

ANALISIS BUTIR SOAL ULANGAN HARIAN IPA MENGGUNAKAN SOFTWARE ANATES PADA PENDEKATAN TEORI TES KLASIK

Joelan Maulidina Fiska¹, Yunin Hidayati², Nur Qomaria³, Wiwin Puspita Hadi⁴

^{1, 2, 3, 4} Program studi Pendidikan IPA, Fakultas Ilmu Pendidikan, Universitas Trunojoyo Madura Bangkalan, 69162, Indonesia
yolanmaulidinaf@gmail.com

Diterima tanggal: 12 Maret 2021 Diterbitkan tanggal: 30 Juli 2021

Abstrak Penelitian bertujuan untuk mengetahui kualitas ulangan harian IPA materi “Tanah dan Kehidupan” melalui pendekatan klasik yang terdiri dari validitas, reliabilitas, tingkat kesukaran, daya beda dan efektivitas pengecoh. Penelitian ini berjenis kuantitatif dengan metode deskriptif. Desain yang digunakan dalam penelitian ini yaitu *ex post facto*. Penelitian dilaksanakan di UPTD SMPN 1 Bangkalan dengan subjek sejumlah 145 peserta didik yang terdiri dari kelas IX-F sampai IX-J Tahun Ajaran 2019/2020. Data dikumpulkan dengan teknik observasi, dokumentasi dan tes kemudian dianalisis dengan *software* anates. Hasil yang diperoleh yaitu validitas masih rendah, reliabilitas yang belum reliabel, tingkat kesukaran dan daya beda yang cukup baik, serta efektivitas pengecoh yang kurang baik.

Kata Kunci: anates, pendekatan klasik, ulangan harian IPA

Abstract *This research aimed to find out the quality of science formative test on “Land and Life” topic by classical approach which consist of validity, reliability, difficulty level, difference power, and effectiveness of distractors. It was quantitative research which is using descriptive method. The research design was ex post facto. This research was implemented in UPTD SMPN 1 Bangkalan which the samples are 145 students from IX-F to IX-J of 2019/2020 school year. The data collection uses observation technique, documentation, and analyzed the data by using anates software. The result obtained was the low validity, reliability which has not been reliable, the level of difficulty and power difference was quite good, and the effectiveness distractors was deficient.*

Keywords: Anates, classical approach, science formative test.

Pendahuluan

Evaluasi merupakan sebuah kegiatan yang tidak akan bisa dipisahkan dalam proses pembelajaran. Evaluasi merupakan kegiatan yang direncanakan dengan baik oleh guru untuk mengetahui kemampuan dari peserta didik yang telah diukur untuk menghasilkan sebuah keputusan (Lubis, Syarifuddin, & Dongoran, 2017). Dengan adanya evaluasi, guru dapat terus meningkatkan kualitas dari pembelajaran yang dilakukan agar kemampuan peserta didik terus meningkat. Peningkatan kemampuan peserta didik akan mempengaruhi kualitas dari suatu pendidikan. Kualitas pendidikan itu sebanding dengan kemampuan peserta didik. Jika kemampuan peserta didik tinggi, maka kualitas pendidikan pun akan baik. Sedangkan apabila kemampuan peserta didik rendah, maka kualitas pendidikan pun akan buruk.

Upaya yang dilakukan untuk meningkatkan kualitas pendidikan selain dari ditingkatkannya kualitas pembelajaran, juga bisa ditempuh dengan ditingkatkannya kualitas asesmennya atau sistem penilaian (Azhary, 2015). Oleh sebab itu peningkatan asesmen juga harus selalu dilakukan. Asesmen merupakan sebuah proses yang dilakukan oleh guru untuk mengambil suatu keputusan terhadap peserta didik (Irmayta, Rudibyani, & Efkar, 2018). Semakin berkualitasnya asesmen maka semakin mudah guru untuk memahami kelebihan maupun kelemahan dari peserta didik.

¹ Corresponding Author

Dengan diketahuinya kelemahan peserta didik, maka guru dapat mengevaluasi proses pembelajaran yang dilakukan dengan memanfaatkan kelebihan dari peserta didik.

Salah satu alat asesmen atau penilaian yang sering digunakan guru dalam proses pembelajaran yaitu tes berupa ulangan harian. Ulangan harian merupakan sebuah kegiatan yang dilakukan guru dalam mengukur pencapaian kemampuan dari peserta didik setelah menyelesaikan satu atau lebih kompetensi dasar (KD) (Magdalena, Nurkhofifah, & Hatta, 2015). Ulangan harian IPA merupakan salah satu tes yang juga sering digunakan dalam mengetahui pencapaian kemampuan peserta didik. Ulangan harian IPA biasanya dilakukan saat guru dan peserta didik telah menyelesaikan suatu kompetensi dasar.

Analisis butir soal merupakan sebuah kegiatan yang wajib dilakukan oleh guru yang berfungsi untuk meningkatkan kualitas butir soal yang telah dibuat (Lestari, Wardana, & Rahmawati, 2019). Kegiatan analisis butir soal lebih baik dilakukan ketika seorang guru hendak mengujikan suatu set soal pada peserta didik. Ketika guru telah melakukan kegiatan analisis butir soal, guru dapat mengetahui apakah butir soal yang telah disusun tersebut dapat menjalankan fungsinya dengan baik atau sebaliknya. Apabila hasilnya kurang bahkan tidak baik sebaiknya perlu dilakukan suatu perubahan terhadap soal-soal yang telah dibuat oleh guru sesuai dengan penyebabnya sedangkan jika hasilnya baik maka soal-soal tersebut dapat disimpan di bank soal untuk digunakan pada tes berikutnya.

