

## Analisis Butir Empiris terhadap Instrumen Penilaian Vokasi dalam Pelatihan Servis dan Konversi Kendaraan Listrik

Muhamad Yasfani Nur<sup>\*1</sup>, Ruffi<sup>2</sup>, Sabariah<sup>3</sup>

<sup>1,2,3</sup>Program Studi Teknologi Pendidikan, Sekolah Pascasarjana, Universitas PGRI Adi Buana  
Surabaya, Indonesia

Email: <sup>1</sup>[yasfani06@gmail.com](mailto:yasfani06@gmail.com), <sup>2</sup>[ruffi.adibuana@gmail.com](mailto:ruffi.adibuana@gmail.com), <sup>3</sup>[sabariah@unipasby.ac.id](mailto:sabariah@unipasby.ac.id)

### Abstrak

Penelitian ini bertujuan mengevaluasi kualitas empiris instrumen asesmen akhir berbentuk pilihan ganda pada pelatihan Servis dan Konversi Kendaraan Listrik. Penelitian menggunakan desain kuantitatif deskriptif dengan menganalisis respons 35 peserta terhadap 25 butir soal lima opsi. Data diolah menggunakan ANATES Versi 4.0.2 dan diverifikasi melalui analisis tingkat butir yang mencakup reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Hasil menunjukkan rata-rata skor mentah peserta sebesar 18,94 dari 25 butir atau 75,77% jawaban benar. Koefisien reliabilitas internal instrumen sebesar 0,44, sehingga konsistensi skor masih tergolong rendah hingga sedang dan perlu ditafsirkan secara hati-hati. Analisis tingkat kesukaran menunjukkan 9 butir berkategori sedang, 13 butir mudah, dan 3 butir sangat mudah, tanpa butir sukar maupun sangat sukar. Analisis daya pembeda menunjukkan 11 butir sangat baik, 1 butir baik, 2 butir cukup, dan 11 butir buruk. Analisis pengecoh memperlihatkan beberapa opsi salah tidak berfungsi, terutama opsi E yang berulang kali tidak dipilih peserta. Butir 6, 7, 12, 13, 19, dan 25 memerlukan revisi besar atau penggantian, sedangkan butir 2, 4, 10, 16, dan 22 memerlukan revisi. Instrumen dapat memberikan informasi awal mengenai hasil belajar peserta, tetapi belum memadai sebagai alat ukur yang stabil. Revisi perlu difokuskan pada penyeimbangan tingkat kesukaran, penguatan daya pembeda, perbaikan pengecoh berbasis miskonsepsi, dan validasi isi oleh ahli.

**Kata kunci:** ANATES, analisis butir soal, asesmen vokasi, pelatihan kendaraan listrik, tes pilihan ganda

## Empirical Item Analysis of a Vocational Assessment Instrument in Electric Vehicle Service and Conversion Training

### Abstract

*This study aims to evaluate the empirical quality of a multiple-choice final assessment instrument used in Electric Vehicle Service and Conversion training. A descriptive quantitative design was applied by analysing the responses of 35 trainees to a 25-item test with five answer options. The data were processed using ANATES Version 4.0.2 and verified through item-level analysis covering reliability, item difficulty, discrimination power, and distractor effectiveness. The results showed an average raw score of 18.94 out of 25 items, equivalent to 75.77% correct responses. The internal reliability coefficient was 0.44, indicating low to moderate score consistency and requiring cautious interpretation. The difficulty analysis identified 9 moderate items, 13 easy items, and 3 very easy items, with no difficult or very difficult items. The discrimination analysis showed that 11 items had very good discrimination, 1 item was good, 2 items were fair, and 11 items were poor. Distractor analysis indicated that several incorrect options did not function effectively, especially option E, which was repeatedly not selected by participants. Items 6, 7, 12, 13, 19, and 25 require major revision or replacement, while items 2, 4, 10, 16, and 22 require revision. The instrument provides useful preliminary information about trainees' learning outcomes but is not yet sufficiently stable as a measurement tool. Future improvement should focus on balancing item difficulty, strengthening discrimination power, revising distractors based on common misconceptions, and conducting expert-based content validation.*

**Keywords:** ANATES, electric vehicle training, item analysis, multiple-choice test, vocational assessment

## 1. PENDAHULUAN

Perkembangan kendaraan listrik telah mengubah kebutuhan kompetensi pada pendidikan dan pelatihan vokasi otomotif. Kompetensi teknisi tidak lagi hanya berpusat pada perawatan mekanis kendaraan bermesin pembakaran internal, tetapi bergeser ke penguasaan sistem baterai, keselamatan tegangan tinggi, elektronika

daya, motor listrik, diagnostik, serta prosedur konversi kendaraan. Perubahan ini sejalan dengan penguatan pendidikan vokasi berbasis capaian yang menuntut asesmen mampu merepresentasikan keterampilan kerja aktual dan kebutuhan industri yang terus berkembang [1]-[4].

