

Item Analysis of Multiple Choice Questions of Nahwu Class XII Madrasah Aliyah Pondok Pesantren Islam Al Mukmin Ngruki Sukoharjo

Analisis Butir Soal Pilihan Ganda Mata Pelajaran Nahwu Kelas XII Madrasah Aliyah Pondok Pesantren Islam Al Mukmin Ngruki Sukoharjo

Istianah¹, Nisa Arrohmah², Nursafitri³, Imam Makruf⁴

^{1,2,3,4} UIN Raden Mas Said, Surakarta, Indonesia

Article Info

Article History

Submitted 04-12-2023

Accepted 05-01-2024

Published 07-01-2024

Keywords:

Question Analysis,
Evaluation, Nahwu,
Arabic Language
Education,
Pesantren Ngruki

Correspondence:

224041011@mhs.uinsaid.ac.id

Abstract

The purpose of this research is to analyze the question items of nahwu class XII of Madrasah Aliyah Pondok Pesantren Islam Al-Mu'min Ngruki. This research uses a quantitative approach and descriptive method. Based on the results of the study, from a total of 30 questions seen from the difficulty of the questions found 3 items categorized as Very Difficult, 2 items categorized as Difficult, 10 items categorized as Moderate, 11 items categorized as Easy, and 4 items categorized as Very Easy. As for the differential power, there is 1 item with a negative index, which means that the item cannot be used, namely number 2, 10 items in the Weak category that cannot be used, 6 items in the Medium category that must be corrected, and 13 items in the Good category, which can be used but requires improvement. Meanwhile, in terms of distractors or question triggers, it can be classified as good if selected $\geq 5\%$ by students. there are 5 items categorized as Not Good because the distractors / triggers do not function properly, namely numbers 9, 24, 25, 27, and 30. And there are 11 items categorized as Less Good, 6 items categorized as Fair, 4 items categorized as Good, and 4 items categorized as Very Good because the distractors / triggers function very well.

Tujuan penelitian ini untuk menganalisis butir soal mata Pelajaran nahwu kelas XII Madrasah Aliyah Pondok Pesantren Islam Al-Mu'min Ngruki. Penelitian ini menggunakan pendekatan kuantitatif dan metode deskriptif. Berdasarkan hasil penelitian, dari total 30 soal dilihat dari kesukaran soal ditemukan 3 butir soal berkategori Sangat Sukar, 2 butir soal berkategori Sukar, 10 butir soal berkategori Sedang, 11 butir soal berkategori Mudah, dan 4 butir soal berkategori Mudah Sekali. Adapun dari segi daya beda, terdapat 1 butir soal yang berindeks Negatif yang maknanya butir soal tersebut tidak dapat digunakan yaitu nomor 2, 10 butir soal berkategori Lemah yang tidak bisa digunakan, 6 butir soal berkategori Sedang yang harus diperbaiki, dan 13 butir soal berkategori Baik yaitu bisa digunakan namun memerlukan perbaikan. Sedangkan dari segi distraktor atau pengecoh soal bisa tergolong baik jika dipilih $\geq 5\%$ oleh siswa. terdapat 5 butir soal berkategori Tidak Baik karena distraktor / pengecoh tidak berfungsi dengan baik yaitu nomor 9, 24, 25, 27, dan 30. Dan terdapat 11 butir soal yang berkategori Kurang Baik, 6 butir soal berkategori Cukup, 4 butir soal berkategori Baik, serta 4 butir soal berkategori Sangat Baik karena distraktor / pengecoh berfungsi sangat baik.

A. PENDAHULUAN

Dalam proses pendidikan terjadi proses belajar mengajar antara guru dan murid yang sistematis, yang melibatkan berbagai macam komponen (Hidayat & Syafe'i, 2018). Masing-masing komponen pengajaran tidak bersifat terpisah atau berjalan sendiri tetapi saling berkesinambungan antara satu dengan yang lainnya. Miladya (2003) mencatat proses belajar mengajar pada dasarnya adalah adanya interaksi antara guru dan murid untuk mencapai tujuan pendidikan. Guru sebagai pengarah dalam proses pendidikan, sedangkan siswa sebagai orang yang mengalami dan terlibat aktif dalam proses pendidikan, dan diharapkan dalam memperoleh perubahan yang lebih baik setelah proses pembelajaran dilaksanakan. Maka, Hidayat & Asyafah (2019) mencatat guru bertugas melakukan suatu kegiatan yang bertujuan untuk melihat sejauh mana ketercapaian proses pembelajaran yang telah dilaksanakan, yaitu dengan penilaian atau evaluasi.