Pada kenyataannya, banyak sekali guru yang masih belum melakukan kegiatan analisis butir soal karena merasa terbebani dalam melakukan proses analisis butir soal sebelum soal tersebut diujikan pada peserta didiknya (Susanto, Rinaldi, & Novalia, 2015). Kebanyakan guru membuat soal secara mendadak saat soal akan diberikan pada peserta didik sehingga guru merasa terbebani jika harus melakukan kegiatan analisis butir soal terlebih dahulu karena terkendala oleh waktu yang ada. Selain itu banyak guru yang menganggap butir soal yang telah dibuatnya sudah memiliki kualitas yang baik sehingga tidak perlu melakukan kegiatan analisis butir soal (Susanto et al., 2015). Hal tersebut terjadi ketika peserta didik saat diberikan suatu set soal oleh guru dan peserta didik tersebut banyak yang mendapat nilai tinggi, sehingga membuat guru merasa bahwa soal yang telah dibuatnya merupakan soal yang berkualitas baik. Bisa saja, peserta didik mendapatkan nilai tinggi tersebut terjadi karena soal yang dibuat oleh guru memiliki kualitas tingkat kesukaran yang terlalu mudah, sehingga peserta dengan mudahnya menjawab soal yang diberikan oleh guru tersebut.

Secara umum soal-soal yang digunakan sebagai tes atau alat penilaian belum diketahui kualitasnya terutama ditinjau dari pendekatan klasik yang terdiri dari aspek validitas, reliabilitas, tingkat kesukaran, daya beda dan efektivitas pengecoh setiap butir soal (Azis, 2016). Sehingga kualitasnya tidak dapat ditingkatkan. Soal yang bermutu atau yaitu soal yang bisa memberikan informasi mengenai pemahaman peserta didik terhadap materi yang diajarkan oleh guru.

Masalah yang juga sering ditemukan yaitu masih banyak guru yang menganalisis kualitas butir soal masih menggunakan cara manual yang meliputi rumus yang rumit dan pastinya membutuhkan waktu yang lama dalam melakukan kegiatan analisis butir soal (Kurniawan et al., 2017). Hal tersebut terjadi karena masih banyak guru yang belum atau tidak mengetahui adanya berbagai macam *software* yang berfungsi dalam memudahkan proses menganalisis butir soal. Anates merupakan salah satu dari berbagai macam *software* yang berfungsi dalam menganalisis kualitas dari setiap butir soal melalui pendekatan klasik atau ditinjau dari aspek validitas, reliabilitas, tingkat kesukaran, daya beda dan efektivitas pengecoh soal dengan cepat. Selain itu kelebihan dari *software* anates yaitu dapat menganalisis butir soal baik dalam bentuk pilihan ganda maupun uraian. Anates sendiri juga berbahasa Indonesia sehingga memudahkan penggunanya dalam mengoprasikannya.

Berdasarkan hasil dari observasi yang dilakukan di UPTD SMPN 1 Bangkalan guru IPA belum pernah melakukan kegiatan analisis butir soal pada soal UH. Sehingga mutu dari UH sendiri belum diketahui dan tidak dapat ditingkatkan. Guru hanya mengetahui peserta didik yang bisa menjawab soal tes yang diberikan guru dan peserta didik yang belum bisa menjawab soal tes yang diberikan oleh guru. Sesuai dengan hal tersebut menandakan bahwa tugas guru belum sesuai

dengan yang seharusnya dilakukan dalam mengukur kemampuan dari peserta didik sesuai dengan hasil belajarnya di kelas.

Agar kualitas dari soal UH yang dibuat oleh guru dapat diketahui kualitasnya baik dari aspek validitas, reliabilitas, tingkat kesukaran, daya beda dan efektivitas pengecoh maka perlu diadakan penelitian dengan judul “Analisis Butir Soal Ulangan Harian IPA Menggunakan *Software Anates*”. *Software anates* digunakan untuk memudahkan peneliti dalam menganalisis butir soal. Tujuan dari penelitian ini yaitu untuk mengetahui kualitas dari aspek validitas, reliabilitas, tingkat kesukaran, daya beda dan efektivitas pengecoh soal.

Hasil dari penelitian diharapkan bisa bermanfaat untuk berbagai pihak. Bagi guru, penelitian ini dapat dijadikan informasi untuk mengetahui kualitas dari soal yang dibuat sehingga guru dapat terus meningkatkan soal sesuai dengan penyebab permasalahannya. Bagi peneliti, penelitian ini berfungsi untuk menambahkan wawasan mengenai *assessment* khususnya pada bidang analisis butir soal.

Metode Penelitian

Jenis pada penelitian ini yaitu kuantitatif dengan menggunakan metode deskriptif. Hal yang akan di deskripsikan dan digambarkan pada penelitian ini yaitu mengenai kualitas dari butir soal ulangan harian IPA kelas IX UPTD SMPN 1 Bangkalan yang ditinjau dari pendekatan klasik. Desain dalam penelitian ini menggunakan *ex post facto*. Penelitian dilakukan bertempat di UPTD SMPN 1 Bangkalan dan dilaksanakan pada 10 dan 12 Maret 2020. Subjek dalam penelitian ini yaitu 145 peserta didik yang terdiri dari peserta didik kelas IX-F sampai IX-J UPTD SMPN 1 Bangkalan Tahun Ajaran 2019/2020 dan objek dalam penelitian ini yaitu 20 butir soal ulangan harian IPA pada materi “Tanah dan Kehidupan” yang dibuat langsung oleh guru pengajar yang bersangkutan. Data yang digunakan pada penelitian ini yaitu soal ulangan harian IPA materi “Tanah dan Kehidupan”, kunci jawaban serta lembar jawaban dari peserta didik. Teknik yang digunakan dalam pengumpulan data yaitu observasi, dokumentasi dan tes.

Metode yang dapat digunakan dalam melakukan analisis butir soal dilakukan secara kuantitatif. Analisis secara kuantitatif biasanya ditinjau dari pendekatan klasik yang meliputi aspek validitas, reliabilitas, tingkat kesukaran, daya beda dan efektivitas pengecoh. Data yang diperoleh kemudian akan dianalisis dengan menggunakan *software anates* versi 4.09. Dalam mengetahui kualitas butir soal yang baik, aspek yang perlu diperhatikan secara urut yaitu dimulai dari validitas, daya beda, tingkat kesukaran, efektivitas pengecoh dan reliabilitas.

