

## Pengembangan instrumen asesmen high order thinking skills (HOTS) berbasis computerized assisted test (CAT) untuk meningkatkan hasil belajar pada mata pelajaran ekonomi

Dinda Asri Sudioanto<sup>1\*)</sup>, Indra Maipita<sup>1</sup>, Muhammad Bukhori Dalimunthe<sup>1</sup>

<sup>1</sup>Universitas Negeri Medan

\*) Correspondence regarding this article should be addressed to: Dinda Asri Sudioanto Author address e-mail: [ssndad18@gmail.com](mailto:ssndad18@gmail.com)

**Abstract:** Penelitian ini bertujuan untuk mengevaluasi kelayakan, efektivitas, dan kepraktisan instrumen asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT) pada mata pelajaran Ekonomi untuk meningkatkan hasil belajar peserta didik. Penelitian ini menggunakan metode Research and Development (R&D) dengan desain Pra-Eksperimen (One Group Pre-Test–Post-Test Design) dan model pengembangan ADDIE. Instrumen pengumpulan data meliputi angket validasi ahli, angket respons guru dan peserta didik, serta uji analisis butir soal yang mencakup uji validitas, reliabilitas, normalitas, dan homogenitas. Pengujian hipotesis dilakukan melalui uji kelayakan, uji efektivitas, dan uji kepraktisan. Hasil penelitian menunjukkan bahwa instrumen asesmen berbasis CAT sangat layak digunakan dengan hasil validasi ahli materi sebesar 94,23%, ahli media 84,62%, ahli desain pembelajaran 88,89%, dan validasi guru mata pelajaran sebesar 94,61%. Uji coba instrumen pada individu (82,41%), kelompok kecil (72,86%), dan kelompok besar (86,96%) juga menunjukkan kategori layak hingga sangat layak. Instrumen ini terbukti efektif meningkatkan hasil belajar peserta didik, ditunjukkan oleh peningkatan nilai rata-rata dari 29,95 (sebelum intervensi) menjadi 80,51 (setelah intervensi), dengan nilai signifikansi uji-t sebesar 0,000 ( $< 0,05$ ). Instrumen ini juga dinyatakan sangat praktis, dengan nilai kepraktisan rata-rata 86,37% dan nilai efektivitas berdasarkan uji Cohen's d sebesar 6,72 (kriteria sangat besar). Penelitian ini memberikan kontribusi nyata dalam bidang pendidikan, khususnya dalam pengembangan instrumen asesmen berbasis teknologi yang mampu mengukur dan mendorong keterampilan berpikir tingkat tinggi peserta didik secara efektif. Integrasi instrumen berbasis CAT menjadi salah satu solusi strategis dalam meningkatkan kualitas asesmen pembelajaran abad ke-21 di sekolah

**Keywords:** HOTS, Computerized Assisted Test (CAT), Hasil Belajar

**Article History:** Received on 01/01/2025; Revised on 30/01/2025; Accepted on 13/02/2025; Published Online: 02/04/2025.



This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2025 by author.

### INTRODUCTION

Perkembangan teknologi telah mendorong transformasi signifikan dalam dunia pendidikan, termasuk dalam aspek pembelajaran dan evaluasi. Salah satu aspek krusial dalam pendidikan abad ke-21 adalah pengembangan keterampilan berpikir tingkat tinggi atau Higher Order Thinking Skills (HOTS), yang mencakup kemampuan menganalisis,

mengevaluasi, dan mencipta. Keterampilan ini menjadi pilar utama dalam taksonomi Bloom hasil revisi oleh Anderson dan Krathwohl, dan dituntut dalam kurikulum Indonesia saat ini untuk menghasilkan peserta didik yang mampu bersaing secara global. Namun, penelitian menunjukkan bahwa asesmen konvensional berbasis kertas dan pensil memiliki berbagai keterbatasan dalam mengukur kemampuan HOTS secara optimal, seperti keterbatasan dalam adaptivitas dan akurasi pengukuran (Istiyono et al., 2020).

Dalam pembelajaran ekonomi, HOTS berperan penting untuk membentuk peserta didik yang mampu menalar, mengambil keputusan, dan memecahkan masalah secara mandiri. Suwartini (2017) menyebutkan bahwa tujuan mata pelajaran ekonomi di SMA adalah mengembangkan kemampuan peserta didik dalam merancang, menalar, mengolah, dan menyaji informasi secara kritis dan kreatif. Namun kenyataannya, hasil belajar peserta didik masih menunjukkan rendahnya kemampuan dalam menyelesaikan soal-soal HOTS. Berdasarkan data observasi awal di SMAN 1 Kualuh Hulu, ditemukan bahwa 81,9% peserta didik kelas XI IPS tidak mencapai nilai KKM dalam tes HOTS mata pelajaran ekonomi. Mereka lebih nyaman mengerjakan soal berbasis hafalan (LOTS) dan cenderung kesulitan menghadapi soal analitis atau aplikatif.

Kelemahan ini tidak hanya bersumber dari kesiapan peserta didik, tetapi juga dari kurangnya ketersediaan instrumen HOTS yang valid dan terstandar. Guru menyatakan bahwa belum banyak soal HOTS yang tersedia, apalagi yang disesuaikan secara kontekstual dengan pembelajaran ekonomi. Hal ini sejalan dengan temuan Budiman dan Jailani (2014) serta Dalman dan Junaidi (2022), bahwa guru sering mengalami kesulitan dalam menafsirkan level kognitif taksonomi Bloom dan mengembangkan soal untuk mengukur keterampilan berpikir tingkat tinggi. Selain itu, Ghani et al. (2017) menegaskan bahwa minimnya instrumen asesmen HOTS turut berkontribusi terhadap rendahnya kemampuan berpikir kritis peserta didik.

