

Analisis Kualitas Butir Soal Penilaian Tengah Semester Berbasis HOTS Menggunakan Program Anates pada Mata Pelajaran PKn di Sekolah Dasar

Alvin Ilham Adestya ^{1*}, Encep Supriatna ²

^{1,2} Universitas Pendidikan Indonesia, Indonesia

* alvinilham1997@gmail.com

Abstrak

Penelitian ini dilatarbelakangi rendahnya hasil belajar siswa dalam PTS khususnya mata pelajaran PKn. Padahal, hasil PTS tersebut merupakan gambaran dari proses pembelajaran yang telah ditempuh. Maka, perlu dilakukan pengujian untuk menggambarkan kualitas soal yang dipakai pada PTS. Penelitian ini bertujuan melakukan analisa kualitas butir meliputi validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh pada butir soal Penilaian Tengah Semester mata pelajaran PKn di SDN Cipocok Jaya 4. Selain itu penulis juga melakukan analisa apakah butir soal tersebut berorientasi HOTS atau tidak. Metode yang digunakan adalah metode kombinasi (*mixed method*). Teknik pengumpulan data pada data kuantitatif dan data kualitatif adalah teknik pengumpulan data dokumentasi. Data yang dikumpulkan berupa lembar jawaban siswa, soal yang diujikan dan kunci jawaban. Analisis data kuantitatif menggunakan Anates untuk menguji validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh pada soal. Lalu data kualitatif untuk telaah HOTS dianalisis dengan reduksi, penyajian data, dan verifikasi. Hasil penelitian ini diperoleh dalam 25 butir soal PTS Mata Pelajaran PKn, hanya terdapat 7 soal yang berorientasi HOTS. Lalu dari segi validitas, reliabilitas, daya pembeda, tingkat kesukaran dan efektivitas pengecoh terdapat 10 butir soal yang berkualitas sangat baik, sebanyak 9 butir soal berkualitas baik, 4 butir soal termasuk dalam soal yang memiliki kualitas sedang, soal yang berkualitas tidak baik berjumlah 2 butir soal, dan tidak terdapat butir soal yang memiliki kualitas sangat tidak baik. Simpulan analisis ini yakni soal PTS dengan kualitas sangat baik dapat disimpan di bank soal, sedangkan soal dengan kualitas baik dan sedang perlu dilakukan revisi.

Keywords: *Analisis, Kualitas Butir Soal, HOTS, Program Anates, Pembelajaran PKn, Sekolah Dasar*

Pendahuluan

Sebagai cara untuk memperoleh pertanggungjawaban guru atas pelaksanaan pembelajaran yang dilaksanakan, evaluasi pembelajaran ialah rangkaian aktivitas metodis atau ilmiah, kontinu yang mengendalikan, memastikan serta berkualitas tidaknya pembelajaran untuk beragam keterampilan belajar berdasarkan indikator tertentu (Kurniawan et al, 2022). Sedangkan pemerolehan dan analisa data untuk memberikan penilaian akan pencapaian proses dan hasil belajar siswa, evaluasi hasil belajar merupakan serangkaian tindakan yang metodis, berkelanjutan, dan menyeluruh (Yazid et al, 2022). Maka, aktivitas ini di dalamnya berisikan penilaian akan hasil belajar maupun penilaian akan proses belajar. Agar proses evaluasi pembelajaran berjalan baik, tentunya diperlukan guru yang profesional. Guru yang profesional pasti memiliki kemampuan membentuk siswa supaya mencapai kriteria dari pendidikan nasional yang tertuang dalam Standar Pendidikan Nasional. Selain membimbing, guru juga bertugas untuk merencanakan dan menilai proses serta hasil belajar yang dilakukan. Untuk menciptakan

guru yang profesional maka diadakan penataran-penataran, penyempurnaan fasilitas diberbagai sektor pendidikan. Langkah tersebut adalah upaya mempertinggi kualitas pendidikan di Indonesia. Melaksanakan evaluasi pembelajaran, semua guru mendapatkan ragam soal dalam buku maupun bank soal maupun membuat soal sendiri. Baik itu soal tes maupun non tes. Soal tes yang digunakan diantaranya berbentuk pilihan ganda, isian singkat dan uraian. Hasil tes yang baik merupakan gambaran dari proses pembelajaran berhasil. Sebaliknya, apabila hasilnya buruk menandakan ada sesuatu yang salah dalam proses pembelajaran tersebut. Guru dan siswa ialah unsur ekosistem pendidikan, oleh karenanya mereka pasti ingin tahu hasil dari proses pembelajaran yang mereka lakukan.

Penelaahan akan berkualitas atau tidaknya instrumen soal merupakan langkah penting yang perlu ditempuh oleh pendidik guna terjadinya peningkatan kualitas soal yang dipergunakan. Proses ini melibatkan pengumpulan, peringkasan, serta pemanfaatan data dari respon siswa berupa jawaban untuk mengambil simpulan terkait efektivitas dan kualitas suatu penilaian (Santosa et al, 2021). Tujuan dari kegiatan penelaahan adalah mengevaluasi dan mengkaji setiap butir soal secara cermat guna memastikan kualitas soal yang tinggi sebelum diaplikasikan saat proses evaluasi proses ajar (Vitalocca et al, 2019). Itu semua dilakukan dengan maksud untuk mengontrol kualitas tes dengan tahapan proses eliminasi maupun perbaikan atau penghapusan terhadap soal-soal yang terbukti tidak layak. Analisis ini juga berfungsi untuk instrumen diagnostik agar pembuat soal memahami di fase apa peserta didik telah menguasai muatan ajar yang telah dilakukan. Soal memiliki kualitas tinggi adalah soal yang bisa menggambarkan informasi tanpa ambiguitas atau multi tafsir selaras dengan tujuan penyusunan tes, termasuk didalamnya mengidentifikasi seorang individu yang mencapai ataupun tidak mencapai kemampuan yang ditargetkan. Suatu soal dianggap bermutu tinggi jika di dalamnya memenuhi kriteria kesesuaian pengukuran (validitas), konsistensi hasil pengujian (reliabilitas), tingkat kesulitan soal, kemampuan membedakan tingkat kemampuan siswa (daya pembeda), serta keefektifan pilihan pengecoh dalam menjaring jawaban yang keliru secara fungsional juga proporsional, serta bisa mengukur kemampuan yang diharapkan secara tepat (Mania et al, 2023). Dengan menganalisis butir soal juga dapat menjadi indikator pencapaian kompetensi siswa setelah proses pembelajaran berlangsung. Hal yang perlu diukur dari suatu soal termasuk didalamnya kesesuaian pengukuran (validitas), konsistensi hasil pengujian (reliabilitas), tingkat kesulitan soal, kemampuan membedakan tingkat kemampuan siswa (daya pembeda), serta keefektifan pilihan pengecoh dalam menjaring jawaban yang keliru secara fungsional.