Dalam konteks tersebut, pelatihan Servis dan Konversi Kendaraan Listrik perlu didukung oleh instrumen penilaian yang mampu mengukur penguasaan kompetensi peserta secara akurat. Instrumen asesmen akhir tidak hanya berfungsi memberikan nilai, tetapi juga menjadi dasar untuk mengidentifikasi ketercapaian pembelajaran, kelemahan peserta, dan kebutuhan perbaikan proses pelatihan. Asesmen vokasi yang baik perlu selaras dengan domain kompetensi, mencakup aspek pengetahuan dan keterampilan yang relevan, serta menghasilkan informasi yang dapat dipertanggungjawabkan untuk pengambilan keputusan pembelajaran [5], [6].

Kualitas instrumen penilaian dapat dievaluasi melalui analisis butir berbasis Teori Tes Klasik. Analisis ini menelaah karakteristik setiap soal berdasarkan data respons peserta, terutama reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Reliabilitas menggambarkan konsistensi skor tes, tingkat kesukaran menunjukkan proporsi peserta yang menjawab benar, daya pembeda menunjukkan kemampuan butir membedakan peserta berperforma tinggi dan rendah, sedangkan pengecoh menunjukkan sejauh mana pilihan salah berfungsi sebagai alternatif jawaban yang masuk akal [7]-[11].

Pada soal pilihan ganda, pengecoh yang tidak berfungsi dapat menurunkan kualitas pengukuran karena jawaban benar menjadi terlalu mudah dikenali. Pengecoh yang baik seharusnya merepresentasikan miskonsepsi, kesalahan prosedural, atau pemahaman yang belum lengkap sehingga dapat memberikan informasi diagnostik bagi instruktur. Kajian terbaru menegaskan bahwa efisiensi pengecoh berkaitan dengan indeks kesukaran dan daya pembeda, sedangkan kompetensi penyusunan opsi jawaban juga menjadi faktor penting dalam mutu soal pilihan ganda. Karena itu, evaluasi pengecoh perlu menjadi bagian penting dalam pengembangan bank soal [11]-[13].

Walaupun analisis butir telah banyak digunakan pada berbagai konteks pendidikan, evaluasi empiris terhadap instrumen spesifik pada pelatihan Servis dan Konversi Kendaraan Listrik masih terbatas. Kebaruan penelitian ini terletak pada fokusnya terhadap instrumen asesmen akhir yang digunakan pada konteks pelatihan vokasi otomotif berbasis kendaraan listrik, yaitu bidang yang relatif baru dan menuntut pembaruan kompetensi secara cepat. Evaluasi terhadap instrumen khusus ini menjadi penting karena keputusan mengenai keberhasilan peserta pelatihan dapat menjadi kurang akurat apabila soal terlalu mudah, tidak mampu membedakan peserta, atau memiliki pengecoh yang tidak bekerja secara efektif.

Urgensi empiris penelitian ini juga berkaitan dengan keberlanjutan program pelatihan. Apabila instrumen yang sama digunakan kembali tanpa pemeriksaan kualitas butir, maka hasil asesmen berisiko memberikan gambaran yang terlalu optimistik terhadap penguasaan kompetensi peserta. Kondisi tersebut dapat menghambat perbaikan pembelajaran karena instruktur tidak memperoleh informasi yang cukup tajam mengenai kompetensi yang belum dikuasai peserta. Oleh karena itu, analisis butir diperlukan sebagai bagian dari penjaminan mutu asesmen dan pengembangan bank soal pada pelatihan kendaraan listrik.

Berdasarkan kesenjangan tersebut, penelitian ini bertujuan mengevaluasi kualitas empiris instrumen asesmen akhir pelatihan Servis dan Konversi Kendaraan Listrik. Secara khusus, penelitian ini menjawab tiga pertanyaan: (1) bagaimana tingkat reliabilitas instrumen tes; (2) bagaimana distribusi tingkat kesukaran dan daya pembeda butir; dan (3) bagaimana efektivitas pengecoh pada soal pilihan ganda. Hasil penelitian diharapkan menjadi dasar revisi instrumen agar penilaian hasil belajar peserta pelatihan menjadi lebih akurat, adil, dan relevan dengan kompetensi kendaraan listrik.

## **2. METODE PENELITIAN**

### **Desain Penelitian**

Penelitian ini menggunakan desain kuantitatif deskriptif untuk mengevaluasi kualitas teknis instrumen penilaian akhir berbentuk pilihan ganda. Desain ini dipilih karena penelitian tidak menguji pengaruh perlakuan pembelajaran, tetapi mendeskripsikan karakteristik empiris butir berdasarkan respons peserta. Indikator yang dianalisis meliputi reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Hasil analisis digunakan untuk menetapkan rekomendasi keputusan terhadap butir, yaitu dipertahankan, direvisi, atau diganti.

### **Konteks Penelitian dan Partisipan**

Penelitian dilaksanakan pada konteks pelatihan vokasi Servis dan Konversi Kendaraan Listrik. Pelatihan ini membahas konsep dasar kendaraan listrik, komponen electric powertrain, sistem baterai, keselamatan tegangan tinggi, dan prinsip dasar konversi kendaraan. Karena materi pelatihan bersifat teknis, instrumen penilaian akhir diharapkan mampu memberikan informasi yang akurat mengenai penguasaan pengetahuan dan kesiapan peserta dalam memahami prosedur kerja yang aman.

