

**Analisis Kualitas Butir Soal Asesmen Sumatif Mata Pelajaran Sejarah
Kebudayaan Islam Kelas XI MAN 3 Cilacap Tahun Ajaran 2024/2025****Mulia Sagesti****Rohmad****UIN Prof. Saifuddin Zuhri Purwokerto**email : muliasagesti223@gmail.com, gomarirohmad@gmail.com**ABSTRAK**

Analisis butir soal merupakan suatu proses evaluasi yang bertujuan untuk menilai kualitas butir soal agar menjadi alat ukur yang akurat. Tujuan penelitian ini untuk menganalisis kualitas butir soal ASAS mata pelajaran SKI dari segi validitas, reliabilitas, tingkat kesukaran, daya beda, efektivitas pengecoh. Tentunya penelitian ini menggunakan bantuan aplikasi *Microsoft Exel* dan ANATES V4. Adapun bentuk soal pilihan ganda terdiri dari 40 soal terdapat 10 soal yang merupakan soal pilihan ganda kompleks yang memiliki lebih dari satu jawaban benar. Oleh karena itu, analisis dilakukan sebanyak tiga kali. Hasil penelitian menunjukkan bahwa Uji validitas kunci jawaban 1 menunjukkan 36 soal (90%) valid dan 4 soal (10%) tidak valid. Kunci jawaban 2 dan 3 menunjukkan hasil sama yaitu 34 soal (85%) valid dan 6 soal (15%) tidak valid. Uji reliabilitas kunci jawaban 1, 2, dan 3 menunjukkan bahwa soal dikatakan reliabel karena memiliki nilai $>0,70$. Analisis tingkat kesukaran kunci jawaban 1 menunjukkan 4 soal sukar, 16 soal sedang dan 20 soal mudah. Kunci jawaban 2 menunjukkan 6 soal sukar, 15 soal sedang, dan 19 soal mudah. Kunci jawaban 3 menunjukkan 4 soal sukar, 15 soal sedang dan 21 soal mudah. Analisis daya beda kunci jawaban 1 menunjukkan 4 soal sangat buruk, 16 soal cukup baik, dan 17 soal sangat baik. Kunci jawaban 2 menunjukkan 10 soal sangat buruk, 17 soal cukup baik, dan 13 soal baik. pada kunci jawaban 3 menunjukkan 8 soal sangat buruk, 16 soal cukup baik, dan 16 soal baik. Analisis efektivitas pengecoh kunci jawaban 1 menunjukkan 18 soal sangat baik, 11 soal baik, 9 soal cukup, dan 2 soal kurang baik. Kunci jawaban 2 menunjukkan 13 soal sangat baik, 15 soal baik, 8 soal cukup, 2 soal kurang baik dan 2 soal sangat tidak baik. Kunci

Article HistoryReceived: Juli 2025
Reviewed: Juli 2025
Published: Juli 2025Plagiarism Checker No
234Prefix DOI : Prefix DOI :
10.8734/Sindoro.v1i2.365**Copyright : Author
Publish by : Sindoro**This work is licensed
under a [Creative
Commons
Attribution-
NonCommercial 4.0
International License](https://creativecommons.org/licenses/by-nc/4.0/)

jawaban 3 menunjukkan 12 soal sangat baik, 13 soal baik, 12 soal baik dan 3 soal kurang baik.

Kata kunci: Analisis Butir Soal, Asesmen Sumatif, Pilihan Ganda Kompleks

ABSTRACT

Item analysis is an evaluation process aimed at assessing the quality of test items to ensure they serve as accurate measurement tools. This study aims to analyze the quality of ASAS test items in the SKI subject in terms of validity, reliability, difficulty level, discriminating power, and distractor effectiveness. The research utilized Microsoft Excel and ANATES V4 software. The test consisted of 40 multiple-choice items, 10 of which were complex multiple-choice items with more than one correct answer. Therefore, the analysis was conducted three times. The results showed that the validity test for answer key 1 revealed that 36 items (90%) were valid and 4 items (10%) were invalid. Answer keys 2 and 3 produced the same results, with 34 items (85%) valid and 6 items (15%) invalid. The reliability test for answer keys 1, 2, and 3 showed that the items were reliable, each having a reliability score above 0.70. The difficulty level analysis for answer key 1 showed 4 difficult items, 16 moderate items, and 20 easy items. Answer key 2 showed 6 difficult items, 15 moderate, and 19 easy. Answer key 3 showed 4 difficult items, 15 moderate, and 21 easy. The discriminating power analysis for answer key 1 revealed 4 very poor items, 16 fair items, and 17 excellent items. Answer key 2 showed 10 very poor items, 17 fair, and 13 good items. Answer key 3 showed 8 very poor items, 16 fair, and 16 good items. The distractor effectiveness analysis for answer key 1 revealed 18 items were very good, 11 good, 9 fair, and 2 poor. Answer key 2 showed 13 very good items, 15 good, 8 fair, 2 poor, and 2 very poor. Answer key 3 showed 12 very good items, 13 good, 12 fair, and 3 poor.

Keywords: Item Analysis, Summative Assessment, Complex Multiple Choice

PENDAHULUAN

Pendidikan merupakan suatu upaya yang dilakukan secara sadar untuk menciptakan lingkungan dan proses belajar yang mendorong peserta

didik agar terlibat aktif dalam mengembangkan potensi diri, kepribadian, kecerdasan, akhlak, serta keterampilan yang dibutuhkan untuk kehidupannya sendiri maupun demi kemajuan masyarakat, bangsa, dan negara¹. Pendidikan memiliki peran yang sangat vital dalam kehidupan manusia². Secara umum, pendidikan terbagi menjadi tiga jenis, yaitu pendidikan formal, nonformal, dan informal³. Pendidikan, peserta didik, dan kurikulum merupakan komponen terpenting dalam sistem pendidikan dan berperan krusial dalam menjamin keberhasilan proses pembelajaran.

