



Analisis Butir Soal Pilihan Ganda pada Siswa SMK Menggunakan Anates

Dian Septy Ikawati^{1✉}, Eka Wahyu Agustin², Luqman Hakim³, Vivi Pratiwi⁴

Universitas Negeri Surabaya, Indonesia^{1,2,3,4}

e-mail : dian.23157@mhs.unesa.ac.id¹, ekawahyu.23139@mhs.unesa.ac.id², luqmanhakim@unesa.ac.id³, vivipratiwi@unesa.ac.id⁴

Abstrak

Penelitian ini dilaksanakan karena masih terdapat kebutuhan untuk meningkatkan kualitas instrumen tes pilihan ganda dalam mengevaluasi kompetensi siswa di bidang akuntansi. Penelitian ini bertujuan menganalisis kualitas instrumen tes pilihan ganda pada materi profil pekerjaan dan peluang usaha di bidang akuntansi. Studi dilaksanakan dengan pendekatan kuantitatif deskriptif menggunakan software ANATES untuk mengevaluasi karakteristik instrumen tes. Analisis komprehensif meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Sampel penelitian terdiri dari 30 siswa program keahlian akuntansi SMK, dengan instrumen penelitian berupa tes pilihan ganda sebanyak 15 butir soal. Pengumpulan data dilakukan melalui tes tertulis dan analisis menggunakan metode kuantitatif untuk menentukan kualitas butir soal. Hasil penelitian menunjukkan reliabilitas tes mencapai 0,76 yang tergolong kategori reliabilitas tinggi. Dari 15 butir soal yang dianalisis, 4 soal (26,67%) dinyatakan layak digunakan tanpa revisi, sedangkan 11 soal (73,33%) memerlukan perbaikan atau revisi untuk meningkatkan kualitas instrumen evaluasi. Temuan penelitian memberikan kontribusi signifikan dalam pengembangan instrumen evaluasi pembelajaran akuntansi di tingkat SMK, dengan implikasi penting bagi perbaikan kualitas tes dan peningkatan proses penilaian hasil belajar.

Kata Kunci: analisis butir soal, software anates, soal pilihan ganda, evaluasi pembelajaran akuntansi

Abstract

This research was conducted due to the continued need to improve the quality of multiple-choice test instruments in evaluating students' competencies in the accounting field. This study aims to analyze the quality of multiple-choice test instruments on job profiles and business opportunity materials in the accounting field. The study was conducted using a descriptive quantitative approach with ANATES software to evaluate test instrument characteristics. A comprehensive analysis encompassed validity, reliability, difficulty level, discriminating power, and distractor effectiveness. The research sample consisted of 30 students from the accounting expertise program at a vocational high school, with research instruments comprising 15 multiple-choice test items. Data collection was performed through written tests and quantitative method analysis to determine item quality. Research results demonstrated a test reliability of 0.76, categorized as high reliability. Of the 15 analyzed test items, 4 items (26.67%) were deemed suitable for use without revision, while 11 items (73.33%) required improvement or revision to enhance evaluation instrument quality. Research findings contribute significantly to developing accounting learning evaluation instruments at the vocational high school level, with important implications for improving test quality and the learning assessment process.

Keywords: item analysis, anates software, multiple-choice test items, accounting learning evaluation

Copyright (c) 2024 Dian Septy Ikawati, Eka Wahyu Agustin, Luqman Hakim, Vivi Pratiwi

✉ Corresponding author :

Email : dian.23157@mhs.unesa.ac.id

DOI : <https://doi.org/10.31004/edukatif.v6i6.7780>

ISSN 2656-8063 (Media Cetak)

ISSN 2656-8071 (Media Online)

PENDAHULUAN

Di era globalisasi saat ini, sumber daya manusia yang berdaya saing sangat dibutuhkan untuk memenuhi tuntutan zaman yang semakin maju. Meskipun penelitian sebelumnya (Dwi Lestari et al., 2024) telah menekankan pentingnya pendidikan dalam pengembangan sumber daya manusia, kajian mendalam tentang kualitas instrumen evaluasi masih memerlukan eksplorasi lebih lanjut. Sebagai contoh, penelitian oleh (Derri Adi Fernando et al., 2023) menunjukkan bahwa instrumen evaluasi yang valid dan reliabel sangat berpengaruh terhadap efektivitas pembelajaran, sementara penelitian lain oleh (Yuniaria Dwi Prastika, 2021) menyoroti adanya kesenjangan dalam validitas soal tes yang dapat mempengaruhi hasil evaluasi. Oleh karena itu, penelitian ini bertujuan mengisi kesenjangan literatur dengan fokus pada analisis validitas butir soal, yang belum secara komprehensif dikaji dalam studi-studi sebelumnya.

Pendidikan abad 21 mengharuskan pendidik untuk berinovasi dalam berbagai aspek pembelajaran. Sejalan dengan pernyataan (Saraswati et al., 2021) tentang pentingnya Higher Order Thinking Skills (HOTS), penelitian ini akan menelaah lebih lanjut bagaimana validitas instrumen tes berperan dalam mengukur kemampuan berpikir kritis, kreatif, dan pengambilan keputusan. Penelitian oleh (Widarta & Artika, 2021) juga menekankan bahwa evaluasi yang efektif dalam pendidikan abad 21 harus dapat mengukur keterampilan tingkat tinggi, bukan hanya pengetahuan dasar. Kontribusi utama penelitian ini terletak pada pendekatan sistematis dalam mengevaluasi kualitas butir soal, yang dapat memberikan wawasan baru dalam praktik evaluasi pendidikan, khususnya dalam konteks HOTS.