Partisipan penelitian adalah 35 peserta pelatihan yang telah menyelesaikan program dan mengikuti asesmen akhir. Jumlah ini sesuai dengan keluaran data mentah ANATES yang mencatat jumlah subjek sebanyak 35

orang. Seluruh respons peserta yang tersedia digunakan dalam analisis sehingga tidak dilakukan pengambilan sampel acak. Dengan demikian, hasil penelitian merepresentasikan kualitas empiris instrumen pada kelas pelatihan ini secara spesifik.

Ukuran sampel sebesar 35 peserta merupakan keterbatasan penting dalam penelitian ini. Dalam pendekatan Teori Tes Klasik, estimasi tingkat kesukaran, daya pembeda, dan reliabilitas dapat menjadi kurang stabil ketika jumlah responden relatif kecil. Oleh karena itu, hasil analisis perlu digeneralisasikan secara hati-hati dan sebaiknya dikonfirmasi kembali melalui uji coba dengan jumlah peserta yang lebih besar pada pelaksanaan pelatihan berikutnya.

### Instrumen Penilaian

Objek analisis adalah instrumen asesmen akhir yang terdiri atas 25 butir soal pilihan ganda dengan lima opsi jawaban pada setiap butir. Setiap soal memiliki satu kunci jawaban, sedangkan pilihan jawaban dikodekan sebagai A, B, C, D, dan E. Sistem penskoran bersifat dikotomis, yaitu jawaban benar diberi skor 1 dan jawaban salah diberi skor 0. Dalam konversi nilai, setiap jawaban benar berbobot 4 poin sehingga skor maksimum adalah 100.

Instrumen dikembangkan berdasarkan materi pelatihan dan tujuan pembelajaran. Domain kompetensi yang direpresentasikan dalam instrumen meliputi konsep dasar kendaraan listrik, komponen powertrain listrik, sistem baterai, keselamatan tegangan tinggi, dan prinsip konversi kendaraan. Penelitian ini belum melakukan validasi isi melalui penilaian ahli, sehingga hasil analisis dibatasi pada bukti empiris dari respons peserta dan belum dapat diperlakukan sebagai bukti validitas yang komprehensif.

Tabel 1. Domain Kompetensi yang Direpresentasikan dalam Instrumen Asesmen Akhir

| Domain kompetensi              | Ruang lingkup penilaian                                                  | Fokus butir yang diharapkan                                          |
|--------------------------------|--------------------------------------------------------------------------|----------------------------------------------------------------------|
| Konsep dasar kendaraan listrik | Prinsip dasar sistem kendaraan listrik                                   | Memahami istilah kunci, komponen utama, dan fungsi sistem            |
| Komponen powertrain listrik    | Motor, kontroler, pengkabelan, dan sistem pendukung                      | Mengidentifikasi fungsi dan hubungan antarkomponen kendaraan listrik |
| Sistem baterai                 | Karakteristik baterai, pengisian daya, proteksi, dan pemeliharaan        | Memahami pengoperasian baterai, faktor risiko, dan aspek keselamatan |
| Keselamatan tegangan tinggi    | Prosedur kerja aman dan pencegahan risiko                                | Memilih tindakan, alat, dan prosedur keselamatan yang tepat          |
| Prinsip konversi kendaraan     | Prosedur dasar konversi kendaraan konvensional menjadi kendaraan listrik | Memahami tahapan konversi, pemilihan komponen, dan logika prosedural |

Tabel 2. Karakteristik Umum Instrumen Asesmen yang Dianalisis

| Komponen               | Deskripsi                                                             |
|------------------------|-----------------------------------------------------------------------|
| Mata pelatihan         | Servis dan Konversi Kendaraan Listrik                                 |
| Jenis asesmen          | Asesmen akhir                                                         |
| Format soal            | Tes pilihan ganda                                                     |
| Jumlah peserta         | 35 peserta pelatihan                                                  |
| Jumlah butir soal      | 25 butir                                                              |
| Jumlah pilihan jawaban | 5 pilihan pada setiap butir                                           |
| Format jawaban         | Pilihan A, B, C, D, dan E dengan satu jawaban benar pada setiap butir |

| Komponen           | Deskripsi                                                               |
|--------------------|-------------------------------------------------------------------------|
| Sistem penskoran   | Skor dikotomis: 1 untuk jawaban benar dan 0 untuk jawaban salah         |
| Konversi skor      | Setiap butir berbobot 4 poin; skor maksimum 100                         |
| Perangkat lunak    | ANATES Versi 4.0.2                                                      |
| Indikator analisis | Reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh |

### Sumber Data dan Prosedur Analisis

Sumber data penelitian adalah file respons peserta yang diolah melalui ANATES Versi 4.0.2. Data memuat kunci jawaban, respons setiap peserta, dan skor setiap butir. Data dikumpulkan melalui teknik dokumentasi terhadap hasil asesmen akhir yang telah tersedia, sehingga penelitian ini tidak melibatkan wawancara, observasi, atau eksperimen baru.