Tugas seorang guru selain mengajar dan membimbing, adalah menilai hasil belajar peserta didiknya. Kegiatan menilai dalam proses pembelajaran disebut juga sebagai evaluasi pembelajaran (Hidayat et al., 2018). Evaluasi diperlukan guna melihat tingkat keberhasilan pelaksanaan pembelajaran yang akan dijadikan dasar pengambilan keputusan akhir. Evaluasi proses menentukan nilai suatu seperti tujuan, kegiatan, keputusan, unjuk kerja, proses, orang, ataupun objek secara sistematis berlandaskan kriteria tertentu (Elis Ratna Wulan & Rusdiana, 2015). Sehingga bisa dikatakan evaluasi hasil belajar merupakan suatu kegiatan terstruktur yang mencakup kegiatan penilaian dan pengukuran hasil belajar peserta didik, yang mana hasilnya dipergunakan untuk pengambilan keputusan atau membuat kebijakan (Taufik Rakhmat & Hidayat, 2022).

Evaluasi sering dilakukan setelah dilaksanakan kegiatan pembelajaran untuk mengukur tingkat ketercapaian tujuan pembelajaran (Sudin, 2011). Peraturan Pemerintah RI Nomor 19 Tahun 2017 tentang Perubahan atas Peraturan Pemerintah Nomor 74 Tahun 2008 tentang Guru Pasal 1 Ayat 1 menyatakan, "Guru adalah pendidik profesional dengan tugas utama mendidik, mengajar, membimbing, mengarahkan, melatih, menilai, dan mengevaluasi peserta didik pada anak usia dini, jalur pendidikan formal, pendidikan dasar, dan pendidikan menengah". Berlandaskan ayat tersebut diketahui salah satu tugas guru yaitu melaksanakan evaluasi dan penilaian terhadap peserta didik. Evaluasi dilaksanakan guna memastikan hasil yang dicapai peserta didik setelah melaksanakan proses pembelajaran. Melalui kegiatan evaluasi, guru bisa

mengetahui keberfungsiaan komponen pembelajaran yang diberikan (Hidayat & Kosasih, 2019).

Kegiatan evaluasi biasanya dilaksanakan dengan memberikan instrument tes berupa pengerjaan soal yang buatan sendiri ataupun berasal dari beberapa buku. Hal ini dikarenakan instrumen tes dianggap lebih efektif dalam menafsirkan tingkat pemahaman peserta didik terkait materi yang diberikan (Ridho, 2018). Menurut Mardapi, tes adalah salah satu cara menaksir tingkat kemampuan manusia secara tidak langsung, melalui respon seseorang terhadap stimulus atau pertanyaan (Mardapi, 2008). Tes pada pembelajaran digunakan untuk mengetahui kemampuan seseorang dan diberikan oleh guru pada penilaian harian, penilaian tengah semester, maupun penilaian akhir semester. Pemberian tes di akhir semester dilakukan guna mengukur tingkat penguasaan materi yang dipelajari selama satu semester (Kartowagiran & Jaedun, 2016).

Dalam penilaian akhir semester, biasanya bentuk tes yang digunakan adalah tes objektif dan subjektif. Soal dalam penilaian tengah semester sebaiknya dapat mengukur penguasaan materi selama satu semester dengan akurat (Elis Ratna Wulan & Rusdiana, 2015). Sebab itu, guru harus melakukan analisis butir soal guna memahami kelayakan soal yang diberikan kepada peserta didik (Widayanti et al., 2021).

Salah satu tujuan dari analisis butir soal adalah untuk mengadakan identifikasi soal-soal yang baik, kurang baik, dan soal yang jelek (Rozak, 2003). Dengan analisis butir soal dapat diperoleh informasi tentang kejelekan sebuah soal dan petunjuk untuk mengadakan perbaikan. Kapan sebuah soal dikatakan baik? Untuk memberikan jawaban terhadap pertanyaan ini, perlu diterangkan tiga masalah yang berhubungan dengan analisis butir soal, yaitu taraf kesukaran, daya pembeda, dan pola jawaban soal (Kaunang, Revoltje, 2010)

Hasil analisis soal akan menentukan butir soal yang layak untuk ditempatkan pada bank soal, direvisi, serta dibuang. Jadi, guru harus mampu membuat soal yang baik dengan menganalisisnya terlebih dahulu sebelum dan setelah diteskan kepada peserta didik (Darodjat et al., 2016). Soal dapat dinyatakan baik jika mencakup karakteristik penilaian butir soal yakni validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan keefektifan pengecoh. Soal dapat dinyatakan valid jika soal yang diberikan dapat mengukur tujuan yang ingin dicapai dari pemberian tes dengan tepat (Sridadi, 2002).

Soal juga harus bersifat reliabel, yang mana jika soal diujikan secara berulang di waktu yang berbeda tetap memperlihatkan ketetapan kesulitan soal. Analisis soal yang dilakukan dengan daya pembeda berguna untuk mengetahui kualitas soal berkriteria jelek, cukup, baik, atau sangat baik (Setiyawan, 2014).

