

Analisis Kualitas Butir-Butir Soal Menggunakan Iteman sebagai Evaluasi Hasil Belajar Pendidikan Agama Islam

Nurlaila¹, Remiswal², Khadijah³

^{1,2,3}*Pascasarjana UIN Imam Bonjol Padang, Indonesia*

Email: nurlaila080600@gmail.com¹, remiswal@uinib.ac.id², khadijahmpd@uinib.ac.id³

Abstrak

Penelitian ini bertujuan untuk mengevaluasi butir soal berdasarkan tingkat kesulitan, daya pembeda, dan efektivitas pengecoh pada ujian ulangan siswa kelas X di SMA Negeri 1 Rao Utara. Pendekatan yang digunakan adalah deskriptif kuantitatif, dengan sampel dari kelas X IPA 1, X IPS 1, dan X IPS 2 yang berjumlah 200 siswa. Analisis dilakukan terhadap 40 butir soal menggunakan ITEMAN untuk menilai tingkat kesulitan, di mana ditemukan 15 soal sulit, 15 soal sedang, dan 10 soal mudah. Evaluasi ini menyarankan untuk meninjau kembali tingkat kesulitan soal. Selanjutnya, analisis terhadap daya pembeda menunjukkan bahwa 7 soal sangat baik, 17 soal baik, 9 soal cukup, dan 7 soal perlu diperbaiki karena dinilai kurang baik. Hasil analisis juga menunjukkan bahwa ada tiga soal yang perlu diganti karena tidak efektif dalam mengecoh siswa dalam menjawabnya.

Kata Kunci: analisis kualitas butir soal, Iteman, evaluasi hasil belajar PAI

PENDAHULUAN

Penilaian adalah kegiatan pengumpulan informasi tentang ketercapaian tujuan pembelajaran (Nuriyah 2016). Penilaian memiliki pengaruh yang signifikan terhadap keberhasilan pendidikan (Setiadi 2016). Kegiatan penilaian ini adalah tugas guru yang diawali dengan perencanaan, melakukan penilaian, dan melaporkan hasil. Dengan demikian, guru memegang peran penting dalam pembelajaran, khususnya dalam kegiatan evaluasi dan penilaian hasil belajar siswa. Penilaian akan memberikan makna jika dapat memberikan informasi yang tepat untuk mengambil kebijakan (Pradani and Efendi 2023).

Dalam dunia pendidikan, evaluasi hasil belajar memegang peranan penting dalam mengukur efektivitas proses pembelajaran dan pencapaian tujuan pendidikan. Salah satu instrumen utama yang digunakan dalam evaluasi ini adalah tes atau butir-butir soal. Kualitas butir soal yang baik akan memberikan informasi yang akurat mengenai kemampuan dan pemahaman siswa terhadap materi yang diajarkan. Dalam konteks pendidikan agama Islam, evaluasi yang efektif sangat penting untuk memastikan bahwa siswa tidak hanya memahami konsep-konsep agama, tetapi juga mampu mengaplikasikannya dalam kehidupan sehari-hari.

Analisis kualitas butir-butir soal merupakan langkah krusial dalam memastikan validitas dan reliabilitas tes. Salah satu alat yang dapat digunakan untuk menganalisis kualitas butir soal adalah Iteman. Iteman adalah perangkat lunak yang dirancang khusus untuk menganalisis item-item tes, memberikan data rinci mengenai tingkat kesulitan, daya beda, dan fungsi distraktor dari setiap butir soal. Dengan menggunakan Iteman, pendidik dapat mengevaluasi dan memperbaiki butir soal, sehingga tes yang digunakan menjadi lebih efektif dalam mengukur kompetensi siswa.

Penelitian ini bertujuan untuk menganalisis kualitas butir-butir soal yang digunakan dalam evaluasi hasil belajar Pendidikan Agama Islam di sebuah sekolah menengah atas. Dengan menggunakan Iteman, penelitian ini akan mengidentifikasi butir soal yang memenuhi kriteria kualitas yang baik serta mengidentifikasi butir soal yang perlu direvisi atau dihapus. Hasil analisis ini

diharapkan dapat memberikan masukan yang berharga bagi para pendidik dalam merancang tes yang lebih valid dan reliabel, serta meningkatkan kualitas pembelajaran Pendidikan Agama Islam secara keseluruhan.

Melalui analisis yang mendalam, penelitian ini juga bertujuan untuk memberikan gambaran mengenai sejauh mana butir-butir soal yang digunakan saat ini mampu mencerminkan kompetensi yang diharapkan dari siswa. Selain itu, penelitian ini akan membahas implikasi dari temuan tersebut terhadap praktik pengajaran dan evaluasi di bidang Pendidikan Agama Islam. Dengan demikian, penelitian ini tidak hanya berkontribusi pada peningkatan kualitas instrumen evaluasi, tetapi juga pada peningkatan kualitas pendidikan secara umum.

METODE

Subjek penelitian adalah siswa kelas X SMA Negeri 1 Rao Utara yang merespons ulangan harian dalam mata pelajaran Pendidikan Agama Islam. Penelitian ini mengadopsi pendekatan deskriptif kuantitatif, di mana peneliti berusaha untuk menggambarkan dan menginterpretasi obyek penelitian. Metode deskriptif ini bertujuan untuk mendapatkan informasi yang diperlukan untuk mendeskripsikan kualitas dari butir-butir soal tes. Data yang diperoleh berupa data numerik yang dianalisis statistik menggunakan program Item and Test Analysis (ITEMAN).

Analisis dilakukan terhadap ujian sekolah mata pelajaran Pendidikan Agama Islam di SMA Negeri 1 Rao Utara. Data dikumpulkan setelah peserta ujian menyelesaikan ujian, pengumpulan data dilakukan dengan menyalin jawaban siswa dari lembar jawaban ke dalam aplikasi Notepad secara berurutan dari nomor 1 hingga nomor 40. Aspek yang akan dievaluasi menggunakan program komputer ITEMAN mencakup validitas isi, reliabilitas, tingkat kesulitan, dan daya pembeda. Validasi dilakukan menggunakan alat evaluasi berupa lembar telaah soal yang mempertanyakan aspek materi, konstruksi, dan bahasa (Magdalena 2021). Sementara itu, reliabilitas, tingkat kesulitan, daya pembeda, dan efektivitas distraktor akan dianalisis melalui jawaban dari setiap butir soal yang dikerjakan peserta ujian, dengan bantuan program komputer ITEMAN. Soal nomor 1 sampai dengan 40 dijadikan sebagai dasar analisis data. Komponen Alpha dari digunakan untuk menentukan hasil reliabilitas. Daya beda soal dilihat berdasarkan koefisien korelasi point biserial, yaitu biser dan point biser. Tingkat kesukaran disesuaikan dengan proportional correct. Sedangkan pengecoh soal dapat ditinjau dari nilai proportional endorsing.

HASIL DAN PEMBAHASAN

Tujuan dari penelitian ini adalah untuk mengevaluasi kualitas beberapa butir soal ulangan mata pelajaran kelas X SMA 1 Rao Utara pada tahun ajaran 2023/2024. Evaluasi kualitas butir soal tersebut akan dilakukan dengan menganalisis tingkat kesulitan, daya pembeda, dan efektivitas pilihan jawaban yang tersedia. Total terdapat 40 butir soal yang disusun oleh guru SMA 1 Rao Utara, dan setiap soal memiliki lima opsi jawaban (A, B, C, D, dan E). Ujian ini diikuti oleh 200 siswa kelas X.

Analisis Butir Soal

Analisis butir soal adalah kegiatan yang dilakukan oleh seorang guru untuk menilai hasil belajar siswa dan meningkatkan kualitas soal yang telah disusun. Kegiatan ini melibatkan pengumpulan, ringkasan, dan penggunaan data dari jawaban siswa untuk membuat

kesimpulan yang tepat. Tujuannya adalah untuk mengevaluasi setiap butir soal agar dapat menghasilkan soal yang berkualitas saat diujikan. Analisis ini juga membantu dalam memperbaiki atau mengganti soal yang tidak memenuhi syarat untuk diujikan. Soal yang berkualitas adalah soal yang dapat memberikan informasi apakah siswa telah memahami materi yang diajarkan atau belum (Fitrianawati 2017)

Teknik Analisis Data

Analisis tingkat kesukaran butir soal

Analisis item atau analisis butir soal adalah mengkaji (menganalisis) tingkatan soal-soal tes agar adapun dari sisi kesukarannya dapat diperoleh tingkatan soal-soal yang mempunyai bobot cukup layak. Butir soal ini harus diketahui tingkatan kesukarannya, tingkatan bisa diketahui melalui jawaban-jawaban siswa tersebut. Ketika apabila sedikitnya siswa dalam menjawab dengan betul maka dikatakan bobot dari soal-soal termasuk sukar(sulit)(Kustati and Amelia 2023). Indeks Tingkat Kesulitan (ITK), juga dikenal sebagai indeks yang mengukur tingkat kesulitan butir soal dalam pengujian siswa, menunjukkan seberapa mudah atau sulitnya soal tersebut. Rentang skor yang digunakan berkisar dari 0,00 hingga 1,00. Tingkat kesulitan soal dibagi menjadi tiga kategori: mudah (0,71-1,00), sedang (0,31-0,70), dan sulit (0,00-0,30). Proses pengukuran tingkat kesulitan ini menggunakan program Iteman. Tingkat kesukaran soal adalah peluang untuk menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Semakin besar indeks tingkat kesukaran yang diperoleh dari hasil hitungan, berarti semakin mudah soal itu(Hendrayadi, Kustati, and Amelia 2024).

(Thorndike and Hagen 1969) memberikan (interpretasi) sebagai berikut:

- a. 0-0,30 = Soal kategori sukar
- b. 0,32-0,70 = Soal kategori sedang
- c. 0,72-1,00 = Soal kategori mudah

Sebanyak 200 peserta didik yang mengikuti tes ulangan harian untuk mata pelajaran Pendidikan Agama Islam, yang terdiri dari tes pilihan ganda dengan 40 butir soal. Setelah tes selesai, dilakukan koreksi dan pemberian skor seperti yang dijelaskan di bawah ini. Dengan menggunakan analisis pendapat Robert L. Thorndike dan Halen.

Penentuan tingkat kesulitan butir soal dilakukan dengan menghitung persentase siswa yang menjawab soal dengan benar. Jika jumlah siswa yang benar menjawab soal tersebut sedikit, ini menunjukkan bahwa soal tersebut sulit, dan sebaliknya. Berdasarkan informasi yang diberikan, terdapat penilaian tentang tingkat kesukaran dari berbagai butir soal dalam sebuah ujian atau tes. Ada tiga kategori kesukaran yang dapat disimpulkan dari tabel 1.

- 1) Butir soal yang sukar (15 butir): Nomor soal: 3, 5, 7, 8, 16, 17, 18, 20, 21, 22, 26, 28, 32, 38, dan 39.
- 2) Butir soal yang mudah (10 butir): Nomor soal: 2, 10, 12, 13, 14, 15, 19, 25, 29, 31, 33, 34, 35, 36, dan 37.
- 3) Butir soal yang mudah (10 butir): Nomor soal: 1, 4, 6, 9, 11, 23, 24, 27, 30, dan 40.

Tabel 1. Tingkat Kesukaran

No soal	Prop Correct	Keterangan	Tindakan	No soal	Prop Correct	Keterangan	Tindakan
1	0.800	Mudah	-	21	0.150	Sukar	Harus diturunkan tingkat kesukarannya
2	0.417	Sedang	-	22	0.117	Sukar	Harus diturunkan tingkat kesukarannya
3	0.167	Sukar	Harus diturunkan tingkat kesukarannya	23	0.867	Mudah	-
4	0.767	Mudah	-	24	0.917	Mudah	-
5	0.150	Sukar	Harus diturunkan tingkat kesukarannya	25	0.367	Sedang	-
6	0.917	Mudah	-	26	0.233	Sukar	Harus diturunkan tingkat kesukarannya
7	0.083	Sukar	Harus diturunkan tingkat kesukarannya	27	0.867	Mudah	-
8	0.067	Sukar	Harus diturunkan tingkat kesukarannya	28	0.300	Sukar	Harus diturunkan tingkat kesukarannya
9	0.783	Mudah	-	29	0.667	Sedang	-
10	0.467	Sedang	-	30	0.833	Mudah	-
11	0.733	Mudah	-	31	0.533	Sedang	-
12	0.600	Sedang	-	32	0.167	Sukar	Harus diturunkan tingkat kesukarannya
13	0.417	Sedang	-	33	0.567	Sedang	-
14	0.383	Sedang	-	34	0.650	Sedang	-
15	0.517	Sedang	-	35	0.600	Sedang	-
16	0.067	Sukar	Harus diturunkan tingkat kesukarannya	36	0.383	Sedang	-
17	0.217	Sukar	Harus diturunkan tingkat kesukarannya	37	0.667	Sedang	-
18	0.217	Sukar	Harus diturunkan tingkat kesukarannya	38	0.133	Sukar	Harus diturunkan tingkat kesukarannya
19	0.533	Sedang	-	39	0.083	Sukar	Harus diturunkan tingkat kesukarannya
20	0.200	Sukar	Harus diturunkan tingkat kesukarannya	40	0.833	Mudah	-

Dari klasifikasi ini, terlihat bahwa butir soal yang dianggap sukar dan sedang memiliki jumlah yang sama, yaitu masing-masing 15 butir soal. Butir soal yang dianggap mudah ada 10 butir soal. Saran untuk menurunkan atau mengevaluasi kembali tingkat kesukaran butir soal adalah langkah yang penting untuk memastikan bahwa semua butir soal memiliki kualitas yang seimbang dan relevan dengan kemampuan peserta ujian. Dengan menyesuaikan tingkat kesukaran butir soal, evaluasi akan membantu memastikan bahwa ujian tersebut memberikan ukuran yang akurat terhadap pemahaman dan kemampuan peserta ujian sesuai dengan harapan yang ditetapkan.

Analisis Daya Pembeda

Analisis daya pembeda dilakukan dengan mengevaluasi butir-butir soal dalam tes untuk menilai sejauh mana kemampuannya dalam mengukur pencapaian siswa. Manfaat dari daya pembeda pada butir-butir soal adalah dua hal. Pertama, meningkatkan kualitas setiap butir soal berdasarkan data empiris yang ada. Dengan ini, kita dapat menentukan

apakah butir-butir soal tersebut sudah memenuhi standar yang diharapkan, perlu direvisi, atau bahkan ditolak. Kedua, membantu mengidentifikasi sejauh mana butir soal dapat mengukur pencapaian siswa, baik dalam pemahaman materi yang diajarkan maupun sebaliknya (Marambaawang, Bano, and Enda 2023).

Daya pembeda merujuk pada sekelompok butir soal yang memiliki indeks daya beda yang dianggap memadai jika minimal mencapai 0,30 (Wijayanti and Mundilarto 2015). Cara untuk mengetahui indeks daya beda pada butir soal adalah dengan menggunakan bobot yang telah ditentukan.

- $D = 0,00-0,20$ adalah kategori dari item yang jelek
- $D = 0,20-0,40$ adalah kategori dari item yang cukup
- $D = 0,40-0,70$ adalah kategori dari item yang sudah baik
- $D = 0,70-1,00$ adalah kategori dari item yang sudah baik sekali

Adapun kriteria soal yang baik memiliki daya beda mulai rentang antara (0,40-0,70).

Tabel 2. Daya Beda

No soal	Biser	Keterangan	Tindakan	No soal	Biser	Keterangan	Tindakan
1	0.426	Baik	-	21	0.428	Baik	-
2	0.347	Cukup	-	22	0.458	Baik	-
3	0.304	Cukup	-	23	0.961	Baik sekali	-
4	0.441	Baik	-	24	1.000	Baik sekali	-
5	0.043	Jeles	Harus diganti	25	0.575	Baik	-
6	0.799	Baik	-	26	0.408	Baik	-
7	0.331	Cukup	-	27	0.976	Baik sekali	-
8	0.354	Cukup	-	28	0.260	Cukup	-
9	0.696	Baik	-	29	0.288	Cukup	-
10	0.499	Baik	-	30	0.736	Baik sekali	-
11	0.800	Baik sekali	-	31	0.614	Baik	-
12	0.452	Baik	-	32	-0.017	Jeles	Harus diganti
13	0.626	Baik	-	33	0.524	Baik	-
14	-0.143	Jeles	Harus diganti	34	0.772	Baik sekali	-
15	0.480	Baik	-	35	0.518	Baik	-
16	0.366	Jeles	Harus diganti	36	0.303	Cukup	-
17	0.411	Baik	-	37	0.730	Baik sekali	-
18	0.037	Jeles	Harus diganti	38	0.515	Baik	-
19	0.389	Cukup	-	39	0.143	Jeles	Harus diganti
20	0.021	Jeles	Harus diganti	40	0.390	Cukup	-

Berdasarkan hasil dari tabel 2, dapat disimpulkan bahwa kategori Butir Soal yang Kualitasnya Sangat Baik (7 butir soal): Nomor soal: 4, 23, 24, 27, 30, 34, dan 37. Butir-butir soal ini terdeteksi memiliki kualitas sangat baik berdasarkan hasil evaluasi atau analisis yang dilakukan. Ini berarti butir soal ini dianggap efektif dalam mengukur kemampuan atau pemahaman yang diuji. Kategori Butir Soal yang Kualitasnya Baik (17 butir soal): Nomor soal: 1, 4, 6, 9, 12, 13, 15, 17, 21, 22, 25, 26, 31, 33, 35, dan 38. Butir-butir soal ini juga menunjukkan kualitas yang baik dalam pengukuran kemampuan atau pemahaman siswa.

Meskipun tidak sebaik yang sangat baik, namun tetap memberikan kontribusi yang positif dalam evaluasi hasil tes. Kategori Butir Soal yang Kualitasnya Cukup (9 butir soal): Nomor soal: 2, 3, 7, 8, 19, 28, 29, 36, dan 40. Butir-butir soal ini memiliki kualitas yang cukup, artinya masih dapat digunakan untuk mengukur sebagian kemampuan atau pemahaman siswa dengan memadai, meskipun mungkin perlu beberapa penyesuaian atau perbaikan. Kategori Butir Soal yang Kualitasnya Jelek (7 butir soal): Nomor soal: 5, 14, 16, 18, 20, 32, dan 39. Butir-butir soal ini memiliki kualitas yang jelek, yang menandakan bahwa mereka tidak efektif dalam mengukur kemampuan atau pemahaman yang diuji. Butir soal ini mungkin perlu direvisi atau bahkan dipertimbangkan untuk penggantian.

Analisis Keberfungsian Pengecoh

Opsi pengecoh yang terbukti efektif dapat disimpan dalam bank soal karena mampu mengecoh siswa dengan baik. Namun, opsi pengecoh yang belum terbukti efektif perlu direvisi atau dihapus.

Tabel 3. Keberfungsian Pengecoh

No soal	Alternatif	Keterangan		Tindakan
1	A	0.800	80%	Berfungsi
	B	0.083	8%	Berfungsi
	C	0.083	8%	Berfungsi
	D	0.017	1,7%	Tidak berfungsi
	E	0.017	1,7%	Tidak berfungsi
2	A	0.067	6,7%	Berfungsi
	B	0.450	45%	Berfungsi
	C	0.417	41,7%	Berfungsi
	D	0.067	6,7%	Berfungsi
	E	0.000	0	Tidak berfungsi
3	A	0.167	16,7%	Berfungsi
	B	0.183	18,3%	Berfungsi
	C	0.100	10%	Berfungsi
	D	0.450	45%	Berfungsi
	E	0.100	10%	Berfungsi
4	A	0.767	76,7%	Berfungsi
	B	0.067	6,7%	Berfungsi
	C	0.100	10%	Berfungsi
	D	0.017	1,7%	Tidak berfungsi
	E	0.050	5%	Berfungsi
5	A	0.450	45%	Berfungsi
	B	0.150	15%	Berfungsi
	C	0.067	6,7%	Berfungsi
	D	0.217	21,7%	Berfungsi
	E	0.117	11,7%	Berfungsi
6	A	0.000	0	Tidak berfungsi
	B	0.017	1,7%	Tidak berfungsi
	C	0.050	5%	Berfungsi
	D	0.917	91,7%	Berfungsi
	E	0.017	1,7%	Tidak berfungsi
7	A	0.050	5%	Berfungsi

No soal		Alternatif		Keterangan	Tindakan
8	B	0.177	17,7%	Berfungsi	-
	C	0.417	41,7%	Berfungsi	-
	D	0.083	8,3%	Berfungsi	-
	E	0.333	33,3%	Berfungsi	-
	A	0.633	63,3%	Berfungsi	-
9	B	0.067	6,7%	Berfungsi	-
	C	0.083	8,3%	Berfungsi	-
	D	0.083	8,3%	Berfungsi	-
	E	0.133	13,3%	Berfungsi	-
	A	0.033	3,3%	Tidak berfungsi	Harus diganti
10	B	0.150	15%	Berfungsi	-
	C	0.033	3,3%	Tidak berfungsi	Harus diganti
	D	0.783	78,3%	Berfungsi	-
	E	0.000	0	Tidak berfungsi	Harus diganti
	A	0.183	18,3%	Berfungsi	-
11	B	0.467	46,7%	Berfungsi	-
	C	0.200	20%	Berfungsi	-
	D	0.067	6,7%	Berfungsi	-
	E	0.083	8,3%	Berfungsi	-
	A	0.100	10%	Berfungsi	-
12	B	0.067	6,7%	Berfungsi	-
	C	0.017	1,7%	Tidak berfungsi	Harus diganti
	D	0.083	8,3%	Berfungsi	-
	E	0.733	73,3%	Berfungsi	-
	A	0.600	60%	Berfungsi	-
13	B	0.200	20%	Berfungsi	-
	C	0.100	10%	Berfungsi	-
	D	0.017	1,7%	Tidak berfungsi	Harus diganti
	E	0.083	8,3%	Berfungsi	-
	A	0.083	8,3%	Berfungsi	-
14	B	0.383	38,3%	Berfungsi	-
	C	0.100	10%	Berfungsi	-
	D	0.417	41,7%	Berfungsi	-
	E	0.017	1,7%	Tidak berfungsi	Harus diganti
	A	0.383	38,3%	Berfungsi	-
15	B	0.383	38,3%	Berfungsi	-
	C	0.117	11,7%	Berfungsi	-
	D	0.083	8,3%	Berfungsi	-
	E	0.033	3,3%	Tidak berfungsi	Harus diganti
	A	0.517	51,7%	Berfungsi	-
16	B	0.067	6,7%	Berfungsi	-
	C	0.033	3,3%	Tidak berfungsi	Harus diganti
	D	0.250	25%	Berfungsi	-
	E	0.117	11,7%	Berfungsi	-
	A	0.067	6,7%	Berfungsi	-
	B	0.817	81,7%	Berfungsi	-

No soal	Alternatif			Keterangan	Tindakan
17	C	0.050	5%	Berfungsi	-
	D	0.033	3,3%	Tidak berfungsi	Harus diganti
	E	0.033	3,3%	Tidak berfungsi	Harus diganti
	A	0.217	21,7%	Berfungsi	-
	B	0.283	28,3%	Berfungsi	-
18	C	0.217	21,7%	Berfungsi	-
	D	0.200	20%	Berfungsi	-
	E	0.083	8,3%	Berfungsi	-
	A	0.217	21,7%	Berfungsi	-
	B	0.117	11,7%	Berfungsi	-
19	C	0.417	41,7%	Berfungsi	-
	D	0.150	15%	Berfungsi	-
	E	0.100	10%	Berfungsi	-
	A	0.533	53,3%	Berfungsi	-
	B	0.150	15%	Berfungsi	-
20	C	0.033	3,3%	Tidak berfungsi	Harus diganti
	D	0.033	3,3%	Tidak berfungsi	Harus diganti
	E	0.250	25%	Berfungsi	-
	A	0.100	10%	Berfungsi	-
	B	0.200	20%	Berfungsi	-
21	C	0.017	1,7%	Tidak berfungsi	Harus diganti
	D	0.617	61,7%	Berfungsi	-
	E	0.067	6,7%	Berfungsi	-
	A	0.267	26,7%	Berfungsi	-
	B	0.183	18,3%	Berfungsi	-
22	C	0.150	15%	Berfungsi	-
	D	0.067	6,7%	Berfungsi	-
	E	0.333	33,3%	Berfungsi	-
	A	0.117	11,7%	Berfungsi	-
	B	0.400	40%	Berfungsi	-
23	C	0.200	20%	Berfungsi	-
	D	0.200	20%	Berfungsi	-
	E	0.083	8,3%	Berfungsi	-
	A	0.867	86,7%	Berfungsi	-
	B	0.050	5%	Berfungsi	-
24	C	0.083	8,3%	Berfungsi	-
	D	0.000	0	Tidak berfungsi	Harus diganti
	E	0.000	0	Tidak berfungsi	Harus diganti
	A	0.917	91,7%	Berfungsi	-
	B	0.000	0	Tidak berfungsi	Harus diganti
25	C	0.033	3,3%	Tidak berfungsi	Harus diganti
	D	0.033	3,3%	Tidak berfungsi	Harus diganti
	E	0.017	1,7%	Tidak berfungsi	Harus diganti
	A	0.367	36,7%	Berfungsi	-
	B	0.250	25%	Berfungsi	-
				Berfungsi	-

No soal		Alternatif		Keterangan	Tindakan
26	C	0.067	6,7%	Berfungsi	-
	D	0.217	21,7%	Berfungsi	-
	E	0.100	10%	Berfungsi	-
	A	0.333	33,3%	Berfungsi	-
	B	0.233	23,3%	Berfungsi	-
	C	0.233	23,3%	Berfungsi	-
27	D	0.200	20%	Berfungsi	-
	E	0.000	0	Tidak berfungsi	Harus diganti
	A	0.867	86,7%	Berfungsi	-
	B	0.017	1,7%	Tidak berfungsi	Harus diganti
	C	0.033	3,3%	Tidak berfungsi	Harus diganti
	D	0.067	6,7%	Berfungsi	-
28	E	0.017	1,7%	Tidak berfungsi	Harus diganti
	A	0.300	3%	Tidak berfungsi	Harus diganti
	B	0.150	15%	Berfungsi	-
	C	0.150	15%	Berfungsi	-
	D	0.233	23,3%	Berfungsi	-
	E	0.167	16,7%	Berfungsi	-
29	A	0.050	5%	Berfungsi	-
	B	0.050	5%	Berfungsi	-
	C	0.100	10%	Berfungsi	-
	D	0.133	13,3%	Berfungsi	-
	E	0.667	66,7%	Berfungsi	-
	A	0.833	83,3%	Berfungsi	-
30	B	0.050	5%	Berfungsi	-
	C	0.033	3,3%	Tidak berfungsi	Harus diganti
	D	0.083	8,3%	Berfungsi	-
	E	0.000	0	Tidak berfungsi	Harus diganti
	A	0.533	53,3%	Berfungsi	-
	B	0.083	8,3%	Berfungsi	-
31	C	0.217	21,7%	Berfungsi	-
	D	0.150	15%	Berfungsi	-
	E	0.017	1,7%	Tidak berfungsi	Harus diganti
	A	0.133	13,3%	Berfungsi	-
	B	0.183	18,3%	Berfungsi	-
	C	0.250	25%	Berfungsi	-
32	D	0.267	26,7%	Berfungsi	-
	E	0.167	16,7%	Berfungsi	-
	A	0.050	5%	Berfungsi	-
	B	0.567	56,7%	Berfungsi	-
	C	0.100	10%	Berfungsi	-
	D	0.250	25%	Berfungsi	-
33	E	0.033	3,3%	Tidak berfungsi	Harus diganti
	A	0.650	65%	Berfungsi	-
	B	0.100	10%	Berfungsi	-
	C	0.100	10%	Berfungsi	-
	D	0.250	25%	Berfungsi	-
	E	0.033	3,3%	Tidak berfungsi	Harus diganti
34	A	0.650	65%	Berfungsi	-
	B	0.100	10%	Berfungsi	-
	C	0.100	10%	Berfungsi	-
	D	0.250	25%	Berfungsi	-
	E	0.033	3,3%	Tidak berfungsi	Harus diganti
	A	0.650	65%	Berfungsi	-

No soal		Alternatif		Keterangan	Tindakan
35	C	0.067	6,7%	Berfungsi	-
	D	0.100	10%	Berfungsi	-
	E	0.083	8,3%	Berfungsi	-
	A	0.600	60%	Berfungsi	-
	B	0.183	18,3%	Berfungsi	-
36	C	0.033	3,3%	Tidak berfungsi	Harus diganti
	D	0.067	6,7%	Berfungsi	-
	E	0.117	11,7%	Berfungsi	-
	A	0.133	13,3%	Berfungsi	-
	B	0.300	30%	Berfungsi	-
37	C	0.383	38,3%	Berfungsi	-
	D	0.100	10%	Berfungsi	-
	E	0.083	8,3%	Berfungsi	-
	A	0.667	66,7%	Berfungsi	-
	B	0.117	11,7%	Berfungsi	-
38	C	0.083	8,3%	Berfungsi	-
	D	0.117	11,7%	Berfungsi	-
	E	0.017	1,7%	Tidak berfungsi	Harus diganti
	A	0.333	33,3%	Berfungsi	-
	B	0.333	33,3%	Berfungsi	-
39	C	0.100	10%	Berfungsi	-
	D	0.133	13,3%	Berfungsi	-
	E	0.100	10%	Berfungsi	-
	A	0.083	8,3%	Berfungsi	-
	B	0.500	50%	Berfungsi	-
40	C	0.133	13,3%	Berfungsi	-
	D	0.217	21,7%	Berfungsi	-
	E	0.067	6,7%	Berfungsi	-
	A	0.833	83,3%	Berfungsi	-
	B	0.017	1,7%	Tidak berfungsi	Harus diganti
	C	0.083	8,3%	Berfungsi	-
	D	0.017	1,7%	Tidak berfungsi	Harus diganti
	E	0.050	5%	Berfungsi	-

Berdasarkan tabel 3, dapat disimpulkan bahwa ada tiga soal yang perlu direvisi karena persentase responden yang memilih opsi jawaban di bawah 5% dan tidak memberikan kebingungan kepada siswa. Soal-soal tersebut terdapat pada nomor-nomor berikut:

- 1) Nomor 1, opsi jawaban D dan E memiliki persentase yang rendah.
- 2) Nomor 2, opsi jawaban E memiliki persentase yang rendah.
- 3) Nomor 4, opsi jawaban D memiliki persentase yang rendah.
- 4) Nomor 6, opsi jawaban A, B, dan E memiliki persentase yang rendah.
- 5) Nomor 9, opsi jawaban A, C, dan E memiliki persentase yang rendah.
- 6) Nomor 11, opsi jawaban C memiliki persentase yang rendah.
- 7) Nomor 12, opsi jawaban D memiliki persentase yang rendah.

- 8) Nomor 13, opsi jawaban E memiliki persentase yang rendah.
- 9) Nomor 15, opsi jawaban C memiliki persentase yang rendah.
- 10) Nomor 16, opsi jawaban D dan E memiliki persentase yang rendah.
- 11) Nomor 19, opsi jawaban C dan D memiliki persentase yang rendah.
- 12) Nomor 20, opsi jawaban C memiliki persentase yang rendah.
- 13) Nomor 22, opsi jawaban D dan E memiliki persentase yang rendah.
- 14) Nomor 24, opsi jawaban B, C, D, dan E memiliki persentase yang rendah.
- 15) Nomor 26, opsi jawaban E memiliki persentase yang rendah.
- 16) Nomor 27, opsi jawaban B, C, dan E memiliki persentase yang rendah.
- 17) Nomor 28, opsi jawaban A memiliki persentase yang rendah.
- 18) Nomor 30, opsi jawaban C memiliki persentase yang rendah.
- 19) Nomor 31, opsi jawaban E memiliki persentase yang rendah.
- 20) Nomor 33, opsi jawaban E memiliki persentase yang rendah.
- 21) Nomor 35, opsi jawaban C memiliki persentase yang rendah.
- 22) Nomor 37, opsi jawaban E memiliki persentase yang rendah.
- 23) Nomor 40, opsi jawaban B dan D memiliki persentase yang rendah.

Soal yang bermasalah

Ada kebingungan dari ITEMAN untuk menentukan kunci jawaban karena nilai biser/poin biser lebih tinggi dibandingkan dengan kunci jawaban yaitu yang terdapat pada soal nomor: 5,8,14,16,18,20,28,32,39.

KESIMPULAN

Berdasarkan analisis tingkat kesukaran butir soal, ditemukan bahwa 15 butir soal memiliki tingkat kesukaran tinggi. Butir soal ini berada pada nomor 3, 5, 7, 8, 16, 17, 18, 20, 21, 22, 26, 28, 32, 38, dan 39. Selain itu, 15 butir soal menunjukkan tingkat kesukaran sedang, yaitu pada nomor 2, 10, 12, 13, 14, 15, 19, 25, 29, 31, 33, 34, 35, 36, dan 37. Sedangkan 10 butir soal memiliki tingkat kesukaran mudah, yang mencakup nomor 1, 4, 6, 9, 11, 23, 24, 27, 30, dan 40.

Dalam hal daya pembeda, terdapat 7 butir soal dengan kualitas sangat baik, yang terdapat pada nomor 4, 23, 24, 27, 30, 34, dan 37. Kemudian, 17 butir soal dikategorikan memiliki kualitas baik, yaitu pada nomor 1, 4, 6, 9, 12, 13, 15, 17, 21, 22, 25, 26, 31, 33, 35, dan 38. Selanjutnya, 9 butir soal dengan kualitas cukup ditemukan pada nomor 2, 3, 7, 8, 19, 28, 29, 36, dan 40. Terakhir, 7 butir soal berada dalam kategori kualitas buruk, yaitu pada nomor 5, 14, 16, 18, 20, 32, dan 39.

Analisis efektivitas pengecoh menunjukkan bahwa beberapa butir soal memiliki pengecoh yang efektif pada pilihan tertentu. Observasi menunjukkan butir-butir soal dengan pengecoh efektif, antara lain pada nomor 1 (D dan E), 2 (E), 4 (D), 6 (A, B, dan E), 9 (A, C, dan E), 11 (C), 12 (D), 13 (E), 15 (C), 16 (D dan E), 19 (C dan D), 20 (C), 22 (D dan E), 24 (B, C, D, dan E), 26 (E), 27 (B, C, dan E), 28 (A), 30 (C), 31 (E), 33 (E), 35 (C), 37 (E), dan 40 (B dan D).

Untuk memastikan validitas dan kualitas butir soal yang dimasukkan ke dalam bank soal, diperlukan analisis lanjutan terhadap setiap butir soal. Hal ini penting untuk memastikan bahwa setiap soal memiliki tingkat kesukaran, daya pembeda, dan pengecoh yang sesuai, sehingga dapat digunakan sebagai instrumen evaluasi yang efektif dan akurat dalam mengukur hasil belajar siswa.

DAFTAR PUSTAKA

- Fitrianawati, Meita. 2017. "Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru Dan Hasil Belajar Peserta Didik," May. <http://publikasiilmiah.ums.ac.id/handle/11617/9117>.
- Hendrayadi, Hendrayadi, Martin Kustati, and Rezki Amelia. 2024. "ANALISIS ULANGAN HARIAN MATA PELAJARAN PAI DI SMA NEGERI 10 PADANG TAHUN PELAJARAN 2023/2024 (TELAAH TERHADAP RELIABILITAS, DAYA BEDA DAN TINGKAT KESUKARAN MENGGUNAKAN SOFTWARE ANATES)." *Jurnal Review Pendidikan Dan Pengajaran (JRPP)* 7 (3): 6954–61.
- Kustati, Martin, and Rezki Amelia. 2023. "Analisis Kualitas Butir-Butir Soal Ulangan Harian Pendidikan Agama Islam Dengan Menggunakan Iteman." *SOKO GURU: Jurnal Ilmu Pendidikan* 3 (3): 142–55.
- Magdalena, Ina. 2021. *Desain Evaluasi Pembelajaran SD*. CV Jejak (Jejak Publisher). [https://books.google.com/books?hl=id&lr=&id=OvQ4EAAAQBAJ&oi=fnd&pg=PP1&dq=Validasi+dilakukan+menggunakan+alat+evaluasi+berupa+lengkap+tel+aah+soal+yang+mempertanyakan+aspek+materi,+konstruksi,+dan+bahasa+\(Nur+giyantor,+2013\)&ots=Vgq8sv9hl4&sig=fDpli6SVEVAn_4pLd8RLSwJKiLA](https://books.google.com/books?hl=id&lr=&id=OvQ4EAAAQBAJ&oi=fnd&pg=PP1&dq=Validasi+dilakukan+menggunakan+alat+evaluasi+berupa+lengkap+tel+aah+soal+yang+mempertanyakan+aspek+materi,+konstruksi,+dan+bahasa+(Nur+giyantor,+2013)&ots=Vgq8sv9hl4&sig=fDpli6SVEVAn_4pLd8RLSwJKiLA).
- Marambaawang, Damianus, Vidriana Oktoviana Bano, and Riwa Rambuhada Enda. 2023. "ANALISIS KUALITAS BUTIR SOAL PENILAIAN AKHIR SEMESTER GASAL TAHUN 2021/2022 MENGGUNAKAN ITEMAN DI SMP NEGERI 1 KAMBERA." *Dharmas Education Journal (DE_Journal)* 4 (1): 233–43.
- Nuriyah, Nunung. 2016. "Evaluasi Pembelajaran: Sebuah Kajian Teori." *Eduksos Jurnal Pendidikan Sosial & Ekonomi* 3 (1). <https://www.jurnal.syekh-nurjati.ac.id/index.php/edueksos/article/view/327>.
- Pradani, Rizka Abri, and Anwar Efendi. 2023. "Analisis Butir Soal Ujian Sekolah Menggunakan Program Iteman (Analysis of School Exam Questions Using the Iteman Program)." *Indonesian Language Education and Literature* 8 (2): 275–89.
- Setiadi, Hari. 2016. "Pelaksanaan Penilaian Pada Kurikulum 2013." *Jurnal Penelitian Dan Evaluasi Pendidikan* 20 (2): 166–78.
- Thorndike, Robert L., and Elizabeth Hagen. 1969. "Measurement and Evaluation in Psychology and Education." <https://eduq.info/xmlui/handle/11515/8429>.
- Wijayanti, Enny, and Mundilarto Mundilarto. 2015. "Pengembangan Instrumen Asesmen Diri Dan Teman Sejawat Kompetensi Bidang Studi Pada Mahasiswa." *Jurnal Penelitian Dan Evaluasi Pendidikan* 19 (2): 129–44.