Kualitas sumber daya manusia suatu masyarakat tercermin dari kualitas sistem pendidikannya. Sistem pendidikan yang berfungsi dengan baik akan menghasilkan generasi yang cakap, berkualitas, dan mampu bersaing secara lebih efektif dalam berbagai bidang kehidupan⁴. Tercapainya hasil belajar siswa merupakan salah satu tolok ukur utama keberhasilan akademis.

Dalam satuan pendidikan, evaluasi merupakan tugas yang krusial dan sangat dibutuhkan dalam satuan pendidikan. Adanya evaluasi untuk mengetahui capaian proses pembelajaran yang telah berjalan dan segala hal yang perlu ditingkatkan pada proses pembelajaran berikutnya. Dalam melakukan kegiatan evaluasi, tentu diperlukan informasi dan data yang menunjang jalannya evaluasi. Evaluasi pembelajaran dilaksanakan secara

¹ Peraturan Presiden Republik Indonesia, “Undang-Undang Nomor 20 Tahun 2003 Tentang Sistem Pendidikan Nasional,” Pub. L. No. 20, 2010 29 (2003).

² Cahyani Khayati, Miftachul Amri, and Syamsul Sodik, “Analysis of Objective Test Type of Multiple Choice in Final Semester Assessment Problems (PAS) Japanese Language Class VI,” *Proceeding of International Conference on Arts and Humanities: International Conference on Cultural Studies and Applied Linguistics (ICCSAL) Analysis*, 2021, 10–14.

³ Raudatus Syaadah et al., “Pendidikan Formal, Pendidikan Non Formal Dan Pendidikan Informal,” *Pema (Jurnal Pendidikan Dan Pengabdian Kepada Masyarakat)* 2, no. 2 (2023): 125–31.

⁴ Leonardus Kaka, Vidriana Oktoviana Bano, and Yohana Njoeroemana, “Efektivitas Analisis Butir Soal Pilihan Ganda Menggunakan Aplikasi Anates Di SMPN2 Kanatang,” *Jurnal Inovasi Penelitian* 1441, no. 1450 (2024): 4–9

terencana, teratur, dan terjadwal. Secara garis besar alat evaluasi dapat diklasifikasikan menjadi dua macam, yaitu tes dan non-tes⁵.

Untuk memperoleh hasil yang maksimal, hal pertama yang harus dilakukan adalah menganalisis kualitas butir soal sebelum soal dibagikan kepada peserta didik. Hal ini diperkuat oleh Permendikbudristek Nomor 21 Tahun 2022 Pasal 9 ayat 6b yang menyatakan bahwa evaluasi analisis butir soal perlu dilakukan oleh guru sebagai pendidik. Selain merefleksikan proses pembelajaran yang sebenarnya, tujuan analisis butir adalah untuk menyempurnakan efektivitas pembelajaran⁶.

Soal yang bermutu adalah soal yang memenuhi prinsip validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh yang proporsional, sehingga dapat menggambarkan keadaan peserta didik pada tingkat kemampuan dan keterampilan tertentu⁷. Kualitas instrumen sangat diperlukan sehingga dapat membedakan peserta didik mana yang sudah menguasai materi ataupun yang belum menguasai materi yang diajarkan oleh guru. Hal ini yang mendasari perlunya melakukan analisis terhadap kualitas butir soal. Namun, dalam praktiknya, tidak semua guru mampu untuk menyusun soal yang mengukur kemampuan peserta didik secara komprehensif. Permasalahan utama yang sering muncul adalah ketidaksesuaian antara soal dengan indikator pencapaian kompetensi, distribusi tingkat kesukaran yang tidak merata serta lemahnya daya pembeda soal.

⁵ Nur Aqilah Pohan et al., "Pengembangan Alat Evaluasi," *Jurnal Pendidikan Tambusai* 7 (2023): 30701–5,

⁶ Kemendikbudristek, "Peraturan Menteri Pendidikan, Kebudayaan, Riset, Dan Teknologi No 21 Tahun 2022 Tentang Standar Penilaian Pendidikan Pada Pendidikan Anak Usia Dini, Jenjang Pendidikan Dasar, Dan Jenjang Pendidikan Menengah," *Kementerian Pendidikan Dan Kebudayaan Riset Dan Teknologi*, 2022, 9.

⁷ Sedy Santosa and Jami Ahmad Badawi, "Analisis Butir Soal Pilihan Ganda Tema Pertumbuhan Dan Perkembangan Makhluk Hidup Kelas III Madrasah Ibtidaiyah," *Jurnal Basicedu* 6, no. 2 (2022): 1678–86.

Pentingnya penelitian ini terletak pada pentingnya pemetaan kualitas soal sebagai dasar perbaikan sistem evaluasi di MAN 3 Cilacap. Analisis kualitas butir soal tidak hanya bermanfaat untuk meningkatkan kualitas penilaian, tetapi juga sebagai pertimbangan bagi guru dan sekolah untuk meninjau kesesuaian antara pembelajaran dan asesmen. Berdasarkan hal tersebut, maka perlu dikaji lebih dalam sebagai bentuk pembuktian nyata yaitu apakah kualitas butir soal Asesmen Sumatif dapat dikatakan baik sebagai alat ukur keberhasilan peserta didik kelas XI MAN 3 Cilacap khususnya pada mata pelajaran Sejarah Kebudayaan Islam.

Pada penelitian ini, peneliti lebih memfokuskan secara mendalam dari kualitas pilihan ganda dengan menggunakan teori klasik sebagai pendekatannya yaitu *Classical Test Theory* (CTT) untuk memperoleh pemahaman yang lebih baik mengenai efektif atau tidaknya soal dalam mengkaji kemampuan siswa, serta memastikan bahwa kualitas soal-soal tersebut mampu mengukur peserta didik secara tepat dan proporsional. Adapun bentuk soal pilihan ganda (*Multiple Choice*) terdiri dari 40 butir soal dengan jumlah *option* sebanyak lima (A, B, C, D, dan E). Dari 40 butir soal terdapat 10 butir soal yang merupakan soal pilihan ganda kompleks dengan jawaban lebih dari satu *option* (*Complex Multiple Choice*).