Validitas

Validitas digunakan untuk mengetahui tingkat keakuratan suatu tes dalam menjalankan fungsinya yaitu sebagai suatu alat ukur (Kurniawan, 2015). Jenis validitas yang digunakan untuk mengukur kemampuan dari hasil belajar peserta didik yaitu validitas isi. Suatu soal dikatakan valid apabila $r_{pbi} \geq r_{tabel}$ (Rahayu & Djazari, 2016). Nilai dari r_{tabel} diperoleh dari hasil taraf signifikansi 5% dengan n nya sesuai dengan jumlah soal tes yang akan diteliti. Pada *anates*, hasil validitas dapat dilihat melalui hasil korelasi skor dengan skor total. Jika hasilnya signifikan, maka soal tersebut valid, sedangkan jika hasilnya tidak signifikan maka soal tersebut tidak valid. Soal valid bisa disimpan dan soal yang tidak valid harus diubah sesuai dengan penyebabnya.

Reliabilitas

Reliabilitas digunakan untuk mengetahui tentang kekonsistenan suatu tes dalam menilai apa yang seharusnya dinilai (Kurniawan, 2015). Suatu tes dikatakan konsisten apabila hasil yang diperoleh mendekati ataupun sama ketika tes tersebut dilakukan secara berkali-kali. Jika koefisien reliabilitas semakin tinggi (mendekati 1), maka semakin tinggi pula kekonsistenan dari suatu tes tersebut. Reliabilitas dikatakan dapat dipercaya jika hasilnya relatif sama jika diujikan secara

berulang kali. Jika uji pada aspek reliabilitas hanya dilakukan satu kali, maka hasilnya masih belum dapat dipercaya. Reliabilitas yang baik apabila termasuk kedalam kategori tinggi yaitu ≥ 70 . Tabel 1 merupakan klasifikasi pada aspek reliabilitas

Tabel 1. Klasifikasi Reliabilitas

Rentang	Kualitas Butir Soal
$r_{11} < 0,20$	Kekonsistennya sangat rendah
$0,20 \leq r_{11} < 0,40$	Kekonsistennya rendah
$0,40 \leq r_{11} < 0,70$	Kekonsistennya cukup
$0,70 \leq r_{11} < 0,90$	Kekonsistennya tinggi
$0,90 \leq r_{11} \leq 1,00$	Kekonsistennya sangat tinggi

(Modifikasi Azis, 2016)

Tingkat Kesukaran

Tingkat kesukaran berguna mengetahui tentang tingkat kesukaran suatu soal yang dibuat (Kurniawan, 2015). Soal yang mudah, tidak merangsang peserta tes dalam memecahkan suatu soal. Sedangkan soal yang susah membuat peserta tes putus asa dalam memecahkan soal tersebut karena soal yang susah terlalu jauh jangkauannya dari pemikiran peserta tes. Soal yang baik apabila memiliki proporsi yang sebanding antara butir soal yang memiliki kualitas mudah: sedang dan sukar yaitu 3:5:2 atau 3:4:3. Jadi dalam membuat soal harus sebanding antara soal yang tingkat kesukarannya rendah, sedang dan tinggi. Tabel 2 merupakan klasifikasi pada aspek tingkat kesukaran.

Tabel 2. Klasifikasi Tingkat Kesukaran

Rentang	Kualitas Butir Soal
$p \leq 15\%$	Sangat sukar
$15\% < p \leq 30\%$	Sukar
$30\% < p \leq 70\%$	Sedang
$70\% < p \leq 85\%$	Mudah
$p > 86\%$	Sangat mudah

(Modifikasi Rahmasari & Ismiyati, 2016)

Daya Beda

Daya beda digunakan untuk mengetahui peserta tes yang telah menguasai kompetensi dan peserta tes yang belum menguasai kompetensi setelah melakukan proses pembelajaran (Kurniawan, 2015). Selain itu daya beda juga bisa dapat mengetahui peserta tes yang memiliki pandai dan kurang pandai. Tabel 3 merupakan klasifikas pada aspek daya beda.

Tabel 3. Klasifikasi Daya Beda

Rentang	Kualitas Butir Soal
$0,40 \leq D \leq 1,0$	Sangat baik
$0,3 \leq D < 0,4$	Baik
$0,2 \leq D < 0,3$	Cukup baik
$D < 0,2$	Tidak baik

(Modifikasi Rahmasari & Ismiyati, 2016)

Efektivitas Pengecoh

Pengecoh sendiri yaitu pilihan jawaban yang salah yang berfungsi untuk mengecoh peserta didik yang belum memahami konsep dalam pembelajaran dengan baik (Iskandar & Rizal, 2017). Adanya pengecoh juga bertujuan untuk mengetahui peserta didik yang memiliki kemampuan tinggi dan yang rendah. Setiap opsi pada butir soal dikatakan berfungsi apabila bertanda (-) kurang baik, (+) baik dan (++) sangat baik, sedangkan untuk yang tidak berfungsi bertanda (--) buruk, (---) sangat buruk dan terdapat blokyang menandakan opsi tersebut tidak berfungsi (Sari & Surtani, 2020). Berikut adalah kategori efektivitas pengecoh pada setiap butir soal sesuai dengan tabel 4.

Tabel 4. Kategori efektivitas pengecoh pada setiap butir soal

Efektivitas Pengecoh	Kategori
Baik	Tigapengecoh berfungsi
Cukup	Dua pengecoh berfungsi
Kurang baik	Satu pengecoh berfungsi
Sangat tidak baik	Semua pengecoh tidak berfungsi

(Modifikasi Rahayu & Djazari, 2016)

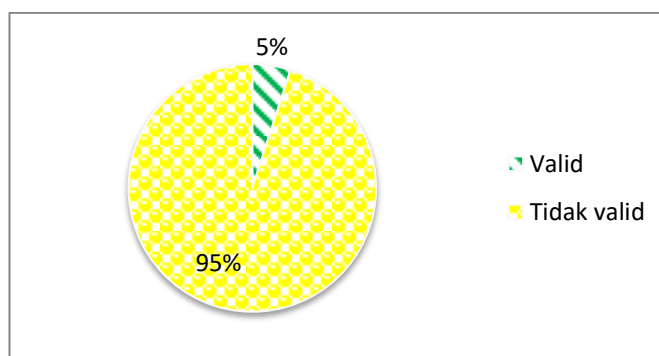
Hasil Penelitian dan Pembahasan

Validitas

Pada *software* anates, validitas bisa dilihat pada korelasi skor butir dengan skor total yang dibagi menjadi dua kategori yaitu signifikan (soal valid) dan tanda (-) yang berarti tidak signifikan (soal tidak valid). Pada *software* anates nilai r_{tabel} langsung diketahui pada bagian bagian korelasi dengan skor total yang terletak pada bawah tabel. Sesuai jumlah soal yaitu sebanyak 20 butir soal dan melihat r_{tabel} pada taraf signifikansi 5%, $n-2$ dari 20 butir soal diperoleh nilai sebesar 0,423. Sehingga hasil penelitian pada aspek validitas berpatokan jika nilai $r_{pbi} \geq 0,423$ maka soal tersebut valid, sedangkan $r_{pbi} < 0,423$ maka soal tidak valid. Berikut merupakan tabel hasil persebaran ulangan harian IPA pada aspek validitas yaitu pada tabel 5.

