



**EVALUASI DAN PENGEMBANGAN INSTRUMEN PENILAIAN KOMPETENSI  
TKJ DI SMKN 1 SAMBENG**

**Muhammad Wildan Habibi<sup>1</sup>, Yusuf Arif Rahman<sup>2</sup>, Briant Tedja<sup>3</sup>, Julio Anggara Putra<sup>4</sup>,  
Rafi Fernandika Putra Tweede<sup>5</sup>, Muhammad Zaky Ardiansyah<sup>6</sup>, Muh Yanuar Ismail  
Akbar<sup>7</sup>**

Universitas Negeri Surabaya

e-mail: [mohammadhabibi@unesa.ac.id](mailto:mohammadhabibi@unesa.ac.id)

Diterima: 24/04/2026 Direvisi: 05/05/2026; Diterbitkan: 11/05/2026

**ABSTRAK**

Penilaian kompetensi kejuruan di SMK memegang peran strategis dalam menjamin kesiapan lulusan memasuki dunia kerja, namun kualitas instrumen yang digunakan kerap belum terstandarisasi secara ilmiah. Tanpa validasi sistematis, instrumen penilaian berisiko menghasilkan data kompetensi yang bias dan tidak dapat dipertanggungjawabkan, sehingga keputusan kelulusan maupun sertifikasi siswa menjadi tidak akurat. Konteks SMK khususnya program Teknik Komputer dan Jaringan (TKJ) menjadikan persoalan ini semakin krusial karena penilaian bersifat praktik dengan rubrik yang kompleks dan rentan terhadap subjektivitas penilai. Penelitian ini bertujuan mengevaluasi validitas dan reliabilitas instrumen penilaian TKJ di SMKN 1 Sambeng, serta merancang instrumen baru berdasarkan temuan evaluasi. Pendekatan kuantitatif dengan desain validasi instrumen digunakan melalui tiga fase: (1) validasi konten oleh lima ahli menggunakan Aiken's V, (2) uji reliabilitas empiris terhadap 30 siswa menggunakan *Cronbach's Alpha* beserta uji korelasi item-total sebagai bukti signifikansi konsistensi internal, dan (3) perancangan serta evaluasi ulang instrumen baru. Hasil menunjukkan instrumen lama memiliki validitas konten di batas minimum ( $CVI = 0,82$ ) dengan enam aspek di bawah ambang batas  $V \geq 0,75$ , serta reliabilitas yang belum memadai ( $\alpha = 0,502$ ). Instrumen baru yang dirancang dengan bobot seimbang (5 poin per sub-aspek) dan deskriptor diskret tiga level (0/3/5) menghasilkan peningkatan signifikan:  $CVI = 0,94$  dan *Cronbach's Alpha* = 0,844 (kategori Sangat Tinggi), dengan seluruh korelasi item-total di atas 0,65 yang mengindikasikan konsistensi internal yang kuat secara statistik. Penelitian ini memberi implikasi praktis bagi guru dan sekolah kejuruan: instrumen tervalidasi dapat langsung diadopsi sebagai acuan penilaian praktik TKJ yang adil dan terstandar, sekaligus menjadi model prosedur validasi yang dapat direplikasi pada program keahlian lain di SMK.

**Kata Kunci:** validitas konten, reliabilitas, instrumen penilaian, Aiken's V, Cronbach's Alpha.

**ABSTRACT**

Vocational competency assessment in vocational high schools (SMK) plays a strategic role in ensuring graduates' readiness to enter the workforce; however, the quality of assessment instruments used is often not scientifically standardized. Without systematic validation, assessment instruments risk producing biased competency data that cannot be scientifically justified, leading to inaccurate decisions regarding student graduation and certification. This issue becomes particularly critical in the context of SMK especially in the Computer and Network Engineering (TKJ) program as assessments are practice-based, involving complex rubrics that are highly susceptible to rater subjectivity. This study aims to evaluate the validity and reliability of the TKJ assessment instrument at SMKN 1 Sambeng and to design a new instrument based on the evaluation findings. A quantitative approach with an instrument



validation design was employed through three phases: (1) content validation by five experts using Aiken's V, (2) empirical reliability testing on 30 students using *Cronbach's Alpha* along with item-total correlation analysis as evidence of internal consistency significance, and (3) design and re-evaluation of the new instrument. The results indicate that the original instrument had content validity at the minimum threshold ( $CVI = 0.82$ ), with six aspects falling below the  $V \geq 0.75$  cutoff, and inadequate reliability ( $\alpha = 0.502$ ). The newly designed instrument featuring balanced scoring weights (5 points per sub-aspect) and a three-level discrete descriptor system (0/3/5) demonstrated significant improvement:  $CVI = 0.94$  and *Cronbach's Alpha* = 0.844 (categorized as Very High), with all item-total correlations exceeding 0.65, indicating statistically strong internal consistency. This study carries practical implications for teachers and vocational schools: the validated instrument can be directly adopted as a fair and standardized reference for TKJ practical assessments, while simultaneously serving as a replicable validation procedure model applicable to other vocational programs within SMK.

**Keywords:** content validity, reliability, assessment instrument, Aiken's V, Cronbach's Alpha.

## PENDAHULUAN

Pendidikan Teknik Jaringan dan Komputer memegang peranan strategis dalam pembelajaran ilmu komputer terapan. Melalui pendidikan ini, peserta didik dibekali keterampilan praktis yang dibutuhkan dalam dunia kerja, meliputi instalasi, konfigurasi, serta pemeliharaan infrastruktur jaringan, sehingga lulusan siap bersaing di industri teknologi informasi dan komunikasi (Sudira, 2021). Hal ini sejalan dengan beberapa studi yang menyatakan penguasaan kompetensi teknik jaringan secara praktis merupakan modal utama lulusan SMK dalam memasuki dunia kerja sekaligus menjadi faktor penentu kesesuaian antara kemampuan lulusan dengan tuntutan industri teknologi jaringan yang terus berkembang (Shuaili, 2025), di mana tanpa kompetensi yang memadai seseorang akan kesulitan untuk bekerja sesuai dengan bidang pendidikannya (Hakim & Kurniawati, 2022). Namun demikian, penilaian keterampilan teknik jaringan menghadirkan tantangan tersendiri yang tidak dijumpai pada penilaian konvensional. Penilaian di bidang ini tidak cukup hanya mengukur pemahaman konseptual, melainkan harus mampu mengevaluasi kemampuan peserta didik dalam memecahkan permasalahan jaringan secara nyata, mengonfigurasi perangkat dengan tepat, serta mengimplementasikan solusi secara akurat. Kompleksitas inilah yang menjadikan penilaian teknik jaringan jauh lebih menantang dibandingkan dengan penilaian berbasis pengetahuan pada umumnya.

Instrumen penilaian adalah alat komprehensif yang digunakan untuk mengumpulkan informasi tentang kemajuan belajar dan pencapaian kompetensi siswa (Khotimah & Hidayat, 2024). Dalam pendidikan teknik jaringan dan komputer, penilaian mendukung pengambilan keputusan instruksional, memberikan umpan balik diagnostik kepada siswa, dan memengaruhi strategi belajar. Untuk memastikan bahwa hasil penilaian secara akurat mewakili kemampuan teknik jaringan siswa, alat penilaian harus memenuhi kriteria kualitas kritis, terutama validitas dan reliabilitas (Rahmah & Festiyed, 2025). Validitas memastikan bahwa instrumen mengukur kompetensi teknik jaringan yang dimaksud seperti pemahaman topologi jaringan, kemampuan konfigurasi perangkat keras, atau pendekatan troubleshooting, sementara reliabilitas menunjukkan konsistensi hasil pengukuran di berbagai konteks dan penilai. Tanpa instrumen yang valid dan reliabel, hasil penilaian dapat menyebabkan kesimpulan yang tidak akurat mengenai keterampilan teknik jaringan siswa (Subando, 2022), sehingga berpotensi salah mengidentifikasi siswa yang kesulitan dan membutuhkan intervensi atau gagal mengenali kemampuan lanjutan yang memerlukan penguasaan. Dalam konteks pendidikan vokasi dan



teknik, penilaian berbasis kompetensi merupakan pendekatan yang paling banyak digunakan untuk mengevaluasi hasil belajar siswa, dan kualitas instrumen penilaian terbukti secara langsung memengaruhi mutu umpan balik yang diterima siswa maupun para pemangku kepentingan pendidikan (Yusop et al., 2022).

Meskipun implementasi teknik jaringan dan komputer dalam kurikulum SMK semakin meningkat, validasi sistematis terhadap instrumen penilaian kompetensi jaringan masih sangat diabaikan (Aliffiansyah, Ekarini, & Djuniadi, 2025). Penelitian sebelumnya telah mendokumentasikan baik kelangkaan instrumen penilaian yang valid dan reliabel dalam pendidikan berbasis kompetensi teknis dan komputer (De Ruiter & Bers, 2021), maupun kurangnya instrumen yang tervalidasi secara luas untuk praktik penilaian digital di dunia pendidikan (Viberg, Pargman, & Lilliesköld, 2024). Kesenjangan validasi ini tetap ada karena beberapa faktor: kompleksitas inheren dalam menilai kompetensi teknik jaringan yang bersifat multidimensi yang mencakup pemahaman konsep jaringan, keterampilan instalasi perangkat keras, kemampuan konfigurasi sistem, dan proses troubleshooting ditambah dengan keterbatasan waktu di lingkungan pendidikan dan kurangnya keahlian psikometrik di kalangan pendidik teknik jaringan. Yusop et al. (2023) menegaskan bahwa terdapat kesenjangan mendasar dalam praktik penilaian TVET, di mana keterampilan vokasional yang diukur kerap tidak selaras dengan capaian pembelajaran yang ditetapkan, dan konsistensi praktik penilaian guru pun masih jauh dari yang diharapkan. Akibatnya, instrumen penilaian sering kali diterapkan tanpa pembuktian empiris yang memadai mengenai kualitas pengukurannya, sehingga berpotensi menyebabkan penilaian keterampilan yang tidak akurat, umpan balik siswa yang menyesatkan, dan evaluasi hasil belajar yang tidak dapat diandalkan. Konsekuensi ini memengaruhi pemetaan tingkat individu siswa, memengaruhi keputusan pengembangan kurikulum, dan berpotensi menghasilkan lulusan yang kurang siap untuk praktik profesional (Villarroel et al., 2024). Masalah ini sangat kritis dalam konteks pendidikan menengah kejuruan, di mana hasil penilaian secara signifikan memengaruhi sertifikasi kompetensi teknik jaringan dan keputusan jalur karier. Oleh karena itu, penelitian yang secara sistematis mengevaluasi validitas dan reliabilitas instrumen penilaian teknik jaringan dan komputer telah menjadi prioritas mendesak (Yusop et al., 2023).

Penelitian ini bertujuan untuk mengevaluasi validitas dan reliabilitas instrumen penilaian teknik komputer dan jaringan di SMK Negeri 1 Sambeng, serta merancang instrumen penilaian baru apabila hasil evaluasi menunjukkan bahwa instrumen yang ada belum memenuhi standar psikometrik.

## **METODE PENELITIAN**

Penelitian ini menggunakan pendekatan kuantitatif dengan desain validasi instrumen (*instrument validation study*) yang bertujuan untuk mengevaluasi kualitas psikometrik instrumen rubrik penilaian kompetensi Teknik Komputer dan Jaringan (TKJ), yaitu Rubrik Penskoran Soal Uraian UKK Instalasi Jaringan LAN yang digunakan di SMKN 1 Sambeng. Penelitian dilakukan dalam tiga fase. Fase pertama dan kedua merupakan fase utama. Apabila hasil evaluasi menunjukkan reliabilitas belum memenuhi standar, maka akan dilanjutkan fase ketiga. Fase pertama adalah validasi konten (*content validity*) melalui penilaian oleh lima orang ahli (*expert judgment*) di bidang TKJ yang berpendidikan minimal S1 dan berpengalaman mengajar minimal lima tahun, di mana setiap aspek rubrik dinilai menggunakan lembar validasi berskala Likert 1–4 pada lima dimensi, yaitu relevansi indikator terhadap kompetensi TKJ, kejelasan rumusan indikator, kelengkapan cakupan aspek teknis, kelayakan penggunaan instrumen, serta keterlaksanaan dan kepraktisan instrumen dalam pembelajaran praktik TKJ.

Data hasil penilaian kemudian dianalisis menggunakan dua indeks validitas yang saling melengkapi: koefisien *Aiken's V* dan *Content Validity Index (CVI)*. *Aiken's V* digunakan untuk mengevaluasi validitas konten pada tingkat item, yakni mengukur sejauh mana setiap butir indikator secara individual dinilai relevan dan representatif oleh para ahli, dengan mempertimbangkan rentang skala penilaian yang digunakan; kriteria penerimaan ditetapkan pada  $V \geq 0,75$ . Sementara itu, CVI digunakan untuk mengevaluasi validitas konten pada tingkat instrumen secara keseluruhan, yakni menghitung proporsi item yang disepakati valid oleh seluruh panel ahli, sehingga memberikan gambaran menyeluruh tentang kelayakan instrumen sebagai satu kesatuan; kriteria penerimaan ditetapkan pada  $CVI \geq 0,80$ . Penggunaan kedua indeks secara bersamaan memungkinkan identifikasi item bermasalah secara spesifik sekaligus penilaian kelayakan instrumen secara holistik.

$$V = \frac{\{\sum S\}}{\{n(c - 1)\}}$$

Validasi isi dilakukan menggunakan koefisien Aiken's V berdasarkan penilaian ahli, karena metode ini secara statistik mengukur sejauh mana para ahli sepakat bahwa item-item instrumen merepresentasikan konstruk yang diukur, dengan rentang nilai 0 hingga 1 (Roebianto et al., 2023; Nurjanah et al., 2023). Fase kedua adalah analisis reliabilitas melalui uji empiris, di mana rubrik yang telah lolos validasi konten digunakan untuk menilai hasil kerja 30 siswa program TKJ yang telah menerima materi instalasi jaringan LAN pada semester genap tahun akademik 2025/2026. Pemilihan subjek menggunakan teknik *purposive sampling* dengan kriteria yang ditetapkan secara eksplisit untuk menjamin relevansi dan kualitas data psikometrik yang diperoleh. Kriteria inklusi subjek meliputi: (1) siswa yang aktif mengikuti pembelajaran, didefinisikan secara operasional sebagai siswa dengan tingkat kehadiran minimal 80% selama semester berjalan dan tidak sedang dalam status skorsing atau cuti akademik; (2) siswa yang telah menyelesaikan seluruh materi praktik TKJ yang relevan dengan aspek-aspek yang diukur dalam instrumen, sehingga dapat dipastikan bahwa variasi skor yang muncul mencerminkan perbedaan kompetensi nyata, bukan ketidaktuntasan materi; dan (3) siswa yang bersedia mengikuti prosedur penilaian secara penuh hingga selesai. Pemilihan subjek yang memenuhi kriteria tersebut secara khusus bertujuan untuk menghindari bias data akibat ketidakhadiran, ketidaktuntasan belajar, maupun data yang tidak lengkap, yang dapat menurunkan estimasi reliabilitas secara artifisial.

Jumlah subjek sebanyak 30 siswa dipilih untuk memenuhi kecukupan analisis psikometrik. Bujang et al. (2018) menyatakan bahwa ukuran sampel sekitar 30 responden dapat digunakan dalam analisis reliabilitas, dengan mempertimbangkan karakteristik data dan kekuatan korelasi antar item. Analisis reliabilitas antar penilai dilakukan pada setiap deskriptor penilaian untuk mengidentifikasi tingkat kesepakatan antar penilai pada masing-masing aspek rubrik. Data skor penilaian per aspek rubrik selanjutnya dianalisis menggunakan *Cronbach's Alpha* untuk mengukur konsistensi internal antar aspek rubrik, dengan nilai minimum yang disyaratkan untuk membuat instrumen dengan reliabilitas tinggi sesuai kriteria Guilford adalah  $\alpha > 0,60$  (Juliani & Erita, 2023).

*Cronbach's Alpha* dihitung menggunakan formula berikut:

$$\alpha = \frac{N * \bar{c}}{\bar{v} + (N - 1) * \bar{c}}$$

di mana N adalah jumlah item dalam instrumen,  $\bar{c}$  adalah rata-rata kovarians antar item, dan  $\bar{v}$  adalah rata-rata varians item. Formula ini dipilih karena secara matematis menangkap proporsi variansi total skor yang dapat dijelaskan oleh konsistensi antar item, sehingga cocok digunakan pada instrumen rubrik praktik dengan skala politomus seperti yang diterapkan dalam

penelitian ini. *Cronbach's Alpha* merupakan salah satu teknik paling umum dan direkomendasikan untuk menilai konsistensi internal instrumen pengukuran (Alizadeh, Shahsavari, & Matourypour, 2025).

## HASIL DAN PEMBAHASAN

### Hasil

#### A. Hasil Fase Pertama: Validitas Konten Instrumen Penilaian yang Ada

Fase pertama penelitian ini bertujuan untuk mengevaluasi validitas konten rubrik penilaian UKK Instalasi Jaringan LAN yang telah digunakan di SMKN 1 Sambeng melalui penilaian lima orang ahli. Setiap aspek rubrik dinilai pada lima dimensi menggunakan lembar validasi berskala Likert 1–4, kemudian dianalisis menggunakan koefisien Aiken's V dengan kriteria penerimaan  $V \geq 0,75$  dan *Content Validity Index* (CVI)  $\geq 0,80$ .

**Tabel 1.** Hasil Analisis Validitas Konten Instrumen Penilaian Lama (Aiken's V)

Soal	Dimensi	Rata-rata Skor	Aiken's V	Kategori
Soal 1	Relevansi Indikator	3.80	0.93	Sangat Valid
Soal 1	Kejelasan Rumusan	3.60	0.87	Sangat Valid
Soal 1	Kelengkapan Aspek Teknis	3.40	0.80	Valid
Soal 1	Kelayakan Instrumen	3.60	0.87	Sangat Valid
Soal 1	Keterlaksanaan	3.40	0.80	Valid
Soal 2	Relevansi Indikator	3.80	0.93	Sangat Valid
Soal 2	Kejelasan Rumusan	3.20	0.73	Cukup Valid
Soal 2	Kelengkapan Aspek Teknis	3.40	0.80	Valid
Soal 2	Kelayakan Instrumen	3.40	0.80	Valid
Soal 2	Keterlaksanaan	3.20	0.73	Cukup Valid
Soal 3	Relevansi Indikator	3.80	0.93	Sangat Valid
Soal 3	Kejelasan Rumusan	3.60	0.87	Sangat Valid
Soal 3	Kelengkapan Aspek Teknis	3.20	0.73	Cukup Valid
Soal 3	Kelayakan Instrumen	3.60	0.87	Sangat Valid
Soal 3	Keterlaksanaan	3.40	0.80	Valid
Soal 4	Relevansi Indikator	3.80	0.93	Sangat Valid
Soal 4	Kejelasan Rumusan	3.40	0.80	Valid
Soal 4	Kelengkapan Aspek Teknis	3.20	0.73	Cukup Valid
Soal 4	Kelayakan Instrumen	3.40	0.80	Valid
Soal 4	Keterlaksanaan	3.20	0.73	Cukup Valid
Soal 5	Relevansi Indikator	3.80	0.93	Sangat Valid
Soal 5	Kejelasan Rumusan	3.60	0.87	Sangat Valid
Soal 5	Kelengkapan Aspek Teknis	3.40	0.80	Valid
Soal 5	Kelayakan Instrumen	3.40	0.80	Valid
Soal 5	Keterlaksanaan	3.20	0.73	Cukup Valid

*Content Validity Index* (CVI) 0.82 (Cukup Valid — di atas batas minimum 0,80)

Berdasarkan Tabel 1, hasil analisis Aiken's V menunjukkan bahwa sebagian besar aspek instrumen penilaian lama memperoleh nilai  $V \geq 0,75$  dan dapat dikategorikan valid. Nilai CVI instrumen lama sebesar 0,82 tepat berada di atas batas minimum yang disyaratkan ( $CVI \geq 0,80$ ). Namun demikian, terdapat enam aspek yang memperoleh nilai  $V = 0,73$  (kategori Cukup Valid dan berada di bawah ambang batas 0,75), yaitu pada dimensi Kejelasan Rumusan dan Keterlaksanaan di beberapa soal. Temuan ini mengindikasikan bahwa meskipun instrumen

lama secara keseluruhan dapat diterima dari sisi validitas konten, terdapat sejumlah aspek yang rumusannya dinilai para ahli kurang jelas dan kurang dapat dilaksanakan secara konsisten dalam penilaian praktik, khususnya pada aspek yang menggunakan sub-kriteria dengan bobot poin tidak seimbang.

**B. Hasil Fase Kedua: Reliabilitas Instrumen Penilaian yang Ada**

Fase kedua bertujuan untuk mengevaluasi reliabilitas instrumen penilaian lama melalui uji empiris menggunakan *Cronbach's Alpha* serta analisis reliabilitas antar penilai menggunakan *Intraclass Correlation Coefficient*. Hasil analisis reliabilitas antar penilai yang melibatkan lima orang penilai (guru TKJ) menunjukkan bahwa nilai *Intraclass Correlation Coefficient* berada pada kisaran 0,50–0,62 untuk *single measures* dan 0,68–0,76 untuk *average measures*. Analisis dilakukan menggunakan model *two-way random effects* dengan tipe *absolute agreement* untuk mengukur tingkat kesepakatan antar penilai terhadap objek yang sama. Berdasarkan kriteria interpretasi menurut Koo & Li (2016), nilai tersebut termasuk dalam kategori cukup, yang menunjukkan bahwa terdapat tingkat konsistensi dan kesepakatan yang cukup memadai antar penilai dalam menggunakan instrumen rubrik yang dikembangkan. Rubrik yang telah melalui validasi konten digunakan untuk menilai hasil kerja 30 siswa aktif program TKJ yang telah menerima materi instalasi jaringan LAN. Data skor penilaian per aspek rubrik dianalisis menggunakan SPSS untuk mengukur konsistensi internal antar aspek.

**Tabel 2.** Hasil Analisis Reliabilitas Instrumen Penilaian Lama (*Cronbach's Alpha*)

Instrumen	<i>Cronbach's Alpha</i>	Kriteria Guilford	Kategori	Keterangan
Rubrik Instalasi Jaringan LAN	0.502	$0,40 < \alpha \leq 0,60$	Sedang	Tidak Memenuhi

Hasil analisis pada Tabel 2 menunjukkan bahwa reliabilitas instrumen penilaian lama belum memenuhi standar yang dipersyaratkan. Rubrik instalasi jaringan LAN memperoleh  $\alpha = 0.502$  yang termasuk kategori Sedang (Juliani & Erita, 2023) dan berada di bawah batas minimum yang disyaratkan ( $\alpha > 0,60$ ).

**Tabel 3.** Korelasi Item-Total Instrumen Penilaian Lama

Soal	Topik	Korelasi Item-Total	Kategori
Soal 1	Server/Switch/Router	0,329	Rendah
Soal 2	Troubleshooting	0,215	Sangat Rendah
Soal 3	Uji Konektivitas	0,615	Sedang
Soal 4	Kabel UTP	0,538	Sedang
Soal 5	Jaringan LAN	0,493	Rendah

Tabel 3 menunjukkan bahwa instrumen penilaian memiliki korelasi item-total yang rendah hingga sangat rendah ( $r = 0,215-0,615$ ), dengan satu soal berada di bawah ambang batas yang disyaratkan ( $r \geq 0,30$ ), yaitu Soal 2 ( $r = 0,215$ ). Saryadi et al. (2025) mengutip Ghozali (2018) bahwa item dinyatakan valid jika nilai korelasi item-total  $\geq 0,30$ . Rendahnya korelasi ini disebabkan oleh distribusi bobot sub-aspek yang tidak merata. Sub-aspek berbobot rendah seperti 'Kejelasan Jawaban' (skor 1–2) cenderung menghasilkan skor seragam karena hampir seluruh siswa memperoleh nilai penuh, sehingga varians item menjadi kecil dan menurunkan nilai *Cronbach's Alpha* secara keseluruhan, sebagaimana terlihat dari data analisis SPSS pada Gambar 1.

Scale Statistics				Reliability Statistics	
Mean	Variance	Std. Deviation	N of Items	Cronbach's Alpha	N of Items
73.64	168.815	12.993	5	.502	5

**Gambar 1.** Hasil Analisis SPSS Statistik Skala Total dan *Cronbach's Alpha* dari 5 Item

Berdasarkan temuan pada fase pertama dan kedua ini dimana validitas konten berada pada batas minimum dan reliabilitas instrumen lama belum memenuhi standar yang ideal, maka penelitian dilanjutkan ke fase ketiga, yaitu perancangan instrumen rubrik penilaian baru yang bertujuan untuk mengatasi kelemahan-kelemahan yang teridentifikasi.

**C. Hasil Fase Ketiga: Perancangan dan Evaluasi Instrumen Penilaian Baru**

Berdasarkan temuan fase pertama dan kedua, dilakukan perancangan instrumen rubrik penilaian baru dengan menerapkan empat prinsip perbaikan utama: (1) penyeimbangan bobot sub-aspek menjadi 5 poin per aspek (4 aspek × 5 poin = 20 poin per soal); (2) penggunaan deskriptor diskret tiga level (0/3/5) yang spesifik untuk meminimalkan ambiguitas penilai; (3) penghapusan sub-aspek 'Kejelasan Jawaban' yang bersifat generik dan menggantinya dengan sub-aspek yang mengukur penguasaan materi secara substantif; serta (4) penyusunan sub-aspek yang mencakup empat level kognitif Bloom revisi Anderson dan Krathwohl, yaitu mengingat (C1), memahami (C2), menerapkan (C3), dan menganalisis (C4), untuk meningkatkan daya beda instrumen (Rohman, 2023). Hasil perancangan instrumen penilaian baru disajikan pada Tabel 4.

**Tabel 4.** Hasil Perancangan Instrumen Penilaian Baru

No	Aspek Penilaian	Deskriptor Skor	Skor Maks
<b>Soal 1: Jelaskan pengertian jaringan LAN dan fungsinya dalam lingkungan perusahaan!</b>			
1	Ketepatan Definisi LAN	0 – Tidak menjelaskan / salah 3 – Menyebut jaringan lokal tanpa batasan area/cakupan 5 – Definisi lengkap: jaringan terbatas satu gedung/kampus, protokol, kecepatan tinggi	5
2	Fungsi Berbagi Sumber Daya	0 – Tidak menyebutkan fungsi 3 – Menyebut 1 fungsi saja 5 – ≥2 fungsi: berbagi file, printer, internet, storage, dengan penjelasan	5
3	Fungsi Komunikasi Internal	0 – Tidak menyebutkan komunikasi 3 – Menyebut komunikasi tanpa contoh konkret 5 – Email internal, chat, VoIP, video conference dalam jaringan lokal	5
4	Relevansi dengan Lingkungan Perusahaan	0 – Tidak menghubungkan ke konteks perusahaan 3 – Menyebut perusahaan tapi hanya 1 manfaat bisnis 5 – Efisiensi operasional + keamanan data + penghematan biaya dalam bisnis	5
<b>Soal 2: Jelaskan perbedaan fungsi server, switch, dan router dalam jaringan LAN!</b>			
1	Fungsi Server	0 – Tidak menjelaskan atau salah total 3 – Menyebut server melayani client tanpa detail jenis layanan 5 – Menyediakan layanan (file, web, DHCP, dll), menyimpan data terpusat, autentikasi	5
2	Fungsi Switch	0 – Tidak menjelaskan atau salah total	5

No	Aspek Penilaian	Deskriptor Skor	Skor Maks
		3 – Menyebut switch untuk menghubungkan perangkat tanpa detail cara kerja	
		5 – Menghubungkan perangkat di LAN, MAC address table, full-duplex, Layer 2	
3	Fungsi Router	0 – Tidak menjelaskan atau salah total	
		3 – Menyebut router untuk internet tanpa penjelasan routing	5
		5 – Menghubungkan antar jaringan (LAN-WAN), routing table, IP forwarding, NAT	
4	Analisis Perbedaan Antar Perangkat	0 – Tidak membedakan, menganggap sama	
		3 – Menyebut perbedaan tapi tidak tepat secara teknis	5
		5 – Perbedaan layer OSI, cakupan jaringan, dan cara pengolahan paket disajikan jelas	
<b>Soal 3: Uraikan langkah-langkah pembuatan kabel UTP straight-through standar T568B!</b>			
1	Persiapan Alat & Bahan	0 – Tidak menyebutkan alat/bahan	
		3 – Menyebut $\leq 2$ alat/bahan utama	5
		5 – Kabel UTP Cat5e/6, RJ-45 konektor, crimping tool, LAN tester, cutter — lengkap	
2	Urutan Warna T568B yang Benar	0 – Urutan salah atau tidak menyebut	
		3 – $\geq 4$ warna benar dari 8 pasang	5
		5 – 8 urutan benar: O-P/O/HB-P/HB/B/HB-P/HB/Cb — semua tepat	
3	Proses Crimping & Pemasangan Konektor	0 – Tidak menjelaskan proses	
		3 – Menjelaskan memasukkan kabel ke RJ-45 tanpa urutan	5
		5 – Kupas $\pm 2.5$ cm, luruskan, urutkan, masukkan ke RJ-45, crimping hingga klik	
4	Pengujian dengan LAN Tester	0 – Tidak menyebut pengujian	
		3 – Menyebut uji pakai tester tanpa interpretasi hasil	5
		5 – Lampu 1-2-3-4-5-6-7-8 menyala berurutan = straight-through berhasil	
<b>Soal 4: Jelaskan tujuan dilakukan pengujian konektivitas (ping) pada jaringan LAN!</b>			
1	Tujuan Utama Perintah Ping	0 – Tidak menjelaskan / salah	
		3 – Menyebut cek koneksi tanpa mekanisme ICMP	5
		5 – Mengirim paket ICMP echo request $\rightarrow$ reply, verifikasi host aktif dan terjangkau	
2	Manfaat Diagnosa Jaringan	0 – Tidak menyebut manfaat diagnosa	
		3 – 1 manfaat diagnosa saja	5
		5 – Cek IP aktif, latency, packet loss, rute jaringan — $\geq 2$ manfaat dengan penjelasan	
3	Interpretasi Output Ping	0 – Tidak menjelaskan output	5

No	Aspek Penilaian	Deskriptor Skor	Skor Maks
		3 – Menyebut berhasil/gagal tanpa parameter 5 – TTL, time (ms), packet sent/received/lost, Request Timed Out — dijelaskan artinya	
4	Penggunaan Troubleshooting dalam	0 – Tidak menghubungkan ke troubleshooting 3 – Menyebut dipakai saat ada masalah tanpa langkah konkret 5 – Ping loopback → default gateway → server: urutan diagnosa sistematis dijelaskan	5
Soal 5: Jelaskan langkah-langkah troubleshooting apabila PC tidak dapat terhubung ke server!			
1	Pemeriksaan Fisik Layer 1	0 – Tidak menyebut pemeriksaan fisik 3 – Menyebut cek kabel tanpa detail indikator 5 – Kabel terpasang, indikator LED NIC/switch menyala, ganti kabel jika rusak	5
2	Pemeriksaan Konfigurasi IP	0 – Tidak menyebut konfigurasi IP 3 – Menyebut cek IP tanpa cara/perintah 5 – ipconfig/ifconfig, IP sekelas dengan server, subnet mask, default gateway benar	5
3	Pengujian Konektivitas Bertahap	0 – Tidak melakukan pengujian bertahap 3 – Ping langsung ke server tanpa urutan 5 – Ping 127.0.0.1 → ping gateway → ping server: urutan dan logika tepat	5
4	Langkah Eskalasi & Solusi Lanjutan	0 – Tidak menyebut langkah lanjutan 3 – Menyebut restart saja 5 – Cek firewall, driver NIC, DHCP server, koordinasi admin jaringan jika perlu	5

### 1. Validitas Konten Instrumen Penilaian Baru

Instrumen rubrik penilaian baru divalidasi oleh panel ahli yang sama menggunakan prosedur identik dengan fase pertama. Hasil analisis Aiken's V disajikan pada Tabel 5.

**Tabel 5.** Hasil Analisis Validitas Konten Instrumen Penilaian Baru (Aiken's V)

Soal	Dimensi	Rata-rata Skor	Aiken's V	Kategori
Soal 1	Relevansi Indikator	4.00	1.00	Sangat Valid
Soal 1	Kejelasan Rumusan	3.80	0.93	Sangat Valid
Soal 1	Kelengkapan Aspek Teknis	3.80	0.93	Sangat Valid
Soal 1	Kelayakan Instrumen	3.80	0.93	Sangat Valid
Soal 1	Keterlaksanaan	3.60	0.87	Sangat Valid
Soal 2	Relevansi Indikator	4.00	1.00	Sangat Valid
Soal 2	Kejelasan Rumusan	3.80	0.93	Sangat Valid
Soal 2	Kelengkapan Aspek Teknis	3.80	0.93	Sangat Valid
Soal 2	Kelayakan Instrumen	3.80	0.93	Sangat Valid
Soal 2	Keterlaksanaan	3.60	0.87	Sangat Valid
Soal 3	Relevansi Indikator	4.00	1.00	Sangat Valid
Soal 3	Kejelasan Rumusan	3.80	0.93	Sangat Valid

Soal	Dimensi	Rata-rata Skor	Aiken's V	Kategori
Soal 3	Kelengkapan Aspek Teknis	3.80	0.93	Sangat Valid
Soal 3	Kelayakan Instrumen	3.80	0.93	Sangat Valid
Soal 3	Keterlaksanaan	3.80	0.93	Sangat Valid
Soal 4	Relevansi Indikator	4.00	1.00	Sangat Valid
Soal 4	Kejelasan Rumusan	3.80	0.93	Sangat Valid
Soal 4	Kelengkapan Aspek Teknis	3.80	0.93	Sangat Valid
Soal 4	Kelayakan Instrumen	3.80	0.93	Sangat Valid
Soal 4	Keterlaksanaan	3.60	0.87	Sangat Valid
Soal 5	Relevansi Indikator	4.00	1.00	Sangat Valid
Soal 5	Kejelasan Rumusan	3.80	0.93	Sangat Valid
Soal 5	Kelengkapan Aspek Teknis	3.80	0.93	Sangat Valid
Soal 5	Kelayakan Instrumen	3.80	0.93	Sangat Valid
Soal 5	Keterlaksanaan	3.80	0.93	Sangat Valid
<i>Content Validity Index (CVI)</i>			0.94 (Sangat Valid — jauh di atas batas 0,80)	

Tabel 5 menunjukkan peningkatan signifikan pada validitas konten instrumen baru. Seluruh 25 aspek memperoleh nilai Aiken's V  $\geq 0,87$  dan masuk dalam kategori Sangat Valid, tidak ada satu pun aspek yang berada di bawah ambang batas 0,75. Nilai CVI instrumen baru sebesar 0,94 meningkat dibandingkan instrumen lama (0,82) dan berada jauh di atas batas minimum yang disyaratkan. Peningkatan ini mencerminkan bahwa deskriptor diskret tiga level yang spesifik dan sub-aspek yang substantif dinilai para ahli lebih representatif dalam mengukur kompetensi teknik jaringan dibandingkan sub-aspek generik seperti 'Kejelasan Jawaban' yang digunakan pada instrumen lama.

## 2. Reliabilitas Instrumen Penilaian Baru

Setelah dinyatakan valid secara konten, instrumen baru diujicobakan secara empiris kepada 30 siswa yang sama untuk memperoleh data reliabilitas. Hasil analisis *Cronbach's Alpha* dan korelasi item-total ada pada Tabel 6 dan Tabel 7.

**Tabel 6.** Perbandingan Reliabilitas Instrumen Penilaian Instalasi LAN Lama dan Baru

Instrumen	Cronbach's Alpha	Kriteria Guilford	Kategori	Keterangan
Rubrik Lama	0.502	$0,40 < \alpha \leq 0,60$	Sedang	Dibawah Batas Minimum
Rubrik Baru	0,844	$0,80 < \alpha \leq 1,00$	Sangat Tinggi	Memenuhi Standar

Tabel 6 menunjukkan peningkatan reliabilitas yang sangat signifikan dari instrumen lama ke instrumen baru. *Cronbach's Alpha* instrumen baru sebesar  $\alpha = 0,844$  masuk dalam kategori Sangat Tinggi berdasarkan kriteria Guilford ( $0,80 < \alpha \leq 1,00$ ) (Juliani & Erita, 2023), jauh melampaui nilai Alpha instrumen lama ( $\alpha = 0.502$ ).

**Tabel 7.** Korelasi Item-Total Instrumen Penilaian Baru

Soal	Topik	Korelasi Item-Total	Kategori
Soal 1	Server/Switch/Router	0.880	Tinggi
Soal 2	Troubleshooting	0.749	Tinggi
Soal 3	Uji Konektivitas	0.824	Tinggi
Soal 4	Kabel UTP	0.653	Tinggi
Soal 5	Jaringan LAN	0.872	Tinggi

Tabel 7 menunjukkan bahwa seluruh lima soal dalam instrumen baru memiliki korelasi item-total yang tinggi ( $r = 0,653-0,880$ ), jauh melampaui ambang batas yang disyaratkan ( $r \geq 0,30$ ). Hal ini mengindikasikan bahwa setiap soal berkontribusi secara konsisten dalam

mengukur konstruk yang sama, yaitu kompetensi teknik jaringan siswa. Soal 1 (Server/Switch/Router) memiliki korelasi item-total tertinggi ( $r = 0,880$ ), menunjukkan bahwa soal ini paling kuat dalam membedakan siswa yang menguasai kompetensi dari yang belum.

### Pembahasan

Temuan penelitian ini menunjukkan bahwa instrumen penilaian lama yang digunakan di SMKN 1 Sambeng memiliki validitas konten yang berada di batas minimum yang dapat diterima ( $CVI = 0,82$ ), namun memiliki kelemahan pada aspek reliabilitas. Kondisi ini konsisten dengan temuan De Ruitter & Bers (2021) yang mendokumentasikan kelangkaan instrumen penilaian kompetensi teknis yang valid dan reliabel, serta Rahmah & Festiyed (2025) yang menegaskan bahwa tanpa instrumen yang valid dan reliabel, hasil penilaian berpotensi menghasilkan kesimpulan yang tidak akurat mengenai kemampuan peserta didik.

Analisis mendalam terhadap penyebab rendahnya reliabilitas instrumen lama mengungkapkan bahwa masalah utama terletak pada desain sub-aspek penilaian. Bobot poin yang tidak merata menyebabkan beberapa sub-aspek, terutama 'Kejelasan Jawaban', terlalu mudah dicapai oleh hampir seluruh siswa. Kondisi ini menghasilkan varians item yang kecil sehingga korelasi antar item menjadi rendah dan nilai Alpha menurun.

Perancangan instrumen baru yang menerapkan bobot seimbang (5 poin per sub-aspek) dan deskriptor diskret tiga level (0/3/5) terbukti efektif mengatasi kelemahan tersebut. Instrumen baru menghasilkan peningkatan CVI dari 0,82 menjadi 0,94, dan peningkatan *Cronbach's Alpha* yang sangat signifikan dari 0,502 menjadi 0,844. Peningkatan ini dicapai melalui perbaikan desain instrumen yang memungkinkan variasi skor yang lebih bermakna antar siswa, sehingga setiap soal mampu membedakan secara lebih tepat antara siswa yang menguasai kompetensi dan yang belum. Peningkatan nilai *Cronbach's Alpha* tersebut juga menunjukkan adanya peningkatan homogenitas konstruk, yang dalam perspektif teori pengukuran klasik (*Classical Test Theory*) mengindikasikan bahwa item-item dalam instrumen telah merepresentasikan domain kompetensi secara lebih konsisten. Temuan ini sejalan dengan Shuaili (2025) yang menegaskan bahwa desain kurikulum dan instrumen penilaian yang terarah merupakan faktor kunci dalam meningkatkan kualitas pengukuran kompetensi di pendidikan vokasi.

Dengan nilai *Cronbach's Alpha* sebesar 0,844 yang masuk kategori Sangat Tinggi berdasarkan kriteria koefisien Guilford (Juliani & Erita, 2023) dan seluruh korelasi item-total di atas 0,65, instrumen penilaian baru ini telah memenuhi standar psikometrik yang dipersyaratkan dan layak digunakan sebagai alat ukur kompetensi teknik jaringan yang andal di SMKN 1 Sambeng.

Terdapat beberapa keterbatasan dalam penelitian ini yang perlu menjadi catatan, yaitu uji reliabilitas empiris dilakukan menggunakan data dari 30 siswa, yang merupakan jumlah sampel relatif kecil untuk menghasilkan estimasi *Cronbach's Alpha* yang sepenuhnya stabil. Meskipun nilai  $\alpha = 0,844$  yang diperoleh termasuk kategori Sangat Tinggi, estimasi koefisien reliabilitas pada sampel kecil cenderung memiliki interval kepercayaan yang lebih lebar sehingga generalisasi hasil ke populasi yang lebih besar perlu dilakukan dengan kehati-hatian.

### KESIMPULAN

Penelitian ini menyimpulkan bahwa instrumen penilaian kompetensi TKJ yang selama ini digunakan di SMKN 1 Sambeng belum memenuhi standar psikometrik yang memadai. Meskipun validitas konten instrumen lama berada pada batas minimum yang dapat diterima ( $CVI = 0,82$ ), reliabilitasnya masih tergolong sedang ( $\alpha = 0,502$ ) dan di bawah ambang batas



yang disyaratkan, sehingga tidak dapat diandalkan sebagai alat ukur kompetensi yang akurat dan konsisten. Kelemahan tersebut bersumber dari desain rubrik yang tidak proporsional, yakni bobot poin antar sub-aspek yang tidak seimbang dan deskriptor yang bersifat kontinum, sehingga menghasilkan varians item rendah dan korelasi antar item yang lemah. Sebagai respons atas temuan tersebut, instrumen baru dirancang dengan prinsip keseimbangan bobot (5 poin per sub-aspek), deskriptor diskret tiga level (0/3/5), dan cakupan empat level kognitif Bloom yang lebih komprehensif. Hasilnya, instrumen baru terbukti valid dengan CVI = 0,94 dan reliabel pada kategori sangat tinggi ( $\alpha = 0,844$ ), dengan seluruh korelasi item-total di atas 0,65, sehingga layak digunakan sebagai alat ukur kompetensi teknik jaringan yang andal di SMKN 1 Sambeng.

Sebagai saran, pengujian instrumen perlu diperluas ke sekolah lain dengan karakteristik yang beragam guna menguji generalisabilitas temuan. Selain itu, integrasi instrumen dengan sistem asesmen digital dan penyelarasan dengan standar sertifikasi kompetensi nasional perlu diupayakan agar instrumen ini dapat berkontribusi pada sistem evaluasi kejuruan yang lebih terstandar dan berkelanjutan.

#### DAFTAR PUSTAKA

- Aliffiansyah, A. N., Ekarini, F., & Djuniadi, D. (2025). Implementasi model Rasch pada analisis kualitas butir soal jaringan dasar pada jenjang pendidikan vokasi. *Jurnal Metaedukasi: Jurnal Ilmiah Pendidikan*, 7(2), 144–154. <https://jurnal.unsil.ac.id/index.php/metaedukasi/article/view/17889/4671>
- Alizadeh, M., Shahsavari, H., & Matourypour, P. (2025). Key points in writing instrument development and psychometric evaluation reports. *Revista Brasileira de Enfermagem*, 78(4), e20250006. <https://doi.org/10.1590/0034-7167-2025-0006>
- Bujang, M. A., Omar, E. D., & Baharum, N. A. (2018). A review on sample size determination for Cronbach's alpha test: A simple guide for researchers. *Malaysian Journal of Medical Sciences*, 25(6), 85–99. <https://doi.org/10.21315/mjms2018.25.6.9>
- De Ruiter, L. E., & Bers, M. U. (2021). The coding stages assessment: Development and validation of an instrument for assessing young children's proficiency in the Scratch Jr programming language. *Computer Science Education*, 32(4), 388–417. <https://doi.org/10.1080/08993408.2021.1956216>
- Ghozali, I. (2018). *Aplikasi analisis multivariate dengan program IBM SPSS 25*. Badan Penerbit Universitas Diponegoro.
- Hakim, D. R., & Kurniawati, D. (2022). The effect of competency and ICT skills on vocational students' work readiness. *Jurnal Pendidikan Ekonomi dan Bisnis (JPEB)*, 10(1), 15–34. <https://doi.org/10.21009/jpeb.010.1.2>
- Juliani, R. P., & Erita, S. (2023). Analisis validitas dan reliabilitas instrumen penilaian kemampuan berpikir kritis dalam konteks sekolah menengah. *JEID: Journal of Educational Integration and Development*, 3(3), 169–179. <https://doi.org/10.55868/jeid.v3i3.313>
- Khotimah, H., & Hidayat, N. (2024). *Evaluasi pembelajaran: Konsep, model, dan pengembangan tes hasil belajar*. Penerbit Litnus.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Nurjanah, S., Istiyono, E., Widiastuti, Y., Iqbal, M., & Kamal, S. (2023). The application of Aiken's V method for evaluating the content validity of instruments that measure the



- implementation of formative assessments. *Journal of Research and Educational Research Evaluation*, 12(2), 125–133. <https://doi.org/10.15294/jere.v12i2.76451>
- Rahmah, S., & Festiyed. (2025). The validity of online assessment instruments for measuring higher order thinking skills of high school students in physics subject. *Jurnal Penelitian Pendidikan IPA*, 11(2), 92–98. <https://doi.org/10.29303/jppipa.v11i2.4993>
- Roebianto, A., Savitri, S. I., Aulia, I., Suciyan, A., & Mubarakah, L. (2023). Content validity: Definition and procedure of content validation in psychological research. *Testing, Psychometrics, Methodology in Applied Psychology*, 30, 5–18. <https://doi.org/10.4473/TPM30.1.1>
- Rohman, S. (2023). Pengembangan instrumen penilaian hasil belajar berbasis taksonomi Bloom. *Al-Ibrah: Jurnal Pendidikan dan Keilmuan Islam*, 8(1), 86–108. <https://doi.org/10.61815/alibrah.v8i1.262>
- Saryadi, S., Listyorini, P. I., Arini, L. D. D., & Pattinama, A. M. (2025). Uji validitas dan reliabilitas pengukuran kepuasan pengguna RME dengan metode EUCS. *Jurnal Riset Rumpun Ilmu Kesehatan*, 4(2), 44–59. <https://doi.org/10.55606/jurrikes.v4i2.5255>
- Shuaaili, A. S. T. A. (2025). Bridging the skills gap: Technical and vocational education and training student competencies and labor market alignment in Oman. *Vocation Technology & Education*, 2(3). <https://doi.org/10.54844/vte.2025.0991>
- Subando, J. (2022). *Validitas dan reliabilitas instrumen non tes*. Penerbit Lakeisha.
- Sudira, P. (2021). *Pendidikan vokasi abad 21: Inovasi, relevansi, dan tantangan*. UNY Press.
- Viberg, O., Pargman, T. C., & Lilliesköld, J. (2024). Exploring teachers' (future) digital assessment practices in higher education: Instrument and model development. *British Journal of Educational Technology*, 55, 2597–2616. <https://doi.org/10.1111/bjet.13462>
- Villarroel, V., Melipillán, R., Santana, J., & Aguirre, D. (2024). How authentic are assessments in vocational education? An analysis from Chilean teachers, students, and examinations. *Frontiers in Education*, 9, 1–12. <https://doi.org/10.3389/feduc.2024.1308688>
- Yusop, S. R. M., Rasul, M. S., Yasin, R. M., & Hashim, H. U. (2022). Assessment approaches and learning outcomes in technical and vocational education: A systematic review using PRISMA. *Sustainability*, 14(9), 5225. <https://doi.org/10.3390/su14095225>
- Yusop, S. R. M., Rasul, M. S., Yasin, R. M., & Hashim, H. U. (2023). Identifying and validating vocational skills domains and indicators in classroom assessment practices in TVET. *Sustainability*, 15(6), 5195. <https://doi.org/10.3390/su15065195>