Analisis dilakukan melalui enam tahap. Pertama, kunci jawaban dan respons peserta dimasukkan ke dalam perangkat lunak. Kedua, jawaban peserta diverifikasi dan diberi skor dikotomis. Ketiga, skor total individu dihitung berdasarkan jumlah jawaban benar dan konversi nilai. Keempat, reliabilitas tes dianalisis untuk melihat konsistensi skor. Kelima, tingkat kesukaran dan daya pembeda butir dihitung. Keenam, efektivitas pengecoh ditinjau dari pilihan jawaban salah yang dipilih atau tidak dipilih peserta.

Indeks kesukaran butir dihitung berdasarkan proporsi peserta yang menjawab benar sebagaimana dirumuskan pada Persamaan (1).

$$P_i = B_i / N \quad (1)$$

Pada Persamaan (1),  $P_i$  adalah indeks kesukaran butir ke- $i$ ,  $B_i$  adalah jumlah peserta yang menjawab benar pada butir ke- $i$ , dan  $N$  adalah jumlah seluruh peserta.

Indeks daya pembeda dihitung dengan membandingkan proporsi jawaban benar pada kelompok atas dan kelompok bawah sebagaimana ditunjukkan pada Persamaan (2).

$$D_i = (B_A / n_A) - (B_B / n_B) \quad (2)$$

Pada Persamaan (2),  $D_i$  adalah indeks daya pembeda,  $B_A$  dan  $B_B$  adalah jumlah peserta yang menjawab benar pada kelompok atas dan bawah, sedangkan  $n_A$  dan  $n_B$  adalah jumlah peserta pada masing-masing kelompok.



Gambar 1. Alur Prosedur Analisis Butir Soal Menggunakan ANATES Versi 4.0.2

Gambar 1 menunjukkan alur analisis yang dimulai dari dokumentasi lembar jawaban peserta, penyusunan kunci jawaban, penskoran dikotomis, input ke ANATES, analisis indikator psikometrik, verifikasi nomor butir, dan penetapan keputusan terhadap setiap butir soal.

Tabel 3. Kriteria Interpretasi Hasil Analisis Butir Soal

| Indikator              | Fokus interpretasi                                               | Implikasi keputusan                                                                    |
|------------------------|------------------------------------------------------------------|----------------------------------------------------------------------------------------|
| Koefisien reliabilitas | Konsistensi skor yang dihasilkan oleh tes                        | Digunakan untuk mengevaluasi stabilitas instrumen secara keseluruhan                   |
| Tingkat kesukaran      | Proporsi peserta yang menjawab benar                             | Butir sedang umumnya diutamakan; butir terlalu mudah atau terlalu sukar perlu ditinjau |
| Daya pembeda           | Kemampuan butir membedakan peserta berperforma tinggi dan rendah | Butir rendah perlu direvisi atau diganti                                               |
| Efektivitas pengecoh   | Fungsi pilihan jawaban salah dalam menarik respons peserta       | Pengecoh yang tidak dipilih perlu diganti dengan alternatif yang lebih masuk akal      |
| Status akhir butir     | Interpretasi gabungan dari indikator butir                       | Butir diklasifikasikan sebagai dipertahankan, direvisi, atau diganti                   |

### 3. HASIL DAN PEMBAHASAN

#### Gambaran Umum Kinerja Tes

Data mentah ANATES mengonfirmasi bahwa instrumen terdiri atas 25 butir pilihan ganda yang dikerjakan oleh 35 peserta pelatihan. Setiap butir memiliki lima opsi jawaban dan satu kunci jawaban. Verifikasi awal dilakukan untuk memastikan bahwa seluruh interpretasi hanya menggunakan nomor butir 1 sampai 25. Nomor butir di luar rentang tersebut diperlakukan sebagai kesalahan pelaporan dan tidak dimasukkan dalam analisis.

Skor mentah peserta berada pada rentang 12 sampai 24 jawaban benar dari 25 butir. Rata-rata skor mentah sebesar 18,94 atau setara dengan 75,77% jawaban benar. Pola ini menunjukkan bahwa tes relatif mudah bagi kelompok peserta yang mengikuti pelatihan. Dalam asesmen akhir pelatihan vokasi, kondisi seperti ini perlu diperhatikan karena tes yang terlalu mudah dapat mengurangi kemampuan instrumen dalam memberikan informasi diagnostik pada rentang kompetensi yang lebih tinggi.

Tabel 4. Ringkasan Deskriptif Kinerja Tes

| Indikator                          | Hasil  |
|------------------------------------|--------|
| Jumlah peserta                     | 35     |
| Jumlah butir soal                  | 25     |
| Skor mentah minimum                | 12     |
| Skor mentah maksimum               | 24     |
| Rata-rata skor mentah              | 18,94  |
| Rata-rata persentase jawaban benar | 75,77% |
| Skor maksimum setelah konversi     | 100    |

#### Reliabilitas Instrumen Tes

Koefisien reliabilitas internal instrumen adalah 0,44. Nilai ini menunjukkan bahwa konsistensi skor tes masih rendah hingga sedang. Temuan ini merupakan aspek kritis karena reliabilitas yang rendah menunjukkan bahwa skor peserta belum cukup stabil untuk dijadikan satu-satunya dasar pengambilan keputusan tentang

penguasaan kompetensi. Dengan kata lain, skor tinggi peserta belum sepenuhnya dapat dipastikan merepresentasikan perbedaan kemampuan yang konsisten antarindividu.

