dari urutan paling bawah diambil 50% sebagai kelompok bawah. Sehingga banyak siswa kelompok atas = banyaknya siswa kelompok bawah yaitu 5 siswa.

Hasil dan Pembahasan

Daya pembeda dapat diukur dengan mengetahui jawaban yang benar untuk semua siswa kelas X, dan untuk mencari daya pembeda soal, pertama dengan membentuk kelompok atas dan kelompok bawah. Karena jumlah siswa ada 10, maka ditentukan 5 siswa kelompok atas yang terdiri dari siswa nomor 1-5 dan 5 siswa kelompok bawah yang terdiri dari siswa nomor 6-10. Setelah itu, jawaban peserta kelompok atas dan kelompok bawah dianalisis menggunakan rumus DB (Daya Pembeda). Hasil dari perhitungan rumus dicocokkan dengan tabel.

Tabel 2. Hasil Perhitungan Daya Pembeda Pada Soal Pilihan Ganda Bahasa Arab di SMA Mamba'ul Hikmah

No	Nama	Nomor Soal										Skor total
		1	2	3	4	5	6	7	8	9	10	
1	Siswa 1 (ZWA)	1	1	0	1	1	1	1	1	1	1	9
2	Siswa 2 (LNH)	1	1	0	1	0	0	1	1	1	0	6
3	Siswa 3 (ML)	1	1	0	0	0	1	1	0	1	1	6
4	Siswa 4 (WAT)	1	1	0	1	1	1	0	1	0	0	6
5	Siswa 5 (NZ)	1	0	0	1	0	1	1	0	1	0	5
6	Siswa 6 (HAM)	1	0	0	0	1	0	0	1	0	1	4
7	Siswa 7 (MAA)	1	0	0	1	0	1	0	0	0	1	4
8	Siswa 8 (RNS)	1	0	0	0	1	0	0	1	0	0	3
9	Siswa 9 (HAH)	1	0	0	1	0	0	0	1	0	0	3
10	Siswa 10 (WDF)	1	0	0	0	0	0	0	1	0	0	2
Jumlah Jawaban Benar		10	4	0	6	4	5	4	7	4	4	
Jumlah Peserta		10	10	10	10	10	10	10	10	10	10	

Berdasarkan hasil perhitungan daya pembeda pada soal pilihan ganda di SMA Mamba'ul Hikmah di Tabel 2 di atas, diperoleh informasi yang menggambarkan perincian kelompok siswa yang terlibat dalam analisis ini. Terdapat dua kelompok siswa yang diidentifikasi berdasarkan nomor soal, yaitu kelompok atas yang terdiri dari siswa dengan nomor 1 hingga 5, dan kelompok bawah yang mencakup siswa dengan nomor 6 hingga 10. Dengan demikian, kelompok siswa atas terdiri dari 5 orang siswa, sedangkan kelompok siswa bawah juga terdiri dari 5 orang siswa.

Pengelompokan ini didasarkan pada nomor soal yang dijawab oleh masing-masing siswa dan akan membantu dalam menganalisis daya pembeda dari setiap soal secara lebih rinci.

Informasi ini akan menjadi dasar penting dalam mengevaluasi sejauh mana setiap soal memiliki kemampuan untuk membedakan antara siswa yang memiliki pemahaman yang lebih baik dengan mereka yang memiliki pemahaman yang lebih rendah terhadap materi bahasa Arab yang diajarkan.

Tabel 3. Hasil Analisis Daya Pembeda Butir Soal Pilihan Ganda

No. soal	B_A	J_A	P_A	B_B	J_B	P_B	$DP = P_A - P_B$	Kriteria
1	5	5	1,00	5	5	1,00	0,00	Buruk
2	4	5	0,80	0	5	0,0	0,80	Sangat baik
3	0	5	0,00	0	5	0,0	0,00	Buruk
4	4	5	0,80	2	5	0,40	0,40	Cukup
5	2	5	0,40	2	5	0,40	0,00	Buruk
6	4	5	0,80	1	5	0,20	0,60	Baik
7	4	5	0,80	0	5	0,0	0,80	Sangat baik
8	3	5	0,60	4	5	0,80	-0,20	Buruk
9	4	5	0,80	0	5	0,0	0,80	Sangat baik
10	2	5	0,40	2	5	0,40	0,00	Buruk

Tabel 3 menggambarkan daftar nilai daya pembeda dari masing-masing butir soal pilihan ganda. Dalam tabel ini, terdapat informasi yang signifikan terkait dengan penilaian daya pembeda setiap soal, yang merupakan indikator penting dalam mengukur sejauh mana sebuah soal mampu membedakan antara siswa yang memiliki pemahaman yang lebih baik dengan mereka yang memiliki pemahaman yang lebih rendah terhadap materi bahasa Arab yang diajarkan.