Untuk menjawab tantangan tersebut, diperlukan pengembangan instrumen asesmen HOTS yang tidak hanya relevan secara pedagogis tetapi juga responsif terhadap perkembangan teknologi. Salah satu pendekatan yang potensial adalah penggunaan Computerized Assisted Test (CAT). Model asesmen ini tidak hanya memudahkan pengelolaan evaluasi tetapi juga memungkinkan desain soal yang adaptif, interaktif, dan lebih efisien (Mulyadi et al., 2021). CAT juga mampu meningkatkan motivasi dan keterlibatan belajar peserta didik karena tampilannya yang dinamis dan berbasis digital (Sulistyaningsih et al., 2023). Penelitian terbaru juga menunjukkan bahwa integrasi HOTS dengan platform digital atau berbasis komputer mampu meningkatkan pemahaman konseptual dan hasil belajar secara signifikan (Alim et al., 2020; Firmansyah et al., 2022).

Namun demikian, terdapat gap yang nyata dalam literatur dan praktik pendidikan, khususnya pada pengembangan instrumen HOTS berbasis CAT yang spesifik untuk mata pelajaran ekonomi. Sebagian besar penelitian sebelumnya masih berfokus pada pengembangan HOTS untuk bidang matematika dan IPA, sementara kajian dalam konteks ekonomi SMA masih sangat terbatas. Padahal, konteks ekonomi menuntut aplikasi konsep dalam kehidupan nyata yang sangat sesuai untuk ditanamkan melalui pendekatan HOTS.

---

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan instrumen asesmen HOTS berbasis Computerized Assisted Test (CAT) yang kontekstual dan valid dalam meningkatkan hasil belajar peserta didik pada mata pelajaran ekonomi. Pengembangan ini diharapkan dapat membantu guru dalam menyiapkan instrumen evaluasi yang sesuai dengan kurikulum dan kebutuhan peserta didik, serta menjadi langkah strategis dalam mendorong pembelajaran ekonomi yang bermakna dan berpikir tingkat tinggi. Oleh karena itu, peneliti mengangkat judul: "Pengembangan Instrumen Asesmen Higher Order Thinking Skills (HOTS) Berbasis Computerized Assisted Test (CAT) untuk Meningkatkan Hasil Belajar pada Mata Pelajaran Ekonomi di Kelas XI IPS SMAN 1 Kualuh Hulu."

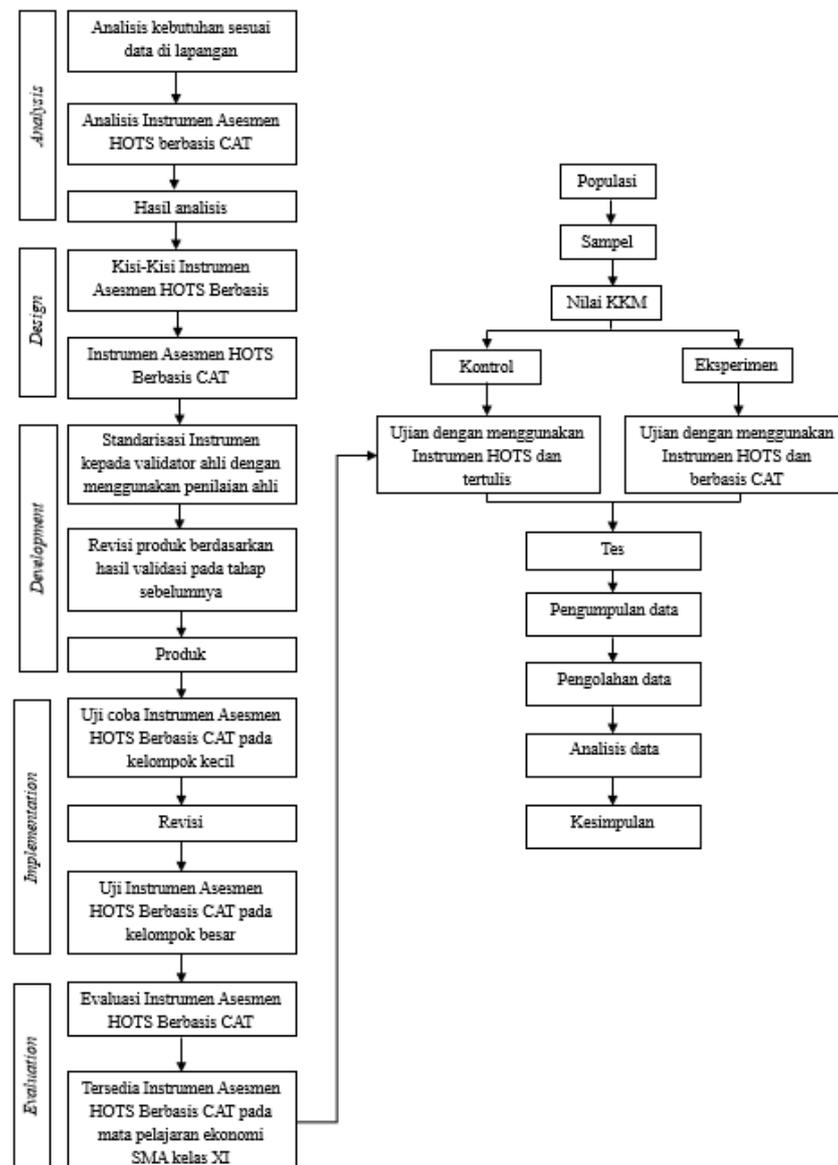
## **METHOD**

Metode penelitian yang digunakan dalam studi ini adalah metode penelitian pengembangan (Research and Development/R&D) yang dikombinasikan dengan pendekatan pra-eksperimen. Tujuan utama dari penelitian ini adalah untuk mengembangkan instrumen asesmen Higher Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT) pada mata pelajaran ekonomi kelas XI IPS. Model pengembangan yang digunakan dalam penelitian ini adalah model ADDIE, yang terdiri dari lima tahapan sistematis, yaitu Analysis, Design, Development, Implementation, dan Evaluation. Pemilihan model ADDIE didasarkan pada kemampuannya dalam menyusun proses pengembangan pembelajaran secara terstruktur dan adaptif.

