

Analisis Butir Soal Bahasa Arab Di SDIT Al Islam Kudus menggunakan Metode Distinguishing, Difficulty, dan Dispersion

Iwan Harsono¹, Munirul Abidin², Danial Hilmi³

^{1,2,3} Universitas Islam Negeri Malik Ibrahim Malang

Info Artikel :

Diterima 26 Juli 2024

Direvisi 26 September, 2024

Dipublikasikan 28 Oktober 2024

Kata Kunci:

Analisis butir soal

Validitas Reliabilitas

Daya pembeda, pengecoh

Keywords:

Analysis of Question Items,

Validity, Reliability

Discriminating Power,

Distraction.

Abstrak

Penelitian ini bertujuan untuk menganalisis butir soal Bahasa Arab di SDIT Al Islam Kudus dengan fokus pada validitas, reliabilitas, tingkat kesulitan, daya pembeda, dan pengecoh butir soal. Metode yang digunakan adalah pendekatan kuantitatif dengan pengumpulan data dari lembar jawaban siswa kelas IV Semester 1. Dari 25 butir soal yang diuji, 21 dinyatakan valid berdasarkan analisis Product Moment, sementara 4 tidak valid. Sebanyak 21 butir soal dinilai reliabel, sementara 4 tidak reliabel berdasarkan koefisien Alpha Cronbach. Tingkat kesulitan menunjukkan bahwa 3 butir soal berada pada tingkat kesulitan sedang dan 22 pada tingkat sukar, tanpa ada yang mudah. Analisis daya pembeda menunjukkan bahwa sebagian besar butir soal memiliki daya pembeda yang baik. Analisis pengecoh menunjukkan bahwa sebagian besar butir soal memiliki pengecoh yang diterima dengan baik, namun beberapa perlu direvisi. Disimpulkan bahwa evaluasi butir soal Bahasa Arab di SD perlu ditingkatkan untuk meningkatkan kualitas pendidikan Bahasa Arab di tingkat SD.

Abstract

This research aims to analyze Arabic language questions at SDIT Al Islam Kudus with a focus on validity, reliability, level of difficulty, distinguishing power and distracting test items. The method used was a quantitative approach by collecting data from the answer sheets of class IV Semester 1 students. Of the 25 questions tested, 21 were declared valid based on Product Moment analysis, while 4 were invalid. A total of 21 questions were rated as reliable, while 4 were not reliable based on Cronbach's Alpha coefficient. The level of difficulty shows that 3 questions are at the medium level of difficulty and 22 are at the difficult level, with none being easy. Discriminative power analysis shows that most of the items have good discriminating power. Distractor analysis showed that most of the items had well-accepted distractors, but some needed to be revised. It is concluded that the evaluation of Arabic language questions in elementary school needs to be improved to improve the quality of Arabic language education at the elementary level.



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:

Iwan Harsono, Munirul Abidin, Danial Hilmi

Email: 230104210076@student.uin-malang.ic.id

Pendahuluan

Sekolah Dasar (SD) memegang peranan krusial dalam membentuk mutu pendidikan anak-anak, khususnya dalam aspek keagamaan dan akademik.¹ Salah satu subjek yang menjadi fokus utama di Sekolah Dasar (SD) adalah Bahasa Arab, yang berperan sebagai fondasi kunci dalam pemahaman ajaran agama Islam serta pengembangan pengetahuan siswa dalam bidang bahasa dan kesusasteraan.² Penilaian efektivitas dan keberhasilan proses pendidikan di Sekolah Dasar (SD) menjadi sangat penting dengan melakukan evaluasi terhadap butir-butir soal Bahasa Arab.³

Butir soal merupakan elemen kunci dalam proses asesmen pembelajaran, yang tidak hanya merefleksikan pemahaman siswa terhadap materi pelajaran, tetapi juga menjadi indikator kualitas pengajaran yang disampaikan oleh pendidik. Evaluasi merupakan suatu proses sistematis penilaian atau peninjauan terhadap suatu entitas atau kegiatan. Dalam pengertian yang lebih sederhana, evaluasi adalah metode yang digunakan untuk menilai atau mengukur sejauh mana suatu proses atau kegiatan berjalan dengan mempertimbangkan berbagai aspek yang terlibat di dalamnya.⁴

Analisis butir tes adalah suatu proses yang bertujuan untuk mengevaluasi kualitas dan karakteristik dari setiap butir tes. Salah satu metode yang lazim digunakan dalam analisis ini meliputi indeks distinguishability, indeks kesulitan, dan indeks dispersi. Indeks distinguishability, atau yang sering disebut sebagai indeks distinguishing, digunakan untuk menilai kemampuan suatu butir tes dalam memisahkan siswa yang memiliki kemampuan tinggi dengan siswa yang memiliki kemampuan rendah. Nilai indeks yang tinggi menunjukkan bahwa butir tes tersebut efektif dalam memisahkan antara kedua kelompok tersebut.