Tabel 5. Validitas Hasil Tes Ulangan Harian

Kategori Soal	Jumlah	Persentase (%)	Nomor Soal
Valid	1	5	15
Tidak Valid	19	95	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20
Jumlah	20	100	



Gambar 1. Diagram validitas soal tes

Gambar 1 menunjukkan hasil Analisis Butir Soal Ulangan Harian IPA Menggunakan *Software* Anates pada aspek validitas. Berdasarkan hasil analisis diketahui bahwa butir soal ulangan harian IPA berkategori valid berjumlah 1 butir soal saja dengan persentase 5%. Soal ulangan harian IPA yang lain berkategori tidak valid dengan jumlah 19 butir soal dengan persentase 95%.

Berdasarkan hasil validitas diketahui bahwa soal ulangan harian IPA yang terdiri dari 20 butir soal hanya diperoleh satu butir soal (5%) yang valid yaitu pada nomor 15 dengan skor $r_{pbi} \geq 0,423$, sedangkan sembilanbelas soal (95%) lainnya yaitu pada nomor 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20 dinyatakan tidak valid dengan skor $< 0,423$. Berdasarkan hasil tersebut dapat disimpulkan ulangan harian IPA yang ditinjau dari aspek validitas memiliki kualitas yang masih sangat rendah (tidak valid). Kualitas validitas yang sangat rendah bisa saja terjadi karena pada saat pembuatan soal ulangan harian guru masih belum membuat kisi-kisi soal terlebih dahulu sehingga pembuatan soal hanya dilakukan sekedarnya sajadan kurang mencakup pokok pembahasan yang telah diajarkan pada proses pembelajaran (Nurhasanah & Ahmad, 2017). Hal

tersebut sesuai dengan kondisi di lapangan yaitu tidak tersedianya kisi-kisi butir soal yang berarti guru tidak memiliki pedoman pada saat pembuatan soal UH IPA tersebut. Kisi-kisi sangat penting dibuat sebelum membuat soal, mengingat kisi-kisi dapat digunakan sebagai pedoman dalam menyusun butir-butir soal yang baik guna mempermudah pekerjaan guru dalam membuat suatu set soal serta juga dapat menghasilkan soal-soal yang dapat memenuhi tujuan dari diadakannya suatu tes tersebut.

Soal yang tidak valid berdampak pada hasil yang didapatkan (Nurhasanah & Ahmad, 2017). Soal yang tidak valid jika digunakan maka hasilnya tidak subjek yang berarti tidak dapat mengukur apa yang hendak diukur. Sehingga tidak akan menggambarkan mengenai kemampuan dari peserta didik yang sesungguhnya. Validitas berhubungan dengan reliabilitas (Handani & Prayitno, 2015). Reliabilitas diperlukan untuk memperoleh hasil yang valid karena soal yang reliabel akan menghasilkan soal yang valid. Soal atau tes sudah valid dan reliabel maka soal atau tes tersebut sudah menunjukkan kesesuaian dari kompetensi yang perlu dikuasai oleh peserta didik sesuai tujuan pembelajaran yang telah ditentukan. Uraian tersebut sesuai dengan pernyataan bahwa soal yang memiliki kualitas baik apabila memenuhi aspek validitas dan reliabilitas (Anita, Tyowati, & Zulfadrial, 2018).

Soal yang valid bisa disimpan pada bank soal untuk digunakan kembali. Butir soal yang tidak valid, maka harus diperbaiki dengan cara disesuaikan dengan teknik penyusunan butir soal maupun dengan indikator pencapaian. Guru juga bisa meminta pendapat dari beberapa para ahli untuk memperoleh soal yang valid.

Reliabilitas

Hasil analisis butir soal ulangan harian IPA menggunakan *software* anates pada aspek reliabilitas bernilai sebesar 0,54. Sesuai kriteria yang digunakan, maka nilai tersebut termasuk dalam kategori cukup yang termasuk dalam rentang $0,40 \leq r_{11} < 0,70$. Nilai reliabilitas pada soal ulangan harian IPA belum memenuhi kriteria yang baik karena kriteria soal yang baik harus memiliki nilai dengan kategori yang tinggi yaitu ≥ 70 . Hasil Analisis Ulangan Harian IPA Menggunakan *Software* Anates dapat dilihat pada tabel 6.

Tabel 6. Reliabilitas Hasil Tes Ulangan Harian

Rata-rata	10,92
Simpangan baku	2,42
KorelasiXY	0,37
Reliabilitas tes	0,54

Nilai reliabilitas yang diperoleh pada soal ulangan harian IPA yang dianalisis menggunakan *software* anates diperoleh sebesar 0,54 yang termasuk kategori cukup dan belum memenuhi kategori yang baik. Pernyataan tersebut sesuai dengan penelitian yang dilakukan oleh (Rahmasari & Ismiyati, 2016) yang menyatakan bahwa jika suatu tes memiliki nilai reliabilitas yang $< 0,70$, maka reliabilitas tersebut masih belum baik dan perlu diadakan perubahan karena soal tes tersebut masih belum reliabel. Jadi dapat disimpulkan bahwa soal ulangan harian IPA tersebut memiliki reliabilitas yang rendah atau tidak reliabel. Suatu tes yang nilai reliabelnya rendah menandakan bahwa soal tes tersebut belum dapat dipercaya, yang berarti kompetensi yang diujikan pada peserta didik tersebut belum atau tidak sebanding dengan kemampuan yang dimiliki oleh peserta didik.

