

Analisis Kualitas Soal Evaluasi pada Buku Ajar PAI dan Budi Pekerti SMKN 3 Kota Serang

Siti Ina Mulyani¹, Umi Syahidah², Salsabila Oktavia³, Saefudin Zuhri⁴, Wahyu Hidayat⁵

¹⁻⁵ UIN Sultan Maulana Hasanuddin Banten, Indonesia

siinamulyani06@gmail.com ^{1*}

Alamat: Jalan Jendral Sudirman No. 30 Panancangan Cipocok Jaya, Sumurpecung, Kec. Serang, Kota Serang, Banten 42118

Korespondensi penulis: siinamulyani06@gmail.com

Abstract. *This research investigates the quality of evaluation questions in the Islamic Religious Education (PAI) and Character Education textbook for vocational high school (SMK) Grade X at SMKN 3 Kota Serang. The analysis focuses on validity, reliability, difficulty level, and discriminating power. The study employed a quantitative descriptive approach, using data collected from the responses of Grade X Accounting students to the evaluation questions provided in the textbook. Findings revealed that, among the 10 questions analyzed, 8 were highly valid, while 2 lacked validity. The reliability of the questions was rated high, with a KR-20 score of 0.74. Regarding difficulty levels, 6 questions were categorized as easy, 4 as moderate, and none as difficult. In terms of discriminating power, 5 questions were classified as excellent, 3 as fair and needing improvement, and 2 as poor and unsuitable for further use. These results highlight the need to revise invalid questions and those with low discriminating power. Additionally, incorporating more challenging questions is recommended to foster students' critical thinking skills. Such measures are critical to enhancing the effectiveness of evaluation questions in assessing students' capabilities comprehensively.*

Keywords: *Evaluation Questions; Discriminating power; Difficulty level.*

Abstrak. Penelitian ini bertujuan untuk menganalisis kualitas soal evaluasi pada buku ajar Pendidikan Agama Islam (PAI) dan Budi Pekerti untuk SMK Kelas X di SMKN 3 Kota Serang berdasarkan kriteria validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Metode yang digunakan adalah deskriptif kuantitatif, dengan data berupa hasil jawaban siswa kelas X Akuntansi terhadap soal evaluasi yang terdapat pada buku ajar tersebut. Hasil analisis menunjukkan bahwa dari 10 soal yang diuji, 8 soal memiliki validitas tinggi, sementara 2 soal dinyatakan tidak valid. Reliabilitas soal berada pada kategori tinggi dengan nilai KR-20 sebesar 0,74. Tingkat kesukaran menunjukkan 6 soal termasuk kategori mudah dan 4 soal kategori sedang, tanpa ada soal dalam kategori sukar. Analisis daya pembeda mengungkapkan 5 soal memiliki kualitas istimewa, 3 soal cukup, dan 2 soal dinyatakan jelek sehingga tidak layak digunakan. Berdasarkan temuan ini, direkomendasikan untuk merevisi soal yang tidak valid dan memiliki daya pembeda rendah, serta menambahkan soal dengan tingkat kesukaran yang lebih tinggi untuk meningkatkan kemampuan analitis siswa. Dengan langkah-langkah tersebut, kualitas soal evaluasi dapat ditingkatkan sehingga lebih efektif dalam mengukur kemampuan siswa secara komprehensif.

Kata Kunci: Analisis Soal; daya pembeda; tingkat kesukaran.

1. LATAR BELAKANG

Pendidikan di Indonesia memiliki tanggung jawab yang besar dalam mendidik anak bangsa sehingga mampu membentuk pribadi yang bukan hanya cerdas dalam intelektualnya namun juga berakhlak mulia. Dalam konteks membentuk generasi yang berakhlak mulia pendidikan agama islam menjadi tonggak utama dalam menanamkan nilai akhlak pada hati setiap anak didiknya. (Rivki dalam Ramadhani & Musyarapah 2024) Untuk mengetahui sejauh mana pendidikan berjalan efektif maka sangat diperlukan evaluasi dalam pembelajaran.