Selanjutnya, indeks kesulitan, atau difficulty index, digunakan untuk mengukur seberapa sulit suatu butir tes dengan membandingkan proporsi siswa yang menjawab dengan benar terhadap jumlah total siswa yang mengikuti tes tersebut. Semakin tinggi nilai indeks kesulitan, maka semakin sulit butir tes tersebut. Terakhir, indeks dispersi, atau dispersion index, digunakan untuk menilai sejauh mana jawaban siswa tersebar secara merata terhadap suatu butir tes. Nilai indeks yang tinggi menunjukkan adanya variasi yang signifikan dalam jawaban siswa terhadap butir tes tersebut.⁵

Dalam konteks ini, "Evaluasi menyeluruh terhadap butir soal Bahasa Arab di Sekolah Dasar" mengacu pada usaha yang komprehensif untuk meninjau dan mengevaluasi kualitas, kesesuaian dengan kurikulum, kebermaknaan, dan keberlakuan butir soal yang diterapkan dalam pembelajaran Bahasa Arab di tingkat Sekolah Dasar.

Tujuan dari evaluasi ini adalah untuk memverifikasi bahwa butir soal yang dipilih mencerminkan materi yang diharapkan, relevan dengan konteks pembelajaran, dan mampu mengukur pemahaman serta keterampilan siswa sesuai dengan standar yang ditetapkan. Melalui proses evaluasi yang teliti terhadap butir soal Bahasa Arab di SD, diharapkan dapat diidentifikasi kekurangan atau kelemahan dalam penyusunan butir soal, yang akan membuka peluang untuk perbaikan yang signifikan guna meningkatkan efektivitas proses pembelajaran. Evaluasi ini juga dimaksudkan untuk memastikan bahwa butir soal yang dipilih memenuhi kebutuhan pembelajaran siswa dan mendukung pencapaian tujuan pendidikan yang telah ditetapkan.

Penelitian sebelumnya MENGANALISIS BUTIR SOAL BAHASA ARAB UJIAN AKHIR MADRASAH BERSAMA DAERAH (UAMBD) MADRASAH IBTIDAIYAH TAHUN 2017- 2018. Tujuannya adalah untuk memberikan sumbangsih bagi peningkatan kualitas pembelajaran bahasa Arab di

¹ Siti Shalihah, "Teori Evaluasi Pengajaran Bahasa Arab di MI," *PRIMARY* 04 No.02, no. Vol. 4 No. 2 (2012): Desember 2012 (Juli 2012), <https://doi.org/10.32678/primary.v12i2.3019>.

² Abu Dharin, *PENDIDIKAN KARAKTER BERBASIS KOMUNIKASI EDUKATIF RELIGIUS (KER) DI MADRASAH IBTIDAIYAH* (Rizquna, 2019).

³ Lady Farah Aziza dan Ariadi Muliensyah, "KETERAMPILAN BERBAHASA ARAB DENGAN PENDEKATAN KOMPREHENSIF," *El-Tsaqafah: Jurnal Jurusan PBA* 19, no. 1 (30 Juni 2020): 56-71, <https://doi.org/10.20414/tsaqafah.v19i1.2344>.

⁴ Astuti Mardiah, *Evaluasi Pendidikan* (Budi Utama, 2022).

⁵ Siti Jubaidah, "Perancangan Tes Profisiensi Bahasa Arab Berbasis Web," *Prosiding Seminar Nasional Bahasa, Sastra, dan Seni* 1, no. https://journal.unj.ac.id/unj/index.php/prosiding_fbs/issue/view/1453 (2021).

Madrasah Ibtidaiyah dan menjadi bahan rujukan bagi penyusunan dan penggunaan soal-soal bahasa Arab pada UAMBD bagi siswa Madrasah Ibtidaiyah di masa yang akan datang. Penelitian ini menggunakan metode deskriptif kualitatif dengan teknik purposive sampling. Implikasi dari penelitian ini adalah perlunya revisi pada butir soal yang dinilai tidak valid, serta butir soal yang memiliki daya pembeda rendah atau tidak ada.