Tinggi rendahnya suatu nilai reliabilitas dapat dipengaruhi oleh beberapa faktor, seperti: penyelenggaraan tes, banyaknya peserta tes dan kesukaran butir soal (Rahmasari & Ismiyati, 2016). Banyak bagian yang tergolong dalam penyelenggaraan tes seperti petunjuk pengerjaan, kurangnya kesiapan peserta didik dalam menjawab tes, tempat tes sendiri sehingga jika beberapa hal tersebut tidak mendukung maka akan mempengaruhi nilai reliabilitas yang dihasilkan. Semakin banyak peserta tes maka mengakibatkan semakin beragam pula jawaban yang dihasilkan dan semakin memperendah nilai reliabilitas. Semakin besar jarak tingkat kesukaran, maka nilai reliabilitas akan semakin rendah karena soal yang susah akan membuat peserta didik merasa putus asa saat

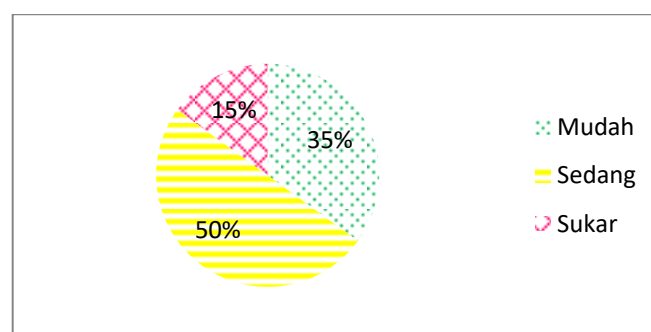
menjawab soal yang diberikan sehingga mereka menjawab soal dengan cara menebak (Anita et al., 2018).

Tingkat Kesukaran

Hasil analisis butir soal ulangan harian IPA menggunakan *software* anates pada aspek tingkat kesukaran dibagi menjadi 3 kategori, yaitu: mudah, sedang dan sukar. Suatu set soal yang baik memiliki proporsi yang seimbang tingkat kesukarannya antara kategori mudah:sedang:sukar dengan perbandingan 3:5:2 (Anita et al., 2018). Berikut merupakan tabel hasil persebaran ulangan harian IPA pada aspek tingkat kesukaran yaitu pada tabel 7.

Tabel 7. Tingkat Kesukaran Hasil Tes Ulangan Harian

Kategori Soal	Jumlah	Persentase (%)	Nomor Soal
Mudah	7	35	1, 2, 4, 7, 15, 17, 20
Sedang	10	50	3, 5, 6, 8, 9, 12, 13, 14, 16, 18
Sukar	3	15	10,11,19
Jumlah	20	100	



Gambar 2. Diagram tingkat kesukaran soal tes

Gambar 2 menunjukkan hasil Analisis Butir Soal Ulangan Harian IPA Menggunakan *Software* Anates pada aspek tingkat kesukaran. Berdasarkan hasil analisis diketahui bahwa butir soal ulangan harian IPA pada aspek tingkat kesukaran yang berketegori mudah berjumlah 7 butir soal dengan persentase 35%. Butir soal yang termasuk kategori sedang berjumlah 10 butir soal dengan persentase 50%. Butir soal yang termasuk kategori sukar berjumlah 3 butir soal dengan persentase 15%. Berdasarkan hasil yang diperoleh dapat diketahui bahwa soal ulangan harian IPA yang terdiri dari 20 butir soal pada aspek tingkat kesukaran menghasilkan perbandingan mudah:sedang:sukar yaitu 3,5:5:1,5. Perbandingan tersebut hampir mendekati proporsi soal pada aspek tingkat kesukaran yang baik yaitu 3:5:2 antara soal yang mudah:sedang:sukar.

Berdasarkan hasil tingkat kesukaran dapat diketahui bahwa soal ulangan harian IPA yang terdiri dari 20 butir soal terdapat 7 butir soal (35%) pada nomor soal 1, 2, 4, 7, 15, 17, 20 termasuk kategori mudah. Butir soal termasuk kategori mudah disebabkan karena banyak peserta didik yang dapat menjawab soal tersebut dengan benar. Butir soal yang termasuk kategori sedang berjumlah 10 butir soal (50%) pada nomor soal 3, 5, 6, 8, 9, 12, 13, 14, 16, 18, soal tersebut termasuk kategori sedang. Butir soal yang termasuk ke dalam kategori sedang sebaiknya disimpan di dalam bank soal untuk digunakan kembali. Butir soal yang baik yaitu butir soal yang termasuk ke dalam kategori sedang (Solichin, 2017). Butir soal yang termasuk kategori sukar berjumlah 3 butir soal (15%) pada nomor soal 10,11,19, soal tersebut termasuk kategori sukar. Butir soal memiliki kategori yang sukar terjadi karena kurangnya pemahaman dari peserta didik terhadap materi yang telah diajarkan atau pun kalimat yang digunakan pada soal terlalu sulit untuk dipahami oleh peserta didik.

Soal yang baik apabila soal tersebut tidak terlalu mudah serta tidak terlalu sukar. Suatu set soal juga harus terdiri dari soal yang sukar dan mudah (Fatimah & Alfath, 2019). Soal yang sukar akan menambah semangat belajar dari peserta didik yang kemampuannya tinggi dan soal yang mudah akan membangkitkan semangat dari peserta didik yang kemampuannya rendah. Soal yang terlalu mudah tidak akan mempertinggi usaha dari peserta didik dalam memecahkan suatu soal,