Tahapan menganalisis instrumen soal penting dilaksanakan untuk mengidentifikasi seberapa baik mutu soal yang dibuat. Suatu tes dikategorikan bermutu jika bisa mengukur apa yang mau diukur. Sulit untuk mengetahui maupun memastikan bahwa soal-soal yang diujikan berkualitas baik. Berdasarkan uraian diatas, penulis menyimpulkan bahwa sangat penting dilakukannya analisa butir soal. Hal ini bertujuan agar diketahuinya kualitas soal sampai bisa dipakai sebagai patokan nantinya. Penulis juga menganggap analisis butir soal ini penting berawal dari hasil observasi yang dilakukan pada kegiatan PPL di SDN Cipocok Jaya 4. Pada observasi tersebut ditemukan 23 siswa dari 26 siswa kelas V mendapatkan nilai PTS (Penilaian Tengah Semester) mata pelajaran PKn tidak mencapai KKM (Kriteria Ketuntasan Minimal = 65). Adapun siswa yang mencapai standar KKM hanya mendapat nilai dibawah angka 7. Penilaian Tengah Semester atau disingkat PTS merupakan salah satu alat evaluasi SDN Cipocok Jaya 4 untuk mengetahui dan mengukur tingkat keberhasilan pembelajaran peserta didik. Soal PTS disusun oleh Dinas Pendidikan dan Kebudayaan Kota Serang. Soal tersebut didistribusikan ke setiap sekolah. Peneliti pun melakukan wawancara kepada guru di SDN Cipocok Jaya 4, dari hasil wawancara tersebut disimpulkan bahwa guru belum pernah melakukan penganalisisan pada

tiap soal sudah dipakai. Berlandaskan temuan ini, kiranya harus dilakukannya analisis agar diketahui kualitas dari butir soal pada PTS yang dilakukan di SDN Cipocok Jaya 4 Tahun Ajaran 2018-2019 baik itu secara kualitatif maupun kuantitatif. Dengan demikian diharapkan dapat mengetahui butir soal tes pilihan ganda yang berorientasi HOTS pada PTS yang diujikan di SDN Cipocok Jaya 4 serta mutu tiap item tes yang dipakai saat PTS memperhatikan berbagai aspek penting dalam analisis butir soal, antara lain kesesuaian pengukuran (validitas), konsistensi hasil pengujian (reliabilitas), tingkat kesulitan soal, kemampuan membedakan tingkat kemampuan siswa (daya pembeda), serta keefektifan pilihan pengecoh dalam menjaring jawaban yang keliru secara fungsional.

Pemanfaatan analisis butir soal dengan bantuan Anates menjadi salah satu metode yang umum digunakan untuk menilai mutu soal secara statistik. Anates merupakan perangkat lunak yang dirancang untuk membantu pendidik dalam mengukur validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta efektivitas pengecoh. Alat ini mempermudah proses evaluasi dengan hasil yang akurat dan efisien, sehingga analisis butir soal dapat dilakukan secara lebih cepat. Melalui proses tersebut, pendidik memperoleh data kuantitatif yang bersifat objektif, yang dapat dijadikan dasar dalam pengambilan keputusan untuk meningkatkan kualitas soal (Alifah et al, 2025). Evaluasi butir soal memiliki peranan signifikan dalam menunjang penerapan Kurikulum Merdeka, yang mengutamakan pembelajaran berbasis proyek serta penilaian autentik. Bagi guru, menjaga mutu soal menjadi salah satu cara untuk memberikan kesempatan kepada siswa dalam mengembangkan kemampuan berpikir analitis, berkreasi, dan memecahkan masalah kemampuan yang akan menjadi bekal berharga di dunia kerja. Dengan demikian, penerapan analisis butir soal di sekolah dasar tidak hanya berfungsi untuk menilai capaian belajar, tetapi juga menjadi langkah strategis dalam meningkatkan mutu pendidikan secara keseluruhan.

Beberapa penelitian mengenai pengujian akan kualitas dari soal menggunakan Anates telah dilakukan. Penggunaan aplikasi Anates dianggap efisien dalam membantu dosen maupun guru melakukan evaluasi terhadap soal tes yang disusun. Tampilan antarmuka yang menggunakan bahasa Indonesia menjadi keunggulan tersendiri dari aplikasi ini (Rizky et al, 2019). Melalui pemanfaatan ANATES, pendidik dapat mendeteksi butir soal yang memiliki tingkat validitas rendah, reliabilitas kurang konsisten, daya pembeda lemah, maupun tingkat kesulitan yang tidak sejalan dengan kemampuan siswa (Khoerunnisa et al, 2025). Output analisis menggunakan Anates dapat dimanfaatkan untuk membangun bank soal atau merevisi bagian soal yang kurang optimal sebelum dipakai kembali. Dengan demikian, mutu evaluasi akan meningkat, sehingga peserta didik memperoleh penilaian yang lebih adil dan tepat (Jatmoko et al, 2025). Ketiga penelitian tersebut menggunakan anates hanya untuk menganalisis kualitas butir soal secara kuantitatif. *Novelty* penelitian ini terletak pada telaah HOTS pada soal PTS secara kualitatif. Dengan demikian, penelitian ini bertujuan melakukan analisa kualitas butir meliputi validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh pada butir soal Penilaian Tengah Semester mata pelajaran PKn di SDN Cipocok Jaya 4 menggunakan Anates. Selain itu penulis juga melakukan analisa apakah butir soal tersebut berorientasi HOTS atau tidak.