Untuk memastikan bahwa proses pendidikan mencapai tujuan yang diharapkan, evaluasi pembelajaran menjadi langkah penting dalam dunia pendidikan. Melalui evaluasi, pencapaian proses pembelajaran yang telah berlangsung dapat diukur, sekaligus menjadi dasar untuk menentukan langkah perbaikan yang diperlukan dalam pembelajaran di masa mendatang (Leonardus Kaka et al., 2024). Melalui evaluasi, guru dapat berkontribusi dalam meningkatkan kemampuan peserta didik dengan memperbaiki kualitas pembelajaran. Peningkatan kemampuan peserta didik secara langsung berdampak pada kualitas pendidikan secara keseluruhan. Terdapat hubungan yang erat antara kemampuan peserta didik dan kualitas pendidikan; jika kemampuan peserta didik tinggi, maka kualitas pendidikan juga akan baik. Sebaliknya, jika kemampuan peserta didik rendah, maka kualitas pendidikan pun akan buruk (Yolanda Farra Alista & Rozana Aufa Syahzanani, 2023).

Metode tes dan non tes adalah dua teknik evaluasi yang digunakan untuk mengukur pencapaian belajar siswa. Di sisi lain, metode pengambilan tes sering kali diterapkan dalam evaluasi praktik pendidikan dengan tujuan utama menilai dan mengukur hasil belajar siswa, khususnya dalam ranah kognitif yang berkaitan dengan penguasaan materi pembelajaran (Puji Rahayu, 2024). Tes adalah alat evaluasi berupa sejumlah pertanyaan yang harus dijawab dengan benar atau salah, atau sejumlah tugas yang harus diselesaikan dengan tingkat keberhasilan atau kegagalan tertentu (Inayah Dzil Izzati Hartono et al., 2024).

Tes pilihan ganda memiliki keunggulan dibandingkan jenis tes lainnya, terutama dalam hal efisiensi dan kemudahan pelaksanaan. Tes ini memungkinkan pengukuran berbagai aspek kemampuan kognitif, seperti pemahaman, analisis, hingga aplikasi konsep, dalam waktu yang relatif singkat. Selain itu, proses koreksi untuk tes pilihan ganda lebih sederhana karena hasilnya dapat dinilai secara objektif (Lasito & Pujiati, 2024). Namun demikian, kualitas evaluasi sangat bergantung pada mutu butir soal yang menyusun tes tersebut. Butir soal yang kurang baik dapat menghasilkan data yang bias dan tidak mencerminkan kemampuan siswa secara akurat. Oleh karena itu, analisis butir soal menjadi langkah esensial untuk memastikan kualitas soal dan meningkatkan validitas hasil evaluasi.

Analisis butir soal adalah serangkaian aktivitas yang dilakukan untuk mengevaluasi kualitas dari suatu item soal. Item soal yang berkualitas baik dapat digunakan kembali dalam penilaian mendatang, sementara item soal yang kurang baik perlu direvisi, dan item soal yang tidak memenuhi standar harus diganti dengan yang baru (Sholeh Mawardi & Fuady, 2023). Kajian oleh (Anazar & Muhammad Munir, 2024) menjelaskan bahwa analisis butir soal yang komprehensif dapat mengidentifikasi kelemahan pada soal dan memberikan arahan

untuk perbaikan instrumen tes. Analisis butir soal bertujuan untuk menilai apakah butir-butir soal tersebut sudah memenuhi kriteria yang baik atau belum. Dengan melakukan analisis butir soal, guru dapat mengidentifikasi kekurangan dalam soal tersebut dan memperoleh informasi yang diperlukan untuk melakukan perbaikan, sehingga kualitas soal dapat ditingkatkan (Ayu, 2024).

Proses analisis butir soal mencakup pengukuran tingkat kesukaran soal, daya pembeda, validitas, serta kualitas pengecoh (distractor). Tingkat kesukaran memberikan informasi mengenai proporsi siswa yang mampu menjawab soal dengan benar, sedangkan daya pembeda mengukur kemampuan soal untuk membedakan siswa yang berkemampuan tinggi dan rendah. Validitas mengacu pada sejauh mana soal mengukur kompetensi yang dimaksud, sementara pengecoh berfungsi untuk memastikan bahwa setiap jawaban alternatif telah dirancang dengan baik.

Anates adalah perangkat lunak yang dirancang khusus untuk menganalisis tes pilihan ganda dan uraian, yang dikembangkan oleh Drs. Kartono, M.Pd. dan Yudi Wibisono, S.T. pada tahun 2004 (Wahyuni Yusuf, 2024). Anates telah menjadi alat bantu yang sangat efektif untuk mendukung proses analisis butir soal. Software ini dirancang untuk mempermudah pendidik dalam mengevaluasi data hasil tes secara cepat dan akurat. Dengan fitur-fitur yang ditawarkan, seperti analisis tingkat kesukaran, daya pembeda, validitas, reliabilitas, hingga kualitas pengecoh, Anates membantu guru dan lembaga pendidikan mendapatkan informasi mendalam mengenai kualitas setiap butir soal. Hasil yang diperoleh dari Anates berupa laporan terperinci dan visualisasi data, yang memudahkan pendidik dalam membuat keputusan berbasis data untuk perbaikan instrumen tes.