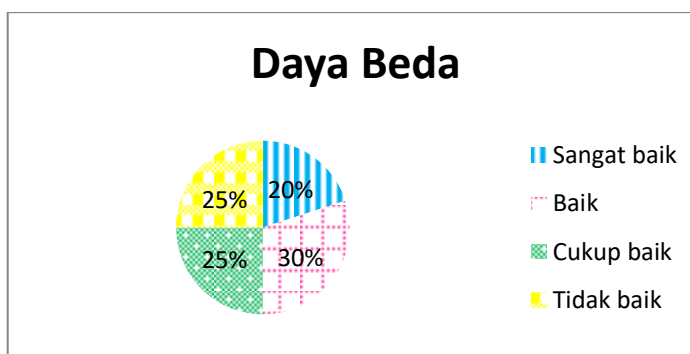
sedangkan soal yang susah membuat peserta didik berputus asa dalam memecahkan soal yang diberikan. Tingkat kesukaran yang dibuat pada setiap butir soal harus disesuaikan dengan tahap perkembangan peserta didik agar suatu tes dapat menjalankan fungsinya secara baik yaitu dapat mengukur tingkat perkembangan maupun kemajuan peserta didik setelah melakukan proses pembelajaran. Perbandingan soal yang baik ketika perbandingan antara kategori yang mudah:sedang:sukar yaitu 3:5:2 atau 3:4:3 (Kurniawan, 2015). Sehingga untuk memperoleh perbandingan yang proporsional maka soal yang mudah dikurangi satu butir soal dan yang sukar ditambah satu butir soal.

Daya Beda

Hasil analisis butir soal ulangan harian IPA menggunakan *software* anates pada aspek daya beda diinterpretasikan kedalam 4 kategori, yaitu: sangat baik, baik, cukup baik dan tidak baik. Hasil dari daya beda bisa bernilai positif dan negatif. Berikut merupakan tabel hasil persebaran ulangan harian IPA pada aspek daya beda yaitu pada tabel 8.

Tabel 8. Daya Beda Hasil Tes Ulangan Harian

Kategori Soal	Jumlah	Persentase (%)	Nomor Soal
Sangat baik	4	20	10, 15, 16, 18
Baik	6	30	2, 4, 7, 8, 14, 19
Cukup baik	5	25	3, 6, 11, 12, 20
Tidak baik	5	25	1, 5, 9, 13, 17
Jumlah	20	100	



Gambar 3. Diagram daya beda soal tes

Gambar 3 menunjukkan hasil Analisis Butir Soal Ulangan Harian IPA Menggunakan *Software* Anates pada aspek daya beda. Berdasarkan hasil analisis dapat diketahui bahwa butir soal ulangan harian IPA pada aspek daya beda yang berketegori sangat baik berjumlah 4 butir soal (20%). Butir soal yang termasuk kategori baik berjumlah 6 butir soal (30%). Butir soal yang termasuk kategori cukup baik berjumlah 5 butir soal (25%) dan yang termasuk kategori tidak baik berjumlah 5 soal (25%). *Software* anates bisa langsung membedakan peserta didik yang memiliki kemampuan tinggi dan rendah yang dapat dilihat pada bagian kelompok unggul dan ansor. Kelompok unggul berisi nama-nama peserta didik yang pandai dan kelompok ansor berisi nama-nama peserta didik yang memiliki kemampuan rendah kurang pandai.

Nilai daya beda yang diperoleh pada soal ulangan harian IPA yang dianalisis menggunakan *software* anates menunjukkan bahwa butir soal yang daya bedanya termasuk dalam kategori sangat baik berjumlah 4 soal (20%) pada soal nomor 10, 15, 16, 18. Butir soal yang daya bedanya termasuk dalam kategori baik berjumlah 6 soal (30%) pada soal nomor 2, 4, 7, 8, 14, 19. Butir soal yang daya bedanya termasuk dalam kategori kurang baik berjumlah 5 soal (25%) pada soal nomor 3, 6, 11, 12, 20 dan butir soal yang daya bedanya termasuk dalam kategori tidak baik berjumlah 5 soal (25%) pada nomor 1, 5, 9, 13, 17.

Daya beda maka suatu soal dapat membedakan antara peserta didik yang pencapaiannya tinggi dan pencapaiannya rendah pada tes pencapaian hasil belajar (Rahmasari & Ismiyati, 2016). Daya beda memiliki keterkaitan dengan tingkat kesukaran yaitu soal dibuat dengan berbagai kategori tingkat kesukaran bertujuan agar guru mengetahui penguasaan materi dari masing-masing

peserta didik. Selain itu keterkaitan antara daya beda dan tingkat kesukaran yaitu soal yang memiliki tingkat kesukaran sangat sukar maka daya bedanya rendah karena peserta didik sedikit yang menjawab soal tersebut dengan benar. Begitu pula dengan soal yang tingkat kesukarannya mudah tidak akan memiliki daya beda karena akan banyak sekali peserta didik yang menjawab soal tersebut (Fatimah & Alfath, 2019).

Butir soal sesuai aspek daya beda memiliki kualitas baik apabila butir soal tersebut lebih banyak dijawab oleh peserta didik pandai daripada yang kurang pandai. Butir soal dikatakan memiliki kualitas yang kurang baik apabila peserta didik yang pandai dan kurang pandai dapat menjawab soal bahkan sebaliknya (Anita et al., 2018). Hal tersebut biasanya terjadi ketika soal tersebut terlalu mudah sehingga semua peserta didik bisa menjawab soal dengan benar atau soal yang diberikan terlalu sukar sehingga peserta didik tidak bisa menjawab soal dengan benar.

Berdasarkan dari pernyataan diatas dapat disimpulkan bahwa kualitas butir soal ulangan harian IPA berdasarkan dari aspek daya beda memiliki kualitas yang cukup baik hal tersebut sesuai dengan penelitian yang dilakukan Bichi (2015) karena terdapat lima dari jumlah keseluruhan butir soal termasuk dalam kriteria yang tidak baik sehingga perlu dilakukan perubahan pada butir soal tersebut. Hal tersebut biasanya terjadi karena soal termasuk soal yang susah, selain itu beberapa materi pada soal tersebut belum diajarkan pada peserta didik (Anggraeni & Darmawan, 2016). Kedua hal tersebut membuat peserta didik menjawab soal secara ragu-ragu dan cenderung menebak dalam memilih jawaban yang tepat atau benar. Akan tetapi berdasarkan analisis butir soal pada aspek tingkat kesukaran, butir soal yang memiliki kualitas sukar hanya 3 butir soal, hal tersebut menyimpang dari alasan yang dipaparkan sebelumnya. Sehingga butir soal yang memiliki daya pembeda kurang baik bisa saja terjadi karena banyak peserta didik yang memiliki kemampuan rendah atau kurang pandai. Hal tersebut sesuai dengan penelitian ini, dari 145 peserta didik yang mengikuti hanya 11 orang peserta didik saja yang nilainya berada atau sama dengan KKM.