Metode

Peneliti mengaplikasikan *mix method* yang dicetuskan oleh (Creswell et al, 2018). Tahapan awal pada proses penelitian ini mencakup pengumpulan data hingga penganalisaan data kuantitatif dengan tujuan mengungkap kualitas atau mutu dari tiap soal *multiple choice* dalam mata pelajaran PKn melalui PTS di SDN Cipocok Jaya 4 ditinjau dari segi analisis kuantitatif yakni kesesuaian pengukuran (validitas), konsistensi hasil pengujian (reliabilitas), tingkat kesulitan soal, kemampuan membedakan tingkat kemampuan siswa (daya pembeda), serta keefektifan

pilihan pengecoh dalam menjangir jawaban yang keliru secara fungsional yang kemudian diinput ke program Anates V4 (Creswell et al, 2018). Lebih lanjut pada tahapan kedua, pemerolehan data dilanjutkan penganalisisan secara kualitatif dengan tujuan apakah butir soal *multiple choice* tersebut berorientasi HOTS atau tidak.

Populasi dalam penelitian ini adalah seluruh siswa SD Negeri Cipocok Jaya 4 yang terdiri dari kelas 1 sampai kelas 6 yang mengikuti PTS, sedangkan sampel dalam penelitian ini adalah seluruh siswa kelas V yang berjumlah 26 orang. Teknik yang dipilih saat pengumpulan data dilaksanakan secara bertahap dan bergantian. Ketika yang dicari adalah data kuantitatif maka memakai dokumentasi sedangkan data kualitatif menggunakan lembar pengamatan. Kedua jenis data tersebut meskipun memakai teknik yang berbeda namun akan saling melengkapi. Data yang dikumpulkan berupa dokumen hasil jawaban siswa, draft soal yang diujikan beserta kunci jawabannya (Yazid et al, 2022). Data kuantitatif, instrumen yang dipakai yakni tiap butir soal dalam lembar soal tersebut. Sedangkan untuk data kualitatif, instrumen yang hendak dipakai yakni lembar pengamatan dengan fungsi menelaah HOTS yang diisi sesuai KKO Taksonomi Bloom yang digunakan dalam soal. Berikut adalah format lembar pengamatan yang digunakan.

Tabel 1. Lembar Telaah HOTS

No Butir Soal	KD	Materi	Indikator	KKO Taksonomi Bloom					
				C1	C2	C3	C4	C5	C6
1									
2									
3									
4									
Dst									

Berdasarkan table diatas, lembar telaah HOTS tersebut, setiap kolom diisi sesuai dengan instruksi yang tertera. Contohnya pada kolom KD diisi sesuai Kompetensi Dasar Mata Pelajaran pada soal tersebut. Untuk kolom C1-C6, diisi sesuai KKO Taksonomi Bloom yang digunakan dalam soal. Contohnya jika intruksi dalam soal tersebut menyebutkan salah satu contoh, maka diisi kolom C1. Data kuantitatif, teknik analisis data pada setiap soal pilihan ganda dengan menganalisa kesesuaian pengukuran (validitas), konsistensi hasil pengujian (reliabilitas), tingkat kesulitan soal, kemampuan membedakan tingkat kemampuan siswa (daya pembeda), serta keefektifan pilihan pengecoh dalam menjangir jawaban yang keliru secara fungsional. Kelima parameter tadi dianalisis melalui program Anates. Anates atau program tersebut diaplikasikan ketika menganalisa item soal uraian maupun bentuk lainnya. Sedangkan untuk data kualitatif mengadposi tahapan analisis dari Miles dan Huberman diantaranya reduksi data, penyajian data dan verifikasi (Sugiyono, 2017).

Penarikan kesimpulan dalam penelitian ini dilakukan secara bertahap berdasarkan hasil analisis tiap aspek kualitas butir soal. Pertama, peneliti meninjau hasil validitas untuk mengetahui jumlah soal yang memenuhi kriteria valid dan tidak valid. Selanjutnya, reliabilitas tes dianalisis untuk menentukan tingkat konsistensi hasil pengukuran. Analisis tingkat kesulitan dilakukan dengan mengelompokkan butir soal ke dalam kategori mudah, sedang, dan sulit, kemudian dilihat apakah proporsinya sudah seimbang. Daya pembeda setiap butir soal juga diidentifikasi untuk mengetahui sejauh mana soal mampu membedakan siswa berkemampuan tinggi dan rendah. Selain itu, fungsi pengecoh dievaluasi dengan melihat persentase peserta yang memilih opsi jawaban salah. Setelah analisis tiap aspek selesai, peneliti menginterpretasikan hasil secara menyeluruh untuk menilai kualitas keseluruhan tes. Pada tahap ini, butir soal diklasifikasikan menjadi tiga kelompok meliputi soal yang dipertahankan, soal yang perlu direvisi, dan soal yang dibuang. Hasil tersebut kemudian dibandingkan dengan kriteria standar kualitas butir soal

sekaligus mempertimbangkan karakteristik soal berbasis HOTS yang menuntut keterampilan berpikir tingkat tinggi. Tahap akhir penarikan kesimpulan dilakukan dengan merumuskan pernyataan yang menjawab rumusan masalah penelitian, termasuk persentase soal berkualitas baik, soal yang harus diperbaiki, dan soal yang harus dihapus.

Hasil

Penelaahan ini ditujukan untuk mengevaluasi mutu/kualitas dari tiap soal Pendidikan Kewarganegaraan (PKn) yang diujikan saat PTS kepada siswa kelas V SDN Cipocok Jaya 4 Tahun Ajaran 2018/2019. Penilaian dilakukan dengan memperhatikan berbagai aspek penting dalam analisis butir soal, antara lain kesesuaian pengukuran (validitas), konsistensi hasil pengujian (reliabilitas), tingkat kesulitan soal, kemampuan membedakan tingkat kemampuan siswa (daya pembeda), serta keefektifan pilihan pengecoh dalam menjaring jawaban yang keliru secara fungsional serta keterkaitan soal dengan indikator HOTS. Analisis dilakukan terhadap 25 butir soal yang digunakan dalam ujian, dengan melibatkan 26 siswa sebagai responden. Instrumen yang dipakai ketika meliti meliputi naskah soal, kunci jawaban, LJK, beserta lembar telaah HOTS. Dokumentasi dipilih sebagai metode pengumpulan data untuk mendapatkan data soal secara langsung, hasil pekerjaan siswa beserta kunci jawabannya. Selanjutnya, jawaban siswa dianalisa menggunakan perangkat lunak Anates versi 4 untuk menilai berkualitas tidaknya masing-masing butir soal merujuk kepada kriteria yang telah ditetapkan. Hasil dari analisis tersebut memberikan gambaran mengenai kualitas soal dari berbagai aspek yang telah disebutkan.