Pada tahap analisis (analysis), peneliti melakukan identifikasi terhadap kebutuhan peserta didik, karakteristik siswa, tujuan pembelajaran, serta analisis kompetensi dasar pada mata pelajaran ekonomi. Informasi diperoleh melalui observasi kelas dan wawancara guru. Selanjutnya, tahap perancangan (design) dilakukan dengan menyusun kisi-kisi instrumen asesmen HOTS berdasarkan taksonomi Bloom revisi (C4–C6), serta perencanaan media berbasis CAT. Tahap pengembangan (development) dilakukan dengan menyusun soal HOTS berbasis stimulus kontekstual dan mengujikannya melalui validasi oleh tiga ahli, yakni ahli materi, ahli media, dan ahli desain pembelajaran. Setelah direvisi berdasarkan saran ahli, produk diuji secara bertahap: uji coba perorangan (3 peserta didik), uji coba kelompok kecil (6 peserta didik berkemampuan tinggi, sedang, rendah), dan uji coba kelompok besar (33 peserta didik kelas XI-F SMAN 1 Kualuh Hulu).

Tahap implementasi (implementation) melibatkan penerapan produk di kelas dalam situasi pembelajaran nyata. Guru menggunakan instrumen asesmen HOTS berbasis CAT dalam proses asesmen hasil belajar peserta didik. Pelaksanaan asesmen dilakukan secara daring dengan dukungan perangkat komputer/laptop yang tersedia di sekolah. Peneliti juga mengamati langsung proses pelaksanaan untuk mengukur kepraktisan dan kendala penggunaan. Setelah pelaksanaan, masuk ke tahap evaluasi (evaluation). Evaluasi dilakukan secara formatif dan sumatif. Evaluasi formatif mencakup hasil validasi ahli dan umpan balik peserta didik, sementara evaluasi sumatif bertujuan untuk mengetahui efektivitas instrumen melalui hasil pretest dan posttest peserta didik. Pengambilan sampel dilakukan menggunakan teknik purposive sampling dengan kriteria peserta didik kelas XI IPS yang telah mempelajari materi ekonomi yang relevan, aktif mengikuti pembelajaran, dan memiliki keterampilan dasar dalam menggunakan komputer. Total peserta untuk uji coba kelompok besar adalah 33 peserta didik, yang

dipilih dengan mempertimbangkan keterwakilan dari berbagai kategori kemampuan akademik. Pemilihan jumlah sampel didasarkan pada pertimbangan kelayakan untuk uji validitas dan efektivitas produk pengembangan awal.



Gambar 1 Prototype Prosedur Penelitian Pengembangan Instrumen Asesmen High Order Thinking Skills (HOTS) Berbasis Computerized Assisted Test (CAT)

Data dalam penelitian ini terdiri atas data kualitatif dan kuantitatif. Data kualitatif diperoleh dari hasil validasi para ahli serta tanggapan peserta didik terhadap produk, sedangkan data kuantitatif berasal dari hasil uji coba instrumen asesmen HOTS. Instrumen pengumpulan data meliputi lembar validasi, lembar tanggapan kepraktisan, dan butir soal HOTS berbasis CAT. Teknik analisis data meliputi analisis deskriptif dan inferensial. Analisis deskriptif digunakan untuk mengukur validitas isi (dari ahli), daya beda soal, tingkat kesukaran, serta kepraktisan produk. Analisis inferensial dilakukan untuk menguji efektivitas produk melalui uji perbedaan rata-rata antara hasil pretest dan

posttest menggunakan paired sample t-test dengan bantuan program SPSS. Uji normalitas dan homogenitas juga dilakukan sebelumnya untuk memastikan bahwa data memenuhi syarat uji statistik parametrik. Dengan pendekatan ini, penelitian diharapkan mampu menghasilkan instrumen asesmen HOTS berbasis CAT yang valid, praktis, dan efektif untuk meningkatkan hasil belajar ekonomi peserta didik.

## RESULTS AND DISCUSSION

### Hasil Pengembangan

Produk yang sudah dibuat harus melalui tahap uji kelayakan, uji keefektifan dan uji kepraktisan. Dimana uji tersebut divalidasi agar produk tersebut layak, efektif dan praktis untuk digunakan. Validasi dilakukan oleh dosen validator yang sudah ahli di setiap bidangnya. Validasi yang dilakukan terdiri dari validasi ahli materi, validasi ahli media, dan validasi ahli desain pembelajaran. Produk juga diberikan kepada guru mata pelajaran ekonomi untuk dinilai dan diberi masukan. Penilaian berupa saran perbaikan menjadi pedoman bagi peneliti untuk melakukan proses perbaikan terhadap instrumen asesmen HOTS berbasis CAT agar menjadi produk yang lebih baik lagi. Untuk mengukur efektivitas instrumen asesmen HOTS berbasis CAT terhadap hasil belajar maka perlu untuk dilakukan uji normalitas, uji homogenitas, uji T-test dan uji Cohen's.