Evaluasi harus dilakukan dalam setiap pembelajaran, begitu juga dalam proses pembelajaran pendidikan agama islam. Proses evaluasi penting dilakukan agar seorang guru dapat mengetahui sejauh mana tingkat keberhasilan siswa dalam memahami pelajaran yang telah disampaikan (Royhanuddin , et al. 2024). Evaluasi di ibaratkan sebagai sebuah cermin yang dapat memperlihatkan bagaimana kualitas suatu pembelajaran. Ketepatan cermin itu begitu bergantung dengan kualitas alat yang akan dipergunakan, kualitas alat yang dimaksud adalah soal evaluasi. Untuk melakukan proses evaluasi maka diperlukan suatu instrumen evaluasi.

Sumber instrument evaluasi yang biasa dipergunakan dalam proses pembelajaran adalah buku ajar. Buku ajar memiliki peran penting dalam menyediakan informasi bagi siswa, di dalam buku ajar bukan hanya memuat tentang pembahasan materi namun juga memuat soal-soal evaluasi. Kualitas soal yang terdapat di dalam buku ajar akan menjadi penentu keberhasilan dalam mengukur sejauh mana siswa dapat memahami pelajaran yang telah diberikan. Namun, seringkali ditemukan bahwa soal yang diujikan disekolah kadang masih kurang bervariasi. Soal yang disajikan kadang hanya dapat memberikan pengukuran pengetahuan pada tingkat rendah saja. Padahal seharusnya, soal yang disajikan untuk tingkat tinggi seperti tingkat SMK adalah soal yang dapat memicu siswa dalam mengembangkan kemampuannya untuk menganalisis serta berfikir kritis.

Maka dari itu menganalisis kualitas soal evaluasi menjadi hal yang harus dilakukan. Penelitian ini bertujuan untuk menganalisis kualitas soal evaluasi pada buku ajar PAI SMKN 3 Kota Serang, dengan tujuan agar dapat mengetahui sejauh mana soal pada buku ajar tersebut telah memenuhi kriteria validitas, reliabilitas, daya beda, dan tingkat kesukaran.

2. METODE PENELITIAN

Jenis penelitian yang digunakan dalam penelitian ini adalah penelitian deskriptif kuantitatif. Metode deskriptif dengan pendekatan kuantitatif adalah metode yang digunakan untuk memberikan gambaran suatu kondisi yang didalamnya berupa angka-angka, yang sedang berlangsung saat ini atau yang pernah terjadi di masa lampau (Sukmadinata 2010). Penelitian ini termasuk penelitian kuantitatif deskriptif namun dalam pelaksanaannya penelitian ini tidak dimaksudkan untuk menolak atau menerima hipotesis, melainkan memberikan penjelasan berdasarkan kenyataan yang terjadi pada objek yang sedang diteliti. (Febrila, Yensy and Susanta 2021)

Penelitian ini dilaksanakan di SMKN 3 Kota Serang yang bertempat di jalan kimara no 1, lebakwangi kecamatan walantaka, kota serang banten. Pengambilan data dilaksanakan ketika sedang melaksanakan program PLP Integratif pada tahun ajaran 2024 pada bulan Juli-Oktober. Adapun subjek pada penelitian ini adalah seluruh lembar jawaban peserta didik pada kelas X Akutansi dengan menggunakan soal evaluasi yang ada pada buku ajar PAI dan budi pekerti untuk SMK Kelas X tahun ajaran 2024 yang terdiri dari 20 peserta didik dalam bentuk google formulir. Sedangkan objek pada penelitian ini merupakan variabel yang diselidiki yaitu kualitas soal evaluasi kelas X tahun ajaran 2024.