Validitas

Validitas merupakan indikator yang menunjukkan karakteristik dari alat yang dipakai beserta kemampuannya saat melakukan pengukuran secara akurat dan sah. Dengan kata lain, ketepatan dan kesahihan hasil pengukuran sangat bergantung pada kemampuan instrumen tersebut dalam menangkap dan merepresentasikan karakteristik yang ingin diungkap. Yang peneliti lakukan saat pengujian validitas yakni perhitungan koefisien korelasi, dibantu oleh aplikasi Anates versi 4. Output dari nilai korelasi tiap butir soal kemudian dibandingkan dengan nilai r_{tabel} dengan derajat kesalahan 5%. Karena saat penelitian dilakukan subjek sejumlah 26 siswa, ditentukan r_{tabel} yakni 0,381. Suatu butir soal diklasifikasikan valid apabila hasil dari r_{hitung} melebihi r_{tabel} ($r_{\text{hitung}} > r_{\text{tabel}}$), sedangkan jika sebaliknya, maka soal tersebut tidak valid. Dari hasil analisa tersebut, diperoleh bahwa sebanyak enam belas butir (64%) tergolong valid, sedangkan sembilan butir (36%) dinyatakan tidak valid.

Tabel 2. Hasil Uji Validitas

Indeks Validitas	No Butir Soal	Jumlah	Presentase
$\geq 0,381$ Valid	1, 3, 4, 5, 6, 7, 12, 14, 15, 16, 17, 18, 20, 21, 23, 25	16	64%
$< 0,381$ Tidak Valid	2, 8, 9, 10, 11, 13, 19, 22, 24	9	36%
Total		25	100%

Hal ini mengindikasikan jika temuan penelitian kurang baik dikarenakan hasil perhitungan secara keseluruhan untuk kesesuaian pengukuran soal sebesar 64%. Soal yang telah diujikan kurang baik untuk mengukur kemampuan siswa dalam mata pelajaran tersebut.

Reliabilitas

Tingkat konsistensi hasil pengujian (reliabilitas) soal pilihan ganda yang dianalisa menggunakan bantuan perangkat lunak Anates Versi 4. Diperoleh melalui tafsiran koefisien reliabilitas, selanjutnya dihubungkan dengan standar kriteria reliabilitas. Output dari analisis menginterpretasikan jika soal-soal tersebut mempunyai angka reliabilitas yang ditunjukkan

dengan angka 0,84. Berpedoman pada kategori indeks reliabilitas, angka ini berada pada rentang 0,80 hingga 1,00, yang mengindikasikan bahwa reliabilitas soal masuk dalam kategorisasi sangat tinggi. Dengan demikian, soal pilihan ganda PKn kelas V di SDN Cipocok Jaya 4 memiliki konsistensi pengukuran yang sangat baik.

Daya Pembeda

Temuan dari analisa pada 25 item yang dipakai saat PTS, tidak ditemukan soal yang termasuk kategorisasi sangat buruk. Sebanyak tujuh butir (28%) tergolong berkategori buruk, enam soal (24%) berada pada kategorisasi cukup, sepuluh soal (40%) termasuk dikategorikan baik, serta dua soal (8%) dikategorikan soal berdaya pembeda sangat baik. Dalam hal daya pembeda, tiap item bila dikategorikan cukup, baik, ataupun sangat baik dianggap memiliki daya pembeda yang mempunyai. Sebaliknya, tiap item yang dikategorisasikan buruk maupun sangat buruk diklasifikasikan berdaya pembeda yang jelek atau tidak mempunyai.

Tabel 3. Hasil Uji Daya Pembeda

Kategori	No Butir Soal	Jumlah	Presentase
Sangat Buruk	-	0	0%
Buruk	3, 4, 8, 9, 10, 13, 19	7	28%
Cukup	1, 2, 6, 11, 22, 24	6	24%
Baik	5, 7, 12, 16, 17, 18, 20, 21, 23, 25	10	40%
Sangat Baik	14, 15	2	8%
Total		25	100%

Berdasarkan table diatas dapat dilihat pada tabel 3, bisa diinterpretasikan yakni 75% dari keseluruhan item soal berdaya pembeda yang tergolong mempunyai, sementara 25% sisanya menunjukkan daya pembeda yang jelek atau kurang mempunyai. Dominasi item dengan daya pembeda rendah berdampak pada ketidakmampuan soal untuk secara efektif membedakan tingkat kemampuan peserta didik. Kondisi ini dapat menyebabkan siswa dengan kemampuan rendah menjawab benar secara kebetulan, sementara siswa dengan kemampuan tinggi justru menjawab salah, sehingga mengurangi fungsi diskriminatif dari soal tersebut.

Tingkat Kesukaran

Hasil telaah yang dilaksanakan terhadap 25 butir soal menggunakan aplikasi Anates Versi 4, diperoleh distribusi tingkat kesukaran sebagai berikut: sebanyak satu soal (4%) tergolong sukar, lima soal (20%) berada pada kategori sedang, lalu mayoritas soal, yaitu sembilan belas soal (76%) termasuk berkategori mudah.