Hasil penilaian daya pembeda diberikan dalam bentuk skor yang dihitung berdasarkan tingkat kesesuaian jawaban siswa dari kelompok atas (PA) dan kelompok bawah (PB). Selisih antara skor PA dan PB ($PA - PB$) menggambarkan nilai daya pembeda (DP) untuk setiap soal. Kriteria penilaian kemudian diberikan berdasarkan skor DP ini.

Misalnya, pada soal nomor 1, nilai DP adalah 0,00, yang mengindikasikan bahwa daya pembeda dari soal tersebut adalah buruk. Di sisi lain, pada soal nomor 2, nilai DP adalah 0,80, menunjukkan bahwa soal tersebut memiliki daya pembeda yang sangat baik. Pada soal nomor 8, nilai DP adalah -0,20, menandakan bahwa soal tersebut memiliki daya pembeda yang buruk dan mungkin perlu diperbaiki.

Soal yang baik adalah soal yang tidak terlalu mudah dan juga tidak terlalu sukar, soal yang mudah tidak mendorong siswa untuk mencoba memecahkan soal dan soal yang sulit menurunkan semangat siswa dalam memecahkan soal atau putus asa²³. Sebaliknya, soal yang

²³ Salmina, M. (2016). No Title. Analisis Soal Ujian Matematika, 4(April 2017), 37-47

terlalu sulit membuat siswa tertekan karena tidak bisa menjawab²⁴. Artinya semakin tinggi daya pembeda soal, semakin banyak peserta dalam kelompok yang dapat menjawab soal dengan benar, dan semakin sedikit peserta dari kelompok bawah yang dapat menjawab soal dengan benar. Hasil dari perhitungan dan pencocokan dengan kriteria daya pembeda adalah sebagai berikut:

Tabel 4. Kriteria Soal dari Hasil Analisis Daya Pembeda

No	Kategori	Nomor soal	Jumlah soal
1	Sangat baik	2, 7, 9	3
2	Baik	6	1
3	Cukup	4	1
4	Buruk	1, 3, 5, 8, 10	5

Berdasarkan hasil analisis daya pembeda yang terdapat pada Tabel 3, terdapat 4 kategori soal dari 10 soal pilihan ganda yang ditampilkan pada Tabel 4. Terdapat 3 butir soal (no. 2, 7, 9) yang memiliki kriteria sangat baik, yang menunjukkan bahwa soal dapat membedakan antara siswa yang memiliki kemampuan tinggi dengan siswa yang berkemampuan rendah, terdapat 1 soal (no. 6) yang memiliki kriteria baik dalam membedakan antara siswa yang berkemampuan tinggi dan rendah, ada 1 soal (no. 4) yang memiliki kriteria cukup untuk membedakan siswa yang berkemampuan tinggi dengan siswa berkemampuan rendah. Dan terdapat 5 butir soal (no. 1, 3, 5, 8, 10) yang memiliki kriteria buruk, sehingga tidak dapat membedakan antara siswa yang berkemampuan tinggi dan siswa yang berkemampuan rendah. Butir soal dengan kategori *No Discrimination* berarti siswa berkemampuan rendah menjawab benar lebih banyak daripada siswa yang berkemampuan tinggi, sehingga tidak cocok untuk digunakan kembali karena gagal membedakan kemampuan siswa.²⁵ Salah satu faktor yang menyebabkan nilai korelasi buruk dan negatif yaitu validitas butir soal yang tidak valid dan koefisien kesukaran setiap butir soal berbeda.

Berdasarkan hasil analisis daya pembeda butir soal pilihan ganda untuk mata Pelajaran bahasa Arab kelas X di SMA Mamba'ul Hikmah Paron Ngawi, peneliti memberikan saran dan rekomendasi kegiatan tindak lanjut sebagai berikut:

1. Butir-butir yang sudah memiliki resolusi yang baik (memuaskan, baik, dan sangat baik) harus ditambahkan di bank soal untuk diberikan lagi dalam ujian nanti, karena kualitasnya sudah cukup.

²⁴ Kadir, A. (2015). MENYUSUN DAN MENGANALISIS TES HASIL BELAJAR Abdul Kadir. Jurnal Al-Ta'dib, 8(2), 70–81.

