



**PENGEMBANGAN INSTRUMEN PENILAIAN PROSES PEMBELAJARAN  
AKUNTANSI BERBASIS *DISCOVERY LEARNING* UNTUK MENGUKUR  
KETERAMPILAN BERPIKIR KRITIS SISWA SMK**

Fadilah Try Mursyid<sup>1\*</sup>, Wahyuni<sup>2</sup>, Muh Ikhsanul Yakin<sup>3</sup>  
Program Studi Pendidikan Ekonomi Universitas Nggusuwaru Bima  
Email Korespondensi : [fadilah.tm@gmail.com](mailto:fadilah.tm@gmail.com)<sup>1\*)</sup>

**Info Artikel**

**Kata Kunci:**

Instrumen Penilaian,  
Pembelajaran Akuntansi,  
*Discovery Learning*,  
Keterampilan Berpikir Kritis

**Abstrak**

Penelitian ini bertujuan mengetahui karakteristik instrumen, mendapatkan instrumen yang valid dan reliabel, menguji kepraktisan instrumen, dan menguji tingkat efektivitas instrumen penilaian berbasis discovery learning dalam mengukur keterampilan berpikir kritis siswa. Penelitian ini merupakan penelitian dan pengembangan (Research and Development) model Borg & Gall. Data pada instrumen penilaian ini dianalisis secara kualitatif dan kuantitatif Mixmethod. Validitas isi instrumen divalidasi oleh 4 ahli dengan teknik walktrough. Instrumen penilaian diujicobakan pada 30 siswa (skala terbatas) dan 100 siswa (skala luas). Validitas isi instrumen dianalisis dengan program SPSS versi 22.0 menghasilkan 16 butir yang mewakili setiap indikator pembelajaran model discovery learning. Reliabelitas dianalisis dengan cronbach alpha menghasilkan nilai reliabelitas sebesar 0.891 (sangat tinggi). Validitas konstruk instrumen penilaian dianalisis dengan menggunakan program smartPLS (partial least squares) versi 4.0. Hasil perhitungan konstruk instrumen didapatkan seluruh indikator memiliki nilai factor loading ( $\lambda$ ) > 0.7, composite reliability > 0.7, cronbach's alpha > 0.7, dan nilai AVE > 0.5. Uji kecocokan model didapatkan nilai RMSEA 0.02 < 0.05, SRMR 0.04 < 0.05, CFI 1.00 > 0.90, dan NNFI 0.95 > 0.90 telah memenuhi kriteria pengujian goodness of fit sehingga model instrumen penilaian dinyatakan fit. Instrumen penilaian memiliki tingkat kepraktisan dan efektivitas dengan kategori "sangat baik" berdasarkan respon pengguna

**I. PENDAHULUAN**

Kurikulum Merdeka yang berlaku saat ini menekankan pentingnya pengembangan keterampilan abad ke-21 yang menuntut peserta didik memiliki berbagai kompetensi yang tidak hanya berorientasi pada penguasaan pengetahuan, tetapi juga pada pengembangan keterampilan berpikir tingkat tinggi. Salah satu keterampilan yang sangat penting untuk dimiliki oleh siswa adalah kemampuan berpikir kritis (*critical thinking*). Kemampuan berpikir kritis memungkinkan siswa untuk menganalisis informasi secara logis, mengevaluasi berbagai alternatif pemecahan masalah, mengambil keputusan berdasarkan bukti yang relevan, serta menyusun argumentasi yang rasional. Dalam konteks pendidikan, kemampuan berpikir kritis menjadi salah satu kompetensi yang harus dikembangkan agar siswa mampu menghadapi tantangan perkembangan ilmu pengetahuan, teknologi, dan dunia kerja yang semakin kompleks (Partono, 2021). Selain dikembangkan melalui pembelajaran, keterampilan abad ke-21 juga dapat dikembangkan melalui penilaian pembelajaran (Chalkiadaki, 2018).

Penelitian ini didasari oleh belum tersedianya instrumen yang dirancang khusus untuk menilai proses pembelajaran akuntansi berbasis model *discovery learning* dalam mengukur keterampilan berpikir kritis siswa di sekolah tingkat SMK. Selama ini, penilaian proses pembelajaran siswa yang dilakukan oleh guru belum terlaksana dengan cukup maksimal. Muh Ikhsanul Yakin (2020), asesmen yang digunakan oleh guru di sekolah masih menggunakan lembar observasi standar dan tes tradisional (*paper and pencil test*) yang tidak mampu mengungkap proses pembelajaran siswa secara kompleks. Seringkali hasil penilaian tidak menggambarkan proses berpikir, penguasaan konsep dan kondisi belajar siswa yang sesungguhnya selama mengikuti kegiatan pembelajaran. Shaklee (1997), dalam mengungkap penguasaan konsep siswa, asesmen tidak hanya mengungkap konsep yang telah dicapai, akan tetapi juga mengetahui proses perkembangan belajar siswa bagaimana suatu konsep diperoleh. Kondisi ini membuat guru mengalami kesulitan dalam memperoleh informasi yang akurat mengenai perkembangan kemampuan berpikir kritis siswa.

Penilaian mata pelajaran akuntansi saat ini di SMK kurang relevan dengan fungsi, tujuan, dan prinsip penilaian yang ditetapkan oleh Kemendiknas dan yang dianjurkan oleh pakar penilaian pendidikan. Masalah tersebut sesuai dengan penelitian yang dilakukan oleh Muh Ikhsanul Yakin dalam *journal of research and educational research evaluation* (2020), bahwa desain penilaian yang digunakan oleh guru belum memenuhi standarisasi penilaian yang sesungguhnya. Selain itu, penilaian masih berfokus pada aspek kognitif serta kompetensi yang ingin dicapai tidak diukur secara tepat sehingga belum aktual menggambarkan secara langsung proses kinerja siswa secara esensial dan mendetail sesuai tujuan pendidikan. Desain penilaian lebih mengutamakan kapasitas hasil belajar bukan proses pemikiran siswa yang dinilai dan diapresiasi dalam kegiatan pembelajaran, hal ini secara tidak langsung akan mematikan daya kritis dan kreatifitas belajar siswa. Resnick (1985), pada hakikatnya asesmen menitikberatkan penilaian pada proses belajar siswa. Wiggins (1984), asesmen merupakan sarana yang secara kronologis dapat membantu guru dalam memonitor siswa dan merupakan bagian dari pembelajaran yang tidak terpisahkan.

Kegiatan penilaian (*assessment*) merupakan kegiatan yang tidak terpisahkan dengan pembelajaran serta dilakukan secara terus menerus dalam setiap proses pembelajaran baik di dalam kelas maupun di luar kelas. Pembelajaran akuntansi merupakan salah satu mata pelajaran yang memiliki peran penting dalam membekali siswa dengan kemampuan memahami, menganalisis, dan menginterpretasikan informasi keuangan. Pembelajaran akuntansi tidak hanya menuntut siswa untuk menguasai konsep dan prosedur pencatatan transaksi keuangan, tetapi juga menuntut kemampuan berpikir kritis dalam mengidentifikasi masalah, menganalisis data keuangan, menginterpretasikan informasi, serta mengambil keputusan yang tepat berdasarkan data yang tersedia. Proses pembelajaran akuntansi seharusnya dirancang untuk memberikan kesempatan kepada siswa dalam mengembangkan kemampuan berpikir kritis melalui berbagai aktivitas pembelajaran yang menantang dan bermakna (Elisabeth, 2026). Permasalahan tersebut menunjukkan adanya kebutuhan untuk mengembangkan instrumen penilaian proses pembelajaran akuntansi yang mampu mengukur keterampilan berpikir kritis siswa pada setiap tahapan model *discovery learning* yang valid, reliabel, praktis dan efektif.

Instrumen penilaian harus memiliki tingkat validitas, reliabilitas, kepraktisan, dan efektifitas yang tinggi untuk mendapatkan hasil penilaian yang akurat. Wiersma (1990), validitas isi sebenarnya mendasarkan pada analisis logika, jadi tidak merupakan suatu

koefisien validitas yang dihitung secara statistika. Validitas isi dilakukan penelaahan kisi-kisi instrumen untuk memastikan relevansi isi instrumen dengan tujuan pengukuran. Allen and Yen (1979), validitas konstruk adalah validitas yang menunjukkan sejauh mana hasil tes mampu mengungkap suatu *trait* atau suatu konstruk teoritik yang hendak diukurnya. Validitas konstruk dalam penelitian ini diuji kecocokan model GOF dengan kriteria *good fit* adalah RMSEA, SRMR, CFI, dan NNFI. Reliabilitas mengacu pada keajegan hasil pengukuran. Reliabilitas model yang baik adalah jika *construct reliability*  $\geq 0,70$  dan *variance extracted*  $\geq 0,5$  (Allen and Yen, 1979).

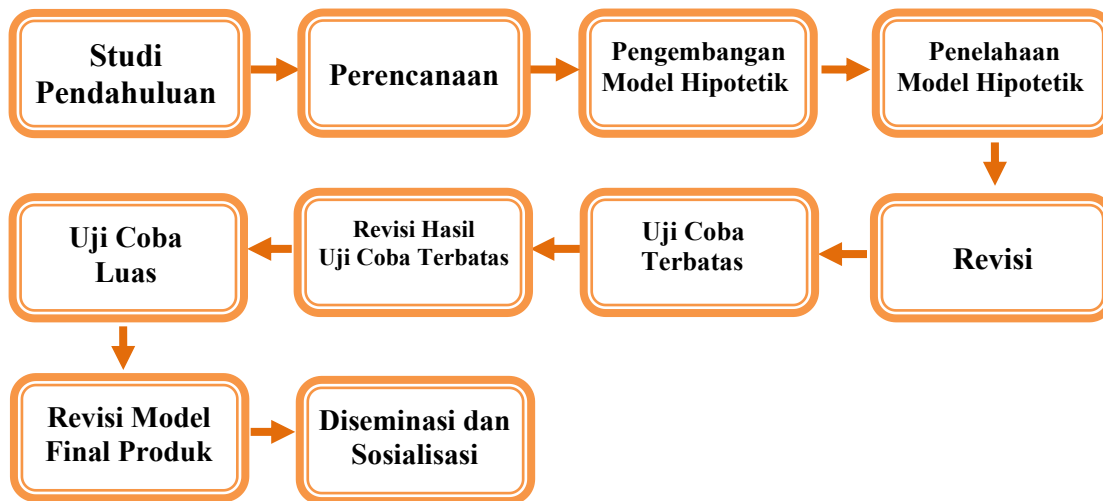
Nieveen (1999), tingkat kepraktisan dilihat dari apakah instrumen mudah digunakan oleh guru dalam melakukan penilaian. Instrumen dikatakan praktis jika pengguna dalam hal ini guru menyatakan bahwa model instrumen dapat diterapkan di lapangan dan tingkat keterlaksanaannya termasuk kategori minimal "baik". Indikator kepraktisan instrumen yaitu: (1) kejelasan petunjuk, (2) bahasa mudah dipahami, (3) waktu yang diperlukan cukup efektif, (4) format instrumen memudahkan pengguna, dan (5) hasil penilaian mudah diolah. Nitko and Brookhart (2007), efektivitas instrumen harus memperhatikan empat hal yaitu: (1) kesesuaian dengan tujuan penilaian, (2) ketepatan pengukuran sesuai kompetensi yang diukur, (3) mendukung pencapaian tujuan pembelajaran, (4) instrumen mampu membedakan kemampuan siswa, dan (5) memberikan informasi yang relevan untuk evaluasi belajar.

Instrumen penilaian yang dikembangkan dalam penelitian ini memuat indikator model *discovery learning* dan indikator kemampuan berpikir kritis. Indikator model pembelajaran *discovery learning* meliputi : (1) simulasi, (2) identifikasi masalah, (3) pengumpulan data, (4) pengolahan data, (5) pembuktian, dan (5) generalisasi. Selain itu, indikator ukur keterampilan berpikir kritis siswa yang dikembangkan dalam penelitian ini yaitu: (1) memberikan penjelasan sederhana, (2) mengembangkan keterampilan dasar, (3) melakukan inferensi, (4) memberikan penjelasan lebih lanjut, dan (5) mengatur strategi dan taktik. Dengan adanya instrumen penilaian yang dikembangkan ini, guru akuntansi dapat melakukan penilaian secara lebih objektif, akurat, dan sistematis terhadap proses pembelajaran yang dijalani siswa.

## **II. METODE PENELITIAN**

### **Jenis dan Desain Penelitian**

Penelitian ini merupakan penelitian pengembangan (*research and development*) yang bertujuan menghasilkan produk berupa instrumen penilaian proses pembelajaran yang dikemas dalam bentuk buku panduan penggunaan instrumen penilaian berbasis model *discovery learning* untuk mengukur keterampilan berpikir kritis siswa SMK. Model pengembangan yang digunakan dalam penelitian ini adalah model pengembangan R&D Borg and Gall (2007). Prosedur dalam pengembangan ini meliputi : (1) studi pendahuluan, (2) perencanaan, (3) pengembangan model hipotetik, (4) penelaahan model hipotetik, (5) revisi, (6) uji coba terbatas, (7) revisi hasil uji coba, (8) uji coba luas, (9) revisi model akhir, (10) desiminasi.



Gambar 1. Alur Langkah Pengembangan Model Borg and Gall

### Populasi dan Sampel

Populasi dalam penelitian ini adalah seluruh siswa Menengah Atas Negeri di Kabupaten Bima. Sampel penelitian yaitu siswa kelas XI SMKN 1 Langgudu dengan 30 sampel untuk uji coba produk skala terbatas dan 100 sampel untuk uji coba produk skala luas. Uji coba prototipe *one-to-one* digunakan untuk mengetahui aspek keterbacaan produk instrumen sebelum ujicoba lapangan dilakukan kepada 3 (tiga) guru akuntansi : 1 (satu) guru berasal dari SMKN 1 Langgudu, 1 (satu) guru dari SMAN 1 Langgudu, dan 1 (satu) guru dari SMAN 2 Langgudu.

### Instrumen Pengumpul Data

Pedoman wawancara digunakan untuk mengumpulkan data analisis kebutuhan. Lembar validasi ahli digunakan untuk mengumpulkan data berupa pendapat dan saran perbaikan terhadap kisi-kisi maupun instrumen penilaian yang dikembangkan dengan skala 1-4. Rubrik penskoran digunakan sebagai panduan terstruktur untuk menilai dan memberi skor terhadap instrumen penilaian berdasarkan kriteria yang telah ditentukan dengan skala likert 1-4. Angket respon guru digunakan untuk mengetahui pendapat, tanggapan, dan saran terkait kepraktisan serta efektivitas instrumen dengan skala 1-4.

### Teknik Analisis Data

Analisis data dilakukan secara kuantitatif dan kualitatif *Mixmethod* sesuai dengan jenis data yang diperoleh pada setiap tahapan pengembangan. Untuk memvalidasi instrumen penilaian dilakukan validasi isi dan validasi konstruk. Validitas isi dianalisis dengan menggunakan SPSS versi 22, sedangkan validitas konstruk dianalisis dengan *confirmatory factor analysis* (CFA) menggunakan software smartPLS versi 4.0. Instrumen dikatakan valid konstruk jika nilai *loading factor* ( $\lambda$ ) > 0.7 dari setiap indikator. Uji reliabilitas instrumen diuji dengan melihat nilai *composite reliability* indikator yang mengukur konstruk dan koefisien *cronbach's alpha*. Nilai *composite reliability* instrumen memuaskan apabila > 0.7.

### III. HASIL DAN PEMBAHASAN

#### Hasil Penelitian

##### Uji Validitas

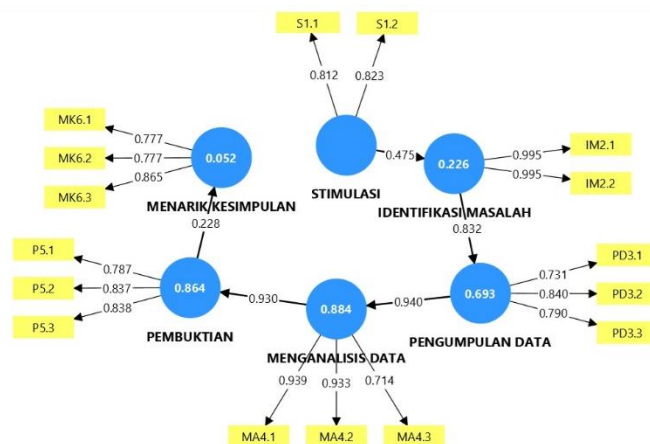
Uji validitas isi bertujuan untuk memastikan instrumen penilaian berbasis model *discovery learning* memiliki tingkat keabsahan yang baik dalam mengukur keterampilan berpikir kritis siswa SMK. Hasil uji validitas isi dengan software SPSS versi 22 yang telah dilakukan melalui metode *expert judgement* oleh 4 ahli menunjukkan 16 butir yang dikembangkan valid isi dengan rentang nilai 0.79 – 0.997 lebih besar dari 0.6. Adapun hasil dari uji validitas isi ditampilkan pada Tabel 1.

**Tabel 1. Hasil Uji Validitas Isi Instrumen Penilaian**

No	Indikator	Item	r Hitung	Kesimpulan
1	Stimulasi	Butir 1	0.872	Valid
		Butir 2	0.813	Valid
2	Identifikasi Masalah	Butir 3	0.997	Valid
		Butir 4	0.965	Valid
3	Pengumpulan Data	Butir 5	0.791	Valid
		Butir 6	0.842	Valid
		Butir 7	0.796	Valid
4	Menganalisis Data	Butir 8	0.983	Valid
		Butir 9	0.943	Valid
		Butir 10	0.819	Valid
5	Pembuktian	Butir 11	0.827	Valid
		Butir 12	0.839	Valid
		Butir 13	0.815	Valid
6	Menarik Kesimpulan	Butir 14	0.872	Valid
		Butir 15	0.977	Valid
		Butir 16	0.841	Valid

*Sumber : Data hasil penelitian (2026)*

Hasil analisis kemudian dilanjutkan dengan uji validitas konstruk dengan menggunakan *Confirmatory Factor Analysis* (CFA) untuk memastikan instrumen penilaian yang terbentuk valid serta memenuhi model fit dari analisis faktor. Hasil analisis CFA membuktikan bahwa seluruh instrumen valid secara konstruk. Berdasarkan output analisis faktor dengan smartPLS 4.0 seluruh indikator instrumen penilaian berbasis *discovery learning* untuk mengukur keterampilan berpikir kritis siswa memiliki nilai *loading factor* > 0.70 sehingga model dinyatakan signifikan atau memenuhi syarat convergent validity.



Gambar 2. Output PLS Nilai Loading Factor ( $\lambda$ ) Instrumen Penilaian

Adapun nilai loading factor hasil analisis pada Gambar 2 diatas, didapatkan nilai loading faktor tertinggi adalah item IM2.1 dan IM2.2 sebesar 0.995. Sedangkan item dengan nilai faktor loading terendah adalah item MA4.3 sebesar 0.714. Untuk lebih jelas bisa dilihat pada Tabel 2 berikut.

Tabel 2. Output Nilai Loading Factor ( $\lambda$ ) Instrumen Penilaian

No	Indikator	Item	Loading Factor ( $\lambda$ )
1	Stimulasi	(S1.1) siswa mengamati stimulus yang diberikan	0.812
		(S1.2) siswa merespon stimulus yang diberikan	0.823
2	Identifikasi Masalah	(IM2.1) siswa mengidentifikasi masalah dengan tepat	0.995
		(IM2.2) siswa menentukan cara penyelesaian masalah	0.995
3	Pengumpulan Data	(PD3.1) siswa mengumpulkan data	0.731
		(PD3.2) siswa menyiapkan laporan sementara	0.840
		(PD3.3) siswa menuliskan laporan sesuai format	0.790
4	Menganalisis Data	(MA4.1) siswa membuat interpretasi	0.939
		(MA4.2) siswa mengolah dan menganalisis interpretasi	0.933
		(MA4.3) siswa menyajikan hasil analisis	0.714
5	Pembuktian	(P5.1) siswa membuktikan hipotesis	0.787
		(P5.2) siswa menjelaskan data yang dikumpulkan	0.837
		(P5.3) siswa mempresentasikan di depan kelas	0.838
6	Menarik Kesimpulan	(MK6.1) siswa menyusun kesimpulan yang sesuai	0.777
		(MK6.2) siswa mengkomunikasikan kesimpulan	0.777
		(MK6.3) siswa membuat laporan akhir	0.865

Sumber : Data hasil output smartPLS (2026)

### Uji Reliabilitas

Uji reliabilitas intraclass correlation coefficient diperoleh nilai koefisien reliabilitas instrumen penilaian proses pembelajaran akuntansi sebesar 0,851. Dari perhitungan menggunakan smartPLS 4.0, didapatkan nilai Cronbach's Alpha > 0.7 dan nilai AVE > 0.5 yang berarti seluruh variabel yang digunakan dinyatakan reliabel.

**Tabel 3. Output smartPLS Nilai Cronbach's Alpha, Composite Reliability dan AVE**

No	Indikator	Cronbach's Alpha	Composite Reliability	AVE	Keterangan
1	Stimulasi	0.835	0.801	0.668	Reliabel
2	Identifikasi Masalah	0.990	0.995	0.990	Reliabel
3	Pengumpulan Data	0.733	0.831	0.621	Reliabel
4	Menganalisis Data	0.827	0.900	0.754	Reliabel
5	Pembuktian	0.768	0.861	0.674	Reliabel
6	Menarik Kesimpulan	0.799	0.848	0.652	Reliabel

*Sumber : Data hasil output smartPLS (2026)*

Berdasarkan data yang disajikan dalam Tabel 3, ditarik kesimpulan bahwa semua konstruk telah memenuhi kriteria reliabel, hal ini dibuktikan dengan nilai composite reliability lebih besar dari 0.7, nilai cronbach's alpha > 0.7, dan memiliki nilai AVE lebih dari 0.5. Perhitungan ini menunjukkan bahwa seluruh indikator instrumen penilaian berbasis model *discovery learning* dalam mengukur keterampilan berpikir kritis siswa SMK memiliki kekonsistenan dalam mengukur enam variabel.

Ghazali (2008), menjelaskan dukungan terhadap kecocokan model yang dikembangkan oleh data empirik dapat dilihat dari beberapa indikator uji kecocokan model yang berbeda. Kriteria yang dipilih untuk menentukan kecocokan model konstruk pada penelitian ini adalah  $RMSEA < 0.05$ ,  $SRMR < 0.05$ ,  $CFI > 0.90$ , dan  $NNFI$  (minimal 0.9). Hasil keempat kriteria tersebut dapat disajikan pada Tabel 4 berikut.

**Tabel 4. Perhitungan Nilai Uji Kecocokan Model Kriteria Goodness Of Fit**

Kriteria Goodness of Fit	Syarat	Nilai Hitung	Pemenuhan
RMSEA	<0.05	0.02	Good Fit
SRMR	<0.05	0.04	Good Fit
CFI	>0.90	1.00	Good Fit
NNFI	>0.90	0.95	Good Fit

*Sumber : Data hasil output smartPLS (2026)*

Hasil evaluasi *goodness of fit* terlihat konstruk sudah memenuhi kriteria *goodness of fit* yang telah ditetapkan. Nilai *probability* pengujian *goodness of fit* menunjukkan  $RMSEA$   $0.02 < 0.05$ , dan  $SRMR$   $0.04 < 0.05$ . Hasil uji kecocokan model seperti  $CFI$   $1.00 > 0.90$ ,  $NNFI$   $0.95 > 0.90$  telah memenuhi kriteria pengujian *goodness of fit* sehingga instrumen penilaian berbasis *discovery learning* untuk mengukur keterampilan berpikir kritis dinyatakan fit.

### **Kepraktisan Instrumen Penilaian**

Angket kepraktisan respon guru diambil untuk menilai kepraktisan instrumen penilaian proses pembelajaran akuntansi berbasis *discovery learning* dalam mengukur keterampilan berpikir kritis siswa SMK yang diperoleh dengan teknik penilaian angket kepraktisan. Kepraktisan instrumen penilaian menunjukkan tingkat kemudahan penggunaan instrumen oleh pengguna. Instrumen yang praktis ditandai dengan petunjuk yang jelas, kemudahan dalam pengisian dan pengolahan data, serta efisiensi waktu pelaksanaan sehingga dapat mendukung proses penilaian secara efektif dan efisien.

**Tabel 5. Hasil Analisis Angket Kepraktisan Instrumen Penilaian**

No	Aspek Kepraktisan Instrumen	$\Sigma$ Per Aspek	Kriteria
1	Kejelasan petunjuk mudah dipahami	18	Sangat baik
2	Bahasa mudah dipahami	19	Sangat baik
3	Waktu yang diperlukan cukup efektif	20	Sangat baik
4	Format instrumen memudahkan user	21	Sangat baik
5	Hasil penilaian mudah diolah	21	Sangat baik
<b>Jumlah Skor</b>		<b>99</b>	<b>Sangat baik</b>

*Sumber : Data hasil penelitian (2026)*

Hasil perhitungan respon guru terhadap kepraktisan instrumen penilaian pada Tabel. 5 di atas, seluruh aspek memiliki nilai kepraktisan sebesar 99 dengan kategori "sangat baik" setelah dikonsultasikan pada kriteria skor yang dibuat oleh peneliti. Hasil perhitungan instrumen penilaian juga diperoleh skor berdasarkan setiap aspek indikator dengan kategori "sangat baik". Hasil perhitungan, instrumen penilaian proses pembelajaran akuntansi praktis digunakan untuk mengukur keterampilan berpikir kritis siswa.

#### **Efektivitas Instrumen Penilaian**

Angket efektivitas respon guru dilakukan untuk melihat efektivitas instrumen penilaian. Efektivitas instrumen penilaian merupakan tingkat keberhasilan instrumen dalam mengukur kompetensi yang menjadi sasaran penilaian dan menghasilkan informasi yang akurat untuk mendukung proses evaluasi pembelajaran. Instrumen yang efektif mampu memberikan gambaran yang objektif mengenai capaian belajar siswa serta menjadi dasar dalam pengambilan keputusan untuk perbaikan proses pembelajaran.

**Tabel 6. Hasil Analisis Angket Efektivitas Instrumen Penilaian**

No	Aspek Efektivitas Instrumen	$\Sigma$ Per Aspek	Kriteria
1	Kesesuaian dengan tujuan penilaian	20	Sangat baik
2	Ketepatan pengukuran sesuai kompetensi yang diukur	21	Sangat baik
3	Mendukung pencapaian tujuan pembelajaran	19	Sangat baik
4	Kemampuan membedakan kemampuan siswa	19	Sangat baik
5	Memberikan informasi yang relevan untuk evaluasi	19	Sangat baik
<b>Jumlah Skor</b>		<b>98</b>	<b>Sangat baik</b>

*Sumber : Data hasil penelitian (2026)*

Hasil perhitungan respon guru terhadap efektivitas instrumen penilaian pada Tabel. 6 bahwa seluruh aspek memiliki nilai efektivitas sebesar 98 dengan kategori "sangat baik" setelah dikonsultasikan pada kriteria skor yang dibuat oleh peneliti. Hasil perhitungan instrumen penilaian juga diperoleh skor berdasarkan setiap aspek indikator dengan kategori "sangat baik". Hasil perhitungan tersebut, instrumen penilaian proses pembelajaran akuntansi efektif dalam mengukur keterampilan berpikir kritis siswa.

#### **Pembahasan**

Produk instrumen penilaian proses pembelajaran akuntansi berbasis *discovery learning* untuk mengukur keterampilan berpikir kritis siswa pertama diujicobakan di SMK Negeri 1 Langgudu pada April 2026. Karakteristik model instrumen penilaian disesuaikan dengan konten materi pelajaran akuntansi di SMK dan dikembangkan sesuai dengan 6 aspek kompetensi model pembelajaran berbasis *discovery learning* yaitu: pemberian rangsangan

(*stimulation*), identifikasi masalah (*problem statement*), pengumpulan data (*data collection*), pengolahan data (*data processing*), pembuktian (*verification*), menarik kesimpulan (*generalization*). Instrumen dikembangkan sesuai dengan model pengembangan R&D Borg and Gall (2007) yang meliputi : (1) studi pendahuluan, (2) perencanaan, (3) pengembangan model hipotetik, (4) penelaahan model hipotetik, (5) revisi, (6) uji coba terbatas, (7) revisi hasil uji coba, (8) uji coba luas, (9) revisi model akhir, (10) desiminasi.

Pertama, pada tahap analisis kebutuhan diperoleh informasi bahwa di sekolah SMK belum tersedianya instrumen yang dirancang khusus untuk menilai proses pembelajaran akuntansi berbasis model *discovery learning* dalam mengukur keterampilan berpikir kritis siswa. Keterbatasan guru dalam mengamati siswa dengan jumlah yang besar menyebabkan guru mengandalkan ingatannya untuk mengingat proses pembelajaran siswa. Permasalahan tersebut menunjukkan adanya kebutuhan untuk mengembangkan instrumen penilaian proses pembelajaran yang mampu mengukur kemampuan berpikir kritis siswa secara lebih autentik dan sesuai dengan tahapan model pembelajaran *discovery learning*.

Kedua, langkah berikutnya adalah penyusunan detail instrumen yang dikembangkan, meliputi spesifikasi instrumen, *prototype* (kisi-kisi) instrumen, skala instrumen, dan teknik penskoran instrumen serta sampel pengambilan data uji coba instrumen penilaian.

Ketiga, setelah kisi-kisi dibuat langkah selanjutnya adalah penyusunan rubrik penskoran dan butir instrumen penilaian. Skala likert digunakan untuk pemberian scoring dengan rentang 1 - 4. Pada draft awal instrumen penilaian terdiri dari 24 butir yang disusun berdasarkan indikator model pembelajaran *discovery learning* dalam mengukur keterampilan berpikir kritis siswa. Butir tersebut dipersiapkan untuk pengembangan selanjutnya sehingga diperoleh butir yang valid, reliabel, praktis, efektif mewakili setiap indikator ukur.

Keempat, setelah instrumen penilaian dibuat langkah selanjutnya adalah melakukan analisis data model hipotetik (*prototyping*). Hal ini untuk memastikan bahwa 6 indikator model pembelajaran *discovery learning* yang dikembangkan memiliki keterwakilan untuk mengukur keterampilan berpikir kritis siswa.

Kelima, setelah prototipe instrumen penilaian dianalisis, langkah selanjutnya adalah melakukan penelaahan model hipotetik (*expert judgment*) dengan teknik walktrough kepada 4 orang ahli (ahli akuntansi, ahli pengukuran, ahli bahasa, dan praktisi guru akuntansi). Hasil validasi ahli menunjukkan bahwa kisi-kisi dan instrumen penilaian sudah memenuhi kriteria "Dapat digunakan tanpa revisi" dengan skor total 12 dari skor total maksimal. Beberapa hal pokok yang menjadi masukan dari ke-4 penilai tersebut diantaranya adalah: (1) hindari penggunaan istilah teknis yang dapat menimbulkan penafsiran ganda, (2) hilangkan butir yang memiliki makna tumpang tindih dengan butir lainnya, (3) perbaiki format dan tata letak instrumen agar lebih rapi dan mudah dibaca.

Keenam, hasil validasi ahli kemudian instrumen diujicobakan secara one-to-one untuk melihat tingkat keterbacaan dengan menggunakan angket. Langkah pada tahap ini dilakukan modifikasi dengan menambahkan uji one-to-one kepada tiga guru akuntansi agar sebelum melakukan ujicoba lapangan baik terbatas maupun luas, peneliti dapat mengetahui aspek keterbacaan instrumen penilaian dan waktu yang dibutuhkan untuk mengerjakan tes berdasarkan respon guru. Hasil ujicoba keterbacaan oleh menunjukkan tingkat keterbacaan instrument penilaian sebesar 92,35%. Hasil ujicoba ini menjadi dasar penentuan kepraktisan dan keefektifan instrumen penilaian proses pembelajaran akuntansi.

Ketujuh, hasil ujicoba one-to-one kemudian menjadi dasar draft baru untuk ujicoba lapangan skala terbatas dan skala luas. Hasil ujicoba lapangan skala terbatas menunjukkan

bahwa instrument terdiri dari 16 butir valid dari 24 butir yang disusun pada awal pengembangan dengan tingkat reliabilitas sebesar 0,891. Tingkat reliabilitas ini termasuk pada tingkat reliabilitas yang tinggi. Hasil uji coba skala terbatas kemudian menjadi dasar penyusunan draft untuk uji-coba lapangan skala luas. Hasil uji coba lapangan skala luas tetap menghasilkan 16 butir instrumen yang valid dengan tingkat reliabilitas sebesar 0,851.

Kedelapan, instrumen penilaian terbukti secara konstruk berdasarkan analisis faktor. Hasil analisis menunjukkan bahwa nilai factor loading ( $\lambda$ ) dari variabel dengan indikator-indikatornya secara statistik signifikan dan menghasilkan model konstruk Goodness of Fit. Hasil uji kecocokan keseluruhan model sudah memenuhi kriteria *goodness of fit* yang telah ditetapkan. Nilai *probability* pengujian *goodness of fit* menunjukkan RMSEA  $0.02 < 0.05$ , dan SRMR  $0.04 < 0.05$ . Hasil uji kecocokan model seperti CFI  $1.00 > 0.90$ , NNFI  $0.95 > 0.90$  telah memenuhi kriteria pengujian *goodness of fit* sehingga model instrumen penilaian berbasis *discovery learning* untuk mengukur keterampilan berpikir kritis dinyatakan fit. Hasil analisis reliabilitas model construct reliability (CR) dan nilai AVE diperoleh hasil bahwa setiap indikator reliabel. Nilai reliabilitas semua konstruk telah memenuhi kriteria reliabel, hal ini dibuktikan dengan nilai composite reliability lebih besar dari 0.7, nilai cronbach's alpha  $> 0.7$ , dan memiliki nilai AVE lebih dari 0.5.

#### IV. KESIMPULAN

Penelitian ini telah menghasilkan produk pengembangan berupa instrumen penilaian proses pembelajaran akuntansi berbasis *discovery learning* untuk mengukur keterampilan berpikir kritis siswa. Hasil penelitian diperoleh kesimpulan bahwa semua indikator signifikan dan dapat mengukur variabel laten yaitu indikator pembelajaran *discovery learning* yaitu: simulasi, identifikasi masalah, pengumpulan data, pengolahan data, pembuktian, generalisasi. Ini bermakna 6 indikator dalam penelitian ini secara signifikan dapat mengukur keterampilan berpikir kritis siswa. Penilaian angket kepraktisan dan efektivitas instrumen menunjukkan nilai kepraktisan dan efektifitas yang baik berdasarkan respon pengguna. Hasil analisis data dihasilkan 16 butir yang valid, maka disusun kedalam sebuah perangkat instrumen penilaian berbasis *discovery learning*. 16 indikator ini, menjadi draft final instrumen penilaian proses pembelajaran akuntansi berbasis *discovery learning* untuk mengukur keterampilan berpikir kritis siswa SMK.

#### DAFTAR PUSTAKA

- Allen, M.J. & Yen, W. (1979). *Introduction to Measurement Theory*. Monterey: Brooks Cole.
- Gall, M.D., Gall, J.P., and Borg, W. R. (2007). *Educational Research: An Introduction* (7<sup>th</sup> Ed.). Boston. MA: A & B Publications
- Ghozali, Imam. (2008). *Structural Equation Modeling: Metode Alternatif dengan Partial Least Square (PLS)*. Semarang: Badan Penerbit Universitas Diponegoro
- Chalkiadaki, A. (2018). A systematic literature review of 21st century skills and competencies in primary education. *International Journal of Instruction*, 11(3), 1–16. <https://doi.org/10.12973/iji.2018.1131a>
- Elisabeth, M.B. et al. (2026). Pengaruh Strategi Pembelajaran Inovatif Terhadap Aktivitas Belajar Akuntansi Lanjutan Siswa SMKN 1 Medan. *Jurnal Ilmiah Pendas*. 1(1), 1-10. <https://journal.unpas.ac.id/index.php/pendas/article/view/42149/22662>
- Nieveen, Nienke. (1999). "Prototyping to Reach Product Quality". In J. Vam Den Akker, R

- Branch, K Gustafson, N Nieveen and Tj. Plomp (Eds). Design Approaches and Tools in Education and Training (hlm. 125-136). Dodrecht: Kluwer Academic Publisher.
- Nitko, A.J., Brookhart, S.M. (2007). Educational Assessment of Student (3<sup>rd</sup> Edition). New Jersey: Pearson Education.
- Partono, P., Wardhani, H. N., Setyowati, N. I., Tsalitsa, A., & Putri, S. N. (2021). Strategi Meningkatkan Kompetensi 4C (Critical Thinking, Creativity, Communication, & Collaborative). *Jurnal Penelitian Ilmu Pendidikan*, 14(1), 41–52. <https://doi.org/10.21831/jpipfip.v14i1.35810>
- Resnick, D. P. & Resnick, L. B. (1985). Standards, Curriculum, and Performance: a Historical and Comparative Perspective. *Educational Researcher* 9, 5 -19.
- Shaklee, B.D. et. al. (1997). Designing and Using Portfolios. United States of America: Allyn & Bacon.
- William Wiersma and Stephen G. Jurs. (1990). Educational Measurement and Testing. Boston London Sydney Toronto: Allyn and Bacon
- Wiggins, G. (1984). A True Test: Toward More Authentic and Equitable Assessment. *Phi Delta Kappan*. 70. (9) 703-713
- Yakin, M., I. Kartono., Susilaningsih, E. (2020). Development of Assessment Instruments The Learning Process Chemistry Based Scientific Approach to High School Students. *Journal of Educational Research and Evaluation*. 9(1), 31-39. DOI:[10.15294/jere.v9i1.30361](https://doi.org/10.15294/jere.v9i1.30361)