Rendahnya reliabilitas dapat dipengaruhi oleh dominasi butir mudah dan sangat mudah. Ketika sebagian besar peserta menjawab butir yang sama dengan benar, variasi respons menjadi terbatas sehingga kontribusi butir terhadap konsistensi total skor menurun. Dampaknya, tes menjadi kurang sensitif dalam membedakan peserta yang benar-benar menguasai kompetensi prosedural dari peserta yang hanya menguasai pengetahuan dasar. Oleh karena itu, instrumen ini belum layak diposisikan sebagai alat ukur final yang stabil tanpa revisi butir dan uji coba ulang.

Temuan reliabilitas ini juga tidak dapat diperlakukan sebagai bukti validitas. Reliabilitas hanya menjelaskan konsistensi skor, sedangkan validitas memerlukan bukti tambahan mengenai kesesuaian isi tes dengan indikator kompetensi, representasi domain materi, dan interpretasi skor. Karena penelitian ini belum melibatkan penilaian ahli, klaim yang dapat diberikan terbatas pada kualitas empiris butir berdasarkan respons peserta.

### Tingkat Kesukaran Butir Soal

Analisis tingkat kesukaran menunjukkan bahwa 9 butir tergolong sedang, 13 butir tergolong mudah, dan 3 butir tergolong sangat mudah. Tidak terdapat butir yang termasuk kategori sukar maupun sangat sukar. Distribusi ini menunjukkan bahwa instrumen belum memiliki keseimbangan tingkat kesukaran yang memadai.

Tabel 5. Distribusi Kategori Tingkat Kesukaran Butir Soal

| Kategori tingkat kesukaran | Jumlah butir | Nomor butir                                    | Interpretasi                 |
|----------------------------|--------------|------------------------------------------------|------------------------------|
| Sedang                     | 9            | 1, 5, 8, 9, 14, 15, 20, 23, 24                 | Sesuai untuk asesmen         |
| Mudah                      | 13           | 2, 3, 4, 6, 10, 11, 16, 17, 18, 19, 21, 22, 25 | Perlu ditinjau               |
| Sangat mudah               | 3            | 7, 12, 13                                      | Revisi besar/ganti           |
| Sukar                      | 0            | -                                              | Tidak ada butir sukar        |
| Sangat sukar               | 0            | -                                              | Tidak ada butir sangat sukar |

Dominasi butir mudah berimplikasi pada kemungkinan terjadinya estimasi penguasaan kompetensi yang terlalu tinggi. Dalam konteks pelatihan kendaraan listrik, asesmen akhir seharusnya tidak hanya mengukur pengenalan istilah, tetapi juga pemahaman prosedural, analisis risiko keselamatan, dan pengambilan keputusan teknis. Oleh karena itu, beberapa butir perlu direvisi dengan menaikkan kompleksitas kognitif dan mengaitkan soal dengan skenario kerja nyata.

### Daya Pembeda Butir Soal

Hasil analisis daya pembeda menunjukkan bahwa 11 butir berkategori sangat baik, 1 butir baik, 2 butir cukup, dan 11 butir buruk. Jumlah butir berkategori buruk yang sama besar dengan jumlah butir sangat baik menunjukkan kualitas instrumen yang belum merata. Butir dengan daya pembeda buruk perlu diprioritaskan karena tidak memberikan informasi yang cukup untuk membedakan peserta berkemampuan tinggi dan rendah.

Butir 7, 12, dan 13 memiliki masalah ganda karena tergolong sangat mudah dan memiliki daya pembeda buruk. Butir 13 bahkan dijawab benar oleh seluruh peserta, sehingga tidak memberi kontribusi terhadap perbedaan kemampuan. Butir 25 juga perlu diperiksa secara khusus karena memiliki indeks daya pembeda negatif, yang dapat mengindikasikan ambiguitas redaksi, kesalahan kunci, atau ketidaksesuaian antara materi soal dan pengalaman belajar peserta.