²⁵ Hanifah, N. (2014). Perbandingan Tingkat Kesukaran, Daya Pembeda Butir Soal dan Reliabilitas Tes Bentuk Pilihan Ganda Biasa dan Pilihan Ganda Asosiasi Mata Pelajaran Ekonomi. SOSIO E-KONS, 6(1), 41–55. <https://doi.org/http://dx.doi.org/10.30998/sosioekons.v6i1.1715>

2. Objek yang masih beresolusi rendah (buruk), ada 2 opsi:
 - a. Diteliti untuk diperbaiki, setelah diperbaiki dapat digunakan dalam tes hasil belajar, yang kemudian dianalisis ulang untuk melihat apakah itu meningkat atau tidak.
 - b. Ditolak atau dikecualikan dan tidak diterbitkan kembali untuk soal pilihan ganda. Terutama untuk judul dengan angka indeks diskriminatif objek negatif sebaiknya tidak perlu lagi diujikan kembali karena item tersebut kualitasnya sangat buruk (kelompok atas lebih banyak menjawab salah dari pada kelompok bawah, bahkan hanya sedikit menjawab salah)²⁶.

Kesimpulan

Berdasarkan analisis daya pembeda pada soal pilihan ganda di SMA Mamba'ul Hikmah, dari 10 soal pilihan ganda ada 3 butir soal yang memiliki kriteria sangat baik, yang menunjukkan bahwa soal dapat membedakan antara siswa yang memiliki kemampuan tinggi dengan siswa yang berkemampuan rendah, terdapat 1 soal yang memiliki kriteria baik dalam membedakan antara siswa yang berkemampuan tinggi dan rendah, ada 1 soal yang memiliki kriteria cukup untuk membedakan siswa yang berkemampuan tinggi dengan siswa berkemampuan rendah. Dan terdapat 5 butir soal yang memiliki kriteria buruk, sehingga tidak dapat membedakan antara siswa yang berkemampuan tinggi dan siswa yang berkemampuan rendah. Butir soal yang memperoleh nilai sangat baik, dan baik dapat dimasukkan ke dalam bank soal. Sedangkan butir soal yang mendapatkan nilai cukup dan buruk, dapat diperbaiki atau dikecualikan pada penerbitan soal di lain waktu.

Penelitian lanjutan perlu difokuskan pada beberapa area untuk meningkatkan kualitas bank soal pilihan ganda. (1) Butir soal dengan kriteria buruk harus menjadi fokus utama dalam penelitian mendalam guna mengidentifikasi faktor penyebab dan potensi perbaikan. (2) Butir soal dengan kriteria cukup dan buruk dapat ditingkatkan melalui restrukturisasi, format yang lebih sesuai, serta pengembangan konten yang relevan. Penelitian ini harus diikuti dengan pengujian ulang untuk mengukur peningkatan daya pembeda. (3) Eksplorasi metode penilaian alternatif yang memperkuat daya pembeda, seperti penggunaan format dan pendekatan evaluasi yang inovatif, diperlukan untuk mencapai hasil yang lebih efektif. (4) Prinsip pengembangan soal yang lebih baik, termasuk konten yang relevan, kejelasan pertanyaan, dan integrasi konteks budaya, harus diperhatikan dalam mengoptimalkan kualitas daya pembeda. Selanjutnya, (5) Evaluasi berkala pada bank soal perlu dilakukan untuk memastikan kesesuaian dan validitas konten.

Dengan mengambil langkah-langkah ini, penelitian lanjutan diharapkan dapat berkontribusi pada pengembangan bank soal yang memiliki daya pembeda yang lebih baik dalam pembelajaran bahasa Arab di SMA Mamba'ul Hikmah Paron Ngawi.

²⁶ Sudjiono, Anas, Pengantar Evaluasi Pendidiksn, Jakarta: Rajawali Press, 2016.