Perlu dilakukan perhatian khusus terhadap butir soal yang bersifat distraktif guna meningkatkan mutu keseluruhan butir soal. Hasil evaluasi secara keseluruhan menunjukkan kebutuhan akan perbaikan pada jenis butir soal tersebut, bahkan memungkinkan untuk merevisi atau mengganti butir soal yang tidak memenuhi standar evaluasi yang diinginkan.⁶⁷

Maka, metode evaluasi komprehensif terhadap butir soal Bahasa Arab di Sekolah Dasar diharapkan dapat memberikan pemahaman yang mendalam mengenai aspek-aspek yang kuat, lemah, dan potensi peningkatan yang dibutuhkan untuk meningkatkan kualitas pendidikan Bahasa Arab di tingkat SD.

Metode

Untuk merespons perumusan masalah yang ada, penelitian ini memutuskan untuk menerapkan metode pendekatan Kuantitatif. Langkah-langkah yang dilakukan meliputi: (1) Perumusan masalah, (2) Analisis, dan (3) Pengujian, sesuai dengan metodologi yang dijelaskan oleh Abdus samad (2021). Penelitian ini menerapkan pendekatan kuantitatif, yang mengacu pada metode penelitian yang berlandaskan pada filsafat positivisme. Pengumpulan data dilakukan melalui instrumen penelitian dan dilanjutkan dengan analisis data yang bersifat kuantitatif dan statistik, dengan tujuan untuk menguji hipotesis yang telah ditetapkan.⁸ Data yang digunakan dalam penelitian ini diperoleh dari soal dan lembar jawaban siswa kelas IV Semester 1 di SDIT Al Islam Kudus yang berjumlah 21 siswa. Data tersebut kemudian disusun menggunakan teknik analisis yang mendalam, diuraikan secara detail berdasarkan validitas, reliabilitas, tingkat kesulitan, daya pembeda, dan juga analisis pilihan ganda distraktif pada setiap butir soal. Proses analisis dilakukan menggunakan perangkat lunak SPSS serta platform Excel.

Hasil dan Pembahasan

1. Validitas Butir Soal

Suatu instrumen dapat dianggap valid dalam menentukan suatu tes jika instrumen tersebut secara akurat dan konsisten mampu mengukur serta menilai apa yang dimaksudkan untuk dinilai.⁹ Dalam konteks ini, untuk menentukan validitas suatu butir soal, peneliti menerapkan rumus analisis validitas yang didasarkan pada daya pembeda. Rumus ini menggunakan indeks sebagai ukuran untuk menilai sejauh mana butir soal tersebut mampu membedakan antara siswa dengan tingkat pengetahuan yang berbeda dalam materi ujian.¹⁰

Dalam hal ini, pengujian validitas dilakukan menggunakan metode Product Moment, yang dapat diuji dengan aplikasi SPSS. Uji validitas Product Moment Pearson Correlation ini menghubungkan atau mengkorelasikan skor masing-masing item atau soal dengan skor total yang diperoleh dari jawaban responden atas kuesioner. Dasar Pengambilan Keputusan Uji Validitas Product Moment yaitu:

- a. Membandingkan nilai r hitung dengan nilai r table
 1. Jika nilai r hitung $>$ r table, maka item soal angket tersebut dinyatakan valid.
 2. Jika nilai r hitung $<$ r table, maka item soal angket tersebut dinyatakan tidak valid.
- b. Membandingkan Nilai Sig. (2-tailed) dengan Probabilitas 0,05

⁶ Deni Maulana dan Anwar Sanusi, "ANALISIS BUTIR SOAL BAHASA ARAB UJIAN AKHIR MADRASAH BERSAMA DAERAH (UAMBD) MADRASAH IBTIDAIYAH TAHUN 2017-2018," *Ta'lim al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab & Kebahasaaraban* 4, no. 1 (15 Juni 2020): 12–24, <https://doi.org/10.15575/jpba.v4i1.8054>.

⁷ Maulana dan Sanusi.

⁸ Sugiyono Sugiyono, *Metode Penelitian Kuantitatif Dan R&D* (Bandung: Alfabeta, 2010).

⁹ Farida Far Ida dan Anna Musyarofah, "Validitas dan Reliabilitas dalam Analisis Butir Soal," *AL-MU'ARRIB: JOURNAL OF ARABIC EDUCATION* 1, no. 1 (2021): 34–44, <https://doi.org/10.32923/al-muarrib.v1i1.2100>.

¹⁰ Sely Syahriani, Abdul Muktadir, dan Endang Widi Winarni, "Analisis Butir Soal Ujian Sekolah Mata Pelajaran Bahasa Indonesia Kelas VI SDN Kota Bengkulu," t.t.