Tabel 6. Distribusi Kategori Daya Pembeda Butir Soal

| Kategori daya pembeda | Jumlah butir | Nomor butir                            | Rekomendasi tindakan                                    |
|-----------------------|--------------|----------------------------------------|---------------------------------------------------------|
| Sangat baik           | 11           | 3, 5, 8, 9, 14, 15, 17, 20, 21, 23, 24 | Dipertahankan dengan peninjauan kecil terhadap pengecoh |
| Baik                  | 1            | 1                                      | Dipertahankan                                           |
| Cukup                 | 2            | 11, 18                                 | Direvisi atau ditinjau kembali                          |
| Buruk                 | 11           | 2, 4, 6, 7, 10, 12, 13, 16, 19, 22, 25 | Direvisi secara substansial atau diganti                |

**Ringkasan Analisis Tingkat Butir**

Tabel 7. Analisis Tingkat Butir Berdasarkan Matriks Respons ANATES (Butir 1-12)

| Butir | Kunci | Benar | Kesukaran           | Daya pembeda       | Pengecoh tidak berfungsi | Keputusan            |
|-------|-------|-------|---------------------|--------------------|--------------------------|----------------------|
| 1     | B     | 19    | 0,54 (sedang)       | 0,33 (baik)        | E                        | Dipertahankan        |
| 2     | B     | 26    | 0,74 (mudah)        | 0,11 (buruk)       | D, E                     | Direvisi             |
| 3     | C     | 27    | 0,77 (mudah)        | 0,44 (sangat baik) | E                        | Tinjauan minor       |
| 4     | D     | 28    | 0,80 (mudah)        | 0,11 (buruk)       | E                        | Direvisi             |
| 5     | D     | 21    | 0,60 (sedang)       | 0,44 (sangat baik) | A, E                     | Dipertahankan        |
| 6     | B     | 27    | 0,77 (mudah)        | 0,00 (buruk)       | A, E                     | Revisi besar         |
| 7     | B     | 33    | 0,94 (sangat mudah) | 0,00 (buruk)       | A, D, E                  | Diganti/revisi besar |
| 8     | C     | 24    | 0,69 (sedang)       | 0,44 (sangat baik) | E                        | Dipertahankan        |
| 9     | A     | 18    | 0,51 (sedang)       | 0,44 (sangat baik) | E                        | Dipertahankan        |
| 10    | D     | 27    | 0,77 (mudah)        | 0,11 (buruk)       | E                        | Direvisi             |
| 11    | C     | 29    | 0,83 (mudah)        | 0,22 (cukup)       | E                        | Revisi minor         |
| 12    | C     | 33    | 0,94 (sangat mudah) | 0,00 (buruk)       | A, D, E                  | Diganti/revisi besar |

Tabel 8. Lanjutan Analisis Tingkat Butir Berdasarkan Matriks Respons ANATES (Butir 13-25)

| Butir | Kunci | Benar | Kesukaran           | Daya pembeda       | Pengecoh tidak berfungsi | Keputusan            |
|-------|-------|-------|---------------------|--------------------|--------------------------|----------------------|
| 13    | C     | 35    | 1,00 (sangat mudah) | 0,00 (buruk)       | A, B, D, E               | Diganti              |
| 14    | B     | 23    | 0,66 (sedang)       | 0,56 (sangat baik) | E                        | Dipertahankan        |
| 15    | D     | 20    | 0,57 (sedang)       | 0,56 (sangat baik) | E                        | Dipertahankan        |
| 16    | B     | 29    | 0,83 (mudah)        | 0,11 (buruk)       | D, E                     | Direvisi             |
| 17    | B     | 29    | 0,83 (mudah)        | 0,44 (sangat baik) | C, E                     | Tinjauan minor       |
| 18    | B     | 31    | 0,89 (mudah)        | 0,22 (cukup)       | D, E                     | Revisi minor         |
| 19    | B     | 31    | 0,89 (mudah)        | 0,00 (buruk)       | A, D, E                  | Revisi besar         |
| 20    | C     | 23    | 0,66 (sedang)       | 0,67 (sangat baik) | E                        | Dipertahankan        |
| 21    | B     | 25    | 0,71 (mudah)        | 0,44 (sangat baik) | E                        | Tinjauan minor       |
| 22    | B     | 31    | 0,89 (mudah)        | 0,11 (buruk)       | D, E                     | Direvisi             |
| 23    | C     | 22    | 0,63 (sedang)       | 0,44 (sangat baik) | E                        | Dipertahankan        |
| 24    | B     | 23    | 0,66 (sedang)       | 0,44 (sangat baik) | E                        | Dipertahankan        |
| 25    | B     | 29    | 0,83 (mudah)        | -0,22 (buruk)      | C, D, E                  | Diganti/revisi besar |

Tabel 7 dan Tabel 8 menunjukkan bahwa butir yang paling stabil adalah butir dengan kesukaran sedang dan daya pembeda baik atau sangat baik, yaitu butir 1, 5, 8, 9, 14, 15, 20, 23, dan 24. Butir-butir tersebut dapat dipertahankan, meskipun pengecoh yang tidak berfungsi tetap perlu diperbaiki agar pilihan jawaban lebih homogen dan meyakinkan.

Sebaliknya, butir 6, 7, 12, 13, 19, dan 25 memerlukan revisi besar atau penggantian. Butir-butir tersebut menunjukkan kombinasi masalah berupa tingkat kesukaran terlalu rendah, daya pembeda buruk, pengecoh tidak berfungsi, atau daya pembeda negatif. Butir 25 harus diperiksa kembali dari sisi kunci jawaban dan redaksi karena peserta berperforma rendah lebih sering menjawab benar dibandingkan peserta berperforma tinggi.