### Hasil Uji Coba Produk

#### Validitas Isi

Validitas isi dilakukan oleh 3 ahli dari ahli materi, ahli media dan ahli design pembelajaran. dimana diperoleh dari ke 60 butir soal yang dikembangkan memperoleh nilai yang baik. Dimana nilai terendah yang di dapat terdapat pada butir soal nomor 1 dengan perolehan skor total 0,6 yang artinya lebih dari 0,4 indeks kesepakatan maka dikatakan validitasnya sedang. Nilai tengah yang diperoleh dari beberapa butir soal seperti contoh butir soal nomor 3,4,5 dan seterusnya memperoleh skor total 0,8 dimana jika dilihat dari indeks kesepakan maka 0,8 dikatakan validitasnya sedang. Kemudian terdapat beberapa butir soal yang memperoleh nilai 0,9-1 seperti butir soal nomor 13,21,30 dan seterusnya jika dilihat dari indeks kesepakatannya maka validitasnya tinggi. Adapun beberapa masukan dan perbaikan yang diberikan oleh validator. Seperti saran dan perbaikan yang diberikan oleh ahli 1 "masih terdapat beberapa soal yang jawabannya masih meragukan, harus ada clue yang menjadi kunci jawaban pada uraian soal sebelumnya". Saran dan perbaikan dari ahli 3 "Pilihan jawaban dengan kategori soal angka maka dibuat berurut dari yang terbesar ke terkecil maupun sebaliknya, serta peringkasan lagi ilustrasi yang terlalu panjang dan rumit". Berikut perbaikan yang telah dilakukan.

#### Ahli Materi

Hasil dari uji ahli materi dapat disimpulkan bahwa total nilai di dapat adalah 98 dengan perolehan 94,23% dan kriteria sangat layak. Dengan saran perbaikan penambahan referensi untuk eksplorasi bahan dan contoh kasus.

## Ahli Media

Hasil dari uji ahli media dapat disimpulkan bahwa diperoleh hasil dengan total skor sebesar 44, persentase sebesar 84,62 dengan kriteria sangat layak. Tidak ada tambahan perbaikan mengenai web CAT, dengan kata lain web CAT sudah layak digunakan dan diujicobakan perorangan dan kelompok kecil.

Tabel 1 Komentar Validator dan Butir Instrumen Yang Perlu Diperbaiki

Nomor Instrumen Asesmen HOTS	Komentar Ahli
4, 24,	Apabila terdapat pilihan jawaban angka maka diurut dari yang terkecil-terbesar atau terbesar-terkecil.
13, 34, 37, 43, 48	- Stimulus terlalu panjang, ringkas narasi yang terlalu panjang agar tidak menyulitkan dan membingungkan peserta didik. - Waktu disesuaikan dengan bentuk instrumen dan tingkat kesulitan instrumen - Perlu menambahkan referensi untuk eksplorasi bahan dan contoh-contoh kasus.
19	Harus ditambahkan unsur <i>social sustainability</i>
25	Diperhatikan kembali rumusan pendekatan perkapita
52	Tidak ada jawaban yang salah. Harus ada clue yang menjadi kunci jawaban pada uraian instrumen sebelumnya.

## Ahli Design Pembelajaran

Hasil dari uji ahli design pembelajaran dapat disimpulkan bahwa pada validasi ahli design pembelajaran memperoleh total skor 64 dengan persentase 88,89% dan kriteria sangat layak. Dengan perbaikan "pemuatan narasi/stimulus yang digunakan terlalu panjang serta beberapa redaksi yang dimuat". Maka perbaikannya adalah meringkas stimulus yang sudah ada menjadi narasi sederhana serta pemilihan topik studi kasus yang lebih ringan pembahasannya. Revisi dan masukan dari ketiga validator ahli akan diselesaikan sebelum diimplementasikan kepada peserta didik kelas XI-A yang dijadikan sebagai uji coba perorangan dan kelompok kecil.

## Penilaian Tanggapan Guru

Hasil tanggapan guru terhadap pengembangan instrumen asesmen HOTS berbasis CAT dilakukan oleh 3 guru mata pelajaran ekonomi kelas XI SMA Negeri 1 Kualuh Hulu. Penilaian ini bertujuan untuk memperoleh informasi terkait kekurangan produk agar instrumen asesmen HOTS berbasis CAT bisa lebih layak digunakan. Hasil dari penilaian tanggapan guru dapat disimpulkan bahwa memperoleh nilai rata-rata sebesar 64,33 dengan rata-rata persentase 94,61% dan kriteria sangat layak. Hal ini menyatakan bahwa instrumen asesmen HOTS berbasis CAT yang ada sudah layak digunakan.

## Uji Coba Perorangan

Uji coba perorangan dilakukan pada satu orang peserta didik dengan tingkat prestasi tinggi, satu orang peserta didik dengan tingkat prestasi sedang dan satu orang peserta didik dengan tingkat prestasi rendah diujikan secara individual. Uji ini dilakukan untuk mengetahui bagaimana peserta didik merespon instrumen asesmen HOTS berbasis

CAT. Hasil dari uji coba perorangan dapat disimpulkan bahwa memperoleh rata-rata keseluruhan 82,41% dalam kriteria sangat layak. Rinciannya pada aspek materi memperoleh nilai rata-rata 80,56% kriteria layak, aspek penyajian memperoleh nilai rata-rata 83,33% kriteria sangat layak dan aspek bahasa memperoleh rata-rata 83,33% dengan kriteria sangat layak.

Tabel 2 Persentase Rata-Rata Hasil Penilaian Instrumen Asesmen High Order Thinking Skills (HOTS) berbasis CAT Uji Coba Perorangan.

Aspek yang Dinilai	Persentase	Kriteria
Aspek Materi	80,56%	Layak
Aspek Penyajian	83,33%	Sangat Layak
Aspek Bahasa	83,33%	Sangat Layak
Rata-Rata	<b>82,41%</b>	<b>Sangat Layak</b>

Sumber data diolah

### Uji Coba Kelompok Kecil

Pada uji coba kelompok kecil terdapat 6 peserta didik kelas XI-A SMA Negeri 1 Kualuh Hulu yang menjalani uji coba kelompok kecil ini. Dimana terdiri dari 2 orang peserta didik berprestasi tinggi, 2 orang peserta didik berprestasi sedang dan 2 orang peserta didik berprestasi rendah. Hasil dari uji coba kelompok kecil dapat disimpulkan bahwa nilai rata-rata keseluruhan 72,72% dengan kriteria layak. Dengan rincian pada aspek materi memperoleh rata-rata 68,57% dengan kriteria layak, aspek penyajian memperoleh rata-rata 70,83% dengan kriteria layak serta aspek bahasa memperoleh rata-rata 79,17% dengan kriteria layak. Daftar lengkap dapat dilihat pada tabel 4.10 dibawah ini.