1. Jika nilai Sig. (2-tailed) dan Pearson Correlation bernilai positif, maka item soal angket tersebut valid.
 2. Jika nilai sig. (2-tailed) < 0,05 dan Pearson Correlation bernilai negative, maka item soal angket tersebut tidak valid.
 3. Jika nilai sig. (2-tailed) > 0,05, maka item soal angket tersebut valid.
- Adapun rumus manual uji Validitas menggunakan Product Moment yaitu:

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

Keterangan

r_{xy} : Koefisien validitas

N : Banyaknya subjek

X : Nilai pembanding

Y : Nilai dari instrument yang akan dicari validitasnya.

Dalam penelitian ini, peneliti menggunakan SPSS untuk menguji validitas butir soal yang ada di SDIT Al Islam Kudus Kelas 4. Adapun langkah-langkah pengujian yang dilakukan meliputi:

Tabel. 1 Analisis Validitas Butir Soal

No Butir Soal	Hasil r Hitung	Hasil r Tabel (0.05)	Kriteria Validasi Butir Soal
1	1.75	0.41	Valid
2	4.08	0.41	Valid
3	0.79	0.41	Valid
4	0.91	0.41	Valid
5	2.83	0.41	Valid
6	3.61	0.41	Valid
7	2.04	0.41	Valid
8	3.04	0.41	Valid
9	0.28	0.41	Tidak Valid
10	0.28	0.41	Tidak Valid
11	0.91	0.41	Valid
12	0.00	0.41	Tidak Valid
13	0.55	0.41	Valid
14	2.04	0.41	Valid
15	1.13	0.41	Valid
16	2.04	0.41	Valid
17	4.17	0.41	Valid
18	3.61	0.41	Valid
19	3.35	0.41	Valid
20	1.67	0.41	Valid
21	4.40	0.41	Valid
22	1.38	0.41	Valid

23	0.28	0.41	Tidak Valid
24	1.13	0.41	Valid
25	1	0.41	Valid

Hasil pengujian validitas yang menggunakan uji Validitas Product Moment pada butir soal Bahasa Arab di kelas 4 SDIT Al Islam Kudus menunjukkan bahwa dari total 21 butir soal, 4 butir soal tidak memenuhi kriteria validitas berdasarkan nilai r hitung. Sebuah butir soal dianggap valid jika nilai signifikansi melebihi 0,41 dengan tingkat kepercayaan 0,05, sedangkan butir soal dianggap tidak valid jika nilai r hitung kurang dari 0,41 dengan tingkat kepercayaan yang sama. Butir soal yang terbukti valid layak digunakan oleh guru dalam proses pembelajaran, sementara butir soal yang tidak valid memerlukan evaluasi dan penyusunan kembali untuk memastikan validitasnya sebelum digunakan dalam pembelajaran.

Menurut Sugiyono, validasi data penelitian adalah proses sistematis yang akurat untuk menilai sejauh mana variabel penelitian mengintegrasikan proses penelitian dan objek penelitian menggunakan data yang berbeda yang disampaikan oleh peneliti. Sementara itu, Utama berpendapat bahwa dukungan data dalam konteks desain penelitian merupakan aspek yang efektif, yang menunjukkan kualitas informasi yang diberikan kepada ilmuwan untuk menjelaskan fenomena dan realitas dunia.¹¹ Keyakinan mengacu pada fakta atau pendapat yang dihasilkan dari penelitian. Oleh karena itu, peneliti menguji validitas butir soal di kelas 4 untuk memastikan keabsahan soal tersebut.

2. Reabilitas Butir Soal

Uji reliabilitas digunakan untuk mengetahui apakah alat pengumpulan data memiliki tingkat ketelitian, kestabilan, atau konsistensi dalam mengukur gejala tertentu pada kelompok individu, meskipun diterapkan pada waktu yang berbeda. Menurut Suharsimi Arikunto, reliabilitas merujuk pada pemahaman bahwa alat tersebut cukup handal untuk digunakan sebagai alat pengumpul data karena kualitasnya yang baik. Alat pengumpul data yang andal akan menghasilkan data yang dapat dipercaya. Tujuan dari reliabilitas adalah untuk memastikan bahwa alat tersebut sangat handal untuk digunakan sebagai alat pengumpul data.