Pada *software* anates peserta didik yang memiliki kemampuan tinggi dan yang rendah dapat dilihat pada bagian kelompok unggul dan ankor. Pembagian kelompok unggul dan kelompok ankor yaitu jika subjek berjumlah ≥ 100 , maka akan dibedakan sebesar 27% yaitu yang terdiri dari 27% kelompok unggul dan 27% kelompok ankor (Solichin, 2017). Apabila subjeknya berjumlah ≤ 50 maka akan dibedakan sebesar 50% yaitu yang terdiri dari 50% kelompok unggul dan 50% kelompok ankor. Penelitian yang baik meskipun jumlah subjeknya kurang dari 100 dapat membedakan kelompok unggul dan ankor sebanyak 27% karena kelompok ini merupakan kelompok yang sensitif atau bisa diandalkan dalam melakukan kegiatan analisis butir soal (Anita et al., 2018). Pada *software* anates tidak hanya membagi peserta didik yang pandai dan kurang pandai saja, tetapi bisa membedakan peserta didik yang memiliki kemampuan sedang. Peserta didik dengan kemampuan sedang yaitu peserta didik jika kemampuannya diukur maka akan menghasilkan nilai yang tidak jauh dari KKM (Rahmasari & Ismiyati, 2016). Soal dengan kategori sangat baik, baik dan cukup bisa disimpan di bank soal, sedangkan yang tidak baik bisa diperbaiki atau diganti dengan soal yang baru.

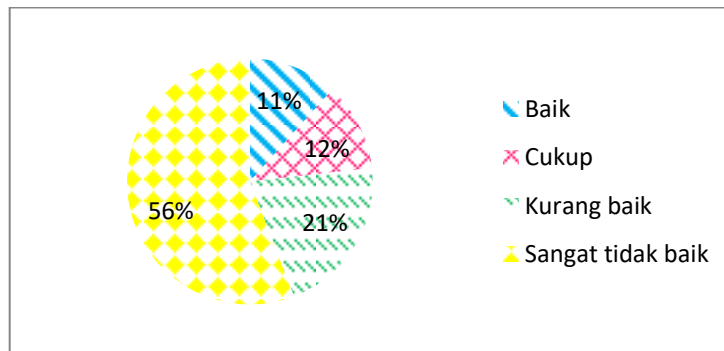
Efektivitas Pengecoh

Hasil analisis butir soal ulangan harian IPA menggunakan *software* anates pada aspek efektivitas pengecoh diinterpretasikan kedalam 5 kategori, yaitu: sangat baik, baik, cukup, kurang baik dan sangat tidak baik. Pada *software* anates sendiri jumlah peserta didik yang menjawab pada setiap opsinya langsung tertera pada masing-masing opsi jawaban, tetapi nama dan kelompok peserta didik yang menjawab setiap opsi jawaban tidak diketahui. Kriteria pada setiap opsi pada setiap butir soal dapat dilihat pada tabel 9.

Tabel 9. Efektivitas Pengecoh Hasil Tes Ulangan Harian

Kategori Soal	Jumlah	Persentase (%)	Nomor Soal
Baik	5	25	2, 4, 8, 16, 20
Cukup	5	25	5, 6, 11, 14, 18
Kurang baik	9	45	1, 3, 7, 9, 10, 12, 13, 15, 17
Sangat tidak baik	1	5	19

Jumlah	20	100
--------	----	-----



Gambar 4. Diagram efektivitas pengecoh soal tes

Gambar 4 menunjukkan hasil Analisis Butir Soal Ulangan Harian IPA Menggunakan *Software Anates* pada aspek efektivitas pengecoh. Berdasarkan hasil analisis diketahui butir soal ulangan harian IPA pada aspek efektivitas pengecoh yang berkategori baik berjumlah 5 butir soal (25%). Butir soal yang termasuk kategori cukup berjumlah 5 butir soal (25%). Butir soal yang termasuk kategori kurang baik berjumlah 9 butir soal (45%) dan yang termasuk kategori sangat tidak baik berjumlah 1 soal (5%).

Efektivitas pengecoh berfungsi dalam mengetahui keefektifan pengecoh dalam menjalankan fungsi ukurnya dan membedakan peserta didik yang sudah memahami konsep dan yang belum memahami konsep (Kurniawan, 2015). Peserta didik yang sudah memahami konsep ditandai dengan peserta didik bisa menjawab soal tes dengan benar sehingga tidak terkecoh dengan pengecoh yang tersedia. Peserta didik yang menjawab soal tes dengan benar maka peserta didik tersebut telah melakukan pembelajaran yang bermakna. Hal tersebut sesuai dengan teori belajar Ausubel yaitu peserta didik telah mengalami pembelajaran bermakna ketika peserta didik bisa menghubungkan pengetahuan yang diberikan guru dengan pengetahuan yang sebelumnya sudah dimiliki dan semua itu disimpan pada memori peserta didik sehingga menghasilkan pemahaman konsep yang utuh menjawab soal tes dengan benar (Gazali, 2016).

Pengecoh yang tidak dipilih sama sekali oleh peserta didik adalah pengecoh yang sangat tidak baik karena pengecoh tersebut terlalu mencolok. Berdasarkan hasil analisis pada aspek efektivitas pengecoh pada 20 butir soal ulangan harian IPA terdapat 5 butir soal (25%) berkategori baik yaitu pada soal nomor 2, 4, 8, 16, 20. Sebanyak 5 butir soal (25%) berkategori cukup yaitu pada nomor 5, 6, 11, 14, 18. Sebanyak 9 butir soal (45%) berkategori kurang baik yaitu pada nomor 1, 3, 7, 9, 10, 12, 13, 15, 17 dan sebanyak 1 soal (5%) berkategori sangat tidak baik yaitu pada soal nomor 19. Hasil dari penelitian tersebut sesuai dengan penelitian yang dilakukan Toksöz & Ertunç (2017) yaitu setengah dari keseluruhan jumlah butir soal memiliki efektivitas pengecoh dalam kategori kurang baik dan sangat tidak baik, hal tersebut terjadi karena efektivitas pengecoh tersebut kurang memiliki daya tarik bagi peserta didik yang belum atau kurang menguasai konsep. Selain itu bisa terjadi saat penyusunan tes masih kurang baik (Aggraeni & Darmawan, 2016). Penyusunan tes dikatakan baik apabila memiliki efektivitas pengecoh yang kerumitan dan panjang pendek kalimatnya relatif sama dengan kunci jawaban dan hal tersebut sudah terpenuhi pada butir soal ulangan harian IPA ini. Sehingga untuk soal yang efektivitas pengecoh berkategori kurang baik dan sangat tidak baik perlu dilakukan perubahan. Perubahan tersebut perlu sekali dilakukan karena efektivitas pengecoh sangat berdampak pada aspek tingkat kesukaran (Iskandar & Rizal, 2017). Ketika efektivitas pengecoh tidak berfungsi dengan baik, maka indeks pada aspek tingkat kesukaran juga menurun sehingga akan meningkatkan peluang peserta didik dalam menjawab soal tes dengan mudah.

Kesimpulan dan Saran

Soal ulangan harian IPA materi tanah dan kehidupan kelas IX UPTD SMPN 1 Bangkalan tahun ajaran 2019/2020 yang dianalisis menggunakan *software* anates pada aspek validitas, reliabilitas, daya beda, tingkat kesukaran dan efektivitas pengecoh secara berturut turut memiliki kualitas rendah, belum reliabel, cukup baik, cukup baik dan kurang baik. Soal yang kurang baik perlu dilakukan perubahan sesuai dengan penyebab yang membuat soal tersebut kurang baik dan soal yang tidak baik sebaiknya diganti dengan soal yang baru. Setelah melakukan tes, guru hendaknya melakukan kegiatan analisis butir soal guna mengetahui kualitas dari soal yang dibuat.

Daftar Pustaka

- Anita, Tyowati, S., & Zulfadrial. (2018). Analisis Kualitas Butir Soal Fisika Kelas X Sekolah Menengah Atas. *Jurnal Pendidikan*, 16(1), 35–47.
- Azhary. (2015). Analisis Assesment Soal Ujian Sekolah Mata Pelajaran Bahasa Indonesia di SMP Negeri 17 Palu. *E-Jurnal Bahasantodea*, 4(1), 39–47.
- Azis. (2016). Analisis Tes Buatan Guru Bidang Studi Matematika Kelas V SD 1 Katobengke. *Edumatica*, 06(01), 15–24.
- Fatimah, L. U., & Alfath, K. (2019). Analisis Kesukaran Soal, Daya Pembeda dan Fungsi Distraktor. *Jurnal Komunikasi Dan Pendidikan Islam*, 8(2), 37–64.
- Gazali, R. Y. (2016). Pembelajaran Matematika yang Bermakna. *Jurnal Pendidikan Matematika*, 2(3), 181–189.
- Handani, H. A., & Prayitno, H. J. (2015). Validitas dan Reliabilitas Soal Tengah Semester Genap Kaitannya dengan Ketercapaian Tujuan Pembelajaran Bahasa Indonesia Kelas VIII A SMP Negeri 2 Banyudono Tahun Ajaran 2013/2014. *University Research Colloquium*, 2(1), 193–206.
- Irmayta, E., Rudibyani, R. B., & Efkar, T. (2018). Pengembangan Instrumen Asesmen Pengetahuan pada Materi Asam Basa Arrhenius. *Jurnal Pendidikan Dan Pembelajaran Kimia*, 7(1), 63–76.
- Iskandar, A., & Rizal, M. (2017). Analisis Kualitas Soal di Perguruan Tinggi Berbasis Aplikasi TAP. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 21(2), 13–23.
- Kurniawan, R., Prakoso, A., Hakim, L., Dewi, R., & Irin, W. (2017). Pemberian Pelatihan Analisis Butir Soal Bagi Guru di Kabupaten Jombang Efektif. *Jurnal Pemberdayaan Masyarakat Madani (JPMM)*, 1(2), 179–193.
- Kurniawan, T. (2015). Analisis Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran IPS Sekolah Dasar. *Journal of Elementary Education*, 4(4), 1–6.
- Lestari, S., Wardana, M. Y. S., & Rahmawati, I. (2019). Analisis Butir Soal Ulangan Akhir Semester Ganjil Mata Pelajaran Matematika Kelas III SD Kecamatan Petarukan Tahun 2018/2019. *Elementary Shcool Jurnal*, 9(2), 118–125.
- Lubis, S. S., Syarifuddin, & Dongoran, H. (2017). Analisis Butir Soal dan Kemampuan Siswa Menjawab Tes UN dan UAS Ganjil Mata Pelajaran Biologi Kelas IX SMAN/MAN di Kota Medan. 5(3), 318–325.

- Magdalena, L., Nurkhofifah, A., & Hatta, M. (2015). Aplikasi Ulangan Harian Menggunakan Systematic Random Sampling Berbasis Web Studi Kasus: SMKN 1 Bulakamba-Brebes. *Jurnal Digit*, 5(2), 133–145.
- Nurhasanah, & Ahmad, H. (2017). Analisis Soal Ujian Tengah Semester (UTS) Mata Kuliah Geometri. *Jurnal Pendidikan PEPATUDZU*, 13(1), 30–42.
- Rahayu, R., & Djazari, M. (2016). Analisis Kualitas Soal Pra Ujian Nasional Mata Pelajaran Ekonomi Akutansi. *Jurnal Pendidikan Akutansi Indonesia*, XIV(1), 85–94.
- Rahmasari, D., & Ismiyati. (2016). Analisis Butir Soal Mata Pelajaran Pengantar Administrasi Perkantoran. *Economic Education Analysis Journal*, 5(1), 317–330.
- Susanto, H., Rinaldi, A., & Novalia. (2015). Analisis Validitas Reabilitas Tingkat Kesukaran dan Daya Beda pada Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran Matematika. *Jurnal Pendidikan Matematika*, 6(2), 203–217.