Metode pengumpulan data yang digunakan dalam penelitian ini yaitu metode dokumentasi. Dokumen yang digunakan dalam penelitian ini adalah soal dan lembar jawaban peserta didik kelas X tahun ajaran 2024 dalam bentuk data drive. Instrumen yang digunakan dalam penelitian ini adalah buku ajar pendidikan agama islam dan budi pekerti untuk SMA/SMK Kelas X tahun ajaran 2024. Teknik analisis data dalam penelitian ini yaitu sebagai berikut:

Uji Validitas

Menguji Validitas Bentuk Tes Objektif (Tes Pilihan Ganda, Menjodohkan dan Melengkapi)

- a. Uji validitas butir soal pilihan ganda menggunakan korelasi point biserial yaitu korelasi antara data interval dan data dikotomi. (Arikunto, Prosedur Penelitian 2019) rumus :

Menghitung rpbis

$$r_{pbis} = \frac{\bar{X}_b - \bar{X}_s}{SD_t} \sqrt{pq}$$

Keterangan :

\bar{X}_b = (mean) rata- rata skor siswa/ peserta tes yang menjawab benar

\bar{X}_s = (mean) rata- rata skor siswa/ peserta tes yang menjawab salah

SD_t = simpangan baku skor total

p = proporsi jawaban benar terhadap semua jawaban siswa

q = 1-p

Kriteria Valid

• $r_{pbis} < 0.3$ **TIDAK VALID**

• $r_{pbis} \geq 0.3$ **VALID**

Uji Reliabilitas

Menghitung reliabilitas tes pilihan ganda menggunakan metode split half

$$KR - 20 = \frac{k}{k-1} \left(1 - \frac{\sum pq}{SD_t^2} \right)$$

Keterangan

k = jumlah butir soal

SD_t^2 = varian skor total

P = proporsi siswa yang menjawab benar

q = 1 - p

Kriteria Reliabilitas

- $KR-20 \geq 0,70$ = Instrumen dapat diandalkan (reliabel)
- $KR-20 < 0,70$ = Instrumen tidak dapat diandalkan (kurang reliabel) (Kaplan and Saccuzzo 2005)

Uji Tingkat Kesukaran

Menentukan indeks kesukaran (IK) soal tipe objektif dapat menggunakan rumus sebagai berikut :

$$TK = \frac{BA + BB}{N}$$

keterangan:

Tingkat Kesukaran

BA = jumlah jawaban benar pada kelompok atas (27%)

BB = jumlah jawaban benar pada kelompok bawah (27%)

N = ukuran kelompok (jumlah peserta kelompok atas dan bawah)

- Indeks Taraf Sukar Butir (item difficulty index)

Indeks Taraf sukar tes dihitung berdasarkan jumlah peserta yang menjawab benar pada kelompok atas dan kelompok bawah yang dirumuskan sbb

- Kriteria Indeks Taraf Sukar Butir Soal (Yadnyawati 2019)

- 0,00 -0,30 = Soal kategori sukar
- 0,31-0,70 = Soal kategori sedang
- 0,71- 1,00 = Soal kategori mudah

Uji Daya Beda

- Daya pembeda butir soal tes mengacu pada kemampuan butir dalam membedakan kemampuan antara peserta tes yang telah menguasai materi dan peserta tes yang tidak/belum menguasai materi yang ditanyakan
- Daya pembeda dinyatakan dalam indeks. Indeks daya pembeda berkisar antara -1,00 sampai dengan +1,00.
- Semakin tinggi indeks daya pembeda soal artinya semakin mampu soal yang bersangkutan membedakan siswa/peserta tes yang telah memahami materi dengan siswa/peserta tes yang belum memahami materi.
- Daya pembeda butir soal tes mengacu pada kemampuan butir dalam membedakan kemampuan peserta tes dalam menjawab butir soal tsb.

Rumus :

$$D = \frac{2(B_A - B_B)}{N}$$

D= daya pembeda

A = jumlah peserta kelompok atas yang menjawab benar

B = jumlah peserta kelompok bawah yang menjawab benar

N = jumlah total peserta tes (Jumlah Peserta Kel. Atas + Kelompok Bawah)

Klasifikasi/Kriteria Indeks Daya Pembeda Soal (Wahyuni and Ibrahim 2012)

> 0,40 : Istimewa (Soal diterima)

0,30-0,39: Baik (Soal diterima)

0,11-0,29: Cukup (Soal diperbaiki)

0,00-0,10: Jelek (Soal tidak dipakai)

Nilai Negatif : Salah kunci atau ada kesalahan

3. HASIL DAN PEMBAHASAN

Berikut ini merupakan hasil analisis dari kualitas soal buku ajar pendidikan agama islam dan budi pekerti tahun pelajaran 2024 yang ditinjau dari validitas, reliabilitas, tingkat kesukaran (TK) dan daya pembeda :