Koefisien alfa Cronbach merupakan statistik yang umum digunakan untuk menguji reliabilitas instrumen survei. Suatu alat penelitian dianggap memiliki tingkat keandalan yang memadai jika koefisien alfa Cronbach bernilai lebih besar atau sama dengan 0,60. Metode ini digunakan untuk mengukur reliabilitas butir soal. Berikut adalah rumus dari Alpha Cornbach secara manual:

$$r_{11} = \left[\frac{k}{(k-1)} \right] \left[1 - \frac{\sum \sigma_b^2}{\sigma_t^2} \right]$$

Keterangan :

- r_{11} = koefisien reliabilitas instrument (total tes)
- k = jumlah butir pertanyaan yang sah
- $\sum \sigma_b^2$ = jumlah varian butir
- σ_t^2 = varian skor total

Dalam penelitian ini, peneliti menggunakan aplikasi SPSS untuk mengukur reliabilitas butir soal dengan menguji menggunakan koefisien alfa Cronbach. Keputusan uji reliabilitas ditentukan dengan ketentuan sebagai berikut:

Jika r hitung lebih dari r tabel berarti item pertanyaan dikatakan reliabel

Jika r hitung kurang dari r tabel berarti item pertanyaan dikatakan tidak reliabel Maka Hasil yang diperoleh peneliti dalam menguji butir soal adalah:

¹¹ Sugi Ono, "Uji Validitas dan Reliabilitas Alat Ukur SG Posture Evaluation," *Jurnal Keterampilan Fisik* 5, no. 1 (4 Mei 2020): 55-61, <https://doi.org/10.37341/jkf.v5i1.167>.

Tabel 2 Reabilitas Butir Soal

No.	No. Item	Statistics Item		
		Prop. Correct	Biser	Point Biser
1	1	0,810	0,793	0,551
2	2	0,667	0,787	0,489
3	3	0,857	0,681	0,518
4	4	0,714	0,842	0,533
5	5	0,619	1,291	0,797
6	6	0,571	1,374	0,853
7	7	0,667	1,202	0,747
8	8	0,286	0,564	0,442
9	9	0,571	1,003	0,623
10	10	0,571	0,731	0,454
11	11	0,286	0,778	0,610
12	12	0,333	0,849	0,623
13	13	0,476	0,538	0,347
14	14	0,667	0,709	0,441
15	15	0,381	0,990	0,689
16	16	0,333	0,585	0,430
17	17	0,429	0,747	0,498
18	18	0,571	1,350	0,838
19	19	0,714	0,763	0,483
20	20	0,571	1,053	0,654
21	21	0,524	1,115	0,703
22	22	0,524	0,560	0,353
23	23	0,429	0,332	0,222
24	24	0,619	-0,001	-0,001
25	25	0,429	0,539	0,360

Berdasarkan analisis menggunakan perangkat lunak SPSS dengan menggunakan Alpha Cronbach, hasil uji reliabilitas butir soal menunjukkan bahwa dari total 21 butir soal, 4 butir soal tidak memenuhi standar reliabilitas yang ditetapkan, sesuai dengan data yang telah dianalisis sebelumnya. Reliabilitas sebagaimana didefinisikan oleh Suharsimi Arikunto, mengacu pada tingkat kepercayaan terhadap suatu alat pengumpul data dalam menghasilkan data yang akurat dan konsisten. Alat yang reliabel mampu menghasilkan data yang terpercaya dan dapat direplikasi. Kepercayaan ini merupakan fondasi penting dalam penelitian, dan untuk itu, reliabilitas bertujuan untuk memastikan bahwa alat tersebut layak dan sesuai untuk digunakan dalam proses pengumpulan data.¹²

3. Tingkat Kesukaran

Tingkat kesulitan soal didefinisikan sebagai penilaian terhadap tingkat kerumitan soal dalam mengukur kemampuan peserta didik. Soal yang ideal memiliki tingkat kesulitan yang proporsional, yang menunjukkan kualitas soal yang baik.¹³ Dalam menganalisis tingkat kesulitan soal, terdapat tiga metode yang umum digunakan. Metode pertama: rumus tingkat Membantu (TK), metode kedua : tabel Tingkat Batas tidak membantu, metode ketiga : Proporsi Jawaban Benar (Proporsi Benar). Penelitian ini menggunakan metode Proporsi

¹² Nilda Miftahul Janna dan H. Herianto, "Konsep Uji Validitas Dan Reliabilitas Dengan Menggunakan SPSS," 22 Januari 2021, <https://doi.org/10.31219/osf.io/v9j52>.

¹³ Maulana dan Sanusi, "ANALISIS BUTIR SOAL BAHASA ARAB UJIAN AKHIR MADRASAH BERSAMA DAERAH (UAMBD) MADRASAH IBTIDAIYAH TAHUN 2017-2018."