Tabel 4. Hasil Uji Tingkat Kesukaran

Kategori	No Butir Soal	Jumlah	Presentase
Sukar (0,00-0,30)	8	1	4%
Sedang (0,31-0,70)	2, 9, 15, 18, 24	5	20%
Mudah (0,71-1,00)	1, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14, 16, 17, 19, 20, 21, 22, 23, 25	19	76%
Total		25	100%

Berdasarkan table diatas, terlihat pada tabel 4, bisa diinterpretasikan tingkat kesukaran pada item soal yang diteliti ternyata tergolong tidak proporsional atau tidak memenuhi komposisi yang sesuai standar. Sebagian besar item ini berada dalam kategori mudah, sementara jumlah soal dengan tingkat kesukaran sedang dan sukar belum mencukupi. Padahal, menurut standar proporsi ideal tingkat kesukaran soal adalah 25% soal dengan tingkat kesukaran mudah, 50% soal dengan tingkat kesukaran sedang, dan 25% soal dengan tingkat kesukaran sukar (2,5:5:2,5). Ketidakeimbangan tersebut menggambarkan bahwa kualitas distribusi tingkat kesukaran soal perlu ditingkatkan agar dapat mencerminkan kemampuan peserta didik secara lebih akurat.

Efektivitas Pengecoh

Merujuk perhitungan dan interpretasi output Anates, ditemukan sebanyak empat soal (16%) berpengecoh dengan efektivitas sangat baik, tiga soal (12%) tergolong baik, 9 butir soal (36%) dinilai cukup efektif, dan sembilan soal lainnya (36%) menunjukkan efektivitas pengecoh yang rendah atau tidak efektif. Pengecoh yang kurang efektif umumnya bersifat terlalu mencolok atau tidak relevan secara konseptual, sehingga gagal menarik perhatian peserta tes yang belum memahami materi. Hal ini menyebabkan fungsi pengecoh sebagai alat untuk mengukur pemahaman siswa menjadi tidak optimal.

Tabel 5. Hasil Uji Pengecoh Soal

Kategori	No Butir Soal	Jumlah	Presentase
Sangat Baik	14, 15, 16, 21	4	16%
Baik	1, 8, 22	3	12%
Cukup	2, 5, 6, 9, 12, 17, 18, 24, 25	9	36%
Kurang Baik	3, 4, 7, 10, 11, 13, 19, 20, 23	9	36%
Total		25	100%

Telaah HOTS

Kemampuan berpikir tingkat tinggi atau HOTS mengacu pada aktivitas kognitif yang melampaui sekadar mengingat atau mengulang informasi. HOTS mencakup kemampuan untuk membangun, memahami, serta mengubah apa yang diketahui dan dialami ataupun dimiliki guna mengambil keputusan dan menyelesaikan masalah dalam konteks baru. Kemampuan ini sangat relevan dalam kehidupan nyata individu dan selalu bertautan dengan proses belajar yang bermakna. Berdasarkan hasil telaah terhadap aspek HOTS dalam butir soal pada PTS PKn, dari keseluruhan 25 soal yang dianalisa berdasarkan kategori dalam Taksonomi Bloom (KKO), ditemukan bahwa 11 soal tergolong dalam kategori C1 (mengingat), 4 soal dalam C2 (memahami), 3 soal dalam C3 (menerapkan), 1 soal dalam C4 (menganalisis), 6 soal dalam C5 (mengevaluasi), dan tidak terdapat soal dalam kategori C6 (mencipta).

Analisis Kualitas Butir Soal

Kesesuaian pengukuran (validitas), konsistensi hasil pengujian (reliabilitas), tingkat kesulitan soal, kemampuan membedakan tingkat kemampuan siswa (daya pembeda), serta keefektifan pilihan pengecoh dalam menjaring jawaban yang keliru secara fungsional dilakukan secara terpadu guna memperoleh gambaran menyeluruh mengenai kualitas butir soal. Penilaian kualitas keseluruhan butir soal ini mengacu kepada kategorisasi yang umum dipakai dalam skala Likert.

Tabel 6. Rekapitulasi Kualitas Butir

Kategori	No Butir Soal	Jumlah	Presentase
Sangat Baik	1, 5, 6, 14, 15, 16, 17, 18, 21, 25	10	40%
Baik	2, 7, 8, 9, 12, 20, 22, 23, 24	9	36%
Sedang	3, 4, 10, 11	4	16%
Tidak Baik	13, 19	2	8%
Sangat Tidak Baik	-	0	0%
Total		25	100%

Rekapitulasi analisis dalam tabel 6, ditemukan bahwa sepuluh soal memiliki kualitas sangat baik, sembilan soal tergolong berkualitas baik, empat soal berada pada kategori kualitas sedang, serta dua soal diklasifikasikan sebagai tidak baik, tidak ditemukan soal yang masuk dalam kategori sangat tidak baik. Sebagai tindak lanjut, semua item yang dikategorikan sangat baik disarankan untuk diinput ke sistem bank soal agar nantinya bisa dipakai kembali dalam

penyusunan alat evaluasi di masa mendatang. Sementara itu, item yang berkategori baik dan sedang harus ditinjau atau direvisi berdasarkan kelemahan atau indikator ketidaksesuaian yang teridentifikasi. Adapun soal yang berkualitas rendah atau tidak baik sebaiknya dieliminasi karena tidak memenuhi standar kelayakan pengukuran.

Pembahasan

Kualitas Butir Soal

Validitas hasil dari tes dipengaruhi oleh tiga elemen meliputi instrumen tes, faktor administrasi dan penilaian, dan faktor yang diperoleh dari respons siswa (Suryadi, 2020). Hasil analisa terhadap soal PTS mata pelajaran PKn kelas V Tahun Ajaran 2018–2019 menyimpulkan bahwa sebanyak sepuluh soal layak untuk diarsipkan dalam bank soal, dikarenakan telah memenuhi seluruh kriteria penilaian yaitu kesesuaian pengukuran (validitas), konsistensi hasil pengujian (reliabilitas), tingkat kesulitan soal, kemampuan membedakan tingkat kemampuan siswa (daya pembeda), serta keefektifan pilihan pengecoh dalam menjaring jawaban yang keliru. Sementara itu, ada sembilan soal yang tergolong bermutu baik, namun memerlukan perbaikan karena hanya mencapai 3 dari 4 kriteria tersebut. Sejumlah empat soal dikategorikan memiliki mutu sedang serta perlu direvisi lebih lanjut, dengan pertimbangan empat soal tersebut hanya mencapai 2 kriteria. Adapun dua soal yang masuk pada kategorisasi tidak baik disarankan untuk jangan digunakan kembali, mengingat rendahnya kualitas berdasarkan keempat aspek penilaian tersebut. Butir berkualitas tinggi ke bank soal supaya bisa dipakai kembali sebagai instrumen ketika dibutuhkan/mendatang merupakan salah satu contoh tindakan lanjutan yang dapat dilakukan. Soal-soal dengan kualitas baik dan sedang alangkah baiknya diperbarui merujuk pada parameter yang tidak tercapai. Sebaiknya buang item yang didapati berkualitas buruk dan sangat buruk (Muluki, 2020).

Berdasarkan temuan tersebut, memberi gambaran peneliti bahwa secara umum soal-soal pada PTS mata pelajaran PKn kelas V Tahun Ajaran 2018–2019 telah mampu berfungsi sesuai dengan tujuan evaluasi pembelajaran. Adapun kegagalan pada beberapa item soal ketika dipakai tidak berfungsi optimal, penyebabnya ialah ketidakterpenuhinya satu atau lebih parameter kualitas soal yang menjadi acuan dalam analisis. Temuan akan tingkat kesukaran adalah temuan dengan hasil terburuk dari telaah ini dengan kontribusi terbanyak. Soal PTS ini dinilai didominasi soal yang terlalu mudah. Dampak negatif dari soal terlalu mudah maupun terlalu rumit merupakan alat evaluasi yang baik. Siswa tidak akan termotivasi untuk bekerja keras menjawab soal yang terlalu sederhana. Di sisi lain, jika pertanyaannya terlalu menantang, anak-anak akan enggan berpikir dan menjawab karena mereka tidak memahaminya. Maka dari itu tingkat kesukaran perlu diperhatikan ketika pembuatan butir soal (Dia et al, 2022). Kemudian kemampuan membedakan tingkat kemampuan siswa (daya pembeda) dan validitas soal. Soal ini terindikasi buruk karena tidak bisa membedakan tinggi rendahnya kemampuan siswa. Kemampuan suatu pertanyaan untuk mengklasifikasikan tinggi rendahnya kemampuan seorang siswa dimaknai daya pembeda. Sebagai patokan seberapa baik suatu item dapat mengklasifikasikan siswa berkemampuan akademik rendah maupun tinggi, penelaahan daya pembeda dilakukan (Arwinda et al, 2024). Untuk memastikan karakteristik dan ketepatannya dalam menilai kemampuan siswa, soal-soal dengan kemampuan membedakan tingkat kemampuan siswa (daya pembeda) sedang maupun buruk nantinya harus direvisi. Sementara itu, soal-soal yang didapati berdaya pembeda baik dan memadai bisa disimpan ataupun digunakan lagi ketika asesmen selanjutnya (Dianova et al, 2024).

Selanjutnya kegagalan validitas artinya item tes yang dipakai dan tidak valid mengindikasikan arah skor total dari siswa tidak sesuai atau tidak sejajar. Lainnya yakni pengecoh pada butir soal juga belum berfungsi dengan baik. Butir soal yang dianggap valid menunjukkan bahwa soal-soal tersebut cukup mewakili pemahaman siswa terhadap tujuan pembelajaran yang ditetapkan. Upaya guru untuk menciptakan perangkat penilaian yang melengkapi kurikulum dan sumber belajar yang digunakan juga tercermin dalam validitas butir soal yang tinggi (Ahsan et al, 2023). Validitas soal teramat penting karena validitas sebuah instrumen butir soal, serta untuk memastikan bahwa hasil pengujian tidak dipengaruhi oleh faktor-faktor lain, uji validitas harus dilakukan sebelum instrumen tersebut digunakan dalam pengujian. Pengujian yang menggunakan butir tes yang tidak valid ataupun tidak reliabel berpotensi berdampak signifikan terhadap hasilnya (Fatayah et al, 2022). Dengan demikian, soal-soal Penilaian Tengah Semester mata pelajaran PKn kelas V Tahun Ajaran 2018–2019 secara keseluruhan dapat dikategorikan mempunyai kualitas yang baik. Simpulan ini berdasar pada temuan bahwa sebanyak 19 butir soal layak disimpan dalam bank soal karena tergolong sangat baik dan baik, sementara 4 butir soal memerlukan revisi, dan 2 butir soal lainnya dinyatakan tidak layak digunakan kembali karena tidak memenuhi standar kualitas. Meskipun berkualitas baik namun pada segi tingkat kesukaran hendaknya dilakukan pengecekan kembali, karena perbandingan tingkat kesukaran masih timbang atau tidak proporsional.

Butir Soal Berorientasi HOTS

Berdasarkan dari hasil telaah HOTS bersumber tiap soal pilihan ganda Pendidikan Kewarganegaraan bisa ditarik kesimpulan yakni dari akumulasi 25 butir soal PG yang dipakai saat PTS, dilihat dari KKO yang digunakan dalam soal (taksonomi Bloom), 11 butir soal dalam kategori C1, 4 butir soal dalam kategori C2, 3 butir soal dalam kategori C3, 1 butir soal dalam kategori C4, 6 butir soal dalam kategori C5 dan tidak ada butir soal dalam kategori C6. Bisa disimpulkan bahwa dalam butir soal PTS Mata Pelajaran PKn kelas V Tahun Ajaran 2018-2019, hanya terdapat 7 soal yang berorientasi HOTS. Hal ini dinilai wajar karena pada satuan pendidikan SD/MI kemungkinan soal tipe HOTS sedikit dibutuhkan untuk melatih siswa berpikir kritis, dan analitis.

Institusi pendidikan tinggi akan semakin dituntut untuk mengembangkan critical thinking serta analitis. Akibatnya, kebutuhan dari pembuatan pertanyaan HOTS akan semakin meningkat di jenjang pendidikan tinggi. Tingkat berpikir yang harus dicapai meningkat seiring dengan jenjang pendidikan. Sebaliknya, kebutuhan akan pertanyaan HOTS dapat menurun seiring dengan menurunnya jenjang pendidikan (Intan et al, 2020). Namun dari banyaknya referensi ilmiah yang ditemukan, soal berorientasi HOTS memiliki segudang manfaat terkhusus bagi siswa SD kelas tinggi. Beberapa keuntungan atau manfaat menggunakan pertanyaan/soal HOTS dalam pembelajaran di sekolah dasar. Meningkatkan kemampuan berpikir kritis siswa merupakan salah satu keuntungan utama, karena mereka memperoleh kemampuan untuk bertanya, mengevaluasi, dan mencari solusi dari berbagai sudut pandang. Selain itu, dengan melatih anak-anak untuk menemukan jawaban yang unik dan kreatif, pertanyaan HOTS mendorong daya cipta anak. Hal ini sangat membantu perkembangan karakter anak, karena mendorong peningkatan kemandirian dan pemikiran kreatif (Putri et al, 2025).

Siswa dapat meningkatkan prestasi akademik, menumbuhkan sikap positif, dan meningkatkan antusiasme belajar dengan HOTS. Oleh karena itu, mengintegrasikan HOTS ke dalam kelas sangat penting sebagai modal awal siswa. Berbekal kemampuan critical thinking, creativity, dan problem solving akan dibutuhkan di dunia yang semakin kompleks saat ini (Zebua,

2024). Siswa didorong untuk menggunakan keterampilan penalaran daripada hanya menghafal, menyalin, atau mengingat melalui pertanyaan berbasis HOTS (Latifah et al, 2023).

Dari banyaknya manfaat soal berorientasi HOTS, faktanya masih ditemukan banyak kendala dalam implementasinya. Salah satunya guru dan siswa belum terbiasa dengan pembelajaran berbasis HOTS terkhusus butir soal yang berorientasi karena baru diterapkan pada tahun 2019. Akibatnya, tidak ada kelompok yang sepenuhnya memahami ide atau penerapan HOTS, meskipun membutuhkan pemikiran tingkat tinggi (Handayani et al, 2021). Selain itu, ketika pengerjaan soal HOTS, siswa kesulitan memahami instruksi pertanyaan, arti gambar, pilihan jawaban, arti kata, mencocokkan daftar kata dalam pertanyaan dengan pilihan jawaban, menawarkan solusi atau saran, menentukan pilihan jawaban, dan berfokus saat mengerjakan pertanyaan (Hariani et al, 2023).

Kesimpulan

Hasil analisis terhadap soal Penilaian Tengah Semester mata pelajaran PKn kelas V Tahun Ajaran 2018–2019 menunjukkan bahwa hanya terdapat 7 butir soal yang berorientasi pada HOTS. Kedua, berdasarkan analisis kualitas soal dari aspek validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh, diperoleh bahwa 10 butir soal (40%) berada dalam kategori sangat baik, 9 butir soal (36%) termasuk kategori baik, 4 butir soal (16%) tergolong sedang, dan 2 butir soal (8%) dinilai tidak baik. Tidak ditemukan butir soal yang masuk dalam kategori sangat tidak baik (0%). Penelitian ini menunjukkan bahwa analisis butir soal dengan Program Anates dapat membantu guru mengevaluasi validitas, reliabilitas, tingkat kesulitan, daya pembeda, dan fungsi pengecoh soal secara objektif.

Implikasinya dapat dimanfaatkan untuk memperbaiki kualitas instrumen penilaian berbasis HOTS, membangun bank soal yang terstandar, serta menjadi bahan pelatihan guru dalam menyusun soal yang selaras dengan tujuan kurikulum. Keterbatasan pada penelitian ini yakni Anates hanya menghitung validitas item tanpa melakukan pengujian terhadap validitas isi. Lalu belum ditemukannya standar yang digunakan dalam menentukan kriteria kualitas butir soal. Oleh karena itu, peneliti mengadaptasi kriteria tersebut dari Skala Likert. Peneliti merekomendasikan agar jumlah soal yang berorientasi pada HOTS diperbanyak atau ditambah jumlahnya, karena jenis soal ini tidak hanya mampu mendorong peningkatan kemampuan berpikir tingkat tinggi siswa, tetapi juga melatih keterampilan dalam mengkonstruksi, memahami, serta mentransformasi pengetahuan dan pengalaman yang telah dimiliki. Kemampuan tersebut sangat penting dalam pengambilan keputusan dan pemecahan masalah pada situasi baru yang erat kaitannya dengan kehidupan sehari-hari. Disamping itu, guru perlu memberikan perhatian lebih terhadap indikator-indikator kualitas soal, agar soal yang disusun memiliki mutu yang optimal. Uji coba soal sebelum pelaksanaan ujian utama juga disarankan sebagai langkah preventif untuk menjamin validitas dan efektivitas instrumen penilaian. Temuan ini juga membuka peluang penelitian lanjutan pada mata pelajaran lain untuk memperkuat praktik asesmen yang mendorong kemampuan berpikir tingkat tinggi siswa.

Acknowledgment

-

Daftar Pustaka

Ahsan, N. H., & Hartoyo, A. (2023). Strategi Pemecahan Masalah Matematika Siswa Sekolah Menengah Pertama. *Jurnal Pendidikan Matematika Indonesia*, 8(2), 137–147.

- Alifah, F., Septianingsih, R., Rohmah, A. T. N., Arif, A., & Hakim, L. (2025). Analisis Butir Soal Melalui Aplikasi Anates Oleh Siswa Smk. *Jurnal Penelitian Pendidikan Indonesia*, 2(2), 235–242.
- Arwinda, P., & Prasetyo, S. (2024). Analisis Akurasi Soal Pilihan Ganda Materi Rangka Manusia pada Ulangan Harian Siswa di Sekolah Dasar. *Nusantara: Jurnal Pendidikan Indonesia*, 4(3), 806-821. <https://doi.org/10.14421/njpi.2024.v4i3-14>
- Creswell, J. W., & Creswell, J. D. (2018). Mixed Methods Procedures. In *Research Defign: Qualitative, Quantitative, and Mixed M ethods Approaches*.
- Dia, E. E., & Syah, A. S. (2022, September). Tingkat Kesukaran Soal Penilaian Tengah Semester Mata Pelajaran Bahasa Indonesia Di Madrasah. In *Prosiding Conference on Research and Community Services* (Vol. 4, No. 1, pp. 14-23).
- Dianova, F. R., & Anwar, N. (2024). Analisis Butir Uji Validitas, Reliabilitas, Tingkat Kesukaran, dan Daya Pembeda Soal Sumatif Bahasa Arab SD Islam. *Jurnal Bahasa Daerah Indonesia*, 1(3), 13-13. <https://doi.org/10.47134/jbdi.v1i3.2863>
- Fatayah, F., Yuliana, I. F., & Mu'fidah, L. (2022). Analisis Validitas Dan Reliabilitas Dalam Mendukung Ketuntasan Belajar Model STEM. *Jurnal Buana Pendidikan*, 18(1), 49-60.
- Handayani, F., & Syukur, M. (2021). Implementasi pembelajaran higher order thinking skill (hots) di ma negeri 1 watansoppeng. *Pinisi Journal of Sociology Education Review*, 1(2), 127-135.
- Hariani, F., Tahir, M., & Oktaviyanti, I. (2023). Analisis Kesulitan Siswa dalam Menyelesaikan Soal HOTS Pada Muatan IPS Kelas V di SDN 12 Ampenan. *Jurnal Ilmiah Profesi Pendidikan*, 8(1), 119-124. <https://doi.org/10.29303/jipp.v8i1.1096>
- Intan, F. M., Kuntarto, E., & Alirmansyah, A. (2020). Kemampuan Siswa dalam Mengerjakan Soal HOTS (Higher Order Thinking Skills) pada Pembelajaran Matematika di Kelas V Sekolah Dasar. *JPDI (Jurnal Pendidikan Dasar Indonesia)*, 5(1), 6-10. <https://doi.org/10.26737/jpdi.v5i1.1666>
- Jatmoko, D., Basuki, B., Pambudi, A., Setiyono, K. A., & Nugroho, B. (2025). Analisis Butir Soal Teknik Dasar Otomotif Berbasis Diagnostik Kognitif dengan Anates Versi 4.02. *Jurnal Riset Rumpun Ilmu Pendidikan*, 4(1), 461-471. <https://doi.org/10.55606/jurripen.v4i1.4812>
- Khoerunnisa, S. F. Z., Chotimah, K. A., Dewi, L. P., Aliandra, A. L., & Putri, H. E. (2025). Pemanfaatan ANATES untuk Menganalisis Validitas dan Reliabilitas Butir Soal Uraian IPAS Materi Perkembangbiakan Tumbuhan. *Jurnal Kolaboratif Sains*, 8(1), 232-242. <https://doi.org/10.56338/jks.v8i1.6688>
- Kurniawan, A., Febrianti, A. N., Hardianti, T., Ichsan, Desy, Risan, R., Sari, D. M. M., Sitopu, J. W., Dewi, R. S., Sianipar, D., Fitriyah, L. A., Zulkarnain, Jalal, N. M., Hasriani, & Hasyim, F. (2022). Evaluasi Pembelajaran. In *Global Eksekutif Teknologi*. Global Eksekutif Teknologi.
- Latifah, A., Husaini, F., & Nisa, A. K. (2023). Pengembangan Instrumen Penilaian Berbasis Hots. *Didaktik: Jurnal Ilmiah Pgsd Stkip Subang*, 9(2), 4486-4496. <https://doi.org/10.36989/didaktik.v9i2.1057>

- Muluki, A. (2020). Analisis kualitas butir tes semester ganjil mata pelajaran IPA Kelas IV MI Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar*, 4(1), 86-96. <https://doi.org/10.23887/jisd.v4i1.23335>
- Mania, S., & Rasyid, M. N. A. (2023). Analisis kualitas butir soal ujian akhir semester mata pelajaran akidah akhlak kelas V MI Radhiatul Adawiyah Makassar. *Inspiratif Pendidikan*, 12(1), 224-240. <https://doi.org/10.24252/ip.v12i1.32040>
- Putri, C. M., Adrias, A., & Syam, S. S. (2025). Mengasah Kecerdasan Anak dengan Soal HOTS: Strategi Efektif untuk Pembelajaran di SD. *Morfologi: Jurnal Ilmu Pendidikan, Bahasa, Sastra dan Budaya*, 3(2), 287-296. <https://doi.org/https://doi.org/10.61132/morfologi.v3i2.1581>
- Rizky, A., Amsul, A. M. Z., Rohmiati, C., Nuraini, F., Fitratama, F. Y., & Fitriawanati, M. (2019). Penerapan Aplikasi Anates Dengan Bentuk Soal Pilihan Ganda di SD Muhammadiyah Gendeng. *Jurnal Elementary*, 2(1), 43. <https://doi.org/10.31764/elementary.v1i2.549>
- Santosa, S., & Badawi, J. A. (2022). Analisis Butir Soal Pilihan Ganda Tema Pertumbuhan dan Perkembangan Makhluk Hidup Kelas III Madrasah Ibtidaiyah. *Jurnal Basicedu*, 6(2), 1678-1686. <https://doi.org/https://doi.org/10.31004/basicedu.v6i2.2206>
- Sugiyono. (2017). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta.
- Suryadi, A. (2020). *Evaluasi Pembelajaran*. CV Jejak.
- Vitalocca, D., & Mardiana, A. N. (2019). Pelatihan analisis butir tes dengan Program ITEMAN pada guru-guru SMK di Kabupaten Sidrap. In *Prosiding Seminar Nasional*. (pp. 120-123). Makassar: Universitas Negeri Makassar.
- Yazid, S. T. M., & Lisnawati, S. (2022). Pengaruh Evaluasi Pembelajaran Daring Terhadap Hasil Belajar Mata Pelajaran Pai Di Sman 5 Bogor. *Inspiratif Pendidikan*, 11(2), 530-543. <https://doi.org/10.24252/ip.v11i2.36016>
- Zebua, N. (2024). Studi literatur: peranan higher order thinking skills dalam proses pembelajaran. *Edukasi Elita: Jurnal Inovasi Pendidikan*, 1(2), 92-100. <https://doi.org/10.62383/edukasi.v1i2.110>