Uji Validitas

Berdasarkan rumus korelasi point biserial yaitu korelasi antara data interval dan data dikotomi didapatkan hasil sebagai berikut :

Tabel 1 Hasil Uji Validitas Soal Evaluasi

	1	2	3	4	5	6	7	8	9	10
P	0,9	1	0,85	0,45	0,45	0,8	0,4	0,9	0,5	0,9
Q	0,1	0	0,15	0,55	0,55	0,2	0,6	0,1	0,5	0,1
Mean b	7,38	7,15	7,58	9	9	7,18	9,12	7,44	8,3	7,5
Mean s	5	0	4,66	5,63	5,63	7	5,83	4,5	6	5
SDt	2,141	2,141	2,141	2,141	2,141	2,141	2,141	2,141	2,141	2,141
\sqrt{pq}	0,3	0	0,357	0,497	0,497	0,4	0,489	0,3	0,5	0,3
r_{pbis}	0,333	0	0,486	0,782	0,782	0,033	0,751	0,411	0,537	0,350
Validitas	V	TV	V	V	V	TV	V	V	V	V

Dapat dilihat bahwa tingkat validitas pada soal evaluasi buku ajar pendidikan agama islam dan budi pekerti memiliki 8 soal yang memenuhi kriteria ≥ 0.3 **VALID** dan 2 soal < 0.3 **TIDAK VALID**.

Uji Reliabilitas

Berdasarkan metode split half dengan rumus KR-20 memberikan hasil analisis reliabilitas pada soal evaluasi buku ajar pendidikan agama islam dan budi pekerti tahun pelajaran 2024 sebagai berikut :

Tabel 2 Tabel Hasil Uji Reliabilitas Soal Evaluasi

	1	2	3	4	5	6	7	8	9	10
P	0,9	1	0,85	0,45	0,45	0,8	0,4	0,9	0,5	0,9
Q	0,1	0	0,15	0,55	0,55	0,2	0,6	0,1	0,5	0,1
Pxq	0,09	0	0,127	0,247	0,247	0,16	0,24	0,09	0,25	0,09
$\sum pq$	1,541									
SDt^2	4,587									

$$KR - 20 = \frac{k}{k-1} \left(1 - \frac{\sum pq}{SD_i^2} \right)$$

$$KR-20 = \frac{10}{10-1} \left(1 - \frac{1,541}{4,587} \right)$$

$$= 1,111 \times 0,665$$

$$= 0,74$$

Reliabilitas menunjukkan angka 0,74 Artinya bahwa tingkat keajegan/konsistensi tes ini adalah tinggi, sehingga skor tes ini dapat dipercaya penggunaannya.

Uji Tingkat Kesukaran

Berdasarkan indeks taraf sukar uji tingkat kesukaran dihitung berdasarkan jumlah peserta yang menjawab benar pada kelompok atas dan kelompok bawah yakni sebagai berikut :

Tabel 3 Hasil Uji Tingkat Kesukaran Soal Evaluasi Buku Ajar

Nomor Soal	Jumlah Benar B _A	Jumlah benar B _B	Jumlah N	Hasil	Tingkatan
1	5	4	10	0,9	Mudah
2	5	5	10	1	Mudah
3	5	3	10	0,8	Mudah
4	5	0	10	0,5	Sedang
5	5	0	10	0,5	Sedang
6	4	4	10	0,8	Mudah
7	5	0	10	0,5	Sedang
8	5	4	10	0,9	Mudah
9	5	1	10	0,6	Sedang
10	5	4	10	0,9	Mudah

Dalam tabel hasil uji tingkat kesukaran mendapatkan hasil 6 soal yang termasuk kedalam kategori mudah sesuai kriteria indeks taraf sukar yakni 0,71- 1,00 (Soal kategori mudah) dan 4 soal termasuk kedalam kategori sedang sesuai 0,31-0,70 (Soal kategori sedang)