Jawaban Benar untuk menganalisis tingkat kesulitan soal.14 sebagai rumus yang digunakan pada aplikasi Ex Excel adalah:

$$\text{Rumus: } \frac{\sum B}{N}$$

- P : Tingkat kesukaran
 $\sum B$: Jumlah Peserta didik yang menjawab benar
N : Jumlah peserta didik

Untuk menafsirkan tingkat kesukaran dapat digunakan criteria sebagai berikut:

- $P > 0,70$ = mudah
 $0,30 \leq p \leq 0,70$ = sedang
 $P < 0,30$ = sukar

Maka dari itu hasil penelitian yang sudah di uji adalah:

Tabel. 3 Hasil Tingkat Kesukaran

No Soal	Hasil	Penafsiran
1	0,33	Sedang
2	0,80	Sedang
3	0,80	Sedang
4	1,19	Sedang
5	0,19	Sukar
6	0,80	Sedang
7	0,76	Sedang
8	0,71	Sedang
9	1,14	Sedang
10	0,42	Sedang
11	0,52	Sedang
12	0,80	Sedang
13	0,28	Sukar
14	0,85	Sedang

¹⁴ Janna dan Herianto, "Konsep Uji Validitas Dan Reliabilitas Dengan Menggunakan SPSS."

15	0,42	Sedang
16	0,76	Sedang
17	0,38	Sedang
18	0,28	Sukar
19	0,33	Sedang
20	1,19	Sedang
21	0,57	Sedang
22	0,33	Sedang
23	0,80	Sedang
24	0,80	Sedang
25	1,19	Sedang
Sukar		
Sedang		

Tabel. 4 Hasil Analisis Tingkat Kesulitan Soal

Analisis Tingkat Kesukaran	Kategori	Jumlah	Keterangan
	Sukar	22	1,2,3,4,6,7,8,9,10,11,12,14,16,17,19,20,21,22,23,24,25
	Sedang	3	5,13,19
	Mudah	0	0

Analisis tingkat kesukaran butir soal menunjukkan bahwa terdapat 3 butir soal yang tergolong dalam kategori sedang, 22 butir soal yang tergolong dalam kategori sukar, dan tidak ada butir soal yang tergolong dalam kategori mudah. Menurut Anas Sudijono, observasi menjadi krusial setelah dilakukan analisis tingkat kesulitan, yang terdefinisikan sebagai berikut: a) Tingkat kesulitan berkorelasi dengan hasil tes pada bagian yang benar (moderat atau sedang); lebih diunggulkan jika item tersebut disimpan dalam bank soal, memungkinkan penggunaan kembali pada tes selanjutnya. b) Penetapan tingkat kesukaran soal didasarkan pada analisis jenis kekuatan. Oleh karena itu, hasil uji daya kesukaran soal butir yang diambil dari SDIT Al Islam Kudus Kelas 4 perlu dievaluasi, mengingat tingkat kesukaran soal yang dominan dalam 25 butir soal yang telah diteliti.

4. Daya Pembeda

Daya pembeda suatu soal adalah kemampuan suatu soal dalam membedakan antara siswa cerdas (berdaya tinggi) dan siswa cerdas (berdaya rendah). rumus berikut dapat digunakan:

$$DP = \frac{JB_A - JB_B}{JS_A} \text{ atau } DP = \frac{JB_A - JB_B}{JS_B}$$

Dengan:

JB_A = jumlah siswa kelompok atas yang menjawab soal dengan benar

JB_B = jumlah siswa kelompok bawah yang menjawab soal dengan benar

JS_A = jumlah siswa kelompok atas

JS_B = jumlah siswa kelompok bawah

Untuk menentukan indeks kekuatan beda (D) pengujian, Metode yang dikemukakan Djiwandono dapat digunakan sebagai berikut:

$D > 0,50$: Baik

$0,20 \leq D \leq 0,50$: Sedang

$> 0,20$: lebih sedikit

0: tidak ada diskriminasi - (negatif): negative

Akan tetapi peneliti memilih menggunakan Excel dalam pengolahan data uji data butir soal daya pembeda. Berikut hasil data butir soal daya pembeda:

Tabel 5 Analisis Tingkat Daya Beda

No Soal	Hasil Daya Beda	Keterangan
1	0,36	Cukup
2	0,44	Baik
3	0,27	Cukup
4	0,35	Cukup
5	0,72	Baik Sekali
6	0,81	Baik Sekali
7	0,63	Baik
8	0,21	Cukup
9	0,62	Baik
10	0,43	Baik
11	0,40	Baik
12	0,50	Baik
13	0,23	Cukup
14	0,44	Baik

15	0,41	Baik
16	0,31	Cukup
17	0,32	Cukup
18	0,81	Baik Sekali
19	0,54	Baik
20	0,62	Baik
21	0,71	Baik Sekali
22	0,14	Jelek
23	0,13	Jelek
24	-0,22	Jelek Sekali
25	0,32	Cukup