METODE

Penelitian ini dilaksanakan dengan tujuan untuk menganalisis dan mengetahui kelayakan butir-butir soal pada elemen Job Profile/Peluang Pekerjaan dan Usaha di bidang Akuntansi. Penelitian ini dilakukan pada peserta didik kelas X di Sekolah Menengah Kejuruan dengan melibatkan sebanyak 30 peserta didik sebagai subjek penelitian. Objek penelitian yang digunakan adalah 15 butir soal pilihan ganda yang dirancang untuk mengukur pemahaman dan kemampuan kognitif level C4-C6 (analisis, evaluasi, dan kreasi) terhadap materi terkait.

Penelitian ini menerapkan metode deskriptif dengan pendekatan kuantitatif. Pendekatan ini dipilih karena analisis terhadap butir soal dilakukan secara sistematis, terencana, dan terstruktur (Ningrum et al., 2023). Penelitian deskriptif kuantitatif bertujuan untuk mengamati, menganalisis, dan menggambarkan data numerik dari objek yang diteliti sesuai dengan kondisi aslinya, kemudian menarik kesimpulan berdasarkan fenomena yang muncul selama proses penelitian. Dari penjelasan tersebut, dapat disimpulkan bahwa penelitian deskriptif kuantitatif adalah metode penelitian yang bertujuan untuk memperoleh nilai variabel secara independen terkait suatu keadaan secara objektif, yang kemudian dijelaskan menggunakan data berbasis angka (Puspa Hanan et al., 2023).

Prosedur penelitian yang dilakukan dalam penelitian ini mencakup beberapa tahapan, yaitu: (1) penyusunan instrumen berdasarkan kisi-kisi soal yang relevan dengan indikator pencapaian kompetensi, (2) pengumpulan data melalui Google Form untuk mempermudah akses dan efisiensi, (3) deskripsi hasil data yang diperoleh, (4) analisis data menggunakan perangkat lunak Anates versi 4.0.9 for Windows, serta (5) penarikan kesimpulan berdasarkan hasil analisis (Ramadhan et al., 2023). Anates adalah perangkat lunak analisis statistik khusus yang dirancang untuk mengolah data hasil tes, termasuk analisis validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh (Wati et al., 2022). Program ini menghasilkan output berupa tabel dan grafik yang mempermudah interpretasi data. Skor peserta didik diinput ke dalam program anates, kemudian dianalisis berdasarkan langkah-langkah baku dalam evaluasi pendidikan.

Hasil analisis disajikan dalam bentuk persentase, yang kemudian dijelaskan sehingga dapat diketahui soal-soal yang berkualitas baik, soal yang memerlukan revisi, serta soal yang tidak disarankan untuk digunakan

kembali (Mustaqim & Hidayu Sulisti, 2024). Output dari Anates dipaparkan secara deskriptif dengan hasil perhitungan yang terstruktur.

HASIL DAN PEMBAHASAN

1. Validitas

Validitas adalah sejauh mana suatu tes benar-benar mengukur apa yang seharusnya diukur. Sebuah data atau informasi dianggap valid jika sesuai dengan fakta atau keadaan yang sesungguhnya (Hera Apriliana Saputri et al., 2023).

Hasil analisis validitas menunjukkan:

Tabel 1. Hasil Analisis Validitas

Kriteria Validitas	No. Butir Soal	Jumlah	Persentase
Tinggi	1, 9	2	13%
Cukup	3, 4, 8, 15	4	27%
Rendah	2, 5, 6, 10, 11	5	33%
Sangat Rendah	7, 12, 13, 14	4	27%

Dari analisis tersebut dapat disimpulkan bahwa kualitas soal dari segi validitas masih perlu banyak perbaikan karena hanya 6 soal (40%) yang memiliki validitas cukup hingga tinggi (butir 1, 3, 4, 8, 9, dan 15). Sedangkan 9 soal lainnya (60%) memiliki validitas rendah (butir 2, 5, 6, 7, 10, 11, 12, 13, 14) hingga sangat rendah atau bahkan tidak dapat dihitung validitasnya. Soal dengan validitas rendah kemungkinan tidak selaras dengan indikator kompetensi yang ingin diukur, yang bisa disebabkan oleh rumusan soal yang tidak presisi, materi yang jauh dari konteks pembelajaran, atau ketidaksesuaian dengan tingkat kemampuan siswa.

Rendahnya validitas soal dapat berdampak negatif pada proses pembelajaran. Guru tidak akan mendapatkan gambaran akurat tentang kemampuan siswa, umpan balik menjadi kurang efektif, dan strategi pengajaran sulit disesuaikan dengan kebutuhan siswa. (Khasanah et al., 2023) menyatakan bahwa butir soal yang memiliki nilai validitas tinggi secara jelas menunjukkan kemampuan dalam mengukur keterampilan siswa dengan tepat dan akurat. Di sisi lain, butir soal yang memiliki validitas sangat rendah memerlukan tindakan lanjutan, karena soal tersebut dianggap tidak valid. Soal-soal dengan validitas rendah dan sangat rendah perlu direvisi.