Tabel 3 Persentase Rata-Rata Hasil Penilaian Instrumen Asesmen High Order Thinking Skills (HOTS) berbasis CAT Uji Coba Kelompok Kecil

Aspek yang Dinilai	Persentase	Kriteria
Aspek Materi	68,57%	Layak
Aspek Penyajian	70,83%	Layak
Aspek Bahasa	79,17%	Layak
<b>Rata-Rata</b>	<b>72,72%</b>	<b>Layak</b>

Sumber data diolah

### Validitas Butir Soal

Validitas butir soal dianalisis menggunakan persamaan product moment dengan bantuan aplikasi SPSS statistic versi 30.0 for windows yang dikonsultasikan dengan rtabel pada taraf signifikan 5%. Instrumen tes HOTS dikatakan valid apabila  $\text{sig} < 0,05$  atau  $\text{rxy} > \text{rtabel}$ . Dari 60 instrumen tes yang telah dikerjakan oleh peserta didik terdapat 4 instrumen tes yang memperoleh hasil tidak valid dengan penjelasan pada instrumen tes HOTS berbasis CAT nomor 39 memperoleh nilai  $V_{hit}$  0,639 dengan nilai  $\text{sig}$  0,064 dimana nilai  $\text{sig} > 0,05$  berarti tidak valid. Selanjutnya pada instrumen tes HOTS berbasis CAT nomor 44 memperoleh nilai  $V_{hit}$  0,512 dengan nilai  $\text{sig}$  0,159 dimana nilai  $\text{sig} > 0,05$  berarti tidak valid. Pada instrumen tes HOTS berbasis CAT nomor 49 memperoleh nilai  $V_{hit}$  0,476 dengan nilai  $\text{sig}$  0,196 dimana nilai  $\text{sig} > 0,05$  berarti tidak valid. Kemudian pada instrumen tes HOTS berbasis CAT nomor 55 memperoleh nilai  $V_{hit}$  0,496 dengan

nilai sig 0,175 dimana nilai sig > 0,05 berarti tidak valid. Penggolongan kriteria item tes diterima, direvisi dan ditolak merujuk pada pendapat (Retnawati, 2016) validitas butir soal dapat di terima apabila butir soal tergolong dalam kategori valid. Butir soal direvisi apabila tergolong dalam kategori valid/tidak valid. Dan selanjutnya butir soal dikatakan ditolak apabila tergolong dalam kategori tidak valid.

Dengan ketentuan kualitas butir soal diterima, direvisi dan ditolak Retnawati (2016) item tes direvisi apabila salah satu atau lebih dari ketiga kriteria karakteristik item tes tidak memenuhi kriteria. Dikarenakan pada keempat butir soal HOTS berbasis CAT hanya pada uji validitas yang tidak memenuhi kriteria maka keempat butir soal HOTS berbasis CAT tersebut dilakukan revisi.

### Reliabilitas

Teknik analisis reliabilitas menggunakan teknik formula alpha Cronbach dengan menggunakan bantuan aplikasi SPSS statistic versi 30.0 for windows. Tinggi rendahnya tingkat reliabilitas dapat ditentukan dengan melihat nilai rxy mendekati 1 dengan kesepakatan nilai 0,70. Dari 60 instrumen tes yang diujikan memperoleh nilai 0,987 dengan kata lain  $0,987 > 0,70$  memiliki reliabilitas sangat tinggi.

Tabel 4 Data Uji Reliabilitas Kelompok Kecil

Cronbach's Alpha	N of Items
,987	60

### Daya Beda

Analisis daya beda butir soal pilihan ganda pada instrumen asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT) dihitung menggunakan bantuan aplikasi AnBuSo (Analisis Butir Soal) yang dikembangkan oleh (Muhson, 2017). Hasil analisis pada 60 item yang diujikan di dapat bahwa 56 item butir soal memiliki daya beda baik sekali dengan rentang nilai 0,5-1,00. Kemudian terdapat 3 butir soal pada nomor 13, 49 dan 55 memiliki daya beda baik dengan rentang nilai 0,3-0,49. Dan terdapat 1 butir soal pada nomor 52 memiliki daya beda tidak baik/jelek dengan rentang nilai 0,00-0,19.

### Tingkat Kesukaran

Soal dapat dikatakan baik apabila tidak terlalu mudah dan tidak terlalu sulit. Sebab soal yang terlalu mudah tidak mampu merangsang peserta didik untuk meningkatkan usaha dalam menyelesaikannya, sebaliknya soal yang terlalu sulit dapat menyebabkan berkurangnya minat peserta didik dalam menyelesaikan soal-soal tersebut. Tingkat kesukaran soal dengan kategori baik berada pada rentang sedang dengan  $0,3 < p \leq 0,7$ . Berdasarkan analisis data, dapat diperoleh instrumen asesmen HOTS berbasis CAT yang berjumlah 60 butir soal dengan kategori sedang berjumlah 19 butir (31,67%). Dan kategori mudah terdapat 41 butir soal 68,33%.

