

Analisis Butir Soal Teknik Dasar Otomotif Berbasis Diagnostik Kognitif dengan Anates Versi 4.02

Dwi Jatmoko¹, Basuki^{2*}, Agung Pambudi³, Kuntho Alief Setiyono⁴, Beni Nugroho⁵

¹⁻⁵Universitas Muhammadiyah Purworejo, Indonesia

Alamat: Jl. K.H. Ahmad Dahlan No. 3 & 6 Purworejo, Indonesia

Korespondensi penulis: basuki@umpwr.ac.id *

Abstract. *The purpose of this study was to determine the quality of the Odd Semester Final Summative Questions for the Automotive Basic Engineering Subject for Class X SMK TKM Pertambangan Kebumen using the Anates V4 Application. The exam question data, answer keys, and all students' answers were collected through data collection documentation techniques. Furthermore, using the Anates version 4.02 program, item analysis will be carried out to assess the validity, reliability, level of difficulty, discriminatory power, and effectiveness of distractors. This research method is quantitative descriptive research, the data analyzed is not used to accept or reject the hypothesis; instead, the results of the analysis show a description of the observed symptoms, which are represented by numbers or coefficients between variables. The results showed that 95% of the odd semester final summative questions for the Automotive Basic Engineering subject for Class X SMK TKM Pertambangan Kebumen using the Anates V4 Application had validity, with 5% being invalid. For reliability, the questions had a reliability coefficient of 0.78, which indicated that it had a high coefficient. Furthermore, based on the level of difficulty, 2% were included in the easy category; 40% were included in the moderate category; 55% are in the difficult category; and 3% are in the very difficult category. In terms of discrimination power, 25% are in the good category; 57% are in the sufficient category; and 18% are in the poor category. All distractors function well on 90% of the questions.*

Keywords: *Anates, Basic automotive engineering, Item analysis*

Abstrak. Tujuan penelitian ini adalah untuk mengetahui kualitas butir soal Sumatif Akhir Semester Gasal Mata Pelajaran Teknik Dasar Otomotif Kelas X SMK TKM Pertambangan Kebumen dengan menggunakan Aplikasi Anates V4. Data soal ujian, kunci jawaban, dan jawaban semua siswa dikumpulkan melalui teknik dokumentasi pengumpulan data. Selanjutnya, menggunakan program anates versi 4.02, analisis butir soal akan dilakukan untuk menilai validitas, reliabilitas, tingkat kesulitan, daya pembeda, dan efektivitas pengecoh atau distractor. Metode penelitian ini adalah penelitian deskriptif kuantitatif, data yang dianalisis tidak digunakan untuk menerima atau menolak hipotesis; sebaliknya, hasil analisis menunjukkan deskripsi gejala yang diamati, yang diwakili oleh angka atau koefisien antar variabel. Hasil penelitian menunjukkan bahwa 95% dari butir soal sumatif akhir semester gasal mata pelajaran Teknik Dasar Otomotif Kelas X SMK TKM Pertambangan Kebumen menggunakan Aplikasi Anates V4 memiliki validitas, dengan 5% tidak valid. Untuk reliabilitas, soal memiliki koefisien reliabilitas sebesar 0,78, yang menunjukkan bahwa itu memiliki koefisien tinggi. Selanjutnya, berdasarkan tingkat kesulitan, 2% termasuk dalam kategori mudah; 40% termasuk dalam kategori sedang; 55% termasuk dalam kategori sukar; dan 3% termasuk dalam kategori sangat sukar. Dari segi daya pembeda, 25% termasuk dalam kategori baik; 57% termasuk dalam kategori cukup; dan 18% termasuk dalam kategori jelek. Seluruh pengecoh berfungsi dengan baik pada 90% soal.

Kata kunci: Anates, Teknik dasar otomotif, Analisis butir soal

1. LATAR BELAKANG

Perkembangan ilmu pengetahuan dan teknologi telah memberikan dampak positif maupun negatif dalam aspek kehidupan manusia, permasalahan yang timbul dapat dipecahkan dengan upaya peningkatan dan penguasaan ilmu pengetahuan dan teknologi. Adanya perkembangan ilmu pengetahuan dan teknologi juga membuat tidak adanya batasan antar manusia untuk berkomunikasi. Oleh karena itu, berkembangnya ilmu pengetahuan dan

teknologi membawa manusia dalam persaingan global. Salah satu cara yang dapat ditempuh oleh suatu negara agar dapat bertahan dalam persaingan global yaitu dengan meningkatkan sumber daya manusia.

Pendidikan merupakan salah satu upaya untuk meningkatkan kualitas sumber daya manusia. Pendidikan adalah suatu hak bagi seluruh masyarakat yang harus dipenuhi dalam kehidupan bermasyarakat, berbangsa, dan bertanah air. Pendidikan yang baik adalah yang memiliki proses yang terarah dan jelas. Melalui pendidikan diharapkan dapat menciptakan sumber daya manusia (SDM) yang memiliki daya saing dan bermutu tinggi. Pengertian tentang pendidikan ini juga sesuai dengan undang-undang Nomor 2 tahun 1989 tentang sistem pendidikan nasional, Bab II Pasal 4 dalam Ngalim Purwanto (2010: 1) dinyatakan : “pendidikan nasional bertujuan mencerdaskan kehidupan bangsa dan mengembangkan manusia Indonesia seutuhnya, yaitu manusia yang beriman dan bertaqwa terhadap Tuhan Yang Maha Esa dan berbudi pekerti luhur, memiliki pengetahuan dan keterampilan, kesehatan jasmani dan rohani, kepribadian yang mantap dan mandiri serta rasa tanggung jawab kemasyarakatan dan kebangsaan.”

Penilaian menurut Yosep Sunu (2014: 18) merupakan seperangkat prosedur yang didesain untuk mengumpulkan informasi tentang perkembangan, kemajuan, dan prestasi atau hasil belajar yang dicapai siswa yang dibandingkan dengan suatu standar atau acuan yang relevan. Penilaian hasil belajar mengajar dapat diketahui dengan menggunakan alat yaitu instrument, dimana pendidik membuat instrumen ini dalam bentuk soal atau tes dengan kriteria soal yang telah ditentukan sehingga dapat menilai hasil belajar mengajar sesungguhnya. Menurut Ahmad Zainuri, dkk (2022: 76) sebuah tes apabila mengandung 4 aspek yaitu validitas, reabilitas, objektivitas, praktikabilitas dan ekonomis.

Keunggulan dan kelemahan setiap bentuk soal nantinya akan menjadi acuan pendidik untuk menyusun soal yang baik. Teknik penyusunannya juga didasarkan pada karakteristik dan bobot soal itu sendiri. Pendidik harus selektif dalam menyusun butir-butir soal sebelum di uji cobakan terhadap peserta didik. Hal ini bisa diketahui melalui hasil evaluasi yang sudah dilaksanakan selama satu semester dimana terjadi ketimpangan diantara nilai 3 mata pelajaran yang sedang diajarkan pada semester gasal tahun ajaran 2023/2024 di SMK TKM Pertambangan Kebumen.

Tabel 1. Nilai Rata-Rata Siswa Pada Tiga Mata Pelajaran

Mata pelajaran	KKM	Nilai rata-rata	
		X OTO 1	X OTO 2
TDO 1	80	78,22	77,60
TDO 2	80	81,98	82,01
TDO 3	80	83,98	84,01

Mata pelajaran TDO 2 dan TDO 3 sudah menunjukkan bahwa nilai siswa dapat mencapai standar ketuntasan belajar minimal yang ditetapkan. Namun berbeda dengan mata pelajaran TDO 1, dimana hanya ada beberapa siswa yang dapat mencapai standar ketuntasan belajar minimal. Analisis butir soal sebagai usaha untuk mengetahui kualitas soal harus dilakukan. Hal ini karena soal tersebut digunakan sebagai pertimbangan dalam menentukan kualitas pendidikan. Berdasarkan permasalahan yang diuraikan, peneliti memandang bahwa analisis butir soal perlu dilakukan untuk mengetahui kualitas perangkat tes, sehingga bisa digunakan sebagai bahan evaluasi pembuatan soal mendatang.

2. KAJIAN TEORITIS

Evaluasi Pembelajaran dalam Konteks Pendidikan Kejuruan

Proses sistematis untuk mengevaluasi dan menilai tingkat pencapaian siswa terhadap tujuan pembelajaran dikenal sebagai evaluasi pembelajaran. Evaluasi sangat penting dalam pendidikan kejuruan seperti SMK karena memastikan bahwa siswa tidak hanya memahami konsep teoretis tetapi juga mampu menerapkannya dalam dunia kerja. Pembelajaran kejuruan menuntut integrasi antara teori dan praktik, sehingga evaluasi harus mempertimbangkan keduanya. Fitriyani et al. (2021) menyatakan bahwa pendekatan evaluasi pendidikan vokasi harus berfokus pada pencapaian semua kemampuan—kognitif, psikomotorik, dan afektif. Selama abad ke-21, evaluasi telah berubah menjadi alat yang digunakan untuk membuat keputusan tentang pengembangan pembelajaran dan untuk mengukur seberapa baik seseorang belajar. Evaluasi berfungsi sebagai umpan balik (feedback) antara pendidik dan siswa. Hal ini sesuai dengan prinsip belajar bebas, di mana evaluasi dimaksudkan untuk mengidentifikasi peluang dan tantangan belajar siswa. Oleh karena itu, evaluasi pembelajaran teknik otomotif harus dirancang secara menyeluruh dan mencakup proses, hasil, dan diagnostik kognitif siswa agar intervensi pembelajaran dapat dilakukan dengan benar dan efektif. (Putri & Firmansyah, 2022).

Analisis Butir Soal: Penguatan Kualitas Instrumen Evaluasi

Analisis butir soal menilai kualitas soal dengan data empirik dari hasil uji coba atau pelaksanaan tes. Tujuan utama dari analisis butir soal adalah untuk menentukan apakah soal tersebut dapat secara valid dan reliabel mengukur kemampuan siswa. Tingkat kesukaran, daya pembeda, validitas empiris, dan efektivitas pengecoh (distraktor) adalah parameter utama yang dianalisis. Menurut Sari & Pramudya (2022), analisis ini memungkinkan guru untuk mengidentifikasi pertanyaan yang baik, sedang, atau perlu perbaikan, sehingga proses evaluasi menjadi lebih objektif dan bermakna. Analisis butir soal juga dapat membantu guru membuat

bank soal yang bagus dan sesuai dengan tujuan pembelajaran. Hal ini penting dalam bidang teknik dasar otomotif, yang membutuhkan evaluasi yang akurat dari penguasaan siswa terhadap ide dan keterampilan awal. Misalnya, pertanyaan yang terlalu mudah atau terlalu sulit tidak akan menunjukkan kemampuan siswa. Selain itu, daya pembeda yang rendah menunjukkan bahwa soal tidak dapat membedakan siswa berkemampuan tinggi dari siswa berkemampuan rendah. Oleh karena itu, kegiatan analisis butir soal menjadi sangat penting untuk siklus evaluasi pembelajaran yang berkelanjutan di SMK.(Rahmatika et al., 2021).

Asesmen Diagnostik Berbasis Kognitif

Asesmen diagnostik berbasis kognitif adalah jenis penilaian yang bertujuan untuk menentukan tingkat pemahaman siswa terhadap ide-ide tertentu dan mengidentifikasi lokasi kesalahan berpikir atau miskonsepsi. Asesmen ini sangat bermanfaat untuk pembelajaran teknik otomotif karena dapat memberikan gambaran tentang pemahaman awal siswa tentang sistem kerja mesin, kelistrikan, dan prosedur perbaikan mobil. Sebagaimana dinyatakan oleh Supprapto & Fitria (2023), evaluasi diagnostik tidak hanya menilai hasil tetapi juga mengungkap proses berpikir siswa. Ini memungkinkan guru untuk mengidentifikasi pola dan taktik yang digunakan siswa untuk menjawab pertanyaan. Metode diagnostik ini memungkinkan guru untuk membuat program penguatan pembelajaran atau remedial yang lebih khusus untuk menyelesaikan masalah siswa. Hal ini memungkinkan pembelajaran yang berbeda, yang berarti pembelajaran yang disesuaikan dengan kebutuhan siswa secara unik. Menurut Gunawan et al. (2022), asesmen diagnostik berbasis kognitif sejalan dengan pendekatan pembelajaran konstruktivistik, yang menekankan pentingnya membangun pemahaman konsep berdasarkan pengalaman dan proses berpikir siswa. Oleh karena itu, langkah penting dalam meningkatkan kualitas asesmen dalam bidang teknik dasar otomotif adalah pembuatan dan analisis butir soal berbasis diagnostik kognitif.

Evaluasi dalam Mata Pelajaran Teknik Dasar Otomotif

Fokus utama program keahlian Teknik Kendaraan Ringan di SMK adalah mata pelajaran Teknik Dasar Otomotif. Pelajaran ini membahas prinsip kerja sistem otomotif, komponen, dan teknik dasar pemeriksaan dan perawatan kendaraan. Evaluasi yang digunakan dalam Teknik Dasar Otomotif harus mampu mengukur penguasaan teori dan kemampuan analisis serta penerapan dalam dunia nyata, mengingat betapa kompleks dan terkaitnya dengan bidang lanjutan. Wibowo dan Hakim (2021) menekankan bahwa taksonomi kognitif yang luas mulai dari tingkat pemahaman hingga analisis dan evaluasi harus dipertimbangkan saat membuat tes evaluasi. Selain itu, karena fokus pembelajaran teknik otomotif adalah kesiapan kerja, evaluasi harus dilakukan dalam konteks industri. Untuk mendekatkan siswa pada realitas

kerja, pendekatan pabrik berbasis pembelajaran (TEFA) dan penilaian proyek sudah digunakan. Namun, ujian tertulis tetap penting sebagai pengukur awal penguasaan konsep di bidang dasar. Oleh karena itu, setiap item soal yang digunakan harus dianalisis secara menyeluruh untuk memastikan bahwa tingkat pemahaman siswa benar-benar dapat diukur dan kekeliruan kognitif yang mungkin terjadi dapat dideteksi (Suhartini et al., 2023).

Pemanfaatan Aplikasi Anates Versi 4.02 dalam Analisis Butir Soal

Perangkat lunak Anates Versi 4.02 banyak digunakan oleh guru dan peneliti pendidikan untuk menganalisis kualitas butir soal secara statistik. Analisis tingkat kesukaran, daya pembeda, efektivitas distraktor, dan pembuatan laporan analisis soal adalah semua fitur yang ditawarkan oleh program ini. Keunggulan utamanya adalah tampilan antarmuka yang sederhana dan proses analisis yang cepat dan akurat. Menurut Putra & Nugroho (2022), menggunakan Anates sangat membantu guru dalam mengevaluasi soal secara empirik dan mendorong pengambilan keputusan berbasis data saat menyusun kembali soal untuk penilaian selanjutnya. Penggunaan Anates menjadi sangat relevan untuk evaluasi pembelajaran Teknik Dasar Otomotif karena guru dapat menilai soal-soal yang diberikan kepada siswa berdasarkan data real. Ini sejalan dengan prinsip pengambilan keputusan berdasarkan data (DDDM) dalam pendidikan. Hasil dari analisis Anates dapat digunakan untuk menyusun bank soal atau memperbaiki bagian soal yang kurang baik sebelum digunakan lagi. Akibatnya, kualitas evaluasi akan meningkat, dan siswa akan memiliki kesempatan untuk dinilai secara adil dan akurat. Selain itu, penelitian ini dapat digunakan bersama dengan metode diagnostik kognitif untuk mengidentifikasi masalah belajar siswa secara lebih mendalam (Kurniawan & Salim, 2023).

3. METODE PENELITIAN

Penelitian ini menggunakan pendekatan penelitian deskriptif kuantitatif digunakan untuk memberikan gambaran sistematis, faktual, dan akurat tentang fakta dan ciri-ciri populasi atau fenomena yang diteliti. Dalam kasus ini, tujuan analisis bukanlah menguji hipotesis secara inferensial; sebaliknya, tujuan analisis adalah untuk menginterpretasikan data numerik hasil pengukuran terhadap masing-masing butir soal secara objektif. Dalam analisis butir soal, metode ini sangat berguna karena dapat menjelaskan karakteristik setiap butir soal dengan mengacu pada parameter statistik seperti efektivitas p, validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Menurut Sugiyono (2021), penelitian deskriptif kuantitatif memungkinkan para peneliti untuk menilai fenomena atau objek berdasarkan data yang dikumpulkan dalam bentuk angka yang kemudian diolah dengan alat statistik.

Penelitian ini menganalisis data yang dikumpulkan menggunakan perangkat lunak Anates Versi 4.02, yang biasa digunakan untuk menilai kualitas butir soal. Tingkat kesukaran soal, daya pembeda, validitas, reliabilitas, dan efektivitas pengecoh atau opsi jawaban adalah semua hasil analisis terperinci yang diberikan oleh aplikasi ini. Akibatnya, diharapkan bahwa hasil penelitian ini akan memberikan gambaran mendalam tentang kualitas soal sumatif yang digunakan dan menjadi dasar untuk perbaikan soal di masa mendatang (Sari & Pramudya, 2022).

Lokasi penelitian berada di SMK TKM Pertambangan Kebumen berlokasi di Kawedusan, Kebumen, Program Keahlian Teknik Ototronik. Program ini berfokus pada penguasaan teknologi otomotif dan sistem kendali elektronik kendaraan. Pilihan ini didasarkan pada fakta bahwa mata pelajaran Teknik Dasar Otomotif (TDO) sangat penting untuk membangun kompetensi awal siswa sebelum mereka memasuki mata pelajaran lanjutan. Selain itu, penilaian akhir semester sekolah ini dapat digunakan sebagai sumber data utama untuk penelitian ini.

Validitas, reliabilitas, tingkat kesulitan, daya pembeda, dan efektivitas pengecoh (distraktor) adalah lima komponen utama analisis kualitas soal dalam penelitian ini. Validitas soal ditentukan oleh seberapa baik setiap bagian memiliki kemampuan untuk menilai kompetensi yang seharusnya diukur. Namun, reliabilitas berkaitan dengan seberapa konsisten hasil tes pada berbagai kelompok. Tingkat kesukaran menunjukkan seberapa mudah atau sulit suatu soal dijawab oleh siswa, dan daya pembeda menunjukkan seberapa mudah atau sulit soal itu dijawab oleh siswa. Selain itu, pengecoh harus efektif karena pengecoh yang tidak berfungsi akan mengurangi kualitas soal. Kelima variabel ini dianggap selaras dengan pendekatan analisis butir soal modern, yang menuntut data empiris untuk mengevaluasi kinerja setiap item dalam suatu tes. Kurniawan dan Salim (2023) menyatakan bahwa kelima variabel tersebut saling berkaitan dan berkontribusi pada pembuatan alat evaluasi yang akurat, dapat diandalkan, dan relevan untuk digunakan dalam pembelajaran dan penilaian siswa.

Ada 76 siswa di jurusan Ototronik SMK TKM Pertambangan Kebumen. Mereka terdiri dari dua kelompok belajar: kelas X OTO 1 dan kelas X OTO 2. Siswa ini telah mengikuti kelas TDO selama satu semester sebelum ujian sumatif di akhir semester gasal. Pemilihan topik ini didasarkan pada kenyataan bahwa siswa kelas X adalah siswa baru di lingkungan kejuruan, sehingga evaluasi pada tahap ini sangat penting untuk mengevaluasi efektivitas awal pembelajaran (Gunawan et al., 2022). Fokus penelitian adalah soal-soal sumatif yang diberikan pada akhir semester gasal. Soal-soal ini berbentuk pilihan ganda dan mencakup kunci jawaban serta hasil respons siswa pada mata pelajaran Teknik Dasar Otomotif (TDO). Tujuan dari soal-

soal ini adalah untuk mengukur seberapa baik siswa memahami materi teknik dasar otomotif. Aplikasi Anates memungkinkan peneliti melakukan analisis kuantitatif dengan menggunakan lembar soal, lembar kunci, dan lembar jawaban siswa. Tujuannya adalah untuk mengetahui sejauh mana kualitas instrumen asesmen memenuhi standar. Putra dan Nugroho (2022) mengatakan bahwa penggunaan data real dari hasil ujian dapat memberikan gambaran yang lebih akurat tentang kinerja soal dan tingkat pemahaman siswa secara keseluruhan.

Definisi Operasional Variabel dan Teknik Analisis Data

Metode kuantitatif dengan analisis statistik deskriptif digunakan dalam penelitian ini untuk mengevaluasi kualitas soal sumatif akhir semester gasal mata pelajaran TDO di kelas X SMK TKM Pertambangan Kebumen. Validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh adalah beberapa metrik yang dievaluasi. Sejauh mana suatu soal mampu mengukur apa yang seharusnya diukur disebut validitas soal. Validitas penelitian ditentukan melalui korelasi point biserial. Koefisien korelasi ini menunjukkan bagaimana skor seluruh ujian berkorelasi dengan skor masing-masing bagian. Nilai korelasi point biserial lebih besar dari r tabel pada taraf signifikansi 5% menandakan bahwa butir soal valid (Retnawati, 2020). Untuk menghitung korelasi point biserial, rumus berikut digunakan:

$$\gamma_{pbi} = \frac{Mp - Mt}{St} \sqrt{\frac{p}{q}}$$

Keterangan:

γ_{pbi} : koefisien korelasi *point biserial*

Mp : rerata skor dari subjek yang menjawab benar

Mt : rerata skor total

St : Standar deviasi dari skor total

P : proporsi peserta didik yang menjawab benar

$$(p = \frac{\text{banyaknya siswa yang benar}}{\text{jumlah seluruh siswa}})$$

q : proporsi peserta didik yang menjawab salah

$$(q = 1 - p)$$

Reliabilitas pada penelitian ini reliabilitas tes secara keseluruhan dihitung dengan rumus Kuder-Richardson (K-R 20) sebagai berikut :

$$r_{11} = \left(\frac{k}{k-1} \right) \left(1 - \frac{\sum pq}{s_t^2} \right)$$

Keterangan :

r_{11} : koefisien reliabilitas

- p : proporsi subjek yang menjawab item yang benar
q : proporsi subjek yang menjawab item dengan salah
 Σpq : jumlah hasil perkalian antara p dan q
K : banyaknya butir
 S^2_t : varians skor total

Tingkat kesukaran digunakan untuk mengetahui level butir soal termasuk dalam tingkat sukar, sedang atau mudah. Tingkat kesukaran dapat dihitung dengan rumus sebagai berikut :

$$P = \frac{B}{JS}$$

Keterangan :

- P : indeks kesukaran / tingkat kesukaran
B : banyaknya siswa yang menjawab soal itu dengan betul
JS : jumlah seluruh siswa peserta tes

Daya pembeda dapat dihitung dengan rumus sebagai berikut:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

Keterangan:

- D : Indeks daya pembeda
 B_A : Banyaknya peserta kelompok atas yang menjawab benar
 B_B : Banyaknya peserta kelompok bawah yang menjawab benar
J : Jumlah peserta tes
 J_A : Jumlah peserta tes dari kelompok atas
 J_B : Jumlah peserta tes dari kelompok bawah
 P_A : $\frac{B_A}{J_A}$ = Proporsi peserta kelompok atas yang menjawab benar
 P_B : $\frac{B_B}{J_B}$ = Proporsi peserta kelompok bawah yang menjawab benar

Efektifitas pengecoh/*distractor* dapat diketahui dengan melihat pada sebaran jawaban para siswa. Pada sebaran jawaban diperoleh dengan menghitung banyaknya *testee* yang memilih jawaban atau yang tidak memilih apapun. Pengecoh yang baik ditandai dengan dipilih oleh sedikitnya 5% dari peserta tes. Teknik pengumpulan data dalam penelitian ini dilakukan melalui metode dokumentasi. Data yang dikumpulkan meliputi daftar nama siswa, lembar soal beserta kunci jawaban, dan hasil jawaban siswa pada Sumatif Akhir Semester Gasal. Teknik analisis data yang digunakan bersifat kuantitatif deskriptif dengan bantuan software Anates V4.0. Aplikasi ini digunakan untuk menghitung seluruh parameter analisis butir soal seperti validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh dengan lebih

efisien dan akurat. Pendekatan ini memungkinkan peneliti memperoleh gambaran kualitas soal secara menyeluruh dan obyektif, sehingga dapat digunakan sebagai dasar perbaikan instrumen evaluasi pembelajaran berikutnya.

4. HASIL DAN PEMBAHASAN

Tujuan dari penelitian ini adalah untuk mengevaluasi kualitas instrumen soal pilihan ganda dalam mata pelajaran Teknik Dasar Otomotif (TDO) kelas X. Hasil dari analisis 60 soal menunjukkan bahwa soal-soal tersebut secara keseluruhan memiliki kualitas yang cukup baik. Dalam hal validitas, 53 soal (88,3%) termasuk dalam kategori valid dan layak digunakan karena telah terbukti dapat mengukur kemampuan siswa secara akurat. Sementara itu, 7 soal (11,7%) dianggap tidak valid dan memerlukan perbaikan untuk dapat digunakan kembali dalam evaluasi pembelajaran. Dari sisi reliabilitas, analisis dengan aplikasi Anates versi 4.0.2 menunjukkan nilai koefisien alpha sebesar 0,78. Instrumen soal memiliki tingkat keterandalan dan konsistensi yang tinggi dalam mengukur kompetensi siswa, karena nilainya berada di atas batas minimum yang ditetapkan ($r \geq 0,70$). Hasil ini menunjukkan bahwa soal-soal dapat memberikan gambaran yang stabil tentang tingkat belajar siswa.

Selain itu, analisis tingkat kesulitan menunjukkan bahwa distribusi tidak proporsional. Soal-soal tidak termasuk dalam kategori sangat mudah, dan hanya satu soal (2%) dianggap mudah. Sebaliknya, 24 soal (setengah persen) termasuk dalam kategori sedang, 33 soal (55%) termasuk dalam kategori sukar, dan 2 soal (3%) dianggap sangat sukar. Jumlah soal dalam kategori mudah idealnya berjumlah 30%, soal dalam kategori sedang idealnya berjumlah 50%, dan soal dalam kategori sukar idealnya berjumlah 20%. Oleh karena itu, untuk menyeimbangkan komposisi tingkat kesulitan ujian, perlu ditambahkan soal dalam kategori mudah dan mengurangi soal yang terlalu sukar. Ini akan membuat tes secara keseluruhan lebih sesuai dengan kemampuan sebagian besar siswa.

Dilihat dari faktor daya pembeda, 15 soal memiliki daya pembeda yang baik, 34 soal cukup, dan 11 soal buruk. Soal-soal dengan daya pembeda yang baik dan cukup, terutama yang dapat dimasukkan ke dalam bank soal, masih dapat digunakan untuk evaluasi di masa mendatang; namun, soal-soal dengan daya pembeda yang rendah harus diperbaiki atau bahkan dihilangkan jika tidak dapat ditingkatkan dalam kualitasnya. Ini karena mereka tidak dapat secara efektif membedakan antara siswa yang berkemampuan tinggi dan rendah.

Terakhir, analisis kualitas pengecoh menunjukkan bahwa sebagian besar soal—sebanyak 54 butir atau 90 persen—memiliki pengecoh yang berfungsi secara efektif, sehingga opsi jawaban yang salah dapat menarik perhatian siswa yang belum memahami materi. Namun, ada 6 butir atau 10 persen soal yang memiliki pengecoh yang tidak efektif. Hal ini mungkin disebabkan oleh jumlah soal yang terlalu lemah untuk dipilih oleh siswa yang kurang menguasai materi dan banyaknya soal yang berada dalam kategori sedang dan sukar, yang membuat peserta cenderung menebak. Secara keseluruhan, soal-soal yang diuji dalam penelitian ini cukup valid, akurat, dan efektif. Namun, untuk meningkatkan kualitas secara keseluruhan, perlu ada peningkatan pada distribusi kesukaran, soal dengan daya pembeda rendah, dan pengoptimalan pengecoh.

5. KESIMPULAN DAN SARAN

Berdasarkan hasil analisis dan diskusi 60 soal pilihan ganda dari mata pelajaran Dasar-Dasar Teknik Otomotif (TDO), kualitas tes secara keseluruhan baik dan layak digunakan; namun, ada beberapa area yang perlu diperbaiki. Banyak soal memenuhi kriteria validitas (88,3% valid), dan sangat reliabel dengan koefisien alpha sebesar 0,78, menunjukkan bahwa alat ini mampu mengukur kompetensi siswa secara konsisten. Namun demikian, ada ketidakseimbangan dalam tingkat kesukaran soal; 55% dari soal kategori sukar tidak sesuai dengan proporsi ideal, dan 2% dan 40% dari soal kategori mudah dan sedang kekurangan. Ini menunjukkan bahwa distribusi soal harus diubah agar lebih sesuai dengan kemampuan mayoritas siswa.

Berdasarkan daya pembeda, 49 soal (81,7%) dianggap cukup dan baik, sementara 11 soal (18,3%) masih buruk dan memerlukan perbaikan atau penggantian. Sehubungan dengan kualitas pengecoh, sebagian besar sudah berfungsi dengan baik sembilan puluh persen, tetapi enam soal dari sepuluh persen memiliki pengecoh yang tidak berfungsi dengan baik yang perlu diperbaiki. Secara keseluruhan, tes ini dapat digunakan untuk mengevaluasi pembelajaran. Namun, soal-soal yang memiliki kualitas rendah harus diperbaiki, terutama yang berkaitan dengan distribusi kesukaran yang tidak adil, daya pembeda yang kurang, dan pengecoh yang tidak efektif.

DAFTAR REFERENSI

- Arifin, Z., & Rahayu, D. P. (2021). Analisis butir soal pilihan ganda berdasarkan tingkat kesukaran, daya pembeda, dan efektivitas distraktor. *Jurnal Kependidikan*, 9(2), 111–121.
- Azwar, S. (2020). *Reliabilitas dan validitas*. Yogyakarta: Pustaka Pelajar.
- Fitriyani, N., Huda, M., & Rasyid, A. (2021). Evaluasi pembelajaran berbasis kompetensi dalam pendidikan kejuruan. *Jurnal Pendidikan Vokasi*, 11(1), 37–47. <https://doi.org/10.21831/jpv.v11i1.38491>
- Gunawan, I. G. P., Kusuma, G. P., & Marhaeni, A. A. I. N. (2022). Penerapan asesmen diagnostik dalam pembelajaran abad 21 di SMK. *Jurnal Evaluasi Pendidikan*, 13(1), 45–55. <https://doi.org/10.21009/010023>
- Kurniawan, M., & Salim, A. (2023). Analisis kualitas soal pilihan ganda berbasis software Anates 4.02 pada mata pelajaran produktif SMK. *Jurnal Penelitian dan Evaluasi Pendidikan*, 28(2), 112–124. <https://doi.org/10.21831/jpep.v28i2.54321>
- Putra, R., & Nugroho, D. (2022). Efektivitas penggunaan software Anates dalam menganalisis soal mata pelajaran produktif di SMK. *Jurnal Teknologi Pendidikan dan Pembelajaran*, 10(1), 56–65. <https://doi.org/10.26858/jtpp.v10i1.34487>
- Rahmatika, D., Cahyani, A., & Fauzan, M. (2021). Validitas butir soal dalam evaluasi pembelajaran kejuruan. *Jurnal Pendidikan Teknik Otomotif*, 9(2), 80–89. <https://doi.org/10.33394/jpto.v9i2.37341>
- Retnawati, H. (2020). *Validitas dan reliabilitas instrumen*. Yogyakarta: UNY Press.
- Sari, M., & Pramudya, I. (2022). Analisis butir soal pilihan ganda menggunakan Anates versi 4.0 sebagai upaya peningkatan kualitas evaluasi belajar. *Jurnal Inovasi Pendidikan*, 7(2), 90–98. <https://doi.org/10.23887/jip.v7i2.40715>
- Sugiyono. (2021). *Metode penelitian kuantitatif, kualitatif, dan R&D*. Bandung: Alfabeta.
- Suhartini, D., Nugroho, Y. W., & Kartini, A. (2023). Pengembangan asesmen berbasis teaching factory pada pembelajaran teknik otomotif. *Jurnal Pendidikan Vokasi Otomotif*, 4(1), 21–30.
- Sukardi. (2021). *Metodologi penelitian pendidikan: Kompetensi dan praktiknya*. Jakarta: Bumi Aksara.
- Suprpto, E., & Fitria, D. (2023). Asesmen diagnostik dalam pembelajaran berbasis kompetensi untuk SMK. *Jurnal Evaluasi dan Asesmen Pendidikan Vokasi*, 5(1), 23–34. <https://doi.org/10.24252/jeapv.v5i1.2023>
- Wibowo, A., & Hakim, L. (2021). Pengembangan evaluasi berbasis kebutuhan dunia kerja pada mata pelajaran teknik dasar otomotif. *Jurnal Pendidikan Teknologi Kejuruan*, 19(2), 145–154. <https://doi.org/10.21009/jptk.192.06>
- Widyastuti, E., & Nurtanto, M. (2020). Analysis of item validity, reliability, and difficulty index in the evaluation of learning outcomes in vocational education. *Journal of Education Research and Evaluation*, 4(1), 58–65.