Temuan validitas soal mengungkapkan masalah mendasar dalam desain instrumen penilaian. Rendahnya validitas bukan sekadar persoalan teknis, tetapi menunjukkan adanya kesenjangan antara tujuan pendidikan dan praktik evaluasi yang dilakukan. Fakta bahwa hanya 40% soal memiliki validitas cukup hingga tinggi secara jelas mengindikasikan perlunya revisi komprehensif terhadap instrumen tes.

Secara teoretis, validitas rendah dapat disebabkan oleh beberapa faktor kritis yang saling terkait. Pertama, terdapat ketidakselarasan antara rumusan soal dengan indikator kompetensi, yang mengakibatkan instrumen penilaian gagal mengukur aspek yang dimaksud. Kedua, kompleksitas materi yang tidak sesuai dengan tingkat kognitif peserta didik dapat menimbulkan kesulitan dalam memahami dan menjawab soal dengan tepat. Ketiga, abstraksi konsep akuntansi yang tidak terjembatani dengan baik dalam instrumen penilaian turut berkontribusi terhadap rendahnya validitas.

Implikasi dari temuan ini sangat signifikan dan berpotensi memberikan dampak negatif yang komprehensif. Soal dengan validitas rendah tidak sekadar gagal mengukur kompetensi, tetapi berpotensi memberikan gambaran yang menyesatkan tentang kemampuan peserta didik. Hal ini dapat berdampak negatif pada proses umpan balik, perencanaan pedagogis, dan strategi pengembangan kompetensi siswa. Jika soal yang tidak valid tetap digunakan, hasilnya menjadi tidak relevan karena tidak dapat mengukur apa yang seharusnya diukur. Akibatnya, kemampuan peserta didik yang sebenarnya tidak dapat tergambarkan secara akurat, sebagaimana ditegaskan oleh (Fiska et al., 2021), sehingga menghasilkan evaluasi yang tidak valid dan berpotensi merugikan proses pendidikan.

2. Reliabilitas

Reliabilitas tes mengacu pada konsistensi skor yang diperoleh oleh individu yang sama ketika tes yang sama diberikan kembali dalam kondisi yang berbeda. Sebuah tes dianggap reliabel jika hasilnya tetap konsisten meskipun dilakukan pengujian ulang pada waktu yang berbeda (Hendrayadi et al., 2024). Hasil menunjukkan bahwa koefisien reliabilitas tes adalah 0,76, yang termasuk dalam kategori reliabel.

Tabel 2. Hasil Analisis Reliabilitas

Rata-Rata	13,43
Simpang Baku	1,7
Korelasi XY	0,62
Realibilitas Tes	0,76

Penghitungan reliabilitas menunjukkan bahwa instrumen ini dapat diandalkan untuk mengevaluasi kemampuan peserta didik. Koefisien reliabilitas sebesar 0,76 menunjukkan tingkat konsistensi yang cukup tinggi. Hal ini mengindikasikan bahwa skor yang diperoleh peserta didik relatif stabil meskipun dilakukan pengujian berulang dengan kondisi serupa. Koefisien reliabilitas tes sebesar 0,76 menunjukkan tingkat konsistensi yang tinggi, sejalan dengan penelitian (Basri & Turmuzi, 2021), yang menyatakan bahwa koefisien reliabilitas $0,60 \leq r < 0,80$ termasuk kategori tinggi.

Koefisien reliabilitas 0,76 tidak sekadar angka statistik, melainkan cerminan konsistensi instrumen dalam mengukur konstruk pengetahuan. Nilai ini menunjukkan bahwa instrumen cukup handal untuk digunakan, namun masih memiliki ruang untuk peningkatan. Dalam konteks pendidikan akuntansi yang dinamis, reliabilitas menjadi prasyarat penting untuk menciptakan sistem penilaian yang adil dan komprehensif.

Reliabilitas tinggi mengindikasikan bahwa perbedaan skor yang diperoleh peserta didik lebih disebabkan oleh perbedaan aktual dalam kemampuan, bukan karena kelemahan instrumen itu sendiri. Hal ini sangat krusial dalam menciptakan mekanisme evaluasi yang transparan dan objektif.

3. Tingkat Kesukaran

Tingkat kesukaran soal merupakan pengukuran yang digunakan untuk menentukan sejauh mana sebuah soal dapat dianggap sukar atau mudah oleh peserta tes, berdasarkan kemampuan mereka dalam menjawab soal tersebut dengan benar (Ryzca Siti Qomariyah et al., 2022). Hal ini sejalan dengan pendapat (Meyke Joice Tresia Rajagukguk & Dorlan Naibaho, 2023) tingkat kesukaran soal adalah menilai tingkat kesukaran soal untuk menentukan mana yang termasuk mudah, sedang, dan sukar. Tujuannya adalah menciptakan keseimbangan antara soal mudah, sedang, dan sukar, dengan tetap memenuhi kualifikasi dan keandalan soal.