Dalam konteks pendidikan modern, salah satu cara untuk membantu guru dalam menganalisis asesmen sumatif menjadi lebih efisien dengan penggunaan *tools* digital. Aplikasi *Microsoft excel* berperan dalam perhitungan validitas dan reliabilitas, dan ANATES V4 memudahkan untuk menganalisis kualitas butir soal seperti tingkat kesukaran, daya beda dan efektivitas pengecoh.

Berdasarkan latar belakang tersebut, peneliti tertarik untuk melakukan penelitian dengan judul “Analisis Kualitas Butir Soal

Asesmen Sumatif Mata Pelajaran Sejarah Kebudayaan Islam Kelas XI MAN 3 Cilacap Tahun Ajaran 2024/2025”.

PEMBAHASAN

Berdasarkan hasil analisis ASAS mata pelajaran SKI kelas XI MAN 3 Cilacap Tahun ajaran 2024/2025 dari 40 soal pilihan ganda, terdapat 10 soal yang merupakan soal pilihan ganda kompleks. Sehingga agar menghasilkan hasil yang akurat, maka analisis dilakukan sebanyak tiga kali dengan menggunakan tiga kunci jawaban berbeda sesuai dengan jumlah jawaban benar pada pilihan ganda kompleks.

1. Validitas

Menurut Messick, Validitas adalah kesesuaian antara interpretasi hasil asesmen dengan tujuan pengukuran yang diinginkan⁸. Zainal Arifin mengklasifikasikan tingkat validitas, soal dikatakan memiliki nilai validitas sangat tinggi apabila berada pada koefisien 0,81 - 1,00 dan soal dianggap tidak valid apabila nilai koefisiennya < 0,00. Hasil uji validitas membuktikan bahwa 36 soal (90%) pada kunci jawaban pertama dan 34 soal (85%) pada kunci jawaban kedua dan ketiga memiliki korelasi validitas sangat tinggi. Soal dikatakan valid apabila r hitung lebih besar dari rtabel maka hasil yang diperoleh signifikan yang artinya soal dinyatakan valid⁹. Sesuai dengan definisi validitas menurut Arikunto, bahwa sebagian besar soal mampu mengukur apa yang seharusnya diukur¹⁰. Namun, terdapat 4 (10%) soal pada kunci jawaban pertama dan 6 (15%) soal pada kunci jawaban kedua dan ketiga yang dinyatakan tidak valid. Soal dinyatakan tidak valid karena, nilai r hitung lebih kecil dari rtabel.

⁸ Rohmad, *Asesmen Pembelajaran Pendidikan Agama Islam Dan Budi Pekerti*.

⁹ Magdalena et al., “Analisis Validitas, Reliabilitas, Tingket Kesukaran, Dan Daya Bada Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan.”

¹⁰ Widodo et al., *Metodol. Penelit.*

Berikut merupakan contoh soal yang dinyatakan tidak valid pada ketiga kunci jawaban, yaitu soal nomor 19, 29, dan 37. Adapun nilai rhitung dari masing-masing soal sangat rendah, yang menunjukkan bahwa jawaban dari peserta didik pada soal tersebut tidak memiliki hubungan yang signifikan dengan skor total.

Adapun faktor yang menyebabkan soal tidak valid diantaranya adalah soal memiliki tingkat kesukaran yang tinggi sehingga menghasilkan daya beda yang rendah sehingga tidak dapat membedakan kemampuan peserta didik dengan baik. Sebagai tindak lanjut, soal yang telah terbukti valid dapat disimpan dalam bank soal untuk digunakan kembali pada tes mendatang. Sementara, soal yang tidak valid perlu direvisi untuk memenuhi standar yang baik.

2. Reliabilitas

Menurut Nunally dan Bernstein, reliabilitas adalah sejauh mana suatu tes dapat menghasilkan hasil yang sama dalam situasi yang serupa¹¹. Dalam penelitian ini, uji reliabilitas digunakan untuk mengetahui konsistensi soal ASAS mata pelajaran SKI kelas XI MAN 3 Cilacap Tahun ajaran 2024/2025.

Dalam penelitian ini, uji reliabilitas menggunakan rumus *Cronbach's Alpha* yang mana menurut Sudijono tes dikatakan memiliki nilai reliabilitas yang baik ketika nilai *Cronbach's Alpha* $\geq 0,70$ ¹². Berdasarkan klasifikasi reliabilitas, soal dikatakan memiliki nilai reliabilitas sangat tinggi apabila berada pada koefisien $\geq 0,90$. Soal dikatakan memiliki nilai reliabilitas tinggi apabila berada pada koefisien 0,80 - 0,89. Soal dikatakan memiliki nilai reliabilitas cukup apabila berada pada koefisien 0,70 - 0,79. Soal dikatakan memiliki nilai

¹¹ Rohmad, *Asesmen Pembelajaran Pendidikan Agama Islam Dan Budi Pekerti*.

¹² Rahman and Nasryah, *Evaluasi Pembelajaran*.

reliabilitas rendah apabila berada pada koefisien 0,60 - 0,69 dan soal dikatakan memiliki nilai reliabilitas sangat rendah apabila berada pada koefisien $< 0,60$ ¹³.

Hasil uji reliabilitas membuktikan bahwa pada ketiga kunci jawaban, soal dikatakan reliabel karena nilai *Cronbach's Alpha* $\geq 0,70$. Adapun nilai *Cronbach's Alpha* untuk masing-masing kunci jawaban adalah pada kunci pertama adalah 0,773. Pada kunci jawaban kedua adalah 0,746 dan pada kunci jawaban 3 adalah 0,739, artinya soal memiliki nilai konsistensi reliabilitas cukup karena berada pada koefisien 0,70 - 0,79.