Dalam melakukan analisis soal ada dua pendekatan yang bisa dilakukan, yaitu pendekatan kualitatif dan pendekatan kuantitatif (Widoyoko, 2012). Analisis soal pendekatan kualitatif dilakukan dengan menelaah soal-soal dan disesuaikan dengan kompetensi dasar, indikator pembelajaran hingga dilanjutkan dari segi aspek materi, konstruksi, dan bahasa sebelum diberikannya kepada peserta didik (Munip, 2017). Sedangkan analisis soal dengan pendekatan kuantitatif dilakukan setelah soal diberikan kepada peserta didik, lalu selanjutnya jawaban peserta didik dianalisis untuk menilai butir soal menggunakan *software computer* (agar lebih mudah dalam perhitungan) (Muradi, 2014).

Penelitian terdahulu yang relevan dengan penelitian ini:

Pertama, penelitian berjudul Analisis Butir Soal Ulangan Akhir Semester Ganjil Mata Pelajaran IPS Kelas IV SDS Sari Putra oleh Ina Magdalena dkk yang publish di Bintang: Jurnal Pendidikan dan Sains Vol. 3 No. 1 tahun 2021. Berdasarkan hasil penelitian, kualitas butir soal UAS ganjil mata pelajaran IPS kelas IV SDS SARI PUTRA ditinjau dari aspek materi, konstruksi, dan Bahasa memiliki validitas isi berkategori sangat tinggi (Magdalena et al., 2021).

Kedua, penelitian berjudul Analisis Butir Soal Pilihan Ganda Penilaian Tengah Semester Pada Pembelajaran Tematik Kelas V Sekolah Dasar Negeri 39 Pontianak Kota oleh Weni Widayanti dkk yang publish di Jurnal DIDIKA : Wahana Imiah Pendidikan Dasar Vol. 7 No. 2 tahun 2021. Berdasarkan hasil penelitian, soal penilaian tengah semester berkualitas baik dari segi validitasnya, dan reliabilitas. Namun Soal penilaian tengah semester belum berkualitas baik dilihat dari tingkat kesukaran, Daya pembeda yang diperoleh soal pada tema 6 dengan kategori "jelek. Soal penilaian tengah semester berkualitas baik dari segi efektivitas pengecoh (Widayanti et al., 2021).

Ketiga, penelitian berjudul Analisis Butir Soal Latihan pada Buku al-Nahw al-Wāḍiḥ untuk Siswa Madrasah Aliyah oleh Bagusradityo Aryobimo yang publish di Al-Ma'rifah: Jurnal Budaya, Bahasa dan Sastra Arab Vol. 18 No. 2 tahun 2021. Berdasarkan hasil penelitian, Hasil butir soal latihan buku al-Nahw al-Wāḍiḥ relatif berubah. Semua itu bisa dipengaruhi oleh faktor-faktor lain, seperti pemahaman para siswa, waktu berjalannya sebuah latihan atau tes, dan lain sebagainya. Dari hasil penelitian ini

peneliti menyarankan agar guru nahwu yang menggunakan buku ajar al-Nahw al-Wāḍiḥ jilid 1 ini untuk melihat terlebih dahulu kemampuan para siswa. Sebab, soal-soal yang baik adalah soal-soal yang sesuai dengan siswa. Jika soal-soal latihan dalam buku nahwu ini terlihat terlalu sulit atau terlalu mudah bagi para siswa, maka memodifikasi soal-soal latihan ini sangat diperlukan. Buku al-Nahw al-Wāḍiḥ ini adalah buku dengan soal-soal latihan yang cukup baik. Hanya saja, kosakata yang ada perlu diterjemahkan terlebih dahulu untuk memperlancar latihan siswa pemula non penutur Arab asli (Aryobimo, 2021).

Berdasarkan penelitian terdahulu kebanyakan analisis soal yang dilakukan masih dalam tahapan analisis soal mata pelajaran ilmu umum seperti IPS dan tematik, adapun yang analisis soal mata pelajaran nahwu berfokus pada buku al-Nahw al-Wāḍiḥ. Sedangkan analisis soal yang peneliti lakukan adalah berfokus pada analisis soal pilihan ganda mata pelajaran nahwu kelas XII Madrasah Aliyah Pesantren Islam Al-Mukmin Ngruki. Tujuan penelitian ini untuk menganalisis butir soal pilihan ganda mata pelajaran nahwu kelas XII Madrasah Aliyah Pesantren Islam Al-Mukmin Ngruki.

B. METODOLOGI PENELITIAN

Jenis penelitian ini ialah kuantitatif deskriptif. Penelitian kuantitatif merupakan penelitian yang dilakukan mulai dari mengumpulkan data lalu menganalisisnya hingga menyampaikan data tersebut dalam bentuk gambar, tabel, grafik, atau yang lain dengan menerapkan prinsip-prinsip objektivitas (Hidayat & Asyafah, 2018).