Tingkat kesukaran soal biasanya dinyatakan dalam bentuk indeks, dengan nilai yang berkisar antara 0,00 hingga 1,00. Semakin tinggi nilai indeks tersebut, semakin mudah soal tersebut dianggap. Jika sebuah soal memiliki indeks kesukaran 0,00, ini berarti tidak ada siswa yang berhasil menjawabnya dengan benar. Sebaliknya, jika indeks kesukaran mencapai 1,00, berarti semua siswa menjawabnya dengan benar (Akhmadi, 2021). Perhitungan indeks tingkat kesukaran dilakukan untuk setiap nomor soal. Secara prinsip, tingkat kesukaran suatu soal ditentukan oleh skor rata-rata yang diperoleh peserta didik pada soal tersebut (Dianova & Anwar, 2024).

Soal yang memiliki tingkat kesukaran terlalu tinggi atau terlalu rendah dianggap tidak layak digunakan. Soal yang terlalu sukar dapat membuat peserta didik kehilangan motivasi untuk menyelesaikannya, sementara soal yang terlalu mudah tidak mampu mendorong kemampuan peserta didik dalam berpikir dan memecahkan masalah (Nurhalimah et al., 2022). Hasil analisis butir soal elemen Job Profile/Peluang Pekerjaan & Usaha Dibidang Akuntansi menunjukkan bahwa:

Tabel 3. Hasil Analisis Tingkat Kesukaran

Kriteria Tingkat Kesukaran	No. Butir Soal	Jumlah	Persentase
Sangat Mudah	2, 3, 4, 5, 6, 7, 10, 11, 12, 13, 14	11	73%
Mudah	1, 8, 9, 15	4	27%

Hasil ini mengindikasikan bahwa soal cenderung terlalu mudah dan perlu direvisi untuk meningkatkan variasi tingkat kesukaran. Dominasi soal dengan kategori sangat mudah (73%) mengindikasikan masalah serius dalam desain instrumen. Tingkat kesukaran yang seragam tidak hanya gagal membedakan kemampuan peserta didik, tetapi juga berpotensi menurunkan motivasi belajar. Instrumen penilaian ideal seharusnya mencerminkan keragaman tingkat kognitif, memberikan tantangan yang proporsional bagi peserta didik dengan berbagai kemampuan.

Tingkat kesukaran soal harus dilihat dalam konteks pendidikan yang lebih komprehensif. Soal yang terlalu mudah tidak hanya membuat siswa merasa bosan, tetapi juga dapat menghambat perkembangan kemampuan berpikir kritis dan analitis. Hal ini sangat krusial dalam bidang akuntansi yang membutuhkan kemampuan kompleks dan adaptif.

4. Daya Pembeda

Dalam soal pilihan ganda, setiap opsi jawaban harus memiliki daya pembeda yang jelas, sehingga dapat membedakan tingkat pemahaman dan kemampuan peserta didik dalam memilih jawaban yang benar (Merlin Purwati et al., 2021). Daya pembeda pada butir soal berguna untuk meningkatkan kualitas soal berdasarkan data empirik, serta untuk mengetahui sejauh mana setiap soal dapat membedakan kemampuan siswa, yaitu antara siswa yang memahami materi dan yang belum memahami materi yang diajarkan (Magdalena et al., 2021). Hasil analisis butir soal elemen Job Profile/Peluang Pekerjaan & Usaha Dibidang Akuntansi menunjukkan bahwa:

Tabel 4. Hasil Analisis Daya Pembeda

Kriteria Daya Pembeda Soal	No. Butir Soal	Jumlah	Persentase
Sangat Baik	1	1	7%
Baik	8, 9, 15	3	20%
Cukup	3, 4, 5,	3	20%
Jelek	2, 6, 7, 10, 11, 12, 13, 14	8	53%

Analisis daya pembeda mengungkapkan keterbatasan instrumen dalam membedakan tingkat pemahaman peserta didik. Hanya 27% soal memiliki daya pembeda baik hingga sangat baik, sementara mayoritas (53%) memiliki daya pembeda yang jelek. Kondisi ini menandakan kegagalan instrumen dalam mendeteksi nuansa perbedaan kemampuan kognitif.

Kelemahan daya pembeda soal berarti instrumen tidak mampu menggambarkan dengan baik perbedaan kemampuan siswa. Dalam pendidikan akuntansi yang memerlukan analisis mendalam, tes seharusnya dapat membedakan dengan jelas tingkat pemahaman siswa mulai dari yang paling rendah hingga paling tinggi.

5. Efektivitas Pengecoh

Efektivitas pengecoh adalah kemampuan dalam mengecoh peserta didik yang kurang cakap dalam memilih jawaban yang tepat. Distraktor berfungsi untuk membingungkan peserta didik jika mereka tidak memilih jawaban yang benar (Sofiani Putri Radja et al., 2023). Analisis efektivitas pengecoh (*Distractor*) adalah penilaian terhadap jawaban pengecoh yang hampir mirip dengan jawaban yang benar, sehingga sering kali dapat menjebak siswa dalam memilih jawaban yang tepat pada soal tes yang diberikan (Firdaus & Sri Rahayu, 2020). Hasil analisis pengecoh menunjukkan bahwa sebagian besar pengecoh belum berfungsi optimal, dengan rincian:

Tabel 5. Hasil Analisis Efektivitas Pengecoh

Kriteria Efektivitas Pengecoh	Jumlah	Persentase
Sangat Baik	2	13,33%
Baik	5	33,33%
Kurang Baik	8	53,33%

Contoh spesifik dari pengecoh yang kurang efektif adalah pada soal-soal dengan jawaban yang hampir serupa, namun tidak cukup membingungkan peserta didik. Misalnya, beberapa pengecoh tidak memiliki relevansi yang cukup dengan materi yang diuji, sehingga peserta didik dengan pemahaman dasar dapat dengan mudah mengabaikan pengecoh tersebut. Untuk meningkatkan kualitas pengecoh, perlu dirancang pilihan

jawaban yang lebih dekat dengan materi pembelajaran dan lebih sulit dibedakan oleh siswa yang tidak sepenuhnya memahami topik yang diuji.

Efektivitas pengecoh yang rendah (hanya 46,66% pengecoh yang baik hingga sangat baik) menunjukkan kelemahan desain soal pilihan ganda. Pengecoh yang tidak efektif tidak sekadar masalah teknis, tetapi mencerminkan kesulitan dalam merancang instrumen yang benar-benar menguji pemahaman konseptual.

Pengecoh yang baik membutuhkan pemahaman mendalam tentang struktur pengetahuan akuntansi, kemampuan untuk merancang alternatif jawaban yang secara konseptual mendekati jawaban benar, namun tetap salah. Hal ini memerlukan keahlian tinggi dari perancang soal dalam menerjemahkan kompleksitas materi ke dalam format pilihan ganda.

Penelitian ini terbatas pada analisis butir soal dalam elemen Job Profile/Peluang Pekerjaan dan Usaha di bidang Akuntansi, sehingga hasilnya tidak dapat digeneralisasi untuk semua aspek pembelajaran akuntansi atau mata pelajaran lainnya. Selain itu, data diperoleh dari satu kali pelaksanaan tes tanpa memperhitungkan faktor lingkungan atau psikologis siswa yang mungkin memengaruhi hasil tes. Analisis hanya mencakup validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh tanpa mengevaluasi konteks implementasi soal dalam pengajaran yang lebih luas.

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode analisis butir soal. Sampel penelitian terdiri dari 15 butir soal pada materi Job Profile/Peluang Pekerjaan & Usaha di Bidang Akuntansi. Teknik pengumpulan data dilakukan melalui tes pilihan ganda, dengan analisis meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Instrumen dianalisis menggunakan perhitungan statistik deskriptif dengan bantuan perangkat lunak statistik untuk mengidentifikasi kualitas butir soal secara komprehensif.

Studi ini memberikan wawasan bagi para pendidik mengenai pentingnya validitas dan reliabilitas dalam pengembangan instrumen evaluasi. Dengan memperbaiki kualitas butir soal berdasarkan analisis empiris, penelitian ini mendukung pengembangan praktik asesmen yang lebih efektif dan akurat. Selain itu, hasil penelitian ini mendorong pendidik untuk lebih kritis dalam menilai instrumen pembelajaran serta membuka peluang untuk penelitian lanjutan yang memperluas cakupan analisis butir soal pada berbagai topik, populasi siswa, dan metode asesmen lainnya.

SIMPULAN

Penelitian ini mengungkap kualitas instrumen tes pilihan ganda pada materi profil pekerjaan dan peluang usaha di bidang akuntansi, yang memberikan kontribusi penting dalam mengembangkan instrumen evaluasi pembelajaran. Melalui analisis komprehensif, penelitian ini menemukan bahwa meskipun instrumen memiliki reliabilitas tinggi, terdapat kebutuhan signifikan untuk perbaikan sistematis dalam validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Temuan penelitian menegaskan pentingnya pengembangan berkelanjutan instrumen evaluasi untuk meningkatkan kualitas penilaian hasil belajar siswa di program keahlian akuntansi. Implikasi praktis dari penelitian ini meliputi perlunya: (1) revisi menyeluruh terhadap butir soal dengan validitas rendah, (2) pengembangan pengecoh yang lebih kompleks dan representatif, serta (3) penyusunan instrumen dengan variasi tingkat kesukaran yang proporsional untuk mengukur kemampuan kognitif siswa secara komprehensif.

DAFTAR PUSTAKA

- Akhmadi, M. N. (2021). *Analisis Butir Soal Evaluasi Tema 1 Kelas 4 SDN Plumbungan Menggunakan Program Anates*.
- Anazar, & Muhammad Munir. (2024). *Analisis Kualitas Butir Soal Ujian Akhir Semester Genap Mata Pelajaran Simulasi dan Komunikasi Digital Kelas X Multi Media di SMK Negeri 4 Kendari Tahun Ajaran 2021/2022*.

- 6780 *Analisis Butir Soal Pilihan Ganda pada Siswa SMK Menggunakan Anates* - Dian Septy Ikawati, Eka Wahyu Agustin, Luqman Hakim, Vivi Pratiwi
DOI: <https://doi.org/10.31004/edukatif.v6i6.7780>
- Ayu, D. (2024). *Analisis Butir Soal pada Mata Pelajaran Pendidikan Agama Islam di Kelas V SDN 3 Tanjung Pura* (Vol. 4).
- Basri, K., & Turmuzi, M. (2021). Analisis Butir Soal Ulangan Semester Ganjil Mata Pelajaran Matematika Kelas VIII SMP pada Tahun Ajaran 2018/2019. *Desember 2021 Journal of Mathematics Education and Application*, 1, 682. <https://mathjournal.unram.ac.id/index.php/Griya/indexGriya>
- Derri Adi Fernando, Hartatiana, & Fajri Ismail. (2023). *Pentingnya Validitas dan Reliabilitas Instrument Evaluasi pada Pembelajaran Pendidikan Agama Islam*.
<https://doi.org/https://doi.org/10.48094/raudhah.v8i3.580>
- 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.
<https://doi.org/10.47134/jbdi.v1i3.2863>
- Dwi Lestari, I., Zain Yumna, R. K., Alda Lusua Dewi, S., Sri, E. M., Nurhalizah, P., Rosalina, S., Studi Pendidikan Biologi, P., Keguruan Dan Ilmu Pendidikan, F., & Sultan Ageng Tirtayasa, U. (2024). Analisis Kualitas Butir Soal HOTS pada Mata Pelajaran Biologi yang Diberikan pada Siswa Kelas 12 dan Dievaluasi Menggunakan Aplikasi Anates. *Journal of Basic Educational Studies*, 4(2), 764–765.
- Firdaus, D., & Sri Rahayu, Y. (2020). *Pengembangan Instrumen Tes Miskonsepsi Menggunakan Four-Tier Multiple-Choice Test Submateri Katabolisme Karbohidrat Kelas XII SMA* (Vol. 9, Issue 2).
- Fiska, J. M., Hidayati, Y., Qomaria, N., & Hadi, W. P. (2021). Analisis Butir Soal Ulangan Harian IPA Menggunakan Software Anates pada Pendekatan Teori Tes Klasik. *Jurnal Natural Science Educational Research*, 4(1), 2021.
- Hendrayadi, Martin Kustati, & Rezki Amelia. (2024). *Analisis Ulangan Harian Mata Pelajaran PAI di SMA Negeri 10 Padang Tahun Pelajaran 2023/2024 (Telaah Terhadap Reliabilitas, Daya Beda dan Tingkat Kesukaran Menggunakan Software Anates)*. Vol. 7 No. 3 (2024): Volume 7 No 3 Tahun 2024.
<https://doi.org/https://doi.org/10.31004/jrpp.v7i3.29380>
- Hera Apriliana Saputri, Zulhijrah, Nabila Joti Larasati, & Shaleh. (2023). *Analisis Instrumen Assesmen : Validitas, Reliabilitas, Tingkat Kesukaran dan Daya Beda Butir Soal*.
<http://journal.stkipsubang.ac.id/index.php/didaktik/article/view/2268/1883>
- Inayah Dzil Izzati Hartono, Andi Besse Tenriawaru, & Kurnia Ningsih. (2024). *Analisis Butir Soal Penilaian Sumatif IPA Menggunakan Anates*. 8, 162–171. <https://doi.org/10.26418/jurnalkpk.v8i2.78282>
- Khasanah, I., Fuady, A., & Islam Malang, U. (2023). Analisis Soal Ulangan Harian Matematika Bentuk Pilihan Ganda. In *Mathema Journal E-ISSN* (Vol. 5, Issue 2).
- Lasito, W. S., & Pujiati, H. (2024). *Pelatihan Membuat Instrumen Penilaian Literasi Membaca Bahasa Inggris untuk Guru SMP Kabupaten Agam, Sumatra Barat*. 9(1). <https://doi.org/10.25047/j-dinamika.v9i1.4413>
- Leonardus Kaka, Vidriana Oktoviana Bano, & Yohana Njoeroemana. (2024). *Efektivitas Analisis Butir Soal Pilihan Ganda Menggunakan Aplikasi Anates di SMPN 2 Kanatang*. Vol 4 No 9: Februari 2024.
<https://ejournal.stpmataram.ac.id/JIP/article/view/3124>
- Magdalena, I., Fauziah, S. N., Faziah, S. N., & Nupus, F. S. (2021). Analisis Validitas, Reliabilitas, Tingkat Kesulitan dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas Iii Sdn Karet 1 Sepatan. In *Bintang : Jurnal Pendidikan dan Sains* (Vol. 3, Issue 2). <https://ejournal.stitpn.ac.id/index.php/bintang>
- Merlin Purwati, L., Arianty, R., Malaka Syakilah, D., Ridlo, S., & Susilaningsih, E. (2021). *Analisis Soal Tes Pilihan Ganda Berbasis Higher Order Thinking Skill menggunakan Aplikasi Anates Windows Versi 4.0.9 For Windows*. www.journal.uniga.ac.id
- Meyke Joice Tresia Rajagukguk, & Dorlan Naibaho. (2023). *Mampu Memilih Soal Berdasarkan Tingkat Kesukaran*. Vol. 2 No. 4 (2023): Oktober : *Jurnal Pendidikan Sosial dan Humaniora*.
<https://publisherqu.com/index.php/pediaqu/article/view/701>

- 6781 *Analisis Butir Soal Pilihan Ganda pada Siswa SMK Menggunakan Anates* - Dian Septy Ikawati, Eka Wahyu Agustin, Luqman Hakim, Vivi Pratiwi
DOI: <https://doi.org/10.31004/edukatif.v6i6.7780>
- Mustaqim, & Hidayu Sulisti. (2024). *Analisis Butir Soal Pas Matematika Peminatan: Daya Pembeda, Tingkat Kesukaran, dan Kualitas Pengecoh. Vol 3 No 1* (2024).
<https://doi.org/https://doi.org/10.24260/add.v3i1.3011>
- Ningrum, W. A., Rahmawati, R. C., Minarti, I. B., & Mekar, R. M. (2023). Analisis Butir Soal Ulangan Harian Pembelajaran IPA pada Kelas VIII Di SMPN 21 Semarang. *Educatio*, 18(1), 91–101.
<https://doi.org/10.29408/edc.v18i1.18291>
- Nurhalimah, S., Hidayati, Y., Rosidi, I., Wiwin, D., & Hadi, P. (2022). Hubungan Antara Validitas Item dengan Daya Pembeda dan Tingkat Kesukaran Soal Pilihan Ganda PAS. In *Jurnal Natural Science Educational Research* (Vol. 4, Issue 3). <https://doi.org/https://doi.org/10.21107/nser.v4i3.8682>
- Puji Rahayu, L. (2024). Kualitas Soal Bahasa Indonesia Kelas XI SMAN 2 Bangko Pusako: Analisis Butir Soal. In *Bahasa dan Sastra* (Vol. 10, Issue 4). Pendidikan. <https://e-journal.my.id/onoma>
- Puspa Hanan, M., Raudhatul Jannah, R., & Alexander Alim, J. (2023). Analisis Butir Soal Matematika Materi KPK dan FPB Berbasis Cerita di SDN 111 Pekanbaru. *Journal of Educational Learning and Innovation*, 3(1), 92. <https://doi.org/10.46229/elia.v3i1>
- Ramadhan, W., Malahati, F., Romadhon, K., & Ramadhan, S. (2023). Analisis Butir Soal Tipe Multiple Choice Questions pada Penilaian Harian Sekolah Dasar. *Tarbiyah Wa Ta'lim: Jurnal Penelitian Pendidikan dan Pembelajaran*, 10(2), 93–105. <https://doi.org/10.21093/twt.v10i2.6155>
- Ryzca Siti Qomariyah, Rr. Ghina Ayu Putri T.K, Dita Refani Putri, Desvita Sania Putri, & Mohammad Reza Triya P. (2022). *Analisis Tingkat Kesukaran dan Daya Pembeda pada Butir Soal Pilihan Ganda Mata Pelajaran Bahasa Indonesia Kelas V Semester 1 SDN Kedungdalem 2. Vol. 1 No. 2* (2022): April-Juni, 74–80. <https://doi.org/https://doi.org/10.47233/jpvt.v1i2.278>
- Saraswati, S., Rodliyah, I., & Rahmawati, N. D. (2021). Analisis Instrumen Penilaian Berbasis Higher Order Thinking Skills pada Mata Kuliah Matematika Lanjut. *Inovasi Matematika (Inomatika)*, 3(2), 138–151. <https://doi.org/10.35438/inomatika>
- Sholeh Mawardi, M., & Fuady, A. (2023). *Analisis Butir Soal Pilihan Ganda Menggunakan Anates pada Penilaian Tengah Semester Kelas VII D SMP Negeri 1 Ngajum Kabupaten Malang. 75*(1).
<http://jurnal.unipasby.ac.id/index.php/whn>
- Sofiani Putri Radja, Vidriana Oktoviana Bano, & Anita Tamu Ina. (2023). Analisis Kualitas Butir Soal Hasil Belajar Peserta Didik Berdasarkan Tingkat Kesukaran, Daya Beda dan Efektivitas Pengecoh di SMAN 1 Pandawai. *Jurnal Edusavana*, 30–41.
- Wahyuni Yusuf, F. (2024). *Analisis Butir Soal Asesmen Sumatif Biologi Materi Perubahan Lingkungan dengan Menggunakan Anates pada Kelas X SMA. 4*(1).
- Wati, I., Novita, Y., & Iqbal Lubis, M. (2022). *Pelatihan Penggunaan Software Anatest dalam Menganalisis Butir Soal Tes Ekonomi bagi Mahasiswa Calon Guru. 63*(2), 2022.
<https://ejournal.anotero.org/index.php/tasnim>
- Widarta, F. O., & Artika, W. (2021). Analisis Bentuk Stimulus, Dimensi Kognitif, dan Karakteristik HOTS pada Instrumen Evaluasi Mata Pelajaran IPA Karya Guru. *Jurnal IPA & Pembelajaran IPA*, 5(3), 197–208. <https://doi.org/10.24815/jipi.v5i3.21429>
- Yolanda Farra Alista, & Rozana Aufa Syahzanani. (2023). *Analisis Butir Soal Ulangan Harian Fisika dengan Pendekatan Teori Tes Klasik Menggunakan Program Anates. Vol. 1 No. 1* (2023): Volume 1 Number 1 Year 2023. <https://proceeding.unesa.ac.id/index.php/siptek/article/view/173>
- Yuniaria Dwi Prastika. (2021). *Pengaruh Validitas, Reliabilitas dan Tingkat Kesukaran terhadap Kualitas Butir Soal Ekonomi Menggunakan Software Anates di SMKN 3 Bangkalan.*