Adapun faktor yang menyebabkan dari ketiga kunci jawaban yang digunakan tidak ada yang memiliki koefisien sangat tinggi diantaranya adalah adanya butir soal yang tidak valid sehingga mempengaruhi nilai reliabilitas. Sebagai tindak lanjut, perlu dilakukan peninjauan ulang terhadap butir soal khususnya terhadap soal yang terbukti tidak valid karena berpotensi menurunkan reliabilitas.

3. Tingkat Kesukaran

Menurut Zainal Arifin, perhitungan tingkat kesukaran adalah pengukuran seberapa besar derajat kesukaran suatu soal. Jika soal memiliki tingkat kesukaran seimbang, maka soal tersebut baik¹⁴. Demikian Arikunto menyampaikan bahwa suatu soal tes hendaknya tidak terlalu sukar dan tidak terlalu mudah¹⁵. Rohmad mengklasifikasikan tingkat kesukaran bahwa soal dikatakan sukar apabila berada pada koefisien $< 0,30$, soal

¹³ Rohmad, *Asesmen Pembelajaran Pendidikan Agama Islam Dan Budi Pekerti*.

¹⁴ Arifin, *Evaluasi Pembelajaran*.

¹⁵ Mochammad Noor Akhmadi, "Analisis Butir Soal Evaluasi Tema 1 Kelas 4 Sdn Plumbungan Menggunakan Program Anates," *Ed-Humanistics : Jurnal Ilmu Pendidikan* 6, no. 1 (April 30, 2021): 799–806.

dikatakan sedang apabila berada pada koefisien 0,30 - 0,70 dan soal dikatakan mudah apabila berada pada koefisien $> 0,70$ ¹⁶.

Hasil analisis tingkat kesukaran membuktikan bahwa terdapat 4 soal (10%) pada kunci jawaban pertama, 6 soal (15%) pada kunci jawaban kedua, dan 4 soal (10%) pada kunci jawaban ketiga memiliki kategori sukar karena berada pada koefisien $< 0,30$. Pada kategori sedang dengan koefisien 0,30 - 0,70 pada kunci jawaban pertama terdapat 16 soal (40%), pada kunci jawaban kedua terdapat 15 soal (37,5%) dan pada kunci jawaban ketiga terdapat 15 soal (37,5%). Sedangkan soal pada kategori mudah dengan koefisien $> 0,70$ pada kunci jawaban pertama terdapat 20 soal (50%), pada kunci jawaban kedua terdapat 19 soal (47,5%) dan pada kunci jawaban ketiga terdapat 19 soal (47,5%).

Dari ketiga kunci jawaban yang dianalisis, dapat disimpulkan bahwa proporsi soal dengan kategori mudah mendominasi. Hal ini menunjukkan bahwa sebagian besar soal kurang efektif dalam membedakan tingkat kemampuan peserta didik. Hal ini sejalan dengan pendapat Arikunto yang menyatakan bahwa soal yang ideal adalah soal dengan tingkat kesukaran sedang sehingga efektif dalam membedakan kemampuan peserta didik secara kognitif.

Berikut merupakan contoh soal yang dinyatakan memiliki tingkat kesukaran kategori mudah pada ketiga kunci jawaban, yaitu soal nomor 1, 4 dan 24. Pada ketiga soal tersebut memiliki indeks kesukaran sangat tinggi sehingga menghasilkan soal dengan kategori mudah.

Adapun faktor yang menyebabkan soal dikategorikan mudah yaitu soal terlalu sederhana, tidak menuntut peserta didik

¹⁶ Rohmad, *Asesmen Pembelajaran Pendidikan Agama Islam Dan Budi Pekerti*.

untuk analisis mendalam. Selain itu, pilihan jawaban terlalu jelas salah sehingga mudah untuk ditebak. Sebagai bentuk tindak lanjut, soal dengan kategori sedang dapat disimpan dalam bank soal untuk digunakan kembali pada tes mendatang. Sementara, soal dengan kategori sukar dapat diperjelas atau dibuat lebih kontekstual, dan soal dengan kategori mudah dapat dimodifikasi dengan menambahkan elemen yang lebih kompleks.

4. Daya Beda

Daya beda butir soal adalah suatu kemampuan soal untuk membedakan peserta didik pada tingkat kemampuan menjawab. Menurut Sudijono, sebuah soal memiliki daya beda apabila dijawab betul oleh kebanyakan peserta didik yang berkemampuan tinggi dan tidak dapat dijawab oleh peserta didik dengan kemampuan rendah. Rohmad mengklasifikasikan kriteria daya beda. Soal dikategorikan sangat buruk apabila berada pada koefisien $< 0,20$. Soal dikategorikan cukup baik apabila berada pada koefisien $0,20 - 0,40$. Soal dikategorikan baik apabila berada pada koefisien $0,41 - 0,70$. Soal dikategorikan sangat baik apabila berada pada koefisien $\geq 0,70$.

Berdasarkan hasil analisis, soal dengan daya beda sangat buruk ditemukan sebanyak 4 soal (10%) pada kunci jawaban pertama, 10 soal (25%) pada kunci jawaban kedua, dan 8 soal (20%) pada kunci jawaban ketiga. Adapun soal dengan daya beda cukup baik sejumlah 16 soal (40%) pada kunci jawaban pertama, 17 soal (42,5%) pada kunci jawaban kedua, dan 16 soal (40%) pada kunci jawaban ketiga. Adapun soal dengan daya beda baik ditemukan sebanyak 17 soal (42,5%) pada kunci jawaban pertama, 13 soal (32,5%) pada kunci jawaban kedua, dan 16 soal (40%) pada kunci jawaban ketiga. Soal dengan daya beda sangat baik hanya ditemukan pada kunci jawaban pertama sebanyak 3 soal (7,5%).

Dari ketiga kunci jawaban yang dianalisis, dapat disimpulkan bahwa sebagian besar berada pada kriteria cukup baik. Hal ini menunjukkan bahwa soal telah mampu membedakan peserta didik berdasarkan tingkat kognitif. Namun, proporsi soal dengan daya beda sangat buruk masih perlu diperhatikan.