### Efektivitas Pengecoh

Pengecoh merupakan opsi atau pilihan jawaban yang lain dari jawaban yang benar. Suatu opsi atau pilihan jawaban disebut efektif jika memenuhi fungsinya atau tujuan disajikannya opsi tersebut tercapai. Hal ini berarti bahwa setiap opsi yang

disajikan masing-masing mempunyai kemungkinan yang sama untuk dipilih, jika tes menjawab soal itu dengan menerka-nerka (spekulasi). Pengecoh dikatakan berfungsi apabila semua pengecoh pada tiap soal berfungsi baik, jika pengecoh ternyata belum berfungsi atau menyesatkan maka pengecoh tersebut perlu direvisi. Hasil dari dilakukannya uji analisis butir soal dan diperoleh hasil bahwa pada butir soal yang melalui revisi pengecoh terdapat pada nomor 7 pilihan E menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 10 pilihan B menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 11 pilihan ACE menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 12 pilihan A menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 15 pilihan D menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 20 pilihan BC menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 21 pilihan B menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 26 pilihan B menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 31 pilihan BD menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 32 pilihan B menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 33 pilihan BC menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 39 pilihan AE menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 45 pilihan B menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 48 pilihan AC menjadi alternatif jawaban yang tidak efektif. Pada soal nomor 52 pilihan B menjadi alternatif jawaban yang tidak baik. Pada soal nomor 55 pilihan B menjadi alternatif jawaban yang tidak efektif.

### Uji Coba Kelompok Besar

#### Respon Peserta Didik

Peserta didik kelas XI-F SMA Negeri 1 Kualuh Hulu yang berjumlah 33 menjalani uji coba kelompok besar. Hasil respon peserta didik pada uji kelompok besar yang dilakukan pada 33 responden kelas XI-F SMA Negeri 1 Kualuh Hulu berdasarkan aspek yang dinilai mendapatkan nilai rata-rata keseluruhan 86,96% dengan kriteria sangat layak. Dengan rincian rata-rata aspek materi 84,1% kriteria sangat layak, aspek penyajian 84,09% kriteria sangat layak dan aspek bahasa 92,68% kriteria sangat layak.

Tabel 5 Persentase Rata-Rata Hasil Penilaian Instrumen Asesmen High Order Thinking Skills (HOTS) berbasis CAT Uji Coba Kelompok Besar

Aspek yang Dinilai	Persentase	Kriteria
Aspek Materi	84,1%	Sangat Layak
Aspek Penyajian	84,09%	Sangat Layak
Aspek Bahasa	92,68%	Sangat Layak
<b>Rata-Rata</b>	<b>86,96%</b>	<b>Sangat Layak</b>

Sumber data diolah

### Hasil Belajar

#### Data Pretest

Pretest dilakukan dengan memberikan instrumen test HOTS yang sudah valid dan reliabel dan layak untuk digunakan sebanyak 60 soal pilihan berganda dan dilakukan dengan ujian menggunakan lembar jawaban berupa kertas. Tabel 6 menunjukkan hasil pretest peserta didik.

Tabel 6 Data Pretest Hasil Peserta Didik

<i>Mean</i>	<i>Std. Deviation</i>	<i>Minimum</i>	<i>Maximum</i>
29,95	9,02	13	47

Sumber: Data diolah peneliti

Dari tabel 6 menunjukkan hasil pretest peserta didik dengan nilai rata-rata 29,95, standar deviasi 9,02 dengan nilai minimum 13 serta nilai maksimum 47.

#### Data Posttest

Data posttest diambil dari hasil belajar setelah mengerjakan instrumen asesmen HOTS berbasis CAT. Pada tahap eksperimen ini menggunakan instrumen asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT) untuk meningkatkan hasil belajar pada mata pelajaran ekonomi di kelas XI IPS SMAN 1 Kualuh Hulu tahun ajaran 2024/2025. Perbedaan pada tahap pretest dengan tahap posttest terletak pada pemberian perlakuan serta pengerjaan dan pelaksanaan ujian berbasis CAT. Pada saat tahap pretest peserta didik tidak diberi perlakuan dan mengerjakan instrumen test manual menggunakan lembar jawaban berupa kertas. Maka dari itu dapat dilihat pada tabel 7.

Tabel 7 Data Posttest Hasil Nilai Peserta Didik

<i>Mean</i>	<i>Std. Deviation</i>	<i>Minimum</i>	<i>Maximum</i>
80,51	5,64	72	92

Sumber: Data diolah peneliti

Dari tabel 7 menunjukkan hasil posttest memperoleh rata-rata 80,51, standar deviasi 5,64, nilai minimal 72 serta nilai maksimal 92. Tabel 4.16 menunjukkan rentang nilai hasil belajar berdasarkan kategori kelulusan KKM yaitu 75. Data selengkapnya dapat di lihat pada tabel 8.

Tabel 8 Rentang Nilai Pelajaran Ekonomi SMAN 1 Kualuh Hulu

<b>Nilai</b>	<b>Kategori</b>	<b>Kelas XI F</b>	
		<b>Frekuensi</b>	<b>Persentase</b>
90-100	Sangat Baik	2	6%
80-89	Baik	17	52%
70-79	Cukup	14	42%
< 69	Kurang	0	0%

Tabel 8 menunjukkan rentang nilai hasil belajar peserta didik dimana pada mata pelajaran ekonomi SMAN 1 Kauluh Hulu menetapkan KKM pada nilai 75. Berdasarkan posttest yang telah dilakukan didapatkan hasil belajar peserta didik dengan menggunakan instrumen asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT) di kelas XI F sebagian besar sudah mencapai nilai KKM. Dimana 2 peserta didik atau sebesar 6% dengan kategori sangat baik, 17 peserta

didik atau sebesar 52% dengan kategori baik, dan 14 peserta didik atau sebesar 42% dengan kategori cukup. Nilai hasil belajar ini disebut nilai posttest.

### Uji Prasyarat

#### Uji Normalitas

Cara melihat data berdistribusi normal dapat melakukan uji normalitas. Dalam penelitian ini teknik Shapiro-Wilk Test dengan jumlah  $n = 33$ . Berikut disajikan hasil dari uji normalitas.

Tabel 9 Tests of Normality

	Tahap	Kolmogorov-Smirnov <sup>a</sup>			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Hasil	Pretest	,118	33	,200*	,959	33	,235
Belajar	Posttest	,152	33	,052	,943	33	,084

\*. This is a lower bound of the true significance.