Fokus penelitian dalam menganalisis butir soal melalui Microsoft Excel 2021 adalah untuk menjadi alat pengukur kemampuan guru kelas XII dalam membuat soal untuk peserta didik. Metode yang dipakai ialah metode deskriptif. Pada penelitian kuantitatif deskriptif, peneliti selaku perencana penelitian, pelaksana penelitian, pengumpul data, penganalisis data dan pelaporan hasil penelitian (Sugiyono, 2011). Data yang akan dikumpulkan berupa penilaian akhir semester peserta didik kelas XII Madrasah Aliyah Pondok Pesantren Ngruki. Partisipan pada penelitian ini adalah analisis butir soal pilihan ganda penilaian akhir semester pada mata pelajaran Nahwu kelas XII Madrasah Aliyah Pondok Pesantren Ngruki. Penelitian ini dilaksanakan pada semester pertama tahun ajaran 2023/2024.

Sampel penelitian berjumlah 25 peserta didik kelas XII Madrasah Aliyah tahun ajaran 2023/2024. Teknik pengumpulan data yang dipakai yaitu dengan observasi dan dokumentasi. Teknik dokumentasinya berupa perangkat soal penilaian akhir semester

pada mata pelajaran Nahwu kelas XII Madrasah Aliyah Pondok Pesantren Ngruki yang berupa lembar soal Penilaian Akhir Semester, kisi-kisi soal, kunci jawaban. Analisis yang dilakukan dalam penelitian ini adalah teknik analisis kuantitatif. Dengan skor benar = 1 skor salah = 0 (dikotomi). Peneliti dalam menganalisis data dan untuk mencari tingkat kesukaran, daya pembeda, efektivitas pengecoh menggunakan aplikasi Microsoft Excel 2021.

C. HASIL DAN PEMBAHASAN

Penilaian Akhir Semester 1 pada mata pelajaran Nahwu kelas XII Madrasah Aliyah di Pondok Pesantren Al Mukmin Ngruki Sukoharjo terdiri dari 30 soal pilihan ganda. Penilaian ini diikuti oleh 25 siswa. Hasil analisis butir soal ditinjau dari beberapa aspek berikut:

1. Tingkat Kesukaran

Tingkat kesukaran butir soal adalah proporsi jumlah testee yang menjawab benar untuk butir soal yang bersangkutan dengan jumlah keseluruhan testee. Tingkat kesukaran ini dinyatakan dalam bentuk Indeks Kesukaran Soal (IKS) yang berkisar 0,00 – 1,00. Semakin besar indeks tingkat kesukaran yang diperoleh dari hasil perhitungan, berarti semakin mudah soal itu. Perhitungan indeks tingkat kesukaran ini dilakukan untuk setiap nomor butir soal. Pada prinsipnya, skor rata-rata yang diperoleh testee pada butir soal yang bersangkutan dinamakan tingkat kesukaran butir soal (Arifin, 2012).

Fungsi tingkat kesukaran butir soal biasanya dikaitkan dengan tujuan tes. Misalnya untuk keperluan ujian semester dipergunakan butir soal yang memiliki tingkat kesukaran sedang, untuk keperluan seleksi dipergunakan butir soal yang memiliki tingkat kesukaran tinggi/sukar, dan untuk keperluan diagnosis biasanya dipergunakan butir soal yang memiliki tingkat kesukaran rendah/mudah (Munip, 2017:272).

Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang peserta didik untuk mempertinggi usaha memecahkannya. Sebaliknya, soal yang terlalu sukar akan menyebabkan peserta didik menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya (Arikunto, 2003). Untuk itulah proses analisis tingkat kesukaran butir soal diperlukan oleh seorang guru.

Rumus yang bisa digunakan untuk mengukur Indeks Kesukaran Soal adalah sebagai berikut:

$$P = B/J$$

Keterangan :

P : Indeks kesukaran

B : Jumlah siswa yang menjawab benar

J : Jumlah siswa

Adapun kriteria penggolongan indeks kesukaran soal adalah sebagai berikut :

Tabel 1. Indeks Kesukaran dengan skala *Likert*

$0,00 \leq IK < 0,20$	Sangat Sukar
$0,20 \leq IK < 0,40$	Sukar
$0,40 \leq IK < 0,60$	Sedang
$0,60 \leq IK < 0,90$	Mudah
$0,90 \leq IK < 1,00$	Mudah Sekali

Dari rumus di atas, didapatkan hasil analisis tingkat kesukaran sebagaimana tabel berikut:

Tabel 2. Distribusi soal berdasarkan indeks kesukaran

NO SOAL	JUMLAH BENAR	JUMLAH SISWA	INDEKS KESUKARAN	KATEGORI SOAL
1	0	25	0	SANGAT SUKAR
2	13	25	0,52	SEDANG
3	9	25	0,36	SUKAR
4	20	25	0,8	SEDANG
5	0	25	0	SANGAT SUKAR
6	18	25	0,72	MUDAH
7	17	25	0,68	MUDAH
8	11	25	0,44	SEDANG
9	23	25	0,92	MUDAH SEKALI