Berikut merupakan contoh soal yang dinyatakan memiliki daya beda sangat buruk yaitu soal nomor 13, 24, dan 31. Dari ketiga soal tersebut memiliki nilai daya beda yang sangat rendah sehingga menghasilkan daya beda yang rendah pula.

Adapun faktor yang menyebabkan rendahnya daya beda antara lain redaksi soal yang tidak jelas, pengecoh tidak berfungsi dengan efektif, atau materi yang tidak dikuasai oleh peserta didik. Sebagai tindak lanjut, soal dengan daya beda $< 0,20$ soal dapat dimodifikasi dengan menambahkan elemen yang lebih kompleks.

5. Hubungan Antara Daya Beda dan Tingkat Kesukaran

Tingkat daya beda tinggi umumnya cenderung memiliki tingkat kesukaran yang sedang atau lebih tinggi¹⁷. Namun, tidak selalu soal yang memiliki tingkat kesukaran yang tinggi menunjukkan daya beda yang tinggi. Bisa saja soal yang tingkat kesukarannya baik atau cukup memiliki daya beda nol, jika kelompok siswa yang pintar atau tinggi dan kelompok siswa yang rendah memiliki proporsi jawaban benar yang sama. Menurut Susetyo, tingkat kesukaran adalah seberapa sukar suatu butir dijawab oleh peserta tes atau responden. Sedangkan daya beda soal adalah kemampuan sesuatu soal untuk membedakan antara peserta didik antar kemampuan tinggi dengan rendah.

¹⁷ Rohmad, *Pengembangan Instrumen Evaluasi Dan Penelitian*.

Berdasarkan hasil analisis menggunakan ketiga kunci jawaban untuk mengetahui hubungan antara tingkat kesukaran dan daya beda pada soal ASAS, maka hasil analisis menunjukkan bahwa soal dengan kategori tingkat kesukaran sedang cenderung memiliki nilai daya beda tinggi dibandingkan soal yang sangat mudah atau sukar.

Berikut merupakan contoh soal yang dinyatakan memiliki tingkat kesukaran sedang cenderung memiliki nilai daya beda tinggi, yaitu soal nomor 10, 15, dan 30. Hal ini sejalan dengan pendapat Arikunto menyampaikan bahwa suatu soal tes hendaknya tidak terlalu sukar dan tidak terlalu mudah¹⁸. Namun, soal dengan daya beda sangat buruk seperti soal nomor 21, 24, dan 26 soal tersebut memiliki tingkat kesukaran mudah sehingga tidak efektif dalam membedakan kemampuan peserta didik.

Dengan demikian, dapat disimpulkan bahwa soal sukar dan mudah masih dapat digunakan apabila memiliki daya beda yang memadai. Adapun tindak lanjut yaitu soal dengan tingkat kesukaran sedang dan daya beda baik dapat disimpan dalam bank soal. Sementara itu, soal dengan daya beda sangat buruk perlu diperbaiki karena dapat menurunkan validitas.

6. Efektivitas Pengecoh

Efektivitas pengecoh adalah opsi jawaban yang salah namun dirancang menyerupai jawaban benar agar dapat menjebak peserta didik yang belum memahami materi dengan baik¹⁹. Menurut Suharsimi Arikunto, secara umum suatu pengecoh dikatakan telah berfungsi dengan baik apabila pengecoh dipilih minimal 5% oleh peserta didik. sedangkan menurut Fernandes,

¹⁸ Mochammad Noor Akhmadi, "Analisis Butir Soal Evaluasi Tema 1 Kelas 4 Sdn Plumbungan Menggunakan Program Anates," *Ed-Humanistics : Jurnal Ilmu Pendidikan* 6, no. 1 (April 30, 2021): 799–806.

¹⁹ Rohmad, *Pengembangan Instrumen Evaluasi Dan Penelitian*.

pengecoh berfungsi dengan baik apabila dipilih minimal 2% oleh peserta didik.

Zainal Arifin mengklasifikasikan efektivitas pengecoh, soal dikatakan memiliki fungsi pengecoh, soal dikatakan memiliki fungsi pengecoh sangat baik apabila berada pada koefisien 76% - 125%,. Soal dikatakan baik apabila berada pada koefisien 51% - 75% atau 126% - 150%. Soal dikatakan kurang baik apabila berada pada koefisien 26% - 50% atau 151% - 175%. Soal dikatakan buruk apabila berada pada koefisien 0% - 25% atau 176% - 200%. Soal dianggap sangat buruk apabila berada pada koefisien lebih dari 200.

Hasil analisis efektivitas pengecoh membuktikan bahwa pada kunci jawaban pertama terdapat 4 soal sangat baik dengan prosentase 12,5%. 10 soal baik dengan prosentase 31,25%. 6 soal kurang baik dengan prosentase 18,75%. 12 soal buruk dengan prosentase 37,5%. Dan 8 soal sangat buruk dengan prosentase 25%. Pada kunci jawaban 2 terdapat 3 soal sangat baik dengan prosentase 9,375%. 5 soal baik dengan prosentase 15,625 %. 9 soal kurang baik dengan prosentasi 28,125%. 11 soal buruk dengan prosentase 34,375%. 12 soal sangat buruk dengan prosentase 37,5%. Pada kunci jawaban 3 terdapat 4 soal sangat baik dengan prosentase 12,5%. 8 soal baik dengan prosentase 25%. 10 soal kurang baik dengan prosentase 31,25%. 14 soal buruk dengan prosentase 43,75%. 4 soal sangat buruk dengan prosentase 12,5%

Berdasarkan hasil analisis terhadap soal ASAS mata pelajaran SKI kelas XI MAN 3 Cilacap Tahun ajaran 2024/2025, efektivitas pengecoh bervariasi pada masing-masing kunci jawaban. Namun, dapat disimpulkan bahwa proporsi soal dengan pengecoh yang dikategorikan “buruk” dan “sangat buruk” tergolong tinggi. Adapun faktor yang menyebabkan

efektivitas pengecoh dikategorikan buruk dan sangat buruk, yaitu sebagian besar pengecoh tidak berhasil menarik perhatian peserta didik. Hal ini bisa disebabkan oleh rumusan pengecoh yang terlalu mudah dikenali sebagai jawaban salah, atau justru terlalu mirip sehingga tidak membingungkan peserta didik. Sebagai tindak lanjut, soal dengan pengecoh yang tidak berfungsi, maka perlu diperbaiki agar lebih menarik bagi peserta didik yang kurang memahami materi.

