



Analisis Kualitas Soal Ujian Akhir Semester dengan Menggunakan Aplikasi Program Anates

Ni Ketut Wedastuti*, Sunismi, Aniesfuady

Pascasarjana Magister Pendidikan Matematika, Universitas Islam Malang

* Korespondensi Penulis. E-mail: ketutweda81@gmail.com

© 2023 JIPM (Jurnal Ilmiah Pendidikan Matematika)

This is an open access article under the CC-BY-SA license

(<https://creativecommons.org/licenses/by-sa/4.0/>) ISSN 2337-9049 (print), ISSN 2502-4671 (online)

Abstrak: Penelitian ini bertujuan untuk (1) menentukan kualitas item soal ujian dalam hal reliabilitas, tingkat kesukaran, dan daya pembeda, (2) menerapkan software anates untuk menghitung reliabilitas, tingkat kesukaran dan daya pembeda. Metode penelitian yang digunakan adalah teknik deskriptif kuantitatif dan teknik dokumentasi. Datanya terdiri dari 31 lembar jawaban siswa pada kelas VIIA SMPN 6 Singosari Malang tahun ajaran 2022/2023. Data-data tersebut dianalisis reliabilitas, tingkat kesukaran dan daya pembeda menggunakan software anates. Hasil penelitian menunjukkan bahwa Reliabilitas soal tergolong baik berdasarkan *software anates* yang menunjukkan nilai koefisien $r_{11} > r_{tabel}$. Terdapat 3 butir soal (15%) baik sekali, 4 butir soal (20%) tergolong baik, 6 butir soal (30%) tergolong sedang, 6 butir soal (30%) tergolong kurang baik dan 1 butir soal (5%) tergolong jelek sekali berdasarkan analisis daya beda. Proporsi soalnya untuk kelompok mudah, sedang dan sukar yaitu -15-60-10, sehingga belum memenuhi kriteria keseimbangan yang proposional yaitu 3 5-2 atau 3-4- 3. Kesimpulannya bahwa tingkat kesukaran butir soal kurang baik karena keseimbangan soal yang mudah sedang dan sukar belum proposional.

Kata kunci: Analisis; Soal Ujian; Anates

Abstract: This study aims to (1) determine the quality of the test items in terms of reliability, difficulty level, and discriminating power, (2) apply Anates software to calculate reliability, difficulty level and discriminating power. The research method used is quantitative descriptive techniques and documentation techniques. The data consists of 31 student answer sheets in class VIIA at SMPN 6 Singosari Malang for the 2022/2023 academic year. The data were analyzed for reliability, level of difficulty and discriminatory power using Anates software. The results showed that the reliability of the questions was good based on the Anates software which showed a coefficient value of $r_{11} > r_{tabel}$. There are 3 items (15%) very good, 4 items (20%) are good, 6 items (30%) are moderate, 6 items (30%) are not good and 1 item (5%) is very bad based on the analysis of differential power. The proportion of questions for the easy, medium and difficult groups is 15-60-10, so that they do not meet the proportional balance criteria, namely 3 5-2 or 3-4- 3. The conclusion is that the difficulty level of the item items is not good because the balance of questions that are easy to medium and difficult yet proportional.

Keywords: Analysis; Examination Questions; Anates

Pendahuluan

Proses pembelajaran merupakan kegiatan utama dan pertama sekolah (Junaedi, 2019). Proses pembelajaran adalah suatu kegiatan interaksi antara pendidik dan peserta

didik dimana akan di akhiri dengan proses evaluasi hasil belajar (Nurfaizah & Oktavia, 2020; Putra et al., 2020; Ratnasari, 2019). Menurut (Hasyim, 2014) proses pembelajaran merupakan suatu proses yang mengandung serangkaian perbuatan guru dan siswa atas dasar hubungan timbal balik yang berlangsung dalam situasi edukatif untuk mencapai tujuan tertentu. Tiga unsur dalam proses pembelajaran adalah tujuan pengajaran (tujuan instruksional), pengalaman belajar (proses belajar mengajar), dan hasil belajar (Harefa, 2020). Kegiatan penilaian dalam dunia pendidikan dikenal dengan istilah kegiatan evaluasi. Evaluasi merupakan bagian dari proses belajar mengajar (Huljannah, 2021; Idrus, 2019; Utami, 2021). Evaluasi hasil belajar dilakukan oleh pendidik untuk memantau proses, kemajuan, pencapaian dan perbaikan hasil belajar peserta didik secara sistematis dan berkesinambungan (Darmawan, 2016; Mahirah, 2017; Siregar, 2017). Evaluasi hasil belajar yang diadakan pada tiap-tiap mata pelajaran, akhir semester, menjadi sangat penting (urgent) kedudukannya dan fungsinya dalam mengukur tingkat kemampuan dan pemahaman siswa. Aktivitas evaluasi sebenarnya harus selalu dilakukan pada saat akhir pelajaran, gunanya untuk menilai sampai seberapa besar tingkat penguasaan ilmu pengetahuan yang diberikan dan diserap siswa. Dalam hal ini, proses persiapan, pembuatan soal, pelaksanaan tes, observasi dan penilaian tes, hendaknya direncanakan secara teratur dan kontinyu sehingga guru dapat benar-benar mengevaluasi dan membimbing perkembangan siswa secara positif sesuai dengan penerapan Kurikulum Tingkat Satuan Pendidikan (KTSP) 2006, 2010 maupun Kurikulum 2013 (Hadi et al., 2019; Magdalena et al., 2021; Riadi, 2017; Solichin, 2017).