10	18	25	0,72	MUDAH
11	13	25	0,52	SEDANG
12	18	25	0,72	MUDAH
13	19	25	0,76	MUDAH
14	19	25	0,76	SEDANG
15	17	25	0,68	MUDAH
16	18	25	0,72	SEDANG
17	20	25	0,8	MUDAH
18	4	25	0,16	SANGAT SUKAR
19	17	25	0,68	SEDANG
20	11	25	0,44	SEDANG
21	5	25	0,2	SUKAR
22	17	25	0,68	SEDANG
23	15	25	0,6	SEDANG
24	23	25	0,92	MUDAH SEKALI
25	22	25	0,88	MUDAH
26	21	25	0,84	MUDAH
27	24	25	0,96	MUDAH SEKALI
28	22	25	0,88	MUDAH
29	22	25	0,88	MUDAH
30	23	25	0,92	MUDAH SEKALI

Dari tabel di atas dapat diketahui bahwa 3 butir soal berkategori Sangat Sukar, 2 butir soal berkategori Sukar, 10 butir soal berkategori Sedang, 11 butir soal berkategori Mudah, dan 4 butir soal berkategori Mudah Sekali. Sesuai yang dikatakan Arikunto (Arikunto, 2003), soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sulit, maka 3 butir soal berkategori sangat sukar dan 4 butir soal berkategori mudah sekali bisa dibuang atau diperbaiki sedang sisanya bisa disimpan di bank soal yang selanjutnya dapat digunakan untuk penilaian/tes berikutnya.

2. Daya Beda

Daya beda butir soal adalah sebuah indeks yang menunjukkan tingkat kemampuan butir soal untuk membedakan kelompok atas / belahan atas (peserta tes

yang berprestasi tinggi) dari kelompok bawah / belahan bawah (peserta tes yang berprestasi rendah). Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi, disingkat D. Berbeda dengan indeks kesukaran, indeks diskriminasi mempunyai tanda negatif yang digunakan untuk menunjukkan kualitas testee yaitu anak-anak berprestasi tinggi disebut bodoh dan anak berprestasi rendah disebut pintar (Arikunto, 2003).

Bagi suatu soal yang dapat dijawab benar oleh peserta didik yang pandai maupun kurang pandai maka soal itu tidak baik karena tidak mempunyai daya pembeda. Demikian pula jika semua peserta didik yang pandai maupun kurang pandai tidak dapat menjawab dengan benar, maka soal tersebut juga tidak baik karena tidak mempunyai daya pembeda. Soal yang baik adalah soal yang dapat dijawab oleh peserta didik yang pandai saja (Arikunto, 2003).(Hidayat, 2017)

Untuk menentukan daya pembeda/D maka perlu membedakan antara belahan atas dengan belahan bawah. Untuk kelompok kecil yang sampelnya kurang dari 100 maka bisa menggunakan cara semua testee dibagi dua sama besar (50% kelompok belahan atas, 50% kelompok belahan bawah). Selanjutnya dapat menggunakan rumus berikut:

$$D = B_A/J_A - B_B/J_B$$

Keterangan :

D : Daya Pembeda

B_A : Siswa yang menjawab benar pada belahan atas

J_A : Seluruh siswa yang berada pada belahan atas

B_B : Siswa yang menjawab benar pada belahan bawah

J_B : Seluruh siswa yang berada pada belahan bawah

Kriteria yang digunakan untuk mengkategorikan butir soal adalah sebagai berikut :

Tabel 3. Indeks Daya Beda

Negatif	Dibuang
0,00 – 0,20	Pembeda lemah, soal ditolak
0.20 – 0,40	Pembeda sedang, soal diperbaiki

0,40 – 0.70	Pembeda baik, soal diterima & diperbaiki
0,70 – 1,00	Pembeda kuat, soal diterima

Dari rumus di atas didapatkan hasil analisis sebagai berikut :

Tabel 4. Distribusi soal berdasarkan indeks daya beda

NO SOAL	BELAHAN ATAS	BELAHAN BAWAH	INDEKS DAYA BEDA	KATEGORI SOAL
1	0	0	0	LEMAH
2	5	7	-0,166666667	DIBUANG
3	8	1	0,583333333	BAIK
4	12	7	0,416666667	BAIK
5	0	0	0	LEMAH
6	11	6	0,416666667	BAIK
7	11	5	0,5	BAIK
8	7	4	0,25	SEDANG
9	11	11	0	LEMAH
10	10	7	0,25	SEDANG
11	10	2	0,666666667	BAIK
12	11	6	0,416666667	BAIK
13	12	7	0,416666667	BAIK
14	12	6	0,5	BAIK
15	10	6	0,333333333	SEDANG
16	12	5	0,583333333	BAIK
17	12	7	0,416666667	BAIK

18	3	1	0,166666667	LEMAH
19	12	5	0,583333333	BAIK
20	9	2	0,583333333	BAIK
21	4	1	0,25	SEDANG
22	11	6	0,416666667	BAIK
23	8	7	0,083333333	LEMAH
24	11	11	0	LEMAH
25	12	9	0,25	SEDANG
26	12	9	0,25	SEDANG
27	12	11	0,083333333	LEMAH
28	11	10	0,083333333	LEMAH
29	11	10	0,083333333	LEMAH
30	11	11	0	LEMAH

Dari tabel di atas dapat disimpulkan bahwa terdapat 1 butir soal yang berindeks Negatif yang maknanya butir soal tersebut tidak dapat digunakan yaitu nomor 2, 10 butir soal berkategori Lemah yang tidak bisa digunakan karena butir soal tidak bisa membedakan peserta didik yang pandai dan kurang pandai, 6 butir soal berkategori Sedang yang harus diperbaiki, dan 13 butir soal berkategori Baik yaitu bisa digunakan namun memerlukan perbaikan. Sesuai perkataan Arikunto, pembeda soal adalah daya soal dalam menyeleksi tingkat kemampuan peserta didik yang pandai dan kurang pandai. Analisis daya beda digunakan untuk memahami kemahiran soal dalam membedakan tingkat kemampuan peserta didik. Untuk butir soal yang berindeks negatif atau lemah (11 butir soal), maka butir soal harus dibuang. Adapun yang berindeks sedang dan baik (19 butir soal) bisa diperbaiki dan selanjutnya bisa disimpan di bank soal.

3. Distraktor

Distraktor atau pengecoh bisa tergolong baik jika dipilih $\geq 5\%$ oleh siswa. Penafsiran efektivitas distraktor butir soal didasarkan pada kriteria menurut *skala Likert* sebagai berikut: (1) Soal yang keempat jawaban pengecohnya berfungsi maka keefektivitas pengecoh berkategori sangat baik. (2) Soal yang ketiga jawaban pengecohnya berfungsi maka efektivitas pengecoh berkategori baik. (3) Soal yang kedua jawaban pengecohnya berfungsi maka efektivitas pengecoh berkategori cukup. (4) Soal yang satu jawaban pengecohnya berfungsi maka efektivitas berkategori kurang baik. (5) Soal yang tidak memiliki jawaban pengecoh maka efektivitas berkategori sangat tidak baik.

Rumus yang digunakan untuk menghitung distraktor adalah sebagai berikut :

$$A / N \times 100\%$$

Keterangan :

A : Jumlah siswa yang memilih opsi jawaban tersebut

N : Jumlah seluruh siswa

Kriteria yang digunakan sebagai berikut :

Tabel 5. Indeks Distraktor

$D \geq 5\%$	Pilihan diterima
$5\% > D > 0$	Pilihan direvisi
$D = 0$	Pilhan ditolak

Dari rumus di atas diperoleh data sebagaimana tabel berikut :

Tabel 6. Distribusi soal berdasarkan indeks distraktor

NO SOAL	a	b	c	d	e	PILIHAN "a"	PILIHAN "b"	PILIHAN "c"	PILIHAN "d"	PILIHAN "e"	KATEGORI
1	4%	0%	4%	92%	0%	DITOLAK	DITOLAK	DITOLAK	DITERIMA	DITOLAK	Kurang Baik
2	40%	52%	4%	4%	0%	DITERIMA	DITERIMA	DITOLAK	DITOLAK	DITOLAK	Kurang Baik
3	20%	28%	36%	16%	0%	DITERIMA	DITERIMA	DITERIMA	DITERIMA	DITOLAK	Baik
4	84%	0%	0%	12%	4%	DITERIMA	DITOLAK	DITOLAK	DITERIMA	DITOLAK	Kurang baik

5	60%	20%	8%	0%	12%	DITERIMA	DITERIMA	DITERIMA	DITOLAK	DITERIMA	Sangat Baik
6	72%	12%	0%	12%	4%	DITERIMA	DITERIMA	DITOLAK	DITERIMA	DITOLAK	Cukup
7	20%	4%	68%	4%	4%	DITERIMA	DITOLAK	DITERIMA	DITOLAK	DITOLAK	Kurang Baik
8	44%	32%	16%	0%	8%	DITERIMA	DITERIMA	DITERIMA	DITOLAK	DITERIMA	Baik
9	4%	92%	0%	0%	4%	DITOLAK	DITERIMA	DITOLAK	DITOLAK	DITOLAK	Tidak Baik
10	72%	4%	8%	0%	16%	DITERIMA	DITOLAK	DITERIMA	DITOLAK	DITERIMA	Cukup
11	12%	52%	12%	16%	8%	DITERIMA	DITERIMA	DITERIMA	DITERIMA	DITERIMA	Sangat Baik
12	12%	4%	72%	12%	0%	DITERIMA	DITOLAK	DITERIMA	DITERIMA	DITOLAK	Cukup
13	4%	8%	0%	12%	76%	DITOLAK	DITERIMA	DITOLAK	DITERIMA	DITERIMA	Cukup
14	16%	76%	4%	4%	0%	DITERIMA	DITERIMA	DITOLAK	DITOLAK	DITOLAK	Kurang Baik
15	68%	32%	0%	0%	0%	DITERIMA	DITERIMA	DITOLAK	DITOLAK	DITOLAK	Kurang Baik
16	12%	12%	4%	0%	72%	DITERIMA	DITERIMA	DITOLAK	DITOLAK	DITERIMA	Cukup
17	4%	8%	80%	4%	4%	DITOLAK	DITERIMA	DITERIMA	DITOLAK	DITOLAK	Kurang Baik
18	16%	24%	16%	24%	20%	DITERIMA	DITERIMA	DITERIMA	DITERIMA	DITERIMA	Sangat Baik
19	24%	0%	4%	68%	4%	DITERIMA	DITOLAK	DITOLAK	DITERIMA	DITOLAK	Kurang Baik
20	16%	8%	16%	44%	16%	DITERIMA	DITERIMA	DITERIMA	DITERIMA	DITERIMA	Sangat Baik
21	0%	12%	20%	64%	4%	DITOLAK	DITERIMA	DITERIMA	DITERIMA	DITOLAK	Cukup
22	12%	68%	8%	12%	0%	DITERIMA	DITERIMA	DITERIMA	DITERIMA	DITOLAK	Baik
23	60%	12%	8%	20%	0%	DITERIMA	DITERIMA	DITERIMA	DITERIMA	DITOLAK	Baik
24	0%	4%	0%	92%	4%	DITOLAK	DITOLAK	DITOLAK	DITERIMA	DITOLAK	Tidak Baik
25	4%	4%	4%	0%	88%	DITOLAK	DITOLAK	DITOLAK	DITOLAK	DITERIMA	Tidak Baik
26	0%	4%	84%	12%	0%	DITOLAK	DITOLAK	DITERIMA	DITERIMA	DITOLAK	Kurang Baik
27	96%	4%	0%	0%	0%	DITERIMA	DITOLAK	DITOLAK	DITOLAK	DITOLAK	Tidak Baik
28	0%	88%	4%	8%	0%	DITOLAK	DITERIMA	DITOLAK	DITERIMA	DITOLAK	Kurang Baik
29	0%	0%	88%	8%	4%	DITOLAK	DITOLAK	DITERIMA	DITERIMA	DITOLAK	Kurang Baik
30	0%	0%	4%	92%	4%	DITOLAK	DITOLAK	DITOLAK	DITERIMA	DITOLAK	Tidak Baik

Dari tabel di atas dapat kita peroleh hasil bahwa terdapat 5 butir soal berkategori Tidak Baik karena distraktor / pengecoh tidak berfungsi dengan baik yaitu nomor 9, 24, 25, 27, dan 30. Dan terdapat 11 butir soal yang berkategori Kurang Baik, 6 butir soal berkategori Cukup, 4 butir soal berkategori Baik, serta 4 butir soal berkategori Sangat Baik karena distraktor/pengecoh berfungsi sangat baik.

D. SIMPULAN

Berdasarkan hasil analisis butir soal pilihan ganda Penilaian Akhir Semester 1 mata pelajaran Nahwu kelas XII Madrasah Aliyah Ponpes Al Mukmin Ngruki Sukoharjo sudah menunjukkan butir soal berkualitas baik, maka bisa disimpulkan bahwa: (1) Butir soal berkualitas baik dilihat dari tingkat kesukarannya yang mana 76,7% berada pada tingkat Mudah, Sedang, dan Sukar, adapun 23,3% sisanya berada pada tingkat Sangat Sukar dan Mudah Sekali yang mana soal pada tingkat ini harus dibuang. (2) Butir soal berkualitas baik dilihat dari daya pembedanya yang mana 63,3% berada pada tingkat Sedang dan Baik, adapun 36,7% sisanya berindeks negatif dan Lemah yang mana butir soal tidak dapat digunakan. (3) Butir soal berkualitas baik namun perlu perbaikan pada distraktor/pengecohnya dikarenakan 53,3% distraktor tidak berfungsi dengan baik.

DAFTAR PUSTAKA

- Arifin, Z. (2012). *Evaluasi Pembelajaran*. Direktorat Jenederal Pendidikan Islam Kementrian Agama.
- Arikunto, S. (2003). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.
- Aryobimo, B. (2021). Analisis Butir Soal Latihan pada Buku al-Nahw al-Wāḍiḥ untuk Siswa Madrasah Aliyah. *Al-Ma'rifah: Jurnal Budaya, Bahasa Dan Sastra Arab*, 18(2), 127–138. <https://doi.org/10.21009/almakrifah.18.02.02>
- Darodjat, Zuchdi, D., & Zamroni. (2016). Model Evaluasi Pembelajaran Akidah dan Akhlak di Madrasah Tsanawiyah (MTs). *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(1), 11–26. <https://doi.org/10.21831/pep.v20i1.7517>
- Elis Ratna Wulan, & Rusdiana. (2015). *Evaluasi Pembelajaran*. Pustaka Ceria.
- Hidayat, T. (2017). *Pola Pendidikan Islam di Pondok Pesantren Mahasiswa Miftahul Khoir Bandung Dalam Membentuk Kepribadian Islami*. Universitas Pendidikan Indonesia.
- Hidayat, T., & Asyafah, A. (2018). Paradigma Islam Dalam Metodologi Penelitian Dan Implikasinya Terhadap Penelitian Pendidikan Agama Islam. *Tadrib*, 4(2), 225–245.
- Hidayat, T., & Asyafah, A. (2019). Konsep Dasar Evaluasi dan Implikasinya dalam Evaluasi Pembelajaran Pendidikan Agama Islam di Sekolah. *Al-Tadzkiyyah : Jurnal Pendidikan Islam*, 19(1), 159–181.
- Hidayat, T., & Kosasih, A. (2019). Analisis Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 22 tahun 2016 Tentang Standar Proses Pendidikan Dasar dan Menengah Serta Implikasinya Dalam Pembelajaran PAI di Sekolah. *Murobbi : Jurnal Ilmu Pendidikan*, 3(1), 45–69.
- Hidayat, T., Rizal, A. S., & Fahrudin. (2018). Pendidikan Dalam Perspektif Islam dan Peranannya Dalam Membina Kepribadian Islami. *Jurnal Mudarrisuna : Media Kajian Pendidikan Agama Islam*, 8(22), 218–244. <https://doi.org/http://dx.doi.org/10.22373/jm.v8i2.33970> Ta
- Hidayat, T., & Syafe'i, M. (2018). Peran Guru Dalam Mewujudkan Tujuan Pembelajaran Pendidikan Agama Islam di Sekolah. *Rayah Al-Islam : Jurnal Ilmu Islam*, 2(1), 101–111.
- Kartowagiran, B., & Jaedun, A. (2016). Model Asesmen Autentik Untuk Menilai Hasil Belajar Siswa Sekolah Menengah Pertama (SMP): Implementasi Asesmen Autentik di SMP. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(2), 132–141.

- Kaunang, Revoltje, O. W. (2010). Menganalisis Butir Soal. *Inovasi*, 7(1), 176–188.
- Magdalena, I., Ningsih, Y. A., Berliana, N., & Gunawan, A. N. (2021). Analisis Butir Soal Ulangan Akhir Semester Ganjil Mata Pelajaran IPS Kelas IV SDS Sari Putra. *Bintang: Jurnal Pendidikan Dan Sains*, 3(1), 15–22.
- Mardapi, D. (2008). *Teknik Penyusunan Instrumen Tes dan Nontes*. Mitra Cendekia.
- Miladya, J. (2003). Evaluasi Pembelajaran Bahasa Arab. *Konferensi Nasional Bahasa Arab I*, 179–187.
- Munip, A. (2017). *Penilaian Pembelajaran Bahasa Arab*. UIN Sunan Kalijaga.
- Muradi, A. (2014). Pendekatan Komunikatif Dalam Pembelajaran Bahasa Arab. *ARABIYAT: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban*, 1(1), 29–48. <https://doi.org/10.15408/a.v1i1.1129>
- Ridho, U. (2018). Evaluasi Dalam Pembelajaran Bahasa Arab. *An Nabighoh Jurnal Pendidikan Dan Pembelajaran Bahasa Arab*, 20(01), 19. <https://doi.org/10.32332/an-nabighoh.v20i01.1124>
- Rozak, P. (2003). *Evaluasi Afektif Dalam Pembelajaran*. 58–77.
- Setiyawan, A. (2014). Faktor-Faktor yang Mempengaruhi Reliabilitas Tes. *Jurnal An Nur*, 6(2), 341–354.
- Sridadi. (2002). Analisis Butir Soal Pilihan Ganda. *Olahraga*, 8(2), 26–37.
- Sudin, M. (2011). Evaluasi Kegiatan Pembelajaran Pada Lembaga Pendidikan Tinggi Islam : Studi Kasus STAIN Jurai Siwo Metro. *Tapis*, 11(02), 117–131.
- Sugiyono. (2011). *Metode Penelitian Pendidikan : Pendekatan Kuantitatif, Kualitatif, dan R & D*. Alfabeta.
- Taufik Rakhmat, A., & Hidayat, T. (2022). Landasan Pedagogik Pendidikan Agama Islam di Sekolah. *Taklim: Jurnal Pendidikan Agama Islam*, 20(1), 13–31.
- Widayanti, W., Bistari, & Suparjan. (2021). Analisis Butir Soal Pilihan Ganda Penilaian Tengah Semester Pada Pembelajaran Tematik Kelas V Sekolah Dasar Negeri 39 Pontianak Kota. *Jurnal DIDIKA : Wahana Imiah Pendidikan Dasar*, 7(2), 279–296.
- Widoyoko, E. P. (2012). *Evaluasi Program Pembelajaran Panduan Praktis Bagi Pendidik dan Calon Pendidik*. Pustaka Pelajar.