7. Analisis Soal Pilihan Ganda Kompleks (*Complex Multiple Choice*)

Dalam evaluasi terhadap instrumen ASAS, ditemukan 10 butir soal yang merupakan pilihan ganda kompleks. Pilihan ganda kompleks dapat membantu peserta didik untuk berfikir kritis karena tidak asal menebak jawaban benar. Hal ini sesuai dengan pendapat Sax, bahwa bentuk soal pilihan ganda kompleks mempunyai tingkat kesulitan yang tinggi, sehingga menuntut peserta didik untuk menguasai dan memahami materi pelajaran secara menyeluruh. Namun, adanya soal pilihan ganda kompleks pada ASAS menunjukkan permasalahan yang signifikan.

Adapun faktor yang menyebabkan soal pilihan ganda kompleks pada ASAS mata pelajaran SKI kelas XI MAN 3 Cilacap Tahun ajaran 2024/2025 menunjukkan permasalahan yang signifikan diantaranya yaitu tidak adanya disertakannya petunjuk atau penanda khusus yang menjelaskan bahwa soal tersebut merupakan jenis pilihan ganda kompleks. Akibatnya, peserta didik memperlakukan soal tersebut sama seperti soal pilihan ganda tunggal, yaitu hanya memilih satu jawaban benar. Ketidaksesuaian ini menyebabkan potensi jawaban peserta didik menjadi tidak akurat dalam merefleksikan pemahaman yang sebenarnya, bahkan bisa menurunkan nilai secara tidak adil. Selain itu, secara teknis, sistem penilaian manual maupun

digital di banyak satuan pendidikan belum sepenuhnya dirancang untuk mendeteksi dan menilai jawaban dengan format lebih dari satu kunci secara objektif dan efisien. Hal ini menjadikan soal pilihan ganda kompleks kurang praktis digunakan dalam asesmen tanpa persiapan dan sosialisasi yang memadai.

Sebagai tindak lanjut, peneliti memutuskan untuk melakukan perbaikan terhadap 10 soal yang merupakan soal kompleks. Perbaikan dilakukan dengan mengubah struktur soal menjadi soal pilihan ganda biasa yang hanya memiliki satu jawaban benar. Langkah ini tidak hanya memperjelas instruksi kepada peserta didik, tetapi memudahkan proses analisis soal.

KESIMPULAN

Berdasarkan hasil penelitian mengenai kualitas butir soal ASAS mata pelajaran SKI kelas XI MAN 3 Cilacap merupakan soal yang diajukan dalam Musyawarah Guru Mata Pelajaran (MGMP) Karisidenan Banyumas. Untuk mengetahui kualitas soal ASAS mata pelajaran SKI kelas XI MAN 3 Cilacap maka perlu dikakukan sebuah analisis terhadap kualitas butir soal ASAS. Adapun komponennya adalah uji validitas, reliabilitas, tingkat kesukaran, daya beda dan efektivitas pengecoh.

Adapun bentuk soal merupakan soal pilihan ganda (*Multiple Choice*) yang terdiri dari 40 pilihan ganda dengan jumlah option sebanyak lima (A, B, C, D, dan E). Hal yang menarik dari penelitian ini adalah, dari 40 butir soal, terdapat 10 butir soal yang merupakan soal pilihan ganda kompleks (*Complex Multiple Choice*) dimana 10 soal tersebut memiliki jawaban lebih dari satu *options*. Maka, analisis dilakukan sebanyak 3 kali pada masing-masing komponen, yaitu menggunakan kunci jawaban pertama sebagai analisis pertama, kunci jawaban kedua sebagai analisis kedua dan kunci jawaban ketiga sebagai analisis ketiga.

Berdasarkan informasi yang diperoleh, ASAS diikuti oleh 125 peserta didik yang terbagi menjadi 4 kelas, yaitu kelas XI F1 sejumlah 21 peserta didik, kelas XI F2 sejumlah 32 peserta didik, kelas XI F3 sejumlah 34 peserta didik, dan kelas XI F4 sejumlah 38 peserta didik. Analisis dilakukan dengan memanfaatkan bantuan aplikasi *Microsoft Exel* untuk uji validitas dan reliabilitas, dan menggunakan bantuan aplikasi ANATES V4 untuk tingkat kesukaran, daya beda, dan efektivitas pengecoh yang didasarkan pada teori klasik.

Berikut adalah hasil analisis pada kualitas butir soal asesmen sumatif mata pelajaran sejarah kebudayaan islam kelas XI tahun ajaran 2024/2025:

1. Analisis Kualitas Butir Soal Asesmen Sumatif Mata Pelajaran Sejarah Kebudayaan Islam Kelas XI MAN 3 Cilacap menurut kunci jawaban 1 mengenai validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh, maka menghasilkan hasil sebagai berikut:
 - a. Validitas

Hasil uji validitas membuktikan bahwa pada kunci jawaban 1 terdapat 36 soal (90%) valid, yang artinya nilai validitas berada pada koefisien $> 0,00$ dan 4 soal (15%) tidak valid, yang artinya nilai validitas berada pada koefisien $< 0,00$.