Evaluasi hasil belajar siswa diujikan melalui ujian akhir semester (UAS). Nilai UAS merupakan gambaran penguasaan kompetensi yang dipelajari siswa dalam menempuh proses pembelajaran di sekolah selama satu semester, sehingga diperlukan soal yang berkualitas baik. Evaluasi pembelajaran dilakukan oleh pendidik dalam bentuk tes. Fungsi tes secara umum yaitu (1) sebagai alat pengukur terhadap perkembangan atas kemajuan peserta didik; dan (2) sebagai alat ukur pengukur keberhasilan program pengajaran (Anates et al., 2014; Khaerudin, 2015). Kualitas tes sebagai salah satu alat evaluasi yang penting untuk diperhatikan. Kualitas tes dapat menjadi pedoman untuk penilaian ketepatan hasil belajar peserta didik. Pendidik perlu melakukan analisis terlebih dahulu sebelum melakukan penilaian hasil belajar. Analisis merupakan kemampuan seseorang untuk merinci atau menguraikan suatu bahan atau keadaan menurut bagian-bagian yang lebih kecil dan mampu memahami hubungan di antara bagian-bagian atau faktor-faktor yang satu dengan faktor-faktor lainnya. Analisis yang perlu dilakukan pendidik adalah analisis butir soal.

Kegiatan analisis butir soal merupakan kegiatan penting dalam penyusunan soal agar diperoleh butir soal yang bermutu. Tujuan kegiatan analisis butir soal adalah mengkaji dan menelaah setiap butir soal agar diperoleh soal yang bermutu sebelum digunakan, meningkatkan kualitas butir tes melalui revisi atau membuang soal yang tidak efektif, serta mengetahui informasi diagnostik pada siswa apakah mereka telah memahami materi yang telah diajarkan. Soal yang bermutu adalah soal yang dapat memberikan informasi setepat tepatnya tentang siswa yang telah menguasai materi dan siswa yang belum menguasai materi (Alpusari, 2014; Fiska et al., 2021; Hasibuan, 2013; Langsa, 2018; Muluki et al., 2020). Fakta di lapangan menunjukkan bahwa analisis butir soal masih jarang dilakukan oleh pendidik. Beberapa alasan yang menyebabkan pendidik tidak melakukan analisis butir soal antara lain: (1) pendidik merasa terbebani dalam proses analisis butir soal sehingga tidak melakukannya, (2) pendidik meyakini bahwa kualitas soal tes yang dibuat sudah baik sehingga tidak perlu melakukan penelaahan lebih lanjut.

Soal yang digunakan untuk ujian semester dibuat oleh guru yang memang ahli di bidangnya namun tidak dibuat berdasarkan analisis yang telah dilakukan. Analisis butir

soal dilakukan oleh sekolah atau masing-masing guru, karena menganalisis butir soal secara keseluruhan sekota Malang membutuhkan biaya yang besar dan waktu yang cukup banyak. Berdasarkan uraian di atas, maka perlu dilakukan analisis butir soal secara kuantitatif untuk mengetahui kualitas soal tersebut. Kualitas soal UAS dapat dilihat dari hasil uji validitas, reliabilitas, tingkat kesukaran dan daya beda. Peneliti tertarik untuk melakukan penelitian dengan judul “Analisis Kualitas Soal Ujian Akhir Semester dengan menggunakan aplikasi program Anates”. Penelitian ini bertujuan untuk (1) mengetahui kualitas butir soal ujian akhir semester ditinjau dari validitas, reliabilitas, tingkat kesukaran, dan daya beda, (2) menerapkan program untuk menghitung validitas, reliabilitas, tingkat kesukaran dan daya beda. Dalam melakukan analisis soal tes dapat dilakukan dengan menggunakan program analisis yang sudah dipercaya kemampuan analisisnya, antara lain program Anates. Anates merupakan program komputer yang digunakan untuk menganalisis butir soal. Program ini belum memasyarakat secara umum, namun untuk dunia pendidikan program ini cukup terkenal. Analisis terhadap butir soal pilihan ganda dapat dilakukan menggunakan program komputer yaitu program Anates.

Berdasarkan observasi yang dilakukan, diketahui bahwa kebanyakan peneliti baik dari kalangan guru maupun mahasiswa cenderung melakukan analisis item secara terpisah-pisah. Sebagai contoh untuk analisis tingkat kesukaran maupun daya beda, masing-masing dianalisis secara terpisah sehingga menyita waktu. Analisis soal dengan menggunakan program Anates dapat menghemat waktu dan efektif, karena dalam satu kali kegiatan menganalisis akan didapatkan beberapa hal, antarlain: jumlah soal yang dianalisis, jumlah siswa, rata-rata jawaban benar, standar deviasi, skor minimum dan maksimum siswa, reliabilitas skor, tingkat kesukaran maupun daya beda.