#### Efektivitas Pengecoh dan Implikasi Pedagogis

Analisis pengecoh menunjukkan bahwa beberapa opsi jawaban salah tidak dipilih oleh satu pun peserta. Opsi E merupakan pengecoh yang paling sering tidak berfungsi. Pola ini menunjukkan bahwa pilihan kelima pada banyak butir kemungkinan terlalu jauh dari konteks soal, terlalu mudah dieliminasi, atau tidak merepresentasikan kesalahan berpikir yang realistis. Pada butir 13, seluruh peserta memilih jawaban benar sehingga semua pengecoh tidak berfungsi.

Secara pedagogis, banyaknya butir mudah dan pengecoh yang tidak berfungsi dapat disebabkan oleh beberapa faktor. Pertama, sebagian soal kemungkinan masih berada pada level mengingat atau mengenali istilah, belum cukup menuntut peserta menerapkan konsep dalam situasi kerja. Kedua, materi yang sangat ditekankan selama pelatihan dapat membuat jawaban benar terlalu mudah dikenali apabila soal hanya mengulang informasi eksplisit dari bahan ajar. Ketiga, pengecoh belum dikembangkan dari miskonsepsi nyata peserta, misalnya

kesalahan memahami prosedur isolasi tegangan tinggi, urutan pemeriksaan baterai, atau pemilihan komponen konversi yang aman. Kajian pembelajaran vokasi juga menunjukkan bahwa pengalaman belajar yang aktif, kontekstual, dan dekat dengan praktik otomotif dapat memperkuat capaian kognitif peserta; oleh sebab itu, butir asesmen perlu dikembangkan dari skenario pembelajaran dan situasi kerja yang benar-benar dialami peserta [14], [15].

Temuan ini menunjukkan bahwa revisi instrumen tidak cukup dilakukan pada redaksi soal, tetapi juga perlu dikaitkan dengan evaluasi proses pembelajaran. Instruktur perlu meninjau apakah materi telah disampaikan secara seimbang antara konsep dasar, prosedur keselamatan, dan penerapan kasus. Bank soal juga perlu dikembangkan dari skenario praktik, kesalahan prosedural yang sering terjadi, serta situasi diagnostik yang menuntut peserta memilih tindakan teknis yang paling tepat. Dengan demikian, revisi butir dapat sekaligus menjadi umpan balik untuk memperbaiki desain pembelajaran pada pelatihan berikutnya.

Berdasarkan gabungan indikator, prioritas revisi diarahkan pada tiga kelompok. Kelompok pertama adalah butir yang perlu dipertahankan dengan peninjauan kecil, terutama butir berkategori sedang dan memiliki daya pembeda baik. Kelompok kedua adalah butir yang perlu direvisi, yaitu butir mudah dengan daya pembeda rendah atau pengecoh lemah. Kelompok ketiga adalah butir yang perlu diganti atau direvisi besar, terutama butir sangat mudah, daya pembeda nol atau negatif, dan pengecoh tidak berfungsi.

#### 4. KESIMPULAN

Hasil analisis menunjukkan bahwa instrumen asesmen akhir pelatihan Servis dan Konversi Kendaraan Listrik telah memberikan informasi awal mengenai capaian peserta, tetapi belum cukup kuat untuk digunakan kembali tanpa perbaikan. Masalah utama instrumen terletak pada rendahnya konsistensi skor, dominasi butir mudah, banyaknya butir dengan daya pembeda buruk, serta pengecoh yang tidak berfungsi. Kondisi tersebut menunjukkan bahwa tes belum sepenuhnya mampu membedakan penguasaan kompetensi peserta secara stabil, terutama pada kompetensi terapan yang penting dalam servis dan konversi kendaraan listrik.

Makna utama temuan ini adalah bahwa perbaikan asesmen perlu dilakukan secara sistematis, bukan sekadar mengganti beberapa soal. Butir yang sangat mudah, tidak membedakan peserta, atau memiliki daya pembeda negatif perlu diganti dengan soal berbasis skenario kerja. Pengecoh perlu dirancang berdasarkan miskonsepsi dan kesalahan prosedural yang mungkin muncul dalam praktik, sedangkan butir yang sudah memiliki kualitas baik dapat dipertahankan dengan peninjauan minor. Selain itu, hasil reliabilitas yang rendah menegaskan perlunya uji coba ulang setelah revisi agar stabilitas skor dapat diperiksa kembali.

Penelitian ini juga menegaskan bahwa statistik butir belum cukup untuk menyatakan instrumen valid secara menyeluruh. Validasi isi oleh ahli bidang kendaraan listrik, ahli evaluasi pembelajaran, dan praktisi pelatihan perlu dilakukan untuk memastikan kesesuaian butir dengan indikator kompetensi. Pada pelaksanaan berikutnya, instrumen sebaiknya diuji pada jumlah peserta yang lebih besar agar estimasi parameter butir lebih stabil. Dengan langkah tersebut, instrumen asesmen dapat dikembangkan menjadi alat evaluasi yang lebih akurat, adil, dan mendukung penjaminan mutu pelatihan kendaraan listrik secara berkelanjutan.

#### DAFTAR PUSTAKA

- [1] D. Iliescu, S. Greiff, and A. Ion, "Evidence based approaches for enhancing vocational education worldwide," *npj Science of Learning*, vol. 10, no. 1, Art. no. 20, 2025, doi: 10.1038/s41539-025-00317-2.
- [2] C. Osatis and C. Asavanirandom, "An exploring human resource development in small and medium enterprises in response to electric vehicle industry development," *World Electric Vehicle Journal*, vol. 13, no. 6, Art. no. 98, 2022, doi: 10.3390/wevj13060098.
- [3] H.-C. Yu, T.-J. Hsueh, T.-Y. Wu, C. Liu, C.-W. Liao, and Y.-K. Fu, "A competency framework for electric vehicle maintenance technicians: Addressing the environmental, social, and governance (ESG) imperatives of the BEV industry," *World Electric Vehicle Journal*, vol. 16, no. 6, Art. no. 314, 2025, doi: 10.3390/wevj16060314.
- [4] R. A. P. Hardiyanta, Hermanto, A. Kurniawan, Purnawan, and I. E. Prakoso, "Development of the automotive technology vocational education curriculum based on current needs of the automotive industry," *Jurnal Pendidikan Vokasi Otomotif*, vol. 6, no. 2, 2024, doi: 10.21831/jpvo.v6i2.73156.
- [5] S. R. M. Yusop, M. S. Rasul, R. M. Yasin, and H. U. Hashim, "Identifying and validating vocational skills domains and indicators in classroom assessment practices in TVET," *Sustainability*, vol. 15, no. 6, Art. no. 5195, 2023, doi: 10.3390/su15065195.

- 
- [6] S. R. M. Yusop, M. S. Rasul, R. M. Yasin, H. U. Hashim, and N. A. Jalaludin, "An assessment approaches and learning outcomes in technical and vocational education: A systematic review using PRISMA," *Sustainability*, vol. 14, no. 9, Art. no. 5225, 2022, doi: 10.3390/su14095225.
- [7] Z. Zhang, J. Tian, Z. Zhao, W. Zhou, F. Sun, Y. Que, and X. He, "Factors influencing vocational education and training teachers' professional competence based on a large-scale diagnostic method: A decade of data from China," *Sustainability*, vol. 14, no. 23, Art. no. 15871, 2022, doi: 10.3390/su142315871.
- [8] M. S. Mukhter, J. N. Yunus, and F. K. C. Yuet, "Evaluation of automotive technology diploma program at Malaysian vocational college: Validity and reliability," *International Journal of Academic Research in Business and Social Sciences*, vol. 13, no. 2, pp. 542-549, 2023, doi: 10.6007/IJARBS/v13-i2/16371.
- [9] K. M. Salleh, N. L. Sulaiman, and G. W. Gloeckner, "Exploring test concept and measurement through validity and reliability process in TVET research: Guideline for the novice researcher," *Journal of Technical Education and Training*, vol. 15, no. 1, 2023, doi: 10.30880/jtet.2023.15.01.022.
- [10] M. R. Ridwan, E. Istiyono, and W. Widiastuti, "Test items analysis of mathematical problem solving ability using a classical test theory approach," *Jurnal Pendidikan MIPA*, vol. 22, no. 1, pp. 98-111, 2021, doi: 10.23960/jpmipa/v22i1.pp98-111.
- [11] S. Nurjanah, M. Iqbal, Zafrullah, M. N. Mahmud, D. S. F. Seran, I. K. Suardi, and L. Arriza, "Psychometric quality of multiple-choice tests under classical test theory (CTT): AnBuso, Iteman, and R," *Jurnal Penelitian dan Evaluasi Pendidikan*, vol. 28, no. 2, pp. 161-172, 2024, doi: 10.21831/pep.v28i2.71542.
- [12] A. A. Rezigalla et al., "Item analysis: The impact of distractor efficiency on the difficulty index and discrimination power of multiple-choice items," *BMC Medical Education*, vol. 24, no. 1, Art. no. 445, 2024, doi: 10.1186/s12909-024-05433-y.
- [13] P. K. Kissi, D. Baidoo-Anu, E. Anane, and R. K. Annan-Brew, "Teachers' test construction competencies in examination-oriented educational system: Exploring teachers' multiple-choice test construction competence," *Frontiers in Education*, vol. 8, Art. no. 1154592, 2023, doi: 10.3389/educ.2023.1154592.
- [14] F. Dahalan, N. Alias, and M. S. N. Shaharom, "Gamification and game based learning for vocational education and training: A systematic literature review," *Education and Information Technologies*, vol. 29, pp. 1279-1317, 2024, doi: 10.1007/s10639-022-11548-w.
- [15] I. D. Azizan, M. Alias, and M. Z. Mustafa, "Effect of game-based learning in vehicle air-conditioning course on cognitive and affective skills of vocational students," *Journal of Technical Education and Training*, vol. 13, no. 3, pp. 146-154, 2021, doi: 10.30880/jtet.2021.13.03.014.