Uji Daya Pembeda

Berdasarkan hasil analisis daya beda dari soal evaluasi buku ajar PAI dan budi pekerti diperoleh hasil sebagai berikut :

Tabel 4 Hasil Uji Daya Pembeda

Nomor Soal	Jumlah Benar B _A	Jumlah benar B _B	Jumlah N	Hasil	Daya Beda
1	5	4	10	0,2	Cukup (Soal diperbaiki)
2	5	5	10	0	Jelek (soal tidak dipakai)
3	5	3	10	0,4	Istimewa(soal dipakai)
4	5	0	10	1	Istimewa(soal dipakai)
5	5	0	10	1	Istimewa(soal dipakai)
6	4	4	10	0	Jelek (soal tidak dipakai)
7	5	0	10	1	Istimewa(soal dipakai)
8	5	4	10	0,2	Cukup (Soal diperbaiki)
9	5	1	10	0,8	Istimewa(soal dipakai)
10	5	4	10	0,2	Cukup (Soal diperbaiki)

Dari uji daya pembeda pada tabel diatas didapatkan hasil 2 soal yang termasuk kedalam kategori jelek dan harus dibuang atau tidak terpakai, 3 kategori sedang namun harus diperbaiki, dan 5 kategori istimewa ini sesuai dengan klasifikasi/kriteria indeks daya pembeda soal.

Pembahasan

Uji Validitas

Uji validitas penting dilakukan dalam mengukur suatu instrument, sehingga dapat diketahui sejauh mana instrument tersebut dapat memberikan pengukuran yang akurat. (Arikunto 2018) Uji validitas digunakan untuk menganalisis kualitas soal evaluasi yang ada pada buku ajar PAI dan budi pekerti pada tahun ajaran 2024 yang dilakukan di SMKN 3 Kota Serang. Penelitian ini berfokus untuk menguji validitas dalam bentuk tes objektif berupa pilihan ganda menggunakan korelasi point biserial yaitu korelasi antara data interval dan data dikotomi.

Jenis-Jenis Validitas

a. Validitas Isi

Validitas isi berperan dalam menunjukkan seberapa baik pertanyaan, tugas, atau item dalam suatu tes atau alat dapat mencerminkan secara menyeluruh dan adil perilaku dari kelompok yang diujikan. Validitas ini menilai seberapa efektif tes dalam mencakup elemen-elemen substansi yang dimaksudkan untuk diukur (Azwar, 1997:74). Validitas isi digunakan untuk menilai seberapa baik seseorang dapat belajar, hasil pembelajaran, atau pencapaian pendidikan.

b. Validitas Konstruk

Validitas Konstruk merujuk pada sejauh mana item-item dalam tes dapat mengukur apa yang benar-benar dimaksudkan untuk diukur, berdasarkan definisi atau konsep tertentu yang telah ditentukan. Validitas konstruk berkaitan dengan objek dan fenomena yang bersifat abstrak, namun gejala-gejalanya dapat diamati dan diukur (Kusaeri and Supranato 2012).

Validitas Konstruk dapat diaplikasikan untuk mengevaluasi sikap, ketertarikan, citra diri, kontrol diri, gaya kepemimpinan, motivasi untuk mencapai prestasi, dan lain-lain, serta dapat mencakup hal-hal yang berhubungan dengan performa maksimal seperti alat untuk mengukur bakat (tes bakat), kecerdasan (kemampuan intelektual), kecerdasan emosional, dan sebagainya.

c. Validitas Kriteria (Criterion-Related Validity)

Kriteria untuk validitas atau validitas empiris ditentukan melalui kriteria internal maupun eksternal. Validitas kriteria diukur dengan melakukan uji coba terhadap tes yang diberikan kepada peserta yang setara dengan yang akan dinilai atau diteliti. Validitas kriteria adalah suatu tolok ukur validitas yang ditetapkan dengan cara membandingkan hasil tes dengan performa spesifik pada ukuran lainnya atau ukuran luar (Sudjana, 1999:15).

d. Validitas Muka (Face Validity)

Validitas Muka (Face Validity). Validitas muka merupakan jenis validitas yang memiliki tingkat signifikansi paling rendah karena hanya berlandaskan pada penilaian sekilas mengenai konten alat ukur. Jika konten alat ukur terlihat sesuai dengan apa yang ingin diukur, maka dapat dinyatakan bahwa validitas muka sudah terpenuhi (Arikunto, 1991:66). Validitas muka juga bisa dianggap sebagai validitas yang lebih rendah dibandingkan dengan validitas isi (Content Validity).