METODE

Jenis penelitian ini adalah penelitian deskriptif. Penelitian ini bertujuan untuk (1) menentukan kualitas item soal ujian dalam hal reliabilitas, tingkat kesukaran, dan daya pembeda, (2) menerapkan software anates untuk menghitung reliabilitas, tingkat kesulitan dan daya pembeda. Pendekatannya berupa pendekatan kuantitatif, dan teknik dokumentasi. Data penelitian berupa 31 lembar jawaban siswa. Penelitian ini dilakukan di SMPN 6 Singosari Malang. Data awal yang digunakan adalah hasil ujian akhir semester untuk mata pelajaran matematika. Penelitian ini dilakukan melalui tiga tahapan yaitu tahap perencanaan, tahap pelaksanaan, dan tahap akhir. Tahap perencanaan meliputi: (a) pemilihan topik penelitian; (b) perumusan masalah penelitian; (c) penelusuran sumber data; (d) menyusun instrumen penelitian. Adapun tahap pelaksanaan, kegiatan yang dilakukan meliputi: (a) mengumpulkan data dengan cara dokumentasi hasil *ujian akhir semester*; (b) mengolah data berupa hasil *ujian akhir semester* kelas VII. Tahap Akhir dari penelitian ini kegiatannya meliputi: (a) menganalisis dan membahas temuan dalam penelitian; (b) membuat kesimpulan dari penelitian yang telah dilakukan. Penelitian ini menggunakan data sekunder. Adapun data dalam penelitian yaitu: (a) dokumen soal dan kunci jawaban soal *ujian akhir semester*; (b) dokumen lembar jawaban dan skor jawaban siswa tentang soal *ujian akhir semester* matematika.

Populasi yang digunakan pada penelitian ini adalah seluruh siswa kelas VII SMPN 6 Singosari Malang sebanyak 216 siswa dan jumlah sampel 31 siswa. Teknik pengumpulan data adalah dokumentasi. Dokumentasi ini diperoleh dari Guru matematika berupa soal, kunci jawaban, lembar jawaban, dan skor jawaban siswa. Pengumpulan data dilakukan dengan memasukkan skor jawaban siswa ke tabel tabulasi jawaban siswa. Analisis data dalam penelitian ini adalah analisis deskriptif kuantitatif. Analisis data yang dimaksud meliputi analisis persentase kesulitan dan analisis letak

kesulitan siswa. Data yang telah terkumpul selanjutnya dianalisis menggunakan aplikasi Anates dengan langkah-langkah yaitu: (a) membuka Program Aplikasi Anates; (b) memasukkan data membuat file baru (database). Pada langkah memasukkan data, kegiatan yang dilakukan meliputi: penyekoran data, skor dibobot, pengolahan data, dan olah semua data secara otomatis.

Data dianalisis menggunakan beberapa rumus validitas, reliabilitas, tingkat kesukaran dan daya beda serta penggunaan *software* anates V4. Bentuk tes yang dianalisis adalah tes 20 soal pilihan ganda atau multiple choice. Validitas adalah indeks yang menunjukkan sejauh mana suatu alat pengukur itu mengukur apa yang ingin diukur (Anates et al., 2014). Analisis validitas dilakukan dengan menggunakan Teknik korelasi point biserial yaitu:

$$r_{pbi} = \frac{M_p - M_t}{SD_t} \sqrt{\frac{p}{q}} \quad (1)$$

Rumus lainnya dengan menggunakan korelasi *Product moment*

$$r_{xy} = \frac{n \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \cdot \sum_{i=1}^n Y_i}{\sqrt{[n \sum_{i=1}^n X_i^2 - (\sum_{i=1}^n X_i)^2][n \sum_{i=1}^n Y_i^2 - (\sum_{i=1}^n Y_i)^2]}} \quad (2)$$

Nilai r_{xy} atau r_{pbi} akan dibandingkan dengan koefisien korelasi *table* nilai "r" *product moment* pada taraf signifikan 5%. Apabila nilai r_{xy} atau r_{pbi} hasil koefisien korelasi lebih besar (>) dari nilai *rtabel*, maka hasil yang diperoleh adalah signifikan, artinya butir soal tes dinyatakan valid. Nilai r_{xy} adalah nilai koefisien korelasi dari setiap butir/ item soal sebelum dikoreksi, kemudian dicari *corrected item-total correlation coefficient* dengan rumus sebagai berikut:

$$r_{x(y-1)} = \frac{r_{xy} S_y - S_x}{\sqrt{S_y^2 + S_x^2 - 2r_{xy}(S_y)(S_x)}} \quad (3)$$

Nilai $r_{x(y-1)}$ akan dibandingkan dengan koefisien korelasi *table* $r_{tabel} = r(\alpha, n-2)$. Jika $r_{x(y-1)} \geq r_{tabel}$, maka instrumen valid. Pada output SPSS, *corrected item-total correlation coefficient* $t \geq r_{tabel}$, maka instrumen valid. Reliabilitas adalah indeks yang menunjukkan sejauh mana suatu alat pengukur dapat dipercaya atau dapat diandalkan. Menurut Muhidin dan Abdurahman dalam (Imron, 2019) suatu instrument pengukuran dikatakan reliabel jika pengukurannya konsisten, cermat dan akurat.. Nilai *koefesien* (r_i) akan dibandingkan dengan koefesien korelasi *tabel* $r_{tabel} = r(\alpha, n-2)$. Jika $r_{11} > r_{tabel}$, maka instrumen reliabel. Pada Output SPSS, jika *Cronbach's Alpha* $> r_{tabel}$, maka instrumen reliabel. Analisis tingkat kesukaran soal artinya mengkaji soal-soal tes dari segi kesulitannya sehingga dapat diperoleh soal-soal mana yang termasuk mudah, sedang, dan sukar. Menurut (Putriani et al., 2020) mengatakan bahwa asumsi yang digunakan untuk memperoleh kualitas soal yang baik, di samping memenuhi validitas dan reliabilitas adalah adanya keseimbangan dari tingkat kesulitan soal tersebut. Keseimbangan yang dimaksudkan adalah adanya soal-soal yang termasuk mudah, sedang dan sukar secara proporsional. Perbandingan antara soal mudah-sedang-sukar bisa dibuat 3-4-3, artinya 30% soal kategori mudah, 40% soal kategori sedang, dan 30% soal kategori sukar. Perbandingan lain yang termasuk sejenis dengan proporsi di atas adalah 3-5-2. Artinya 30% soal kategori mudah, 50% soal kategori sedang,

dan 20% soal kategori sukar (Butir et al., 1990; Padang, 2021). Penafsiran tingkat kesukaran butir tes digunakan kriteria menurut *Witherington* dalam (Fiska et al., 2021) seperti Tabel 1 berikut.