Referensi

- Ainin, M. (2016). Kesahihan Dalam Penyusunan Tes Bahasa Arab Di Madrasah/Sekolah. *Prosiding Konferensi Nasional Bahasa Arab*, 1(2).
- Ambiya, I. Z., Mulyawan, S., & Saefuloh, H. (2022). Analisis Soal Ujian Mata Pelajaran Bahasa Arab Kelas 12 di Madrasah Aliyyah. *EL-IBTIKAR: Jurnal Pendidikan Bahasa Arab*, 11(1), 70-87.
- Amelia, M. A. 2017. *Analisis Soal Tes Hasil Belajar High Order Thinking Skills (HOTS) Matematika Materi Pecahan Untuk Kelas 5 Sekolah Dasar*. *Jurnal Penelitian*, 20(2), Retrieved from <http://ejournal.usd.ac.id/index.php/JP/article/viewFile/869/685>
- Amrullah, A. F., & Hum, S. (2021). *Manajemen Kurikulum Pembelajaran Bahasa Arab*. Prenada Media.
- Arifianto, M. L., Ainin, M., Ahsanuddin, M., Irhamni, I., Fitria, N., Nikmah, K., & Anwar, M. S. (2021). Evaluasi Pembelajaran dan Pengembangan Tes Interaktif Bahasa Arab.
- Fitriatun, A., & Sukanti. 2016. *Analisis Validitas, Reliabilitas dan Butir Soal Latihan Ujian Nasional Ekonomi Akuntansi di MAN Maguwaharjo*. *Jurnal Kajian Pendidikan Akuntansi Indonesia*, 8(3).
- Hanifah, N. 2014. Perbandingan Tingkat Kesukaran, Daya Pembeda Butir Soal dan Reliabilitas Tes Bentuk Pilihan Ganda Biasa dan Pilihan Ganda Asosiasi Mata Pelajaran Ekonomi. *SOSIO E-KONS*, 6(1) <https://doi.org/http://dx.doi.org/10.30998/sosioekons.v6i1.1715>
- Hajaroh, S. (2016). Perbandingan Reliabilitas Butir Soal Bahasa Arab Antara Pilihan Ganda dengan Menjodohkan pada Tes Buatan Guru. *EL-HIKMAH: Jurnal Kajian dan Penelitian Pendidikan Islam*, 10(2), 157-178.
- Ida, F. F., & Musyarofah, A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *AL-MU'ARRIB: JOURNAL OF ARABIC EDUCATION*, 1(1), 34-44. <https://doi.org/10.32923/al-muarrif.v1i1.2100>
- Kamila, I. Z. (2022). *EVALUASI PEMBELAJARAN BAHASA ARAB DENGAN PENILAIAN AUTENTIK DI KELAS XI IPS MADRASAH ALIYAH AL AMIRIYYAH BLOKAGUNG BANYUWANGI TAHUN AJARAN 2021-2022* (Doctoral dissertation, INSTITUT AGAMA ISLAM DARUSSALAM BLOKAGUNG BANYUWANGI).
- Kusumawati, M., & Hadi, S. 2018. *An analysis of Multiple Choice Questions (MCQs): Item and Test Statistics From Mathematics Assessments in Senior High School*. *Research and Evaluation in Education*, 4(1). <https://doi.org/10.21831/reid.v4i1.20202>
- Munip, A. (2017). *Penilaian Pembelajaran Bahasa Arab*. Yogyakarta: FITK UIN Sunan Kalijaga.

- Nashrullah, M. (2021). Penggunaan Media Google Form Dalam Evaluasi Pembelajaran Bahasa Arab (Pilihan Ganda). *Naskhi: Jurnal Kajian Pendidikan dan Bahasa Arab*, 3(1), 30-40.
- Pribadi, W. S. (2014). *Mobile computer based test menggunakan metode item response theory untuk evaluasi pembelajaran Bahasa Arab* (Doctoral dissertation, Universitas Islam Negeri Maulana Malik Ibrahim).
- Purwanto. 2010. Evaluasi Hasil Belajar . Yogyakarta: Pustaka Belajar Sudjana, N. 2012. Penilaian Hasil Proses Belajar Mengajar. Bandung: Remaja Rosdakarya Offset.
- Rahmasari, D., & Ismiyati. 2016. Analisis Butir Soal Mata Pelajaran Pengantar Administrasi perkantoran. *Economic Education Analysis Journal*, 5(1).
- Rohman, F. (2014). Strategi Pengelolaan Komponen Pembelajaran Bahasa Arab. *Arabiyat: Jurnal Pendidikan Bahasa Arab dan Kebahasaaraban*, 1(1), 63-78.
- Rukajat, A. (2018). *Teknik evaluasi pembelajaran*. Deepublish.
- Salmina, M. 2016. No Title. Analisis Soal Ujian Matematika, 4(April 2017), 37-47 Kadir, A. 2015. *Menyusun Dan Menganalisis Tes Hasil Belajar*. Jurnal Al-Ta'dib, 8(2)
- Septiana, N. 2016. Analisis Butir Soal Ulangan Akhir Semester (Uas) Biologi Tahun Pelajaran 2015/2016 Kelas X Dan Xi Pada Man Sampit. *Jurnal Pendidikan Sains Dan Matematika*, 4(2).
- Sudjiono, Anas. Pengantar Evaluasi Pendidikan, Jakarta: Rajawali Press, 2016.
- Tanduklangi, A., & Amri, C. (2019). *Manajemen Sumber Daya Pembelajaran Bahasa Berbantuan Komputer: Computer Assisted Language Learning*. Deepublish.
- Widodo, H. (2021). *Evaluasi Pendidikan*. UAD PRESS.
- Yusrizal. 2016. Analysis of Difficulty Level of Physics National Examination's Questions. *Jurnal Pendidikan IPA Indonesia*, 5(1). <https://doi.org/10.15294/jpii.v5i1.5803>