Faktor-Faktor Yang Mempengaruhi Validitas

- a. Petunjuk yang tidak jelas
- b. Penggunaan kosa kata dan struktur kalimat yang sulit
- c. Ambiguitas
- d. Alokasi waktu yang tidak cukup
- e. Penekanan yang berlebihan terhadap aspek tertentu
- f. Kualitas butir tes yang tidak memadai untuk mengukur hasil belajar
- g. Susunan tes yang jelek
- h. Tes terlalu pendek
- i. Penyusunan butir tes yang tidak runtut
- j. Pola jawaban yang mudah ditebak

Uji Reliabilitas

Uji reliabilitas mengacu pada sejauh mana indikator dapat dipercaya atau diabaikan. Ini menunjukkan bahwa hasil percobaan tetap konsisten bahkan ketika diulang dua kali atau lebih. Masalah yang sama diselesaikan dengan menggunakan alat yang sama. Alat ukur dianggap dapat diandalkan jika menghasilkan hasil yang konsisten meskipun pengujian berulang. Kuesioner dianggap dapat diandalkan jika nilainya konsisten dan stabil sepanjang waktu. Kuesioner, sebagai alat, membutuhkan keandalan yang tinggi. Reliabilitas hanya dapat dinilai jika variabel yang dimaksud valid. Untuk memastikan keandalan, penting untuk menilai validitas terlebih dahulu. Jika pertanyaan tidak valid, pertanyaan tersebut tidak boleh diikuti dengan pengujian keandalan (Reliabilitas) (Sugiono 2007)

Reliabilitas adalah ukuran seberapa stabil dan konsisten responden dalam memberikan jawaban terkait pertanyaan yang merupakan bagian dari dimensi variabel, yang disusun dalam bentuk kuesioner. Pengujian reliabilitas dapat dilakukan secara bersamaan untuk semua pertanyaan dalam lebih dari satu variabel, tetapi sebaiknya pengujian reliabilitas dilakukan

secara terpisah untuk tiap variabel di lembar kerja yang berbeda agar dapat mengidentifikasi konstruk variabel yang tidak memiliki reliabilitas.

- a. Reliabilitas Internal: Untuk mengevaluasi reliabilitas internal, metode yang paling sering digunakan adalah koefisien alpha Cronbach. Metode ini mengukur sejauh mana elemen-elemen dalam alat ukur saling konsisten.
- b. Reliabilitas Eksternal: Dalam menguji reliabilitas eksternal, salah satu metode yang dapat dipergunakan adalah koefisien test retest, di mana alat ukur diuji pada dua waktu yang berbeda untuk menilai konsistensi hasil pengukuran.
- c. Reliabilitas Test-Retest: Pengujian reliabilitas test-retest dilaksanakan dengan menerapkan alat ukur pada dua waktu yang berbeda untuk mengevaluasi konsistensi hasil pengukuran yang diberikan oleh alat ukur tersebut.

Uji Tingkat Kesukaran

Uji tingkat kesukaran merupakan suatu metode yang digunakan untuk mengevaluasi kualitas soal secara efektif. Sebelum melaksanakan uji tingkat kesukaran, penting untuk terlebih dahulu melakukan uji validitas dan reliabilitas.

(S. Azwar 2006) menjelaskan bahwa tingkat kesukaran suatu butir soal dapat diukur melalui proporsi peserta tes yang menjawabnya dengan benar dibandingkan dengan total peserta tes. Dengan kata lain, semakin banyak peserta yang berhasil menjawab butir soal tersebut dengan benar, semakin tinggi indeks tingkat kesukarannya, yang menandakan bahwa soal itu relatif lebih mudah. Sebaliknya, jika hanya sedikit peserta yang berhasil menjawab dengan benar, maka soal tersebut dianggap semakin sulit.