Tabel 1. Interpretasi tingkat kesukaran butir tes

Interval	Interpretasi
0,00 - 0,30	Soal Sukar
0,31 - 0,70	Soal Sedang
0,71 - 1,00	Soal Mudah

Cara lain yang dapat ditempuh untuk menafsirkan analisis tingkat kesukaran butir soal dalam anates adalah sesuai Tabel 2 berikut.

Tabel 2. Interpretasi tingkat kesukaran anates

Interval	Interpretasi
0% - 15%	Sangat Sukar
16% - 30%	Sukar
31% - 70%	Sedang
71% - 85%	Mudah
86% - 100%	Sangat Mudah

Menganalisis daya pembeda artinya mengkaji soal-soal tes dari segi kesanggupan tes tersebut dalam membedakan siswa yang termasuk ke dalam kategori lemah/ rendah dan kategori kuat/ tinggi prestasinya (Darmawan, 2016). Penentuan daya beda butir soal pada Anates dapat diketahui dalam tabel daya pembeda pada kolom DP persen. Butir soal yang memiliki indeks daya beda $\geq 0,30$ dinyatakan baik dan butir soal yang indeks daya beda $< 0,30$ dinyatakan tidak baik. Daya pembeda butir soal memiliki manfaat yaitu untuk meningkatkan mutu setiap butir soal melalui data empiriknya dan untuk mengetahui seberapa jauh masing-masing butir soal dapat membedakan kemampuan siswa, yaitu siswa yang telah memahami atau belum memahami materi yang diajarkan pendidik (Arif et al., 2014). Sebagai catatan, dalam keadaan di mana jumlah *testee* adalah cukup besar (100 orang atau lebih), daya pembeda item cukup dihitung berdasar 27% *testee* kelompok atas dan 27% dari *testee* kelompok bawah, sedangkan *testee* yang terletak diantara kedua ujung ekstrem itu tidak perlu diikutsertakan dalam perhitungan analisis. Dalam menginterpretasi indeks daya pembeda menggunakan interpretasi menurut Anas Sudijono dalam (Susanto et al., 2015) yang dapat dilihat pada Tabel 3 berikut.

Tabel 3. Interpretasi Indeks Daya Pembeda Butir

Daya Pembeda	Klasifikasi	Interpretasi
0,70 - 1,00	Excellent	Baik Sekali
0,40 - 0,69	Good (baik)	Baik
0,20 - 0,39	Satisfactory (Memuaskan)	Cukup
0,00 - 0,19	Poor (lemah)	Kurang Baik
Bertanda Negatif	-	Jelek Sekali

HASIL PENELITIAN DAN PEMBAHASAN

Hasil data yang peneliti peroleh berupa soal dan hasil jawaban tes pada mata pelajaran matematika kelas VII tahun 2022/2023 di SMP 6 Singosari Malang, ditunjukkan pada Tabel 4 berikut.

Tabel 4. Jawaban Siswa

No	No. Butir Baru --->	Skor	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Urut	No. Butir Asli ---->		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
	Kunci ----->		B	D	B	A	C	D	A	B	A	D	C	C	B	B	B	D	A	B	B	C
	Kode Subyek																					
1	ANA	11	1	1	1	1	1	1	1	1	1	-	-	1	-	-	-	-	1	-	-	-
2	AM	13	1	1	1	1	1	1	1	1	1	1	1	-	-	1	1	-	-	-	-	-
3	AHAA	13	1	-	1	-	1	1	1	1	1	1	1	1	1	-	-	-	1	1	-	-
4	ANF	13	-	1	1	1	1	1	1	1	1	1	1	-	1	-	1	-	-	-	1	-
5	ARP	17	1	1	1	1	-	1	1	1	1	1	1	1	1	1	1	-	-	1	1	1
6	AN	14	1	1	1	1	1	1	1	1	1	-	1	1	-	-	-	-	1	-	1	1
7	AI	12	1	1	1	1	1	1	1	1	-	1	1	1	-	-	-	-	-	-	1	-
8	ADN	9	1	1	-	1	1	1	1	1	-	-	1	-	-	1	-	-	-	-	-	-
9	AR	9	1	1	-	-	1	1	-	1	1	-	1	-	-	-	-	1	-	-	-	1
10	AS	12	1	1	-	1	1	1	1	1	1	1	1	-	-	-	-	-	-	-	1	1
11	CB	12	1	1	1	1	1	1	1	1	1	1	1	-	-	-	-	-	-	-	-	1
12	DS	12	1	1	1	1	1	1	1	1	1	1	1	-	-	-	1	-	-	-	-	-
13	DN	12	-	-	1	1	1	1	1	1	1	1	1	-	-	-	-	-	1	-	1	1
14	EM	12	1	1	1	1	1	1	1	1	1	-	1	-	-	1	-	-	-	1	-	-
15	FDA	9	1	1	-	-	-	-	-	1	1	1	1	-	1	-	-	-	-	1	1	-
16	HD	6	1	1	1	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	1	1
17	IA	15	1	1	1	1	1	-	-	-	1	1	1	1	-	1	1	1	1	-	1	1
18	MAF	15	1	1	-	-	1	-	1	-	1	1	1	1	1	1	1	-	1	1	1	1
19	MDN	7	-	-	-	1	1	-	-	-	-	-	1	-	-	-	1	1	1	-	-	1
20	MHP	6	1	1	-	1	1	-	-	-	-	-	-	1	-	-	-	-	-	1	-	-
21	MA	9	-	1	-	1	1	1	1	1	1	-	-	-	-	-	-	-	1	-	-	1
22	MF	9	1	-	-	-	-	-	-	1	1	1	-	-	1	1	1	-	-	1	1	-
23	NA	8	1	1	-	-	-	-	-	-	1	1	-	1	-	-	-	-	1	-	1	1
24	RA	10	1	1	-	1	1	-	-	-	1	1	1	1	1	-	-	-	-	1	-	-
25	RSB	8	1	-	-	1	1	-	-	-	1	1	-	1	-	1	-	-	1	-	-	-
26	RD	8	1	-	-	-	-	-	-	-	1	1	-	1	1	1	1	-	-	-	1	1
27	SA	11	-	1	1	1	1	-	-	-	1	1	1	1	-	1	-	-	-	-	1	1
28	SAA	11	1	1	-	1	1	1	1	1	1	-	-	-	-	1	1	-	-	-	1	-
29	SH	7	1	1	-	1	-	-	-	1	1	-	-	-	1	-	-	-	-	-	-	1
30	TAF	11	1	1	1	1	-	-	1	-	1	-	-	1	-	1	1	-	1	-	1	-

Berdasarkan Tabel 4 tersebut, nilai maksimum atau nilai tertinggi didapatkan oleh 1 siswa yang mendapatkan nilai 17. Nilai minimum atau nilai terendah didapatkan oleh 6 orang siswa yang mendapatkan nilai 6. *Mean* atau rata-rata skor yang diperoleh untuk mata pelajaran matematika adalah 10,65. Jumlah siswa yang memperoleh nilai lebih besar dari *mean* adalah 17 orang. Analisis reliabilitas pada penelitian ini menggunakan *software* Anates.

Berdasarkan hasil analisis menggunakan *Software anates memiliki indeks reliabilitas yaitu 0,68*. Hasil analisis menunjukkan bahwa seluruh soal reliabel. Analisis tingkat kesukaran artinya mengkaji soal-soal tes dari segi kesulitannya sehingga dapat diperoleh soal-soal mana yang termasuk dalam kategori mudah, sedang dan sukar. Soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sukar. Tingkat kesukaran soal dipandang dari kemampuan siswa dalam menjawabnya, bukan dilihat dari sudut guru sebagai pembuat soal. Kesalahan yang dilakukan siswa saat menyelesaikan soal bisa karena siswa tidak memahami maksud soal, salah memilih rumus, tidak menggunakan rumus, tidak mengikuti prosedur penyelesaian soal, kurang berlatih dalam menghadapi soal, tidak dapat menyelesaikan soal dengan mudah, kurang teliti dan cepat dalam menjawab soal, (Saputri et al., 2021). Analisis tingkat kesukaran dalam penelitian ini menggunakan *software* Anates. Hasil analisis Tingkat Kesukaran disajikan pada Gambar 2 berikut.

No Butir Baru	No Butir Asli	Jml Betul	Tkt. Kesukaran(%)	Tafsiran
1	1	26	83,87	Mudah
2	2	25	80,65	Mudah
3	3	15	48,39	Sedang
4	4	23	74,19	Mudah
5	5	22	70,97	Sangat Mudah
6	6	17	54,84	Sedang
7	7	17	54,84	Sedang
8	8	19	61,29	Sedang
9	9	27	87,10	Sangat Mudah
10	10	19	61,29	Sedang
11	11	19	61,29	Sedang
12	12	15	48,39	Sedang
13	13	9	29,03	Sukar
14	14	12	38,71	Sedang
15	15	10	32,26	Sedang
16	16	4	12,90	Sangat Sukar
17	17	12	38,71	Sedang
18	18	8	25,81	Sukar
19	19	16	51,61	Sedang
20	20	15	48,39	Sedang

Gambar 2. Hasil Analisis Tingkat Kesukaran

Hasil analisis yang dilakukan terhadap 20 butir item tes hasil belajar dengan menggunakan software anates, menunjukkan bahwa terdapat 2 butir soal (10%) dalam kategori sangat mudah, 3 butir soal (15%) dalam kategori mudah, 12 butir soal (60%) dalam kategori sedang, 2 butir soal (10%) dalam kategori sukar, dan 1 butir soal (5%) dalam kategori sangat sukar. Butir soal yang termasuk dalam kategori mudah yaitu butir soal nomor 1, 2, dan 4. Butir-butir item yang termasuk kategori item berkualitas baik, dalam arti derajat kesukaran itemnya cukup atau sedang (tidak terlalu sukar dan tidak terlalu mudah) yaitu butir soal nomor 3, 6, 7, 8, 10, 11, 12, 14, 15, 17, 19, dan 20. Adapun butir soal yang termasuk kategori sukar adalah butir soal nomor 13, dan 18. Rumus ini mempunyai pendekatan dengan cara mencari persentase siswa yang menjawab salah. Artinya semakin banyak siswa yang menjawab salah, maka derajat tingkat kesukaran butir soal akan semakin sukar. Penggunaan *software* anates dalam menentukan tingkat kesukaran menghasilkan nilai yang tidak jauh berbeda dengan rumus. Perbedaan hanya terletak pada penafsiran yang cenderung lebih banyak dalam *software* anates. Berdasarkan pembahasan di atas dapat disimpulkan bahwa butir-butir soal matematika yang berbentuk *multiple choice* memiliki kualitas soal yang kurang baik karena belum memiliki keseimbangan. Asumsi yang digunakan untuk memperoleh kualitas soal yang baik, di samping memenuhi validitas dan reliabilitas adalah adanya keseimbangan dari tingkat kesulitan soal tersebut. Keseimbangan yang dimaksud adalah jumlah soal-soal yang tergolong mudah, sedang dan sukar seimbang. Tingkat kesukaran soal perlu dilihat dari kemampuan siswa dalam menjawab soal yang diberikan, bukan dilihat dari sudut pandang guru sebagai pembuat soal. Perbandingan antara soal yang mudah-sedang-sukar dapat dibuat 3-4-3 atau 3-5-2, yang diartinya adalah 30% soal kategori mudah, 40% soal kategori sedang, dan 30% soal kategori sukar; atau 30% soal kategori mudah, 50% soal kategori sedang dan 20% soal kategori sukar. Dalam penelitian ini terdapat lima kategori, yaitu sangat mudah, mudah, sedang, sukar, dan sangat sukar dengan proporsi soal 10-15-60-10-5. Perbandingan antara soal yang mudah-sedang-sukar yaitu 15-60-10. Sehingga belum memenuhi keseimbangan 3-4-3 atau 3-5-2.

Butir-butir soal yang masuk dalam kategori sedang sebaiknya segera dicatat dalam buku bank soal. Selanjutnya, butir-butir soal tersebut dapat dikeluarkan lagi dalam tes hasil belajar pada waktu yang akan datang. Untuk butir soal yang sukar ada tiga kemungkinan tindak lanjut, yaitu: (1) butir soal tersebut dibuang dan tidak dikeluarkan lagi dalam tes hasil belajar yang akan datang, (2) diteliti ulang, dilacak dan ditelusuri sehingga dapat diketahui faktor apa yang menyebabkan butir item yang bersangkutan sulit dijawab oleh *testee*. Setelah dilakukan perbaikan, butir-butir soal tersebut dikeluarkan lagi dalam tes hasil

belajar yang akan datang, (3) butir soal yang sukar sebaiknya digunakan dalam tes seleksi yang sifatnya ketat karena kondisi tersebut mendukung penggunaan soal yang sukar dengan asumsi bahwa *testee* yang berkemampuan cukup memadai akan lolos dalam seleksi yang diadakan. Butir soal yang mudah juga memiliki tiga kemungkinan tindak lanjut yaitu: (1) Butir soal tersebut dibuang dan tidak dikeluarkan lagi dalam tes hasil belajar yang akan datang. (2) diteliti ulang dan ditelusuri secara cermat guna mengetahui faktor yang menyebabkan butir soal tersebut dapat dijawab betul oleh seluruh *testee*; setelah dilakukan perbaikan butir soal yang bersangkutan dikeluarkan lagi pada tes berikutnya untuk mengetahui apakah derajat kesukaran butir soal menjadi lebih baik daripada sebelumnya atau tidak, (3) butir-butir soal yang mudah mempunyai manfaat yaitu butir-butir soal dapat dimanfaatkan pada tes seleksi yang bersifat longgar, dalam arti bahwa sebagian besar dari *testee* akan dinyatakan lulus dalam tes seleksi tersebut. Dalam kondisi ini, pemberian butir soal yang mudah akan memberikan kesempatan bagi banyak *testee* untuk lolos dalam tes seleksi atau ujian yang diadakan.

Analisis daya beda artinya mengkaji soal-soal tes dari segi kesanggupan tes tersebut dalam membedakan siswa yang termasuk kedalam kategori rendah dan kategori tinggi. Daya pembeda butir adalah kemampuan suatu butir tes untuk dapat membedakan antara *testee* yang berkemampuan tinggi dan berkemampuan rendah. Analisis daya beda pada penelitian ini menggunakan rumus dan *software* Anates. Hasil analisis daya beda disajikan pada Gambar 3. berikut.

No Butir Baru	No Butir Asli	Kel. Atas	Kel. Bawah	Beda	Indeks DP (%)
1	1	7	7	0	0,00
2	2	7	5	2	25,00
3	3	7	1	6	75,00
4	4	6	5	1	12,50
5	5	7	3	4	50,00
6	6	6	1	5	62,50
7	7	7	0	7	87,50
8	8	6	2	4	50,00
9	9	7	6	1	12,50
10	10	7	4	3	37,50
11	11	8	1	7	87,50
12	12	6	5	1	12,50
13	13	4	2	2	25,00
14	14	4	2	2	25,00
15	15	5	1	4	50,00
16	16	1	1	0	0,00
17	17	4	4	0	0,00
18	18	3	1	2	25,00
19	19	6	3	3	37,50
20	20	4	5	-1	-12,50

Gambar 3 Hasil Analisis Daya Beda

Hasil analisis dengan menggunakan *software* anates, berdasarkan daya pembeda diperoleh informasi bahwa soal dalam kategori baik sekali yaitu 3 butir soal atau 15%, soal dalam kategori baik yaitu 4 butir soal atau 20%, soal dalam kategori sedang yaitu 6 butir soal atau 30%, soal dalam kategori kurang baik yaitu 6 butir soal atau 30%, dan sisanya 1 butir soal atau 5% tergolong kategori jelek sekali. Butir soal yang termasuk dalam kategori baik sekali adalah butir soal nomor 3, 7 dan 11. Butir soal yang termasuk dalam kategori baik adalah butir soal nomor 5,6, 8 dan 15. Butir soal yang dikategorikan sedang adalah butir soal nomor 2, 10, 13, 14, 18 dan 19. Selanjutnya, butir soal yang termasuk dalam kategori kurang baik adalah soal nomor 1, 4, 9 12, 16 dan 17, sedangkan butir soal yang dikategorikan jelek sekali adalah butir soal nomor 20. Berdasarkan analisis yang peneliti lakukan dengan menggunakan *software* anates, diperoleh informasi bahwa soal tersebut terdapat 20% butir dalam kategori baik atau 4 butir soal, 30% kategori sedang atau 6 butir soal, 30% butir soal dalam kategori kurang baik atau sebanyak 6 butir soal dan 5% butir soal dalam kategori Jelek sekali atau 1 butir soal. Penggunaan *software* anates dalam menentukan

indeks daya beda suatu butir soal menurut peneliti adalah dapat untuk digunakan karena *software anates* memiliki hasil yang tidak jauh berbeda dengan rumus. *Database software* tidak mengalami pembaharuan dan tidak adanya interpretasi yang digunakan pada *software anates* dalam memberikan penafsiran daya beda suatu butir soal sehingga peneliti harus memberikan penafsiran secara mandiri terhadap butir soal yang dianalisis. Adapun tindak lanjut atas hasil analisis mengenai daya pembeda item tes hasil belajar tersebut adalah soal yang memiliki daya pembeda yang baik hendaknya dimasukkan dan dicatat dalam buku bank soal. Butir soal tersebut dapat dikeluarkan lagi dalam tes selanjutnya karena kualitasnya sudah cukup memadai. Butir soal yang daya pembedanya masih rendah memiliki dua kemungkinan tindak lanjut yaitu: (1) ditelusuri dan diperbaiki sehingga nanti dapat diajukan lagi dalam tes hasil belajar yang akan datang, serta perlu dianalisis kembali apakah daya pembeda meningkat atau tidak, (2) dibuang dan tidak dikeluarkan kembali pada tes yang akan datang. Khusus butir soal yang angka indeks diskriminannya bertanda negative sebaiknya tidak perlu dikeluarkan pada tes hasil belajar selanjutnya atau direvisi kembali.

SIMPULAN

Berdasarkan hasil penelitian dan pembahasan, dapat diperoleh kesimpulan yaitu Reliabilitas soal tergolong baik berdasarkan perhitungan menggunakan beberapa rumus dan *software anates* yang menunjukkan nilai koefisien $r_{11} > r_{tabel}$. Berdasarkan analisis yang peneliti lakukan dengan menggunakan *software anates*, bahwa Terdapat 3 butir soal (15%) baik sekali, 4 butir soal (20%) tergolong baik, 6 butir soal (30%) tergolong sedang, 6 butir soal (30%) tergolong kurang baik dan 1 butir soal (5%) tergolong jelek sekali berdasarkan analisis daya beda. Proporsi dari soal tersebut untuk kelompok mudah, sedang dan sukar yaitu 15-60-10. Sehingga tingkat kesukaran butir soal kurang baik karena keseimbangan soal yang mudah sedang dan sukar belum memenuhi keseimbangan yang proposional yaitu 3 5-2 atau 3-4-3.

DAFTAR RUJUKAN

- Alpusari, M. (2014). 106 Analisis Butir Soal, Program Komputer Anates Versi 4.0 For Windows Mahmud Alpusari. *Jurnal Primary Program Studi Pendidikan Guru Sekolah Dasar Fakultas Keguruan dan Ilmu Pendidikan* 3, 106-115.
- Anates, A., Dalam, V., Interdiana, A., Sari, C., & Herawati, M. (2014). Menganalisis Butir Soal. *Faktor Jurnal Ilmiah Kependidikan*. I(2), 203-214.
- Arif, M., Madura, U. T., & Wibisono, Y. (2014). Penerapan aplikasi anates bentuk soal pilihan ganda. *Jurnal Ilmiah Edutic*. 1(1), 1-9.
- Darmawan, C. A. (2016). Analisis Kualitas Soal Try Out Ujian Nasional. *Jurnal Pendidikan dan Pembelajaran Matematika (JP2M)* 2(1), 20-34.
- Fiska, J. M., Hidayati, Y., Qomaria, N., Hadi, W. P., Pendidikan, F. I., Trunojoyo, U., & Bangkalan, M. (2021). Analisis butir soal ulangan harian ipa menggunakan software anates pada pendekatan teori tes klasik. *Journal Natural Science Education Research*. 4(1).
- Hadi, K., Dazrullisa, Manurung, B., & Hasruddin. (2019). Analisis Daya Beda Soal Tes Berpikir Tingkat Tinggi Berbasis Kearifan Lokal pada Materi Keanekaragaman Hayati Kelas X SMA. *BIONatural*, 6(2), 85-91.

- Harefa, D. (2020). Peningkatan Hasil Belajar Ipa Fisika Siswa Pada Model Pembelajaran Prediction Guide. *Indonesian Journal of Education and Learning*, 4(1), 399. <https://doi.org/10.31002/ijel.v4i1.2507>
- Hasibuan, E. S. B. (2013). Analisis Soal Ulangan Tengah Semester Bahasa Indonesia Kelas XII MAS Raudhatul Ulum. *Jurnal Pendidikan dan Pembelajaran*. 2 (4), 1-17
- Hasyim, M. (2014). Penerapan Fungsi Guru Dalam Proses Pembelajaran. *Auladuna*, 1(2), 265–276. <http://103.55.216.56/index.php/auladuna/article/view/556>
- Huljannah, M. (2021). Pentingnya Proses Evaluasi Dalam Pembelajaran Di Sekolah Dasar. *Educator (Directory of Elementary Education Journal)*, 2(2), 164–180. <https://doi.org/10.58176/edu.v2i2.157>
- Idrus. (2019). Evaluasi Dalam Proses Pembelajaran. *Jurnal Managemen Pendidikan Islam*. 2, 920–935.
- Imron. (2019). Analisa Pengaruh Kualitas Produk Terhadap Kepuasan Konsumen Menggunakan Metode Kuantitatif Pada CV. Meubele Berkah Tangerang. *Indonesian Journal on Software Engineering (IJSE)*, 5(1), 19–28. <https://doi.org/10.31294/ijse.v5i1.5861>
- Junaedi, I. (2019). Proses pembelajaran yang efektif. *Journal of Information System, Applied, Management, Accounting and Research*, 3(2), 19–25.
- Khaerudin. (2015). Kualitas Instrumen Tes Hasil Belajar. *Jurnal Madaniyah*, 2, 212–235.
- Langsa, F. I. (2018). *Al-Khawarizmi: Jurnal Pendidikan dan Pembelajaran Matematika ISSN 2549-3906 E-ISSN 2549-3914*. 2(2), 139–151.
- Magdalena, I., Fauziah, S. N., Faziah, S. N., & Nopus, F. S. (2021). Analisis Validitas, Reliabilitas, Tingkat Kesulitan dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan. *BINTANG : Jurnal Pendidikan Dan Sains*, 3(2), 198–214.
- Mahirah, B. (2017). Evaluasi belajar peserta didik (siswa). *Idaarrah: Jurnal Manajemen Pendidikan*, 1(2), 257–267.
- Muluki, A., Bundu, P., & Sukmawati, I. (2020). Analisis Kualitas Butir Tes Semester Ganjil Mata Pelajaran IPA Kelas IV MI Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar*, 4(1), 86. <https://doi.org/10.23887/jisd.v4i1.23335>
- Nurfaizah, S., & Oktavia, P. (2020). Proses Pembelajaran Siswa Sekolah Dasar di MI Nurul Hikmah. *As-Sabiqun: Jurnal Pendidikan Islam Anak Usia Dini* 2(1), 43–48. <https://doi.org/10.36088/assabiqun.v2i1.621>
- Padang, U. N. (2021). *Natural Science: Jurnal Penelitian Bidang IPA dan Pendidikan IPA*. 7(1), 28–35.
- Putria, H., Maula, L. H., & Uswatun, D. A. (2020). Analisis Proses Pembelajaran dalam Jaringan (DARING) Masa Pandemi Covid- 19 Pada Guru Sekolah Dasar. *Jurnal Basicedu*, 4(4), 861–870. <https://doi.org/10.31004/basicedu.v4i4.460>
- Putriani, D., Turahmah, F., Sunarti, N. R., Ismarliana, P., & Walid, A. (2020). Analisis Butir Soal UAS Biologi 2018/2019 Kelas X dan XI SMAN 11 Kota Bengkulu Tahun 2018/2019. *Journal of Biology Learning*, 2(1), 1–7. <https://doi.org/10.32585/.v2i1.559>
- Ratnasari, K. I. (2019). Proses Pembelajaran Inquiry Siswa MI untuk Meningkatkan Kemampuan Matematika. *Auladuna : Jurnal Prodi Pendidikan Guru Madrasah Ibtidaiyah*,

1(1), 100–109. <https://doi.org/10.36835/au.v1i1.166>

- Riadi, A. (2017). Kompetensi Guru dalam Pelaksanaan Evaluasi Pembelajaran. *Ittihad Jurnal Kopertais Wilayah XI Kalimantan*, 15(28), 52–67. <https://doi.org/10.30998/rdje.v6i1.4371>
- Saputri, R. R., Sa'dijah, C., & Chandra, T. D. (2021). Analisis Kemampuan Representasi Matematis Siswa Tipe Camper Dalam Pemecahan Masalah Matematika Di Era Pandemi Covid-19. *JIPM (Jurnal Ilmiah Pendidikan Matematika)*, 10(2), 178. <https://doi.org/10.25273/jipm.v10i2.8694>
- Siregar, R. L. (2017). Evaluasi Hasil Belajar Pendidikan Islam. *HIKMAH: Jurnal Pendidikan Islam*, 6(1), 59–75.
- Solichin, M. (2017). Analisis Daya Beda Soal, Taraf Kesukaran, Validitas Butir Tes, Interpretasi Hasil Tes dan Validitas Ramalan dalam Evaluasi Pendidikan. *Dirasat: Jurnal Manajemen & Pendidikan Islam*, 2(2), 192–213. www.depdiknas.go.id/evaluasi-proses-
- Susanto, H., Rinaldi, A., & Novalia. (2015). Analisis Validitas Reabilitas Tingkat Kesukaran dan Daya Beda pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika. *Al-Jabar: Jurnal Pendidikan Matematika*, 6(2), 203–217.
- Utami, L. W. S. (2021). Penggunaan Google Form Dalam Evaluasi Hasil Belajar Peserta Didik Di Masa Pandemi C0Vid-19. *TEACHING : Jurnal Inovasi Keguruan Dan Ilmu Pendidikan*, 1(3), 150–156. <https://doi.org/10.51878/teaching.v1i3.453>