b. Reliabilitas

Hasil uji reliabilitas membuktikan bahwa pada kunci jawaban 1 nilai reliabilitas adalah 0,773, hal ini menunjukkan bahwa soal memiliki reliabilitas dengan kategori cukup karena berada pada koefisien 0,70 - 0,79.

c. Tingkat kesukaran

Hasil analisis tingkat kesukaran berdasarkan tabel 4.6. membuktikan bahwa pada kunci jawaban 1 terdapat 4 soal dengan kategori sukar dengan prosentase 10%. 16 soal dengan kategori sukar dengan prosentase 40%, dan 20 soal dengan kategori mudah dengan prosentase 50%.

d. Daya beda

Hasil analisis daya beda membuktikan bahwa pada kunci jawaban 1 terdapat 4 soal sangat buruk dengan prosentase 10%. 16 soal cukup baik dengan prosentase 40%. 17 soal baik dengan prosentase 42,5%, dan 3 soal sangat baik dengan prosentase 7,5%.

e. Efektivitas pengecoh

Hasil dari analisis efektivitas pengecoh soal ASAS mata pelajaran SKI kelas XI MAN 3 Cilacap dengan kunci jawaban 1 terdapat 18 soal dengan efektivitas pengecoh sangat baik dengan prosentase 45%. 11 soal dengan efektivitas pengecoh baik dengan prosentase 27,5%. 9 soal dengan efektivitas pengecoh cukup dengan prosentase 22,5%, dan 1 soal dengan efektivitas pengecoh kurang baik dengan prosentase 5%.

2. Analisis Kualitas Butir Soal Asesmen Sumatif Mata Pelajaran Sejarah Kebudayaan Islam Kelas XI MAN 3 Cilacap menurut kunci jawaban 2 mengenai validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh, maka menghasilkan hasil sebagai berikut:

a. Validitas

Hasil uji validitas pada kunci 2 membuktikan bahwa terdapat 34 soal (85%) valid, yang artinya nilai validitas berada pada koefisien $> 0,00$ dan 6 soal (15%) tidak valid, yang artinya nilai validitas berada pada koefisien $< 0,00$.

b. Reliabilitas

Hasil uji reliabilitas pada kunci jawaban 2 nilai reliabilitas adalah 1,025, yang artinya reliabilitas sangat tinggi karena berada pada koefisien $\geq 0,90$.

c. Tingkat kesukaran

Hasil analisis tingkat kesukaran kunci jawaban 2 terdapat 6 soal dengan kategori sukar dengan prosentase 15%. 15 soal dengan kategori sedang dengan prosentase 37,5%, dan 19 soal dengan kategori mudah dengan prosentase 47,5%.

d. Daya beda

Daya beda membuktikan bahwa pada kunci jawaban 2 terdapat 10 soal sangat buruk dengan prosentase 25 %. 17 soal cukup baik dengan prosentase 42,5%. 13 soal baik dengan prosentase 32,5%.

e. Efektivitas Pengecoh

Hasil dari analisis efektivitas pengecoh soal ASAS mata pelajaran SKI kelas XI MAN 3 Cilacap dengan kunci jawaban 2 terdapat 13 soal dengan efektivitas pengecoh sangat baik dengan prosentase 32,5%. 15 soal dengan efektivitas pengecoh baik dengan prosentase 37,5. 8 soal dengan efektivitas pengecoh cukup dengan prosentase 20%. 2 soal dengan efektivitas pengecoh kurang baik dengan prosentase 5%, dan 2 soal dengan efektivitas pengecoh sangat tidak baik dengan prosentase 5%.

3. Analisis Kualitas Butir Soal Asesmen Sumatif Mata Pelajaran Sejarah Kebudayaan Islam Kelas XI MAN 3 Cilacap menurut kunci jawaban 3 mengenai validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh, maka menghasilkan hasil sebagai berikut:

a. Validitas

Hasil uji validitas pada kunci 3 membuktikan bahwa terdapat 34 soal (85%) valid, yang artinya nilai validitas berada pada koefisien $> 0,00$ dan 6

soal (15%) tidak valid, yang artinya nilai validitas berada pada koefisien < 0,00.

b. Reliabilitas

Hasil uji reliabilitas pada kunci jawaban 3 nilai reliabilitas adalah 0,739, yang artinya reliabilitas cukup karena berada pada koefisien 0,70 - 0,79.

c. Tingkat Kesukaran

Hasil analisis tingkat kesukaran membuktikan bahwa pada kunci jawaban 3 terdapat 4 soal dengan kategori sukar dengan prosentase 10%. 15 soal dengan kategori sedang dengan prosentase 37,5%, dan 21 soal dengan kategori mudah dengan prosentase 52,5%.

d. Daya Beda

Daya beda membuktikan bahwa pada kunci jawaban 3 terdapat 8 soal sangat buruk dengan prosentase 20%. 16 soal cukup baik dengan prosentase 40%. 16 soal baik dengan prosentase 40%.

e. Efektivitas Pengecoh

Hasil dari analisis efektivitas pengecoh soal ASAS mata pelajaran SKI kelas XI MAN 3 Cilacap dengan kunci jawaban 3 terdapat 12 soal dengan efektivitas pengecoh sangat baik dengan prosentase 30%. 13 soal dengan efektivitas pengecoh baik dengan prosentase 32,5%. 12 soal dengan efektivitas pengecoh cukup dengan 30%, dan 3 soal dengan efektivitas pengecoh kurang baik dengan prosentase 7,5%.

4. Hubungan Tingkat Kesukaran dan Daya Beda

Tingkat daya beda tinggi umumnya cenderung memiliki tingkat kesukaran yang sedang atau lebih tinggi. Namun, tidak selalu soal yang memiliki tingkat kesukaran yang tinggi menunjukkan daya beda yang tinggi. Jika kita melihat hubungan antara tingkat kesukaran dan daya pembeda, maka daya pembeda memiliki peranan yang lebih penting.

Pada hasil analisis baik menggunakan kunci jawaban 1, 2, maupun 3 tidak semua soal ASAS dengan kategori daya beda sangat buruk memiliki nilai tingkat kesukaran dengan kategori sukar.

5. Analisis Soal Pilihan Ganda Kompleks (*Complex Multiple Choice*)

Dari hasil analisis terhadap soal ASAS mata pelajaran SKI kelas XI MAN 3 Cilacap terdiri dari 40 soal pilihan ganda. 10 diantaranya adalah soal kompleks (*Complex Multiple Choice*) dimana soal ini memiliki lebih dari satu jawaban benar. Oleh karena itu, penyusunan soal pilihan ganda kompleks perlu disertai dengan petunjuk dan apabila diperlukan, maka dilakukan penyederhanaan soal menjadi pilihan ganda biasa. Dalam penelitian ini, peneliti melakukan langkah perbaikan pada 10 soal agar soal tetap memiliki satu jawaban benar.

REFERENSI

- Arifin, Z. (2012). *Evaluasi Pembelajaran*. Direktorat Jendral Pendidikan Islam Kementerian Agama.
- Kaka, L., Bano, V. O., & Njoeroemana, Y. (2024). Efektivitas Analisis Butir Soal Pilihan Ganda Menggunakan Aplikasi Anates Di SMPN2 Kanatang. *Jurnal Inovasi Penelitian*, 1441(1450), 4-9. <https://stp-mataram.e-journal.id/JIP/article/view/3124/2473>
- Kemendikbudristek. (2022). Peraturan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi No 21 Tahun 2022 Tentang Standar Penilaian Pendidikan pada Pendidikan Anak Usia Dini, Jenjang Pendidikan Dasar, dan Jenjang Pendidikan Menengah. *Kementerian Pendidikan Dan Kebudayaan Riset Dan Teknologi*, 9.
- Khayati, C., Amri, M., & Sodiq, S. (2021). Analysis of Objective Test Type of Multiple Choice in Final Semester Assessment Problems (PAS) Japanese Language Class VI. *Proceeding of International Conference on Arts and Humanities: International Conference on Cultural Studies and Applied Linguistics (ICCSAL) Analysis*, 10-14.
- Magdalena, I., Fauziah, S. N., Faziah, S. N., & Nopus, F. S. (2021). Analisis Validitas, Reliabilitas, Tingkat Kesukaran, dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan. *BINTANG : Jurnal Pendidikan Dan Sains*, 3(2), 206. <https://ejournal.stitpn.ac.id/index.php/bintang>
- Mochammad Noor Akhmadi. (2021). ANALISIS BUTIR SOAL EVALUASI TEMA 1 KELAS 4 SDN PLUMBUNGAN MENGGUNAKAN PROGRAM ANATES. *Ed-Humanistics : Jurnal Ilmu Pendidikan*, 6(1), 799-806. <https://doi.org/10.33752/ed->

- Pohan, N. A., Mawaddah, T., Batubara, I. H., & Rahmah, M. F. (2023). pengembangan alat evaluasi. *Jurnal Pendidikan Tambusai*, 7, 30701-30705.
<https://doi.org/https://doi.org/10.31004/jptam.v7i3.11968>
- Rahman, A. A., & Nasryah, C. E. (2019). Evaluasi Pembelajaran. In Funky (Ed.), *Uwais Inspirasi Indonesia* (1st ed.). Uwais Inspirasi Indonesia.
<https://repository.bbg.ac.id/handle/837>
- Rohmad. (2017). *Pengembangan Instrumen Evaluasi dan Penelitian* (Y. AF (ed.)). Kalimedia.
- Rohmad. (2025). Asesmen Pembelajaran Pendidikan Agama Islam dan budi Pekerti. In *Zahira Media Publisher* (1st ed., Vol. 11, Issue 1). Zahira Media Publisher.
http://scioteca.caf.com/bitstream/handle/123456789/1091/RED2017-Eng-8ene.pdf?sequence=12&isAllowed=y%0Ahttp://dx.doi.org/10.1016/j.regsciurbeco.2008.06.005%0Ahttps://www.researchgate.net/publication/305320484_SISTEM_PEMBETUNGAN_TERPUSAT_STRATEGI_MELESTARI
- Santosa, S., & Badawi, J. A. (2022). Analisis Butir Soal Pilihan Ganda Tema Pertumbuhan dan Perkembangan Makhluk Hidup Kelas III Madrasah Ibtidaiyah. *Jurnal Basicedu*, 6(2), 1678-1686. <https://doi.org/10.31004/basicedu.v6i2.2206>
- Syaadah, R., Ary, M. H. A. A., Silitonga, N., & Rangkuty, S. F. (2023). Pendidikan Formal, Pendidikan Non Formal Dan Pendidikan Informal. *Pema (Jurnal Pendidikan Dan Pengabdian Kepada Masyarakat)*, 2(2), 125-131.
<https://doi.org/10.56832/pema.v2i2.298>
- Undang-Undang Nomor 20 Tahun 2003 Tentang Sistem Pendidikan Nasional, Pub. L. No. 20, 2010 29 (2003). [https://piaud.uin-suka.ac.id/media/dokumen_akademik/43_20210506_Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional.pdf](https://piaud.uin-suka.ac.id/media/dokumen_akademik/43_20210506_Undang-Undang%20Nomor%2020%20Tahun%202003%20tentang%20Sistem%20Pendidikan%20Nasional.pdf)
- Widodo, S., Ladyani, F., Asrianto, L. O., Rusdi, Khairunnisa, Lestari, S. M. P., Wijayanti, D. R., Devriany, A., Hidayat, A., Dalfian, Nurcahyati, S., Sjahriani, T., Armi, Widya, N., & Rogayah. (2023). Metodologi Penelitian. In *Cv Science Techno*

Sindoro

CENDIKIA PENDIDIKAN

ISSN: 3025-6488

Direct. CV Science Techno Direct.

Vol. 17 No 2

Palagiarism Check 02/234/67/78

Prev DOI : 10.9644/sindoro.v3i9.252