(Sirait 2009) berpendapat bahwa tingkat kesulitan suatu butir soal dapat dipengaruhi oleh kompleksitas pokok soal itu sendiri serta kondisi pilihan jawaban yang disajikan. Dalam hal ini, luasnya pertanyaan dapat menimbulkan kebingungan bagi peserta tes, sementara pilihan jawaban yang homogen atau kalimat soal yang sulit dipahami dan memiliki makna ganda dapat semakin menyulitkan pemahaman peserta.

Uji Daya Beda

Uji daya beda adalah satu proses analisis yang sangat penting agar dapat mengetahui sejauh mana suatu instrumen dapat membedakan siswa yang berkemampuan tinggi dan juga rendah. Penelitian ini dianalisis menggunakan pembagian pada kelompok atas dan kelompok bawah. Dengan adanya uji daya beda dapat diketahui apakah soal tersebut berfungsi dengan

baik sebagai suatu alat evaluasi yang efektif untuk memberikan identifikasi terhadap kemampuan siswa. Dalam penelitian ini menggunakan rumus daya beda yakni :

Soal yang memiliki kategori istimewa dianggap efektif sehingga mampu membedakan antara siswa yang memiliki kemampuan tinggi ataupun kemampuan rendah sehingga soal yang memiliki kategori istimewa dianggap sebagai soal yang berkualitas. Dan juga sebaliknya apabila soal memiliki kategori jelek ataupun cukup menunjukkan bahwa siswa yang memiliki kemampuan rendah justru akan sering menjawab benar dibandingkan dengan siswa yang berkemampuan tinggi sehingga soal-soal tersebut harus dibuang ataupun direvisi. Analisis daya beda memiliki prosedur dalam menganalisisnya yakni :

- a. Untuk awal dilakukan sebuah pengelompokan siswa yang dibagi menjadi dua kelompok yakni kelompok atas dan kelompok bawah
- b. Penghitungan dilakukan dengan menggunakan rumus
$$D = \frac{2(B_A - B_B)}{N}$$
- c. Berdasarkan hasil dari perhitungan maka setiap soal akan diklasifikasikan berdasarkan kriteria indeks daya beda sehingga dapat ditentukan kualitasnya.

4. KESIMPULAN

Penelitian ini menunjukkan bahwa kualitas soal evaluasi pada buku ajar Pendidikan Agama Islam (PAI) dan Budi Pekerti untuk SMK Kelas X di SMKN 3 Kota Serang umumnya sudah memenuhi kriteria yang baik, meskipun masih ada beberapa kekurangan yang perlu diperbaiki. Dari segi validitas, sebanyak 8 soal dinyatakan valid, sementara 2 soal lainnya belum memenuhi standar. Hal ini mengindikasikan bahwa sebagian besar soal mampu mengukur kemampuan siswa sesuai tujuan pembelajaran, meskipun terdapat beberapa yang memerlukan perbaikan karena kurang relevan atau tidak sesuai dengan indikator yang diharapkan.

Tingkat reliabilitas soal juga tergolong tinggi, dengan nilai KR-20 sebesar 0,74, yang mencerminkan konsistensi pengukuran yang baik. Kendati demikian, peningkatan reliabilitas dapat dilakukan dengan memperbaiki soal-soal yang kurang valid atau memiliki daya pembeda yang rendah. Berdasarkan analisis tingkat kesukaran, mayoritas soal berada dalam kategori mudah (6 soal), sementara 4 lainnya tergolong sedang. Hal ini menunjukkan bahwa instrumen evaluasi cenderung mengukur kemampuan dasar siswa, namun kurang menantang bagi siswa dengan kemampuan yang lebih tinggi. Oleh karena itu, penambahan soal dengan tingkat kesukaran yang lebih tinggi disarankan untuk mendorong keterampilan berpikir kritis dan analitis siswa. Dari aspek daya pembeda, ditemukan bahwa 5 soal termasuk kategori istimewa karena mampu membedakan siswa berkemampuan tinggi dan rendah dengan baik. Sementara

itu, 3 soal berada pada kategori cukup dan memerlukan revisi, sedangkan 2 soal dikategorikan jelek dan sebaiknya tidak digunakan. Soal dengan daya pembeda rendah perlu diperbaiki agar evaluasi dapat memberikan hasil yang lebih akurat dalam mengidentifikasi kemampuan siswa.

Secara keseluruhan, kualitas soal evaluasi pada buku ajar PAI dan Budi Pekerti sudah cukup baik dari sisi validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Namun, beberapa aspek masih perlu ditingkatkan, seperti memperbaiki soal yang tidak valid, meningkatkan tingkat kesukaran soal, dan menyempurnakan daya pembeda. Langkah-langkah ini penting untuk memastikan soal dapat digunakan secara efektif dalam mengukur kemampuan siswa secara holistik.

Saran

Untuk meningkatkan kualitas evaluasi, penting untuk merevisi soal-soal yang dinyatakan tidak valid atau memiliki daya pembeda rendah, sehingga pengukuran kemampuan siswa menjadi lebih akurat. Selain itu, diperlukan penambahan soal dengan tingkat kesukaran yang lebih tinggi untuk merangsang kemampuan berpikir kritis dan analitis siswa. Soal-soal dengan daya pembeda yang tergolong buruk juga perlu diperbaiki agar lebih efektif dalam membedakan siswa berdasarkan tingkat pemahaman mereka. Upaya ini tidak hanya akan meningkatkan mutu evaluasi, tetapi juga mendukung proses pembelajaran yang lebih optimal dalam mencapai tujuan pendidikan secara keseluruhan.

DAFTAR PUSTAKA

- Arikunto, Suharsimi. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara , 1991.
- Azwar, Saifuddin. *Reliabilitas, Validitas, Interpretasi Dan Komputasi*. Yogyakarta : Liberty, 2006 .
- Azwar, Syaifuddin . *Realiabilitas Dan Validitas*. Yogyakarta: Pustaka Pelajar, 1997.
- Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara , 2018.
- Febrila, Lilia Gina, Nurul Astuty Yensy, And Agus Susanta. "Analisis Tingkat Kognitif Dan Kualitas Soal Penelitian Akhir Semester Ganjil Kelas Vii Smp It Darul Fikri Arga Makmur Tahun Ajaran 2020/2021." *Jurnal Penelitian Pembelajaran Matematika Sekolah (Jp2ms)*, 2021.
- Kaplan, R.M , And D.P. Saccuzzo. *Psychological Testing Principles*,. As: Wadsworth, 2005.
- Kusaeri, And Supranato. *Pengukuran Dan Penilaian Pendidikan* . Bandung : Remaja Rosdakarya , 2012.

Prosedur Penelitian . Jakarta : Rineka Cipta, 2019.

Ramadhani, Novia, And Musyarapah. "Tujuan Pendidikan Islam Dalam Membentuk Generasi Berakhlak Mulia." *Jurnal Pendidikan Nusantara*, 2024.

Royhanuddin , Fauzan , Zulhimma, Dakran , And Wahyu Ari Anto Harahap. "Evaluasi Pembelajaran Pada Mata Pelajaran Pendidikan Agama Islam Dan Implikasinya Terhadap Motivasi Belajar Siswa Man 1 Padangsidempuan." *Jurnal Komunikasi Dan Media Pendidikan*, 2024.

Sirait, Bistok. *Bahan Pengajaran Untuk Mata Kuliah Evaluasi Hasil Belajar Siswa* . Jakarta : Depdikbud , 2009 .

Sudjana, Nana. *Penilaian Hasil Proses Belajar Mengajar* . Bandung : Remaja Rosdakarya , 1999.

Sugiono. *Statistika Untuk Penelitian* . Bandung : Alfabeta , 2007 .

Sukmadinata , N.S. *Metode Penelitian Pendidikan* . Bandung: Pt Remaja Rosdakarya Offset , 2010 .

Wahyuni, Sri, And Abd Syukur Ibrahim. *Asesmen Pembelajaran Bahasa*. Bandung: Refika Aditama, 2012.

Yadnyawati, Ida Ayu Gde. *Evaluasi Pembelajaran*. Denpasar: Unhi Press, 2019