Dari tabel 9 dapat disimpulkan nilai pretest kelas XI-F dimana  $\text{sig } 0,235 > 0,005$  dan nilai posttest dimana  $\text{sig } 0,084 > 0,05$ . Berdasarkan data tersebut maka kelas XI-F datanya memiliki nilai berdistribusi normal.

#### Uji Homogenitas

Selanjutnya pengujian dilakukan dengan menguji homogenitas guna melihat varians yang sama atau berbeda pada tahap pretest dan posttest. Dikatakan mempunyai nilai varians yang sama atau berbeda (homogen) apabila taraf  $\text{sig} > 0,05$ , uji homogenitas ini menggunakan bantuan aplikasi SPSS 27 for window. Data selengkapnya dapat dilihat pada tabel 10:

Tabel 10 Hasil Uji Homogenitas

		Tests of Homogeneity of Variances			
		Levene Statistic	df1	df2	Sig.
Hasil	Based on Mean	1,363	1	63	,247
Belajar	Based on Median	1,183	1	63	,281
	Based on Median and with adjusted df	1,183	1	57,559	,281
	Based on trimmed mean	1,450	1	63	,233

Hasil uji homogenitas pada tabel 10 menunjukkan hasil  $\text{sig } 0,247 > 0,05$  maka data memiliki varians yang homogen.

#### Uji Hipotesis

#### Uji Kelayakan

Uji kelayakan instrumen asesmen yang telah dikembangkan diambil dari nilai validitas ahli materi, validitas ahli media, validitas ahli design pembelajaran, angket respon guru, uji perorangan, uji kelompok kecil dan uji kelompok besar. Dibawah ini merupakan tabel hasil uji kelayakan terhadap instrumen asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT). Lembar validasi para ahli,

tanggapan guru dan respon peserta didik selanjutnya dianalisis persentase kelayakannya (Ernawati & Sukardiyono, 2017). Rumus persentase dapat dihitung dengan rumus sebagai berikut:

$$\text{Persentase Kelayakan Instrumen} = \frac{\text{Jumlah Validator}}{\text{Jumlah Validator}} \times 100\%$$

Setelah di dapat hasil presentasinya lalu kita masukkan ke rentang kriteria skor kelayakan. Kriteria instrumen asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT) dapat dikatakan layak digunakan jika seluruh persentase validator ditotalkan lalu dibagi dengan jumlah validator dan mendapatkan persentase sebesar “ $\geq 61\%$ ”. Berikut hasil persentase kelayakan instrumen:

Tabel 11 Rekapitulasi Penilaian Kelayakan Pengembangan Instrumen Asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT)

Validator, Guru, Peserta Didik	Persentase Kelayakan
Ahli Materi	94,23%
Ahli Media	84,62%
Ahli Design Pembelajaran	88,89%
Guru Mata Pelajaran	94,61%
Uji Coba Perorangan	82,41%
Uji Coba Kelompok Kecil	72,86%
Uji Coba Kelompok Besar	86,96%
<b>Rata-Rata</b>	<b>86,37%</b>

Berdasarkan tabel 11, jika dirata-ratakan maka total persentase dari seluruh validator ditambah dengan persentase tanggapan guru mata pelajaran ditambah dengan respon peserta didik pada uji perorangan, uji kelompok kecil dan uji kelompok besar dibagi dengan jumlah yang menilai maka memperoleh hasil rata-rata 86,37% dan mendapatkan kriteria “sangat layak”.

#### Uji Efektifitas

Dalam uji ini menggunakan uji hipotesis untuk melihat apakah instrumen asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT) efektif dilihat dari hasil belajar ekonomi peserta didik kelas XI-F SMAN 1 Kualuh Hulu. Langkah yang pertama kali dilakukan adalah menentukan  $H_0$  dan  $H_a$  terlebih dahulu. Rumusan hipotesisnya sebagai berikut:

$H_0: \mu_1 = \mu_2$  ini berarti tidak terdapat perbedaan rata-rata antara tahap kontrol dengan tahap eksperimen pada kelas XI-F

$H_a: \mu_1 \neq \mu_2$  ini berarti terdapat perbedaan rata-rata antara tahap kontrol dengan tahap eksperimen pada kelas XI-F

Langkah selanjutnya menentukan tingkat sig ( $\alpha$ ). Taraf signifikansi adalah 0,05. Lalu dilakukan uji varian. Pada tahap ini sudah dilakukan uji varian dan dinyatakan bersifat homogen dan berdistribusi normal. Software yang digunakan software SPSS versi 27 for windows dalam pengujian perbedaan rata-rata kelas XI-F tahap eksperimen melalui nilai posttest. Berikut hasilnya di tabel 12.

Tabel 12 Paired Samples Test

		Paired Differences						
				95% Confidence Interval of the Difference				
Pair		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	Sig. (2-tailed)	
1	Pretest - Posttest	50,576	8,526	1,484	53,599	47,553	34,077	2 ,000

Dari tabel 12 nilai  $34,077 > 1,692$  maka  $H_0$  ditolak. Nilai sig  $0,000 < 0,05$  kesimpulannya adalah adanya perbedaan hasil belajar tahap kontrol dan eksperimen yang diberi treatment pemberian kisi-kisi dan tips mengerjakan soal-soal HOTS dan dengan pelaksanaan ujian manual dengan ujian CAT. Maka kesimpulannya adalah instrumen asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT) yang dikembangkan efektif untuk digunakan.

### Uji Kepraktisan

Uji kepraktisan instrumen asesmen yang telah dikembangkan diambil dari nilai validitas ahli materi, validitas ahli media, validitas ahli design pembelajaran, angket respon guru, uji perorangan, uji kelompok kecil dan uji kelompok besar. Dibawah ini merupakan tabel hasil uji kepraktisan terhadap instrumen asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT). Lembar validasi para ahli, tanggapan guru dan respon peserta didik selanjutnya dianalisis persentase kepraktisannya (Ernawati & Sukardiyono, 2017). Rumus persentase dapat dihitung dengan rumus Sugiyono (2019) sebagai berikut:

$$P = \frac{F}{N} \times 100\%$$

Keterangan:

P= Persentase nilai kepraktisan

F= Perolehan skor

N = Skor ideal

Setelah di dapat hasil presentasenya lalu kita masukkan ke rentang kriteria skor kepraktisan. Kriteria instrumen asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT) dapat dikatakan praktis digunakan jika seluruh persentase validator ditotalkan lalu dibagi dengan jumlah validator dan mendapatkan persentase sebesar " $\geq 61\%$ ".

Berdasarkan tabel 13 jika dirata-ratakan maka total persentase dari seluruh validator ditambah dengan persentase tanggapan guru mata pelajaran ditambah dengan respon peserta didik pada uji perorangan, uji kelompok kecil dan uji kelompok besar

dibagi dengan jumlah yang menilai maka memperoleh hasil rata-rata 86,37% dan mendapatkan kriteria “sangat praktis”.

Tabel 13 Rekapitulasi Penilaian Kepraktisan Pengembangan Instrumen Asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT)

Validator, Guru, Peserta Didik	Persentase Kelayakan
Ahli Materi	94,23%
Ahli Media	84,62%
Ahli Design Pembelajaran	88,89%
Guru Mata Pelajaran	94,61%
Uji Coba Perorangan	82,41%
Uji Coba Kelompok Kecil	72,86%
Uji Coba Kelompok Besar	86,96%
<b>Rata-Rata</b>	<b>86,37%</b>

### Uji Effect Size

Untuk melihat seberapa besar pengaruh dari treatment yang telah diberikan pada tahap eksperimen maka dapat dilihat dari effect size. Maka dilakukannya uji cohen's untuk melihat effect size tersebut. Data selengkapnya dapat dilihat pada lampiran 14 dan 15.

Berdasarkan perhitungan cohen's maka diperoleh nilai cohen's d sebesar 6,72 yang artinya berdasarkan kriteria cohen's d nilai yang diperoleh lebih besar dari 1,30 dengan kriteria “sangat besar”. Instrumen asesmen High Order Thinking Skills (HOTS) berbasis Computerized Assisted Test (CAT) memiliki efek pengaruh yang sangat besar terhadap hasil belajar.

### Pembahasan

Berdasarkan Berdasarkan hasil observasi awal dan data empiris yang dikumpulkan, ditemukan bahwa peserta didik kelas XI SMAN 1 Kualuh Hulu cenderung kesulitan dalam menyelesaikan soal-soal yang dirancang untuk mengukur Higher Order Thinking Skills (HOTS). Ketika dihadapkan dengan soal-soal HOTS, peserta didik menunjukkan kecenderungan untuk enggan membaca secara menyeluruh dan gagal memahami konteks soal yang diberikan. Akibatnya, jawaban yang diberikan seringkali tidak tepat, dan ini berdampak langsung terhadap rendahnya hasil belajar mereka. Kondisi ini diduga kuat dipengaruhi oleh minimnya pembiasaan peserta didik terhadap soal-soal HOTS, serta belum tersedianya instrumen asesmen yang dirancang secara sistematis untuk mengembangkan keterampilan berpikir tingkat tinggi.

Instrumen HOTS yang dikembangkan dalam penelitian ini berbasis pada prinsip-prinsip Computerized Assisted Test (CAT), dengan pendekatan berbasis taksonomi Bloom revisi Anderson dan Krathwohl (2001). Taksonomi ini membagi proses kognitif ke dalam enam level, yaitu: mengingat (C1), memahami (C2), menerapkan (C3), menganalisis (C4), mengevaluasi (C5), dan mencipta (C6). Dalam konteks pengembangan instrumen HOTS, fokus utama diarahkan pada tiga level tertinggi, yakni analisis,

evaluasi, dan penciptaan, karena ketiganya mencerminkan keterampilan berpikir tingkat tinggi yang menjadi tuntutan kurikulum Merdeka dan pembelajaran abad ke-21.

Pengembangan instrumen asesmen HOTS berbasis CAT ini bertujuan untuk memberikan alternatif asesmen yang tidak hanya mengevaluasi hasil belajar, tetapi juga melatih peserta didik untuk berpikir secara kritis, logis, dan reflektif. Instrumen ini memiliki karakteristik unik, antara lain: berbasis stimulus kontekstual, mengandung skenario kehidupan nyata, serta menggunakan item pilihan ganda yang mengharuskan peserta didik mengolah informasi sebelum menjawab. Format seperti ini secara signifikan berbeda dari soal LOTS yang hanya menekankan pada hafalan dan pemahaman dasar.

Efektivitas instrumen terbukti melalui data kuantitatif yang menunjukkan adanya peningkatan signifikan dalam nilai hasil belajar peserta didik setelah diberikan perlakuan. Rata-rata nilai sebelum perlakuan adalah 29,95, sedangkan setelah menggunakan instrumen berbasis CAT, nilai meningkat menjadi 80,51. Kenaikan sebesar 50,56 poin ini menunjukkan bahwa peserta didik tidak hanya mengalami peningkatan dalam aspek kognitif, tetapi juga menunjukkan bahwa mereka mampu menyesuaikan diri dengan pola soal yang menantang secara intelektual. Hal ini diperkuat dengan hasil uji paired sample t-test yang menunjukkan nilai signifikansi  $0,000 < 0,05$ , sehingga dapat disimpulkan bahwa terdapat perbedaan yang signifikan antara hasil belajar sebelum dan sesudah perlakuan.

Beberapa faktor pendukung efektivitas instrumen ini antara lain Desain instrumen yang berbasis teknologi digital (CAT) memberikan pengalaman belajar yang lebih interