

Analisis Butir Soal Penilaian Tengah Semester Genap Mata Pelajaran Pendidikan Agama Islam Kelas V di SDN 1 Bumiharjo

Mila Ningrum Masitoh, Imam Subarkah, Rose Kusumaning Ratri

Institut Agama Islam Nahdlatul Ulama Kebumen

E-mail: milaningrum420@gmail.com

Abstract

This study aims to determine the quality of the items about the Even Midterm Assessment of Islamic Religious Education class V subjects at SDN 1 Bumiharjo. This study used evaluative analysis methods. Data collection techniques use documentation. The data analysis techniques used are calculating empirical validity, reliability, item difficulty, item differentiating power, and distractor functions in multiple choice question types through counter software, namely using ANATES software version 4.0 accompanied by data presentation and drawing conclusions. The results showed that in terms of validity, overall the question already has a good rational validity, this is based on the results of a high content and construct validity analysis. Meanwhile, empirical validity shows that as many as 2 items of multiple-choice questions or 12.50% are declared valid, fill-in-the-blank questions as many as 9 questions or 75% are declared valid, and 5-point description questions or 83.33% are declared valid from all questions, both multiple choice, fill in and description. From the reliability of the questions, the form of the fill-in-the-blank questions includes a high reliability coefficient of 0.78 and for multiple-choice questions and descriptions are still relatively poor, the reliability coefficients are still low, namely 0.53 and 0.63. In terms of the difficulty of the questions, the multiple-choice type questions are still relatively not good, because the questions are still relatively too easy, namely as many as 10 (55.56%) multiple-choice questions. Then for the fill form questions, it is quite good, because the questions stated are sufficient as many as 10 (83.33%) of the total fill questions. The question of the type of description is quite good, because 3 (50%) of the questions from the fill-in-the-blank questions are classified as having a sufficient level of difficulty. from the distinguishing power of the items, in the poor category multiple choice type questions there are 3 (18.75%) question items, satisfactory there are 9 (50%) question items, good there are 4 (22.22%) question items, and excellent there are 2 (9.03%) question items from the entire multiple choice type questions. Then for the fill type question, the poor category has only 1 (8.34%) questions, satisfactory there are 4 (33.33%) question items, and good there are 7 (58.33%) question items from all items about the fill type. Then for the question of the type of long description, the poor category has 2 (33.33%) question items, satisfactory there are 2 (33.33%) question items, and good there are 2 (33.33%) question items from the entire description type question. From the distractor function, the multiple choice type question has a fairly good distrktor function, because based on the analysis, it can be concluded that there are 11 questions or 61.11% of the distractor function is able to perform its function properly.

Keywords: *Question Item Analysis, Validity, Reliability, Difficulty Level, Differentiating Power, Distractor Function.*



Abstrak

Penelitian ini bertujuan untuk mengetahui kualitas butir soal Penilaian Tengah Semester Genap mata pelajaran Pendidikan Agama Islam kelas V di SDN 1 Bumiharjo. Penelitian ini menggunakan metode analisis evaluatif. Teknik pengumpulan data menggunakan dokumentasi. Teknik analisis data yang digunakan adalah penghitungan validitas empirik, reliabilitas, tingkat kesukaran item, daya pembeda item, dan fungsi distraktor pada jenis soal pilihan ganda melalui *software* penghitung yaitu menggunakan *software* ANATES versi 4.0 disertai penyajian data dan penarikan kesimpulan. Hasil penelitian menunjukkan bahwa dari segi validitas, secara keseluruhan soal sudah memiliki validitas rasional yang baik hal tersebut berdasarkan pada hasil analisis validitas isi dan konstruk yang tinggi. Sedangkan secara validitas empirik menunjukkan bahwa sebanyak 2 butir soal pilihan ganda atau sebesar 12,50% dinyatakan valid, soal isian sebanyak 9 butir soal atau 75% dinyatakan valid, dan soal uraian 5 butir soal atau sebesar 83,33% dinyatakan valid dari keseluruhan soal baik pilihan ganda, isian maupun uraian. Dari reliabilitas soal, bentuk soal isian termasuk koefisien reliabilitas yang tinggi yakni sebesar 0,78 dan untuk soal pilihan ganda dan uraian masih tergolong kurang baik memiliki koefisien reliabilitas masih rendah yakni 0,53 dan 0,63. Dari segi tingkat kesukaran soal, soal jenis pilihan ganda masih tergolong belum baik, karena soal masih tergolong terlalu mudah yaitu sebanyak 10 (55,56%) keseluruhan soal pilihan ganda. Kemudian untuk soal bentuk isian sudah termasuk cukup baik, karena soal yang dinyatakan cukup sebanyak 10 (83,33%) dari keseluruhan soal isian. Soal jenis uraian tergolong sudah cukup baik, karena 3 (50%) soal dari soal isian tergolong memiliki tingkat kesukaran yang cukup. dari daya pembeda item, pada soal jenis pilihan ganda kategori *poor* terdapat 3 (18,75%) butir soal, *satisfactory* terdapat 9 (50%) butir soal, *good* terdapat 4 (22,22%) butir soal, dan *excellent* terdapat 2 (9,03%) butir soal dari keseluruhan soal jenis pilihan ganda. Kemudian untuk soal jenis isian, kategori *poor* terdapat hanya 1 (8,34%) butir soal, *satisfactory* terdapat 4 (33,33%) butir soal, dan *good* terdapat 7 (58,33%) butir soal dari keseluruhan butir soal jenis isian. Kemudian untuk soal jenis uraian panjang, kategori *poor* terdapat 2 (33,33%) butir soal, *satisfactory* terdapat 2 (33,33%) butir soal, dan *good* terdapat 2 (33,33%) butir soal dari keseluruhan soal jenis uraian. dari fungsi distraktornya, soal jenis pilihan ganda mempunyai fungsi distrktor yang cukup baik, karena berdasarkan analisis dapat disimpulkan bahwa terdapat 11 soal atau 61,11% fungsi distraktornya sudah mampu menjalankan fungsinya dengan baik.

Kata kunci: *Analisis Butir Soal, Validitas, Reliabilitas, Tingkat Kesukaran, Daya Pembeda, Fungsi Distraktor.*



PENDAHULUAN

(Purwanto, 2014:18) menyatakan bahwa pendidikan merupakan sebuah proses kegiatan yang disengaja atas input siswa untuk menimbulkan suatu hasil yang diinginkan sesuai dengan tujuan yang ditetapkan. Oleh karena itu, untuk menciptakan pendidikan yang bermutu perlu adanya kegiatan evaluasi. Evaluasi adalah sebuah proses pengumpulan data untuk menentukan sejauh mana, dala hal apa, dan bagaimana tujuan pendidikan sudah tercapai serta dapat digunakan untuk membuat keputusan (Arikunto, 2004:3). Evaluasi pendidikan mengkhususkan pada evaluasi hasil belajar atau penilaian hasil belajar. Penilaian hasil belajar adalah suatu kegiatan untuk memberikan berbagai informasi secara berkesinambungan dan menyeluruh tentang proses dan hasil yang telah dicapai siswa (Pramono, 2014:11). Pada dasarnya penilaian hasil belajar digunakan guru untuk mengetahui hasil pembelajaran yang dilakukan. Maka dari itu, penilaian hasil belajar penting dalam menentukan kualitas pendidikan.

Berdasarkan Peraturan Pemerintah Nomor 57 Tahun 2021 tentang Standar Pendidikan Nasional, khususnya pasal Pasal 42 ayat 2 menyatakan bahwa evaluasi hasil belajar peserta didik sebagaimana dimaksud pada ayat (1) dilakukan untuk memantau proses, kemajuan, dan perbaikan hasil belajar peserta didik secara berkesinambungan dan menilai pencapaian hasil peserta didik sesuai dengan standar kompetensi lulusan.¹

Dari peraturan tersebut sudah jelas bahwa seorang pendidik mempunyai kewajiban untuk mengadakan penilaian hasil belajar peserta didik, supaya dapat mengetahui sajumlah mana kemajuan hasil peserta didik dalam kurun waktu tertentu, sehingga pendidik dapat mengambil langkah yang tepat untuk menangani permasalahan yang dialami peserta didik dalam pembelajaran. Evaluasi hasil belajar dapat dilakukan melalui tes. Tes berupa soal berfungsi sebagai alat ukur penialian hasil belajar. Soal yang bermutu adalah soal yang dapat meberikan informasi tepat tentang peserta didik yang menguasai materi dan yang belum. Maka dari itu, pendidik diharuskan mengadakan analisis terhadap butir soal untuk mengetahui tingkat kualitas soal yang digunakan sebagai penilaian perkembangan hasil belajar peserta didik sebelum melakukan penilaian.

(Sudijono, 2011:93-97) meyatakan bahwa kriteria instrumen penilaian hasil belajar yang baik adalah instrumen penilaian yang bersifat valid atau memiliki validitas,

¹ Peraturan Pemerintah Nomor 57 Tahun 2021 tentang Standar Pendidikan Nasional



instrumen memiliki reliabilitas yang tinggi, instrumen penilaian memiliki daya pembeda dan tingkat kesukuran yang baik, serta pada soal pilihan ganda memiliki distraktor yang berfungsi dengan baik.

Penilaian dapat diambil dari tes sumatif berupa Penilaian Tengah Semester (PTS) dan Penilaian Akhir Semester (PAS). Penilaian Tengah Semester adalah kegiatan yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik setelah melaksanakan 8 – 9 minggu kegiatan pembelajaran. Bentuk tes yang digunakan biasanya berupa tes objektif, esai, dan jawaban pendek yang memiliki variasi masing-masing.

Mata pelajaran Pendidikan Agama Islam merupakan mata pelajaran wajib di SDN 1 Bumiharjo. Dalam hasil wawancara dengan Ibu Mahmudatul Baroroh, S. Pd. I, mengatakan bahwa instrumen penilaian tengah semester genap dikembangkan oleh tim Kelompok Kerja Guru pada tingkat kecamatan dan selama melakukan Penilaian Tengah Semester (PTS) belum pernah melakukan analisis butir soal terhadap instrumen penilaian tersebut, sehingga kualitas soal yang diujikan belum diketahui apakah soal tersebut sudah termasuk kategori soal yang baik atau belum untuk dijadikan instrumen penilaian.² Analisis butir soal Penilaian Tengah Semester Genap mata pelajaran Pendidikan Agama Islam bertujuan untuk mengetahui kualitas butir soal mata Penilaian Tengah Semester Genap mata pelajaran Pendidikan Agama Islam (PAI) kelas V SDN 1 Bumiharjo. Dilakukan analisis terhadap butir soal Penilaian Tengah Semester Genap Mata Pelajaran Pendidikan Agama Islam Kelas V diharapkan mampu mengidentifikasi butir-butir soal yang baik, kurang baik, dan buruk, sehingga pendidik dapat melakukan perbaikan terhadap soal yang kurang baik dan buruk untuk penilaian selanjutnya dan kegiatan tersebut dapat dijadikan bahan evaluasi pendidik dalam pengembangan instrumen penilaian hasil belajar peserta didik.

METODE PENELITIAN

Jenis metode dalam penelitian ini adalah evaluatif. Subjek penelitian SDN 1 Bumiharjo meliputi Guru Pendidikan Agama Islam (PAI), dan peserta didik kelas V . Sedangkan, objek penelitian ini adalah 36 butir soal Penilaian Tengah Semester Genap

² Mahmudatul Baroroh, Guru Pengampu Mata Pelajaran Pendidikan Agama Islam SDN 1 Bumiharjo, "Wawancara tentang analisis butir soal", 8 Februari 2022.

Mata Pelajaran Pendidikan Agama Islam di SDN 1 Bumiharjo terdiri dari 18 soal pilihan ganda, 12 soal isian, dan 6 soal uraian. Teknik pengumpulan data berupa dokumentasi. Analisis data dalam penelitian ini dilakukan dengan cara penghitungan melalui *software* penghitung yakni menggunakan *software* Anates versi 4.0.

HASIL PENELITIAN DAN PEMBAHASAN

Dalam soal penilaian tengah semester genap mata pelajaran Pendidikan Agama Islam merujuk pada 3 Kompetensi Dasar berdasarkan materi pokok yang diajarkan selama separuh semester yaitu:

- a. KD 3.1 Memahami makna Q.S *at-Tiin* dan Q.S *al- Ma'un*, materi pokok yang diajarkan adalah belajar Q.S. *al- Ma'un*.
- b. KD 3.3 Memahami nama-nama Rasul Swt dan Rasul Ulul Azmi, materi pokok yang diajarkan adalah Rasul Ulul Azmi.
- c. KD 3.8 Memahami makna hidup sederhana dan ikhlas, materi yang diajarkan yaitu mari hidup sederhana dan ikhlas.

Soal yang dibuat berdasarkan masing-masing Kompetensi Dasar (KD) terdiri dari 6 soal pilihan ganda, 4 isian (uraian pendek), dan 2 uraian panjang sehingga total butir soal menjadi 36 soal. Bobot penilaian pada masing-masing jenis soal yakni untuk soal jenis pilihan ganda per item soal diberi skor 5 point, untuk soal jenis isian (uraian pendek) per item soal diberi skor 10 point, dan soal jenis uraian panjang per item diberi skor 15 point. Dalam penelitian ini menyajikan dan menjabarkan hasil perhitungan validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan fungsi distraktor pada soal pilihan ganda sebagai berikut:

1. Validitas

- a. Validitas Rasional

Validitas rasioanal dapat ditinjau dari dua segi yaitu dari segi validitas isi dan validitas konstruknya. (Suprananta, 2005) dalam (Budi Utomo, 2018) validitas isi merupakan suatu alat ukur yang dikatakan valid apabila sudah sesuai dengan isi kurikulum yang akan diukur. (Wulan & Rusdiana, 2013) dalam (Nizzary & Kholiq, 2021) menjelaskan bahwa validitas isi dapat tercapai apabila butir-butir soal disusun berdasarkan kisi-kisi yang telah dibuat.

Menurut (Sudijono, 2011:166), validitas konstruk merupakan kevalidan soal yang dilihat dari segi susunan, kerangka serta rekaannya. (Nizzary & Kholik, 2021) menyatakan bahwa validitas konstruk dapat diketahui dengan cara memerinci dan memasang setiap tujuan instruksional khusus dengan aspek kognitifnya.

Mengacu pada analisis validitas isi yang dikemukakan oleh (Wulan & Rusdiana, 2013) dalam (Nizzary & Kholik, 2021) yaitu melalui pencocokan antara butir soal dengan kisi-kisi yang telah dibuat. Dalam hal ini penulis membandingkan antara soal dengan unsur yang terdapat dalam kisi-kisi salah satunya adalah Kompetensi Dasar.

Tabel 1 Analisis Validitas Isi Soal

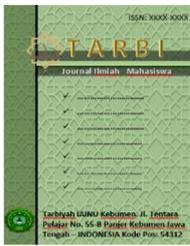
Komptensi Dasar	Meteri Pokok	Nomor Soal
3.1 Memahami makna <i>Q.S. at-Tiin</i> dan <i>Q.S. al-Ma'un</i> dengan benar	<i>Q.S. al-Ma'un</i>	1,2,3,4,5,6,7, 8,9,10,11,12
3.3 Memahami nama-nama Rasul Allah Swt. dan Rasul Ulul Azmi	Rasul Ulul Azmi	13,14,15,16,17, 18,19,20,21,22,24
3.8 Memahami makna sederhana dalam kehidupan sehari-hari	Mari Hidup Sederhana dan Ikhlas	25,26,27,28,29,30 31,33,34,35,36

Dari tabel di atas, dapat dijelaskan bahwa secara umum soal PTS genap mata pelajaran Pendidikan Agama Islam kelas V SDN 1 Bumiharjo sudah sesuai dengan Kompetensi Dasar (KD) yang terdapat dalam kisi-kisi. Butir-butir soal yang ada sudah sesuai dengan KD 3.1, 3.3, dan KD 3.8. Sesuai yang dimaksud adalah soal yang diujikan menunjukkan pengukuran terhadap kompetensi siswa pada KD tersebut dengan perincian di dalam tabel di atas, seperti contoh soal dibawah ini:

6. وَلَا يَحْضُ عَلَىٰ طَعَامِ الْمَسْكِينِ

Terjemah dari ayat di atas

- a. maka celakalah orang yang salat
- b. orang yang menghardik anak yatim
- c. tahukah kamu orang yang mendustakan agama



d. dan tidak menganjurkan memberi makan orang miskin

Soal tersebut mengukur pada kompetensi dasar 3.1 pada materi pokok Q.S. Al-Maun khususnya pada makna Q.S Al-Maun. Kemudian, soal yang mengukur kompetensi terkait KD 3.3 pada materi pokok Rasul Ulul Azmi dapat dilihat pada soal pilihan ganda no. 16 sebagai berikut:

16. Perhatikan nama-nama Rasul berikut!

1. Nabi Nuh AS	4. Nabi Isa AS
2. Nabi Ibrahim AS	5. Nabi Muhammad SAW
3. Nabi Musa AS	6. Nabi Yusuf AS

Yang termasuk Rasul yang bergelar Ulul Azmi ditunjukkan nomor

- a. 1, 2, 3, 4 dan 5
- b. 2, 3, 4, 5 dan 6
- c. 1, 2, 3, 4 dan 6
- d. 1, 2, 3, 6 dan 5

Sedangkan soal yang terkait dengan KD 3.8 dapat dilihat dari contoh soal pada soal uraian no. 35 dan 36, sebagai berikut:

- 35. Sebutkan 2 manfaat hidup sederhana!
- 36. Sebutkan 2 keuntungan hidup sederhana!

Berdasarkan hasil dan pembahasan di atas dapat disimpulkan bahwa validitas isi soal PTS genap mata pelajaran Pendidikan Agama Islam kelas V SDN 1 Bumiharjo sudah sesuai. Hal ini sesuai dengan apa yang disampaikan oleh (Suprananta, 2005) dalam (Utomo, 2018) soal dikatakan memiliki validitas isi apabila sudah sesuai dengan isi kurikulum yang termuat dalam kompetensi dasar. Soal PTS genap mata pelajaran Pendidikan Agama Islam kelas V SDN 1 Bumiharjo mempunyai validitas isi yang tinggi hal tersebut sesuai dengan yang dikemukakan oleh (Utomo, 2018) soal sesuai kurikulum atau alat ukur yang digunakan merujuk pada kurikulum dinyatakan memiliki validitas yang tinggi.

Sedangkan untuk analisis validitas konstruk mengacu pada analisis yang dilakukan (Nizzary & Kholik, 2021) dengan memerinci dan memasangkan antara tujuan instruksional khusus dengan aspek kognitifnya, analisisnya sebagai berikut:

Tabel 2. Analisis Validitas Konstruk Soal

Materi Pokok	Aspek Kognitif			Validitas Konstruk
	C1	C2	C3	
<i>Q.S. al-Ma'un</i>	1,2,7,8,9	3,4,6,10,12	5,11	Soal-soal memiliki validitas konstruk ditinjau dari aspek kognitif pada masing-masing soal yang bersesuaian dengan kompetensi dasar yang hendak dicapai.
Rasul Ulul Azmi	-	13,14,15 16,17,18 19,20,21 22,24	11,23	Soal-soal memiliki validitas konstruk ditinjau dari aspek kognitif pada masing-masing soal, pada KD ini didominasi kompetensi kognitif tingkat dua (C2)
Mari Hidup Sederhana dan Ikhlas	25,26,27 31,32,35	28,29,30 34	36	Soal-soal memiliki validitas konstruk ditinjau dari aspek kognitif pada masing-masing soal yang bersesuaian dengan kompetensi dasar yang hendak dicapai

Berdasarkan tabel di atas dapat disimpulkan bahwa soal-soal PTS genap mata pelajaran Pendidikan Agama Islam kelas V SDN 1 Bumiharjo memiliki validitas konstruk yang baik, merujuk pada apa yang dikemukakan oleh (Nizzary & Kholik, 2021) setiap butir soal dinyatakan memiliki validitas konstruk apabila soal-soal dibuat berdasarkan pada kompetensi kognitif yang hendak dicapai siswa berdasar pada kompetensi dasar pada masing-masing materi pokok yang diajarkan.

b. Validitas Empirik

Proses menentukan validitas butir soal baik pilihan ganda, isian, maupun uraian dapat menggunakan teknik korelasi sebagai teknik analisisnya (Rusmayani, 2019). Berdasarkan yang dikemukakan oleh (Rusmayani, 2019) item tes hasil belajar dapat dikatakan valid, jika skor-skor pada butir soal yang bersangkutan memiliki kesesuaian arah dengan skor totalnya, atau ada korelasi positif yang signifikan antara skor butir item dengan skor totalnya (*Point biserial*). Pemberian interpretasi terhadap r_{pbi} ini menggunakan db sebesar (N-

nr), yaitu: $16 - 2 = 14$. Derajat kebebasan (db) sebesar 14 lalu dikonsultasikan kepada tabel nilai “r” *product moment*, pada taraf signifikansi 5% dan taraf signifikansi 1%. Hasilnya r_{tabel} atau r_t pada taraf signifikansi 5% yaitu sebesar 0,497 dan r_{tabel} atau r_t pada taraf signifikansi 1% yaitu sebesar 0,623. Dari hasil perhitungan di dapatkan butir soal yang tidak valid sebanyak 20 butir soal, sedangkan yang valid 16 butir soal. Berikut data soal yang valid dan tidak valid.

Tabel 3. Hasil Analisis Validitas Empirik Soal

No	Validitas	Jumlah Butir Soal			Nomor Soal		
		PG	Isian	Uraian	PG	Isian	Uraian
1.	Valid	2	9	5	5, 15	7,8,20,21 22,31,32 33,34	11,12,23 35,36
2.	Tidak Valid	16	3	1	1,2,3,4 6,13,14 16,17,18 25,26,27,28	9,10,19	24

Dari tabel di atas didapatkan informasi bahwa sebanyak 2 butir soal pilihan ganda atau sebesar 12,50% dinyatakan valid, soal isian sebanyak 9 butir soal atau 75% dinyatakan valid, dan soal uraian 5 butir soal atau sebesar 83,33% dinyatakan valid dari keseluruhan soal baik pilihan ganda, isian maupun uraian. Sedangkan untuk soal yang dinyatakan tidak valid jenis soal pilihan ganda sebanyak 16 soal atau sebesar 87,50%, isian sebanyak 3 soal atau sebesar 25% , dan uraian hanya 1 soal sebesar 16,67% dari keseluruhan soal baik jenis pilihan ganda, isian maupun uraian.

2. Reliabilitas

Reliabilitas merupakan koefisien yang menunjukkan sejauh mana alat penilaian dapat dipercaya, artinya apabila suatu alat penilaian digunakan berulang-ulang menilai sesuatu yang sama, maka hasil relatif konsisten (Farida & Musyarofah, 2021). Menurut (Sudijono, 2011:209) soal dikatakan mempunyai reliabilitas yang tinggi apabila koefisien reliabilitas tes yang sedang diuji reliabilitasnya lebih dari 0,70. Apabila lebih kecil dari 0,70 berarti tes yang sedang diuji reliabilitasnya dinyatakan belum memiliki reliabilitas yang tinggi (*unreliable*). Adapun perhitungan reliabilitas tes PTS Genap dari hasil analisis melalui Anates dapat dilihat pada tabel berikut:

Tabel 4. Hasil Analisis Reliabilitas Soal

Jenis Soal	Koefisien Reliabilitas	Kategori
Pilihan Ganda	0,53	Reliabilitas Rendah
Isian	0,78	Reliabilitas Tinggi
Uraian	0,63	Reliabilitas Rendah

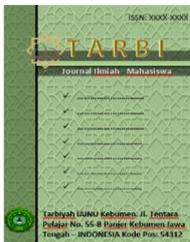
Dari hasil perhitungan di atas dapat didapatkan kesimpulan bahwa soal jenis pilihan ganda dan uraian dari segi reliabilitas tes masih dianggap rendah atau *unreliable* sedangkan untuk jenis soal isian sudah dianggap memiliki reliabilitas yang tinggi atau sudah *reliable*. Menurut (Nurhasah & Ahmad, 2017) dalam (Rusmayani, 2019) menjelaskan bahwa reliabilitas soal akan semakin tinggi jika soal-soal didominasi oleh butir soal dengan kategori sedang, sehingga dapat disimpulkan soal isian memiliki reliabilitas yang tinggi karena jenis soal isian didominasi oleh soal-soal kategori sedang.

3. Tingkat Kesukaran Soal

Menurut (Rahmaini & Taufiq, 2018) soal yang baik adalah soal yang tingkat kesukaran tidak terlalu mudah dan tidak terlalu sukar. Untuk penggolongan tingkat kesulitan berdasarkan pada nilai derajat kesukaran dari setiap item butir soal yaitu besarnya indeks kesukaran item kurang dari 0,30 dianggap terlalu sukar, diantara 0,30 sampai 0,70 dianggap cukup, dan lebih dari 0,70 dianggap terlalu mudah (Sudijono, 2011:372). Adapun hasil analisis taraf kesukaran butir soal dari program Anates sebagai berikut:

Tabel 5. Hasil Analisis Tingkat Kesukaran Soal

No	Jenis Soal	Kategori	Jumlah	Nomor Soal
1.	Pilihan Ganda	Terlalu Mudah	10	1,2,4,5,16,17 25,26,27,28
		Cukup	7	3,6,13,14,15 18,28
		Terlalu Sukar	1	29
2.	Isian	Terlalu Mudah	2	10,33
		Cukup	10	7,8,9,19,20,21 22,31,32,34
3.	Uraian	Terlalu Mudah	3	12,35,36
		Cukup	3	11,23,24



Dari hasil analisis di atas didapatkan informasi berdasarkan apa yang dikemukakan oleh (Rahmaini & Taufiq, 2018) yakni untuk jenis soal pilihan ganda sebanyak 10 butir soal atau sebesar 55,56% termasuk kategori terlalu mudah, 7 soal atau sebesar 38,89% termasuk kategori sedang dan 1 soal atau sebesar 5,55% dari keseluruhan soal pilihan ganda. Sedangkan, untuk soal isian sebanyak 2 soal atau sebesar 16,67% termasuk kategori terlalu mudah dan sebanyak 10 soal atau sebesar 83,33% termasuk kategori cukup dari keseluruhan total jenis soal isian. Untuk soal uraian sebanyak 3 soal atau sebesar 50% termasuk kategori terlalu mudah dan 3 soal atau sebesar 50% termasuk kategori cukup.

4. Daya Pembeda Item

Menurut (Rusmayani, 2019) analisis daya pembeda bertujuan untuk mengetahui kesanggupan soal dalam membedakan peserta didik yang tergolong mampu dengan siswa yang tergolong lemah prestasinya. Menurut (Dewi dkk, 2018) analisis daya pembeda item sangat penting dilakukan untuk mengetahui tingkat keberhasilan soal dalam mengukur kemampuan siswa yang sebenarnya. Soal yang memiliki daya pembeda yang baik, maka semakin baik pula soal tersebut dalam mengukur kemampuan siswa, begitupun sebaliknya soal yang tidak memiliki daya pembeda yang baik menyebabkan soal tidak dapat mengukur kemampuan siswa yang sebenarnya.

Merujuk yang dikemukakan oleh (Sudijono, 2011:389) untuk penggolongan daya pembeda berdasarkan kriteria angka indeks diskriminasi item jika kurang dari 0,20 termasuk kategori *poor* yang mana butir soal bersangkutan daya pembedanya lemah dianggap tidak memiliki daya pembeda yang baik. Jika angka indeks pembeda itemnya sebesar 0,20 – 0,40 termasuk kategori *satisfactory* yang mana butir soal tersebut telah memiliki daya pembeda yang cukup (sedang). Jika angka indeks diskriminasi itemnya 0,40 – 0,70 butir soal termasuk kategori *good* yang mana butir item bersangkutan telah memiliki daya pembeda yang baik. Besar angka indeks diskriminasi item antara 0,70 – 1,00 butir item soal termasuk kategori *excellent* dimana butir soal bersangkutan telah memiliki daya pembeda yang baik sekali, sedangkan jika besar angka diskriminasi itemnya negatif (-) maka butir item

bersangkutan daya pembedanya jelek sekali. Berikut analisis daya pembeda soal PTS genap mata pelajaran Pendidikan Agama Islam kelas V SDN 1 Bumiharjo:

Tabel 6. Hasil Analisis Daya Pembeda Soal Pilihan Ganda

No	Angka Indeks Diskriminasi Item	Interpretasi
1	0,25	<i>Satisfactory</i>
2	0,25	<i>Satisfactory</i>
3	0,25	<i>Satisfactory</i>
4	0,00	<i>Poor</i>
5	0,75	<i>Excellent</i>
6	0,50	<i>Good</i>
13	0,50	<i>Good</i>
14	0,75	<i>Excellent</i>
15	0,50	<i>Good</i>
16	0,25	<i>Satisfactory</i>
17	0,00	<i>Poor</i>
18	0,25	<i>Satisfactory</i>
25	0,25	<i>Satisfactory</i>
26	0,25	<i>Satisfactory</i>
27	0,00	<i>Poor</i>
28	0,25	<i>Satisfactory</i>
29	0,25	<i>Satisfactory</i>
30	0,50	<i>Good</i>

Tabel 7. Hasil Analisis Daya Pembeda Soal Isian

No	Angka Indeks Diskriminasi Item	Interpretasi
7	0,53	<i>Good</i>
8	0,53	<i>Good</i>
9	0,18	<i>Poor</i>
10	0,35	<i>Satisfactory</i>
19	0,53	<i>Good</i>
20	0,53	<i>Good</i>
21	0,53	<i>Good</i>
22	0,36	<i>Satisfactory</i>
31	0,36	<i>Satisfactory</i>
32	0,53	<i>Good</i>
33	0,35	<i>Satisfactory</i>
34	0,53	<i>Good</i>

Tabel 8. Hasil Analisis Daya Pembeda Soal Uraian

No	Angka Indeks Diskriminasi Item	Interpretasi
11	0,33	<i>Satisfactory</i>
12	0,33	<i>Satisfactory</i>
23	0,75	<i>Good</i>
24	0,58	<i>Good</i>
35	0,16	<i>Poor</i>
36	0,16	<i>Poor</i>

Berdasarkan tabel di atas, dari analisis butir soal tes PTS Genap dengan program Anates, untuk interpretasi 18 soal pilihan ganda ada 3 butir soal atau 18,75% dari keseluruhan soal pilihan ganda dengan daya pembeda yang lemah, soal-soal tersebut tidak dapat diperbaiki dan harus diganti. Terdapat 9 butir soal atau 50% dari keseluruhan soal pilihan ganda yang termasuk butir soal dengan daya pembeda kategori cukup, tetapi belum begitu baik sehingga butir soal masih harus diperbaiki atau direvisi. Lalu, dalam kategori *good* ada 4 soal atau 22,22% dari soal pilihan ganda secara keseluruhan, butir soal bersangkutan telah memiliki daya pembeda yang baik dan sisa butir-butir soal pilihan ganda yang dinyatakan mempunyai kategori *excellent*, berjumlah 2 soal atau 9,03% dari keseluruhan soal pilihan ganda, kedua butir soal ini termasuk soal yang baik sekali sehingga benar-benar mampu membedakan antara siswa yang mempunyai kemampuan tinggi dengan siswa yang mempunyai kemampuan rendah. Secara keseluruhan soal bentuk pilihan ganda berdasarkan indeks daya pembeda yang diperoleh pada masing-masing item butir soal termasuk belum cukup baik, karena ada 3 soal atau 18,75% dari keseluruhan soal yang perlu dibuang, dan masih ada butir soal yang perlu diperbaiki yaitu sebanyak 9 soal atau 50% dari keseluruhan soal pilihan ganda.

Kemudian, analisis daya pembeda pada soal PTS Genap jenis isian dengan menggunakan bantuan program Anates, interpretasi 12 soal isian terdapat 1 soal atau 8,34% dari soal isian secara keseluruhan yaitu soal nomor 9 dengan daya pembeda yang lemah, maka soal tersebut harus dibuang. Soal berkategori *satisfactory* terdapat 4 soal atau 33,33% dari keseluruhan soal isian. Keempat soal tersebut termasuk soal kategori cukup baik, akan tetapi soal masih harus diperbaiki.

Lalu, terdapat 7 soal termasuk kategori *good* atau 58,33% dari keseluruhan soal isian. Ketujuh soal tersebut dianggap sudah baik, dan sudah mampu membedakan antara kelompok siswa yang berkeampuan tinggi dan siswa berkemampuan rendah. Secara keseluruhan soal PTS Genap jenis uraian pendek atau isian berdasarkan indeks diskriminasi item atau daya pembeda butir soal yang dihasilkan sudah baik, namun masih ada yang harus diperbaiki.

Untuk butir soal uraian berdasarkan analisis butir tes dengan program Anates, yang berkategori *poor* ada 2 soal atau 33,33% dari keseluruhan soal uraian. Butir soal uraian yang berkategori *satisfactory* ada 2 soal atau 33,33% dari keseluruhan soal uraian. Soal-soal tersebut termasuk kategori cukup baik tetapi belum memuaskan dan soal tersebut harus direvisi dan tidak perlu dibuang. Kemudian, butir soal uraian yang berkategori *good* ada 2 soal atau 33,33% dari keseluruhan soal uraian, maka soal tersebut sudah baik, sudah mampu memmbedakan antara siswa berkemampuan tinggi dengan siswa berkemampuan rendah. Dapat disimpulkan bahwa soal jenis uraian dari segi daya pembedanya masih belum baik, karena terdapat dua soal yang harus diganti dan terdapat dua soal yang harus direvisi atau ditingkatkan kualitasnya serta hanya terdapat dua soal dengan kategori baik.

5. Fungsi Distraktor (Pengecoh)

(Thoha, 2001) dalam (Yani dkk, 2013) tujuan pemakaian distraktor adalah mengecahkan mereka yang kurang mampu untuk dapat dibedakan dengan yang mampu. (Sudijono, 2011:411) menjelaskan bahwa distraktor dinyatakan telah menjalankan fungsinya dengan baik apabila distraktor tersebut sekurang-kurangnya sudah dipilih oleh 5% dari seluruh peserta tes yang ada. Dalam output analisis program Anates, berdasarkan hasil analisis dapat diketahui bahwa terdapat 4 opsi alternatif jawaban, 3 pilihan yang difungsikan sebagai distraktor dan 1 pilihan yang merupakan kunci jawaban soal. Berikut adalah hasil analisis terhadap fungsi distraktor pada soal bentuk pilihan ganda:

Tabel 9. Fungsi Distraktor Soal Pilihan Ganda

No Soal	Kunci Jawaban	Fungsi Distraktor	
		Baik	Tidak Baik
1	D	B,C	A
2	B	A	C,D

3	A	C,D	B
4	A	-	B,C,D
5	D	A	B,C
6	D	B,C	A
13	C	A,B	D
14	D	A,B,C	-
15	C	B,D	A
16	A	D	B,C
17	B	-	A,C,D
18	C	A,B,D	-
25	A	C	B,D
26	D	A,B	C
27	C	-	A,B,D
28	A	B,C,D	-
29	C	A,B,D	-
30	C	A,B,D	-

Berdasarkan analisis fungsi distraktor pada soal pilihan ganda dapat terdapat 2 klasifikasi diantaranya:

Tabel 10. Efektifitas Fungsi Distraktor Soal Pilihan Ganda

Fungsi Distraktor	Nomor Soal
Efektif	1, 3, 6, 13, 14, 15, 18, 26, 28 29, 30
Tidak Efektif	2, 4, 5, 16, 17, 25, 27

Data tersebut menunjukkan bahwa dari 18 soal pilihan ganda PTS genap mata pelajaran Pendidikan Agama Islam kelas V SDN 1 Bumiharjo terdapat 11 soal atau 61,11% fungsi distraktornya sudah mampu menjalankan fungsinya dengan baik. Merujuk pada apa yang dikemukakan (Sudijono, 2011:417) bahwa pengecoh yang sudah menjalankan fungsinya dengan baik dapat dipakai lagi pada penilaian mendatang, sedangkan pengecoh yang belum berfungsi dengan baik sebaiknya diganti dengan pengecoh lain.

KESIMPULAN

Berdasarkan hasil analisis dan penelitian tentang “Analisis Butir Soal Penilaian Tengah Semester Genap Mata Pelajaran Pendidikan Agama Islam Kelas V di SDN 1 Bumiharjo Tahun Ajaran 2021/2022”, maka penulis dapat menyimpulkan bahwa:

1. Ditinjau dari segi validitas, secara keseluruhan soal sudah memiliki validitas rasional yang baik hal tersebut berdasarkan pada hasil analisis validitas isi dan



konstruk yang tinggi. Sedangkan secara validitas empirik menunjukkan bahwa sebanyak 2 butir soal pilihan ganda atau sebesar 12,50% dinyatakan valid, soal isian sebanyak 9 butir soal atau 75% dinyatakan valid, dan soal uraian 5 butir soal atau sebesar 83,33% dinyatakan valid dari keseluruhan soal baik pilihan ganda, isian maupun uraian. Sedangkan untuk soal yang dinyatakan tidak valid jenis soal pilihan ganda sebanyak 16 soal atau sebesar 87,50%, isian sebanyak 3 soal atau sebesar 25% , dan uraian hanya 1 soal sebesar 16,67% dari keseluruhan soal baik jenis pilihan ganda, isian maupun uraian.

2. Bila ditinjau dari reliabilitas soal, bentuk soal isian termasuk soal yang baik karena mempunyai koefisien reliabilitas yang tinggi (reliabel) yakni sebesar 0,78 dan untuk soal pilihan ganda dan uraian masih tergolong kurang baik karena koefisien reliabilitasnya masih rendah yakni 0,53 dan 0,63.
3. Ditinjau dari segi tingkat kesukaran soal, soal jenis pilihan ganda masih tergolong belum baik, karena sebagian besar soal masih tergolong terlalu mudah yaitu sebanyak 10 (55,56%) soal dari keseluruhan soal pilihan ganda. dan 1 soal dinyatakan sukar sedangkan untuk kategori cukup hanya ada 7 (38,89%) soal dari keseluruhan soal pilihan ganda. Kemudian untuk soal bentuk isian sudah termasuk cukup baik, karena jumlah soal yang dinyatakan cukup sebesar 10 (83,33%) soal dari keseluruhan soal isian dan hanya ada 2 (16,67%) yang termasuk kategori terlalu mudah. Soal jenis uraian tergolong sudah cukup baik, karena 3 (50%) soal dari soal isian tergolong memiliki tingkat kesukaran yang cukup, sedangkan sisanya yaitu 3 (50%) soal termasuk kategori terlalu mudah.
4. Ditinjau dari daya pembeda item, pada soal jenis pilihan ganda kategori *poor* terdapat 3 (18,75%) butir soal, kategori *satisfactory* terdapat 9 (50%) butir soal, kategori *good* terdapat 4 (22,22%) butir soal, dan kategori *excellent* terdapat 2 (9,03%) butir soal dari keseluruhan soal jenis pilihan ganda. Kemudian untuk soal jenis isian atau uraian pendek, kategori *poor* terdapat hanya 1 (8,34%) butir soal, kategori *satisfactory* terdapat 4 (33,33%) butir soal, dan kategori *good* terdapat 7 (58,33%) butir soal dari keseluruhan butir soal jenis isian atau uraian pendek. Kemudian untuk soal jenis uraian panjang, kategori *poor* terdapat 2 (33,33%) butir

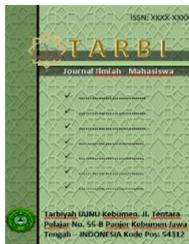


soal, kategori *satisfactory* terdapat 2 (33,33%) butir soal, dan kategori *good* terdapat 2 (33,33%) butir soal dari keseluruhan soal jenis uraian panjang.

5. Ditinjau dari fungsi distraktornya, soal jenis pilihan ganda mempunyai fungsi pengecoh yang cukup baik, karena berdasarkan analisis dapat disimpulkan bahwa terdapat 11 soal atau 61,11% fungsi distraktornya sudah mampu menjalankan fungsinya dengan baik.

DAFTAR PUSTAKA

- Arikunto, S. (2003). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Dewi, S. Sacita, Rachmaniah M. Haristuti, Arfiati Ulfa Utami. (2018). *Analisis Tingkat Kesukaran dan Daya Pembeda Soal Olimpiade Matematika (OMI) Tingkat SMP Tahun 2018*.
- Farida, Anna Musyarofah. (2021). *Validitas dan Reliabilitas dalam Analisis Butir Soal*. Jurnal Pendidikan Bahasa Arab. Vol. I. No. 1
- Nizary, M.A, A. N. Kholik. (2021). *Validitas Instrumen Assesmen (Analisis Validitas Isi Dan Konstruk Instrumen Assesmen Buku Pelajaran Al Quran Hadis Kelas 6 Madrasah Ibtidaiyah Materi Surat Ad Dhuha Bab VI)*. Jurnal Pendidikan Bahasa Arab. Vol. 2. No. 1.
- Peraturan Pemerintah Nomor 57 Tahun 2021 tentang Standar Pendidikan Nasional.
- Pramono, S. (2014). *Panduan Evaluasi Kegiatan Belajar-Mengajar*. Yogyakarta: Diva Press.
- Purwanto. (2014). *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka pelajar.
- Rahmaini, A, Aditya Nur Taufiq. (2018). *Analisis Butir Soal Pendidikan Agama Islam di SMK N 1 Sedayu Tahun Ajaran 2017/2018*. Jurnal MUDARRISUNA. Vol. 8. No.1
- Ratnawulan, Elis and H.A. Rusdiana. (2014). *Evaluasi Pembelajaran: Dengan Pendekatan Kurikulum 2013*, Bandung: Pustaka Setia.
- Rusmayani. (2020). *Analisis Butir Soal Penilaian Akhir Semester Genap Mata Pelajaran Pendidikan Agama Islam di SMP Bintang Persada Tabanan-Bali*.
- Sudijono, A. (2011). *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada.
- Toha, M. Chabib. (2001). *Teknik Evaluasi Pendidikan*. Jakarta: PT RajaGrafindo Persada.



- Utomo, B. (2018). *Analisis Validitas Isi Butir Soal sebagai Salah Satu Upaya Peningkatan Kualitas Pembelajaran di Madrasah Berbasis Nilai-Nilai Islam*. Jurnal Pendidikan Matematika. Vol 1. No 2.
- Yani, A, Ali Fikri Asri, Ahmad Burhan. (2013). *Analisis Tingkat Kesukaran, Daya Pembeda dan Fungsi Distraktor Soal Ujian Semester Ganjil Mata Pelajaran Produktif di SMK Negeri 1 Indralaya Utara Tahun Pelajaran 2012/2013*. Jurnal Pendidikan.