Tabel 6 Analisis daya beda

	Kategori	Jumlah	Keterangan
Analisis Daya Pembeda	Baik Sekali	4	5,6,18,21
	Baik	10	2,7,9,10,11,12,14,15,19,20
	Cukup	8	1,3,4,8,13,16,17,25
	Jelek	2	22,23
	Jelek Sekali	1	24

Berdasarkan hasil uji Butir soal dengan daya beda menggunakan excel dapat disimpulkan bahwa butir soal daya beda dibagi menjadi 5 kriteria yaitu Baik sekali, Baik, cukup, jelek, dan jelek sekali. Butir soal dengan daya pembeda ini menurut hasil ada 4 butir soal dengan keterangan nomor soal 5,6,18,21 dinyatakan kategori baik sekali, 10 butir soal dengan keterangan nomor 2,7,9,10,11,12,14,15,19,20 dinyatakan kategori baik, 8 butir soal dengan keterangan nomor 1,3,4,8,13,16,17,25 dinyatakan kategori cukup, 2 butir soal dengan keterangan nomor 22,23 dinyatakan kategori jelek, sedangkan ada 1 jumlah butir soal dengan keterangan nomor 24 dinyatakan kategori jelek sekali. Dalam hal ini peneliti menyarankan untuk mengevaluasi beberapa butir soal yang dinyatakan cukup, jelek bahkan jelek sekali.

5. Analisis Pengecoh

Pengecoh adalah alternatif jawaban (opsi) yang ada pada bentuk soal pilihan ganda. Butir soal yang baik, pengecohnya akan dipilih secara merata oleh peserta didik yang menjawab salah. Sebaliknya, butir soal yang kurang baik, pengecohnya akan dipilih secara tidak merata. Pengecoh soal ini diberikan guru kepada murid agar murid lebih teliti dengan butir soal yang telah diberikan. Rumus pengecoh butir soal ini dapat dilihat dari:

$$IP = \frac{P}{(N-B)/(n-1)} \times 100\%$$

Keterangan:

- IP : Banyaknya peserta tes yang memilih pengecoh tertentu
- N : Banyaknya seluruh peserta tes
- B : Banyaknya peserta tes yang menjawab benar butir soal yang bersangkutan
- n : Banyaknya alternatif jawaban

Akan tetapi peneliti memilih menggunakan Excel dalam pengolahan data uji data butir soal pengecoh. Berikut hasil data butir soal pengecoh:

Tabel 7 Butir Soal Daya Pengecoh

No Item	kunci	Jawaban				Status Soal
		A	B	C	D	
1	D	0	4	0	17	Dapat diterima
2	B	1	14	2	4	Dapat diterima
3	A	18	3	0	0	Dapat diterima
4	B	3	15	3	0	Dapat diterima
5	C	2	5	13	1	Dapat diterima
6	A	12	5	3	1	Dapat diterima
7	B	4	14	3	0	Dapat diterima
8	B	9	6	2	4	Soal Sebaiknya Direvisi
9	B	7	12	1	1	Dapat diterima
10	D	4	4	1	12	Dapat diterima
11	A	6	2	9	4	Soal Sebaiknya Direvisi
12	C	9	3	7	2	Soal Sebaiknya Direvisi
13	C	9	0	10	2	Dapat diterima
14	A	14	1	5	1	Dapat diterima
15	C	5	5	8	3	Dapat diterima
16	C	9	2	7	2	Soal Sebaiknya Direvisi
17	C	7	5	9	0	Dapat diterima
18	A	12	4	5	0	Dapat diterima
19	C	4	1	15	1	Dapat diterima
20	B	2	12	3	4	Dapat diterima
21	D	3	5	2	11	Dapat diterima
22	A	11	6	1	3	Dapat diterima
23	B	7	9	1	4	Dapat diterima
24	D	5	2	1	13	Ditolak / Jangan digunakan
25	C	4	2	9	6	Dapat diterima

Berdasarkan hasil uji butir soal daya pengecoh dapat dinyatakan bahwa dari 21 siswa dengan jumlah butir soal bahasa arab 25 maka dapat disimpulkan bahwa soal nomor 1,2,3,4,5,6,7,9,10,13,14,15,17,18,19,20,21,22,23,25 dapat diterima dan siswa mampu mengerjakan dengan baik. Sedangkan butir soal nomor 8,11,12,16,24 dapat direvisi oleh tim pendidik dikarenakan dari uji validitas, reabilitas, tingkat kesukaran, daya pembeda, serta daya pengecoh menyatakan bahwa butir soal tersebut harus direvisi.

Kesimpulan

Hasil pengujian validitas yang diuji menggunakan uji Validitas Product Moment pada butir soal Bahasa Arab di SDIT Al Islam Kudus kelas 4 adalah jumlah validitas soal ada 21 soal dan 4 soal tidak valid berdasarkan hasil r hitung. Dinyatakan valid jika signifikan butir soal melebihi 0,41 dengan tumpuan 0,05, dan dinyatakan tidak valid jika hasil r hitung kurang dari 0,41 dengan tumpuan 0,05. Dalam hal ini butir soal layak digunakan guru dikelas karena

dinyatakan valid, lalu untuk butir soal yang tidak valid maka akan dilakukan evaluasi perombakan soal yang tidak valid menjadi valid.

Hasil uji Reabilitas menggunakan SPSS dengan Alpha Cornbach yang digunakan peneliti dalam menguji reabilitas butir soal ialah 21 soal dinyatakan Reabilitas dan 4 soal dinyatakan tidak Reabilitas berdasarkan data yang sudah diujikan. Hasil uji tingkat kesukaran butir soal maka dinyatakan bahwa tingkat kesukaran butir soal dengan kategori sedang 3 butir soal, dengan kategori sukar 22 soal, dan dengan kategori mudah 0 soal. Hasil uji Butir soal dengan daya beda menggunakan excel dapat disimpulkan bahwa butir soal daya beda dibagi menjadi 5 kriteria yaitu Baik sekali, Baik, cukup, jelek, dan jelek sekali. Hasil uji butir soal daya pengecoh dapat dinyatakan bahwa dari 21 siswa dengan jumlah butir soal bahasa arab 25 maka dapat disimpulkan bahwa soal nomor 1,2,3,4,5,6,7,9,10,13,14,15,17,18,19,20,21,22,23,25 dapat diterima dan siswa mampu mengerjakan dengan baik, selain dari itu butir soal tidak dapat diterima.

Referensi

- Aziza, Lady Farah, dan Ariadi Muliansyah. "KETERAMPILAN BERBAHASA ARAB DENGAN PENDEKATAN KOMPREHENSIF." *El-Tsaqafah: Jurnal Jurusan PBA* 19, no. 1 (30 Juni 2020): 56–71. <https://doi.org/10.20414/tsaqafah.v19i1.2344>.
- Dharin, Abu. *PENDIDIKAN KARAKTER BERBASIS KOMUNIKASI EDUKATIF RELIGIUS (KER) DI MADRASAH IBTIDAIYAH*. Rizquna, 2019.
- Ida, Farida Far, dan Anna Musyarofah. "Validitas dan Reliabilitas dalam Analisis Butir Soal." *AL-MU'ARRIB: JOURNAL OF ARABIC EDUCATION* 1, no. 1 (2021): 34–44. <https://doi.org/10.32923/al-muarrif.v1i1.2100>.
- Janna, Nilda Miftahul, dan H. Herianto. "Konsep Uji Validitas Dan Reliabilitas Dengan Menggunakan SPSS," 22 Januari 2021. <https://doi.org/10.31219/osf.io/v9j52>.
- Jubaidah, Siti. "Perancangan Tes Profisiensi Bahasa Arab Berbasis Web." *Prosiding Seminar Nasional Bahasa, Sastra, dan Seni* 1, no. https://journal.unj.ac.id/unj/index.php/prosiding_fbs/issue/view/1453 (2021).
- Mardiah, Astuti. *Evaluasi Pendidikan*. Budi Utama, 2022.
- Maulana, Deni, dan Anwar Sanusi. "ANALISIS BUTIR SOAL BAHASA ARAB UJIAN AKHIR MADRASAH BERSAMA DAERAH (UAMBD) MADRASAH IBTIDAIYAH TAHUN 2017-2018." *Ta'lim al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab & Kebahasaaraban* 4, no. 1 (15 Juni 2020): 12–24. <https://doi.org/10.15575/jpba.v4i1.8054>.
- Ono, Sugi. "Uji Validitas dan Reliabilitas Alat Ukur SG Posture Evaluation." *Jurnal Keterampilan Fisik* 5, no. 1 (4 Mei 2020): 55–61. <https://doi.org/10.37341/jkf.v5i1.167>.
- Shalihah, Siti. "Teori Evaluasi Pengajaran Bahasa Arab di MI." *PRIMARY* 04 No.02, no. Vol. 4 No. 2 (2012): Desember 2012 (Desember 2012). <https://doi.org/10.32678/primary.v12i2.3019>.
- Sugiyono, Sugiyono. *Metode Penelitian Kuantitatif Dan R&D*. Bandung: Alfabeta, 2010.
- Syahriani, Sely, Abdul Muktadir, dan Endang Widi Winarni. "Analisis Butir Soal Ujian Sekolah Mata Pelajaran Bahasa Indonesia Kelas VI SDN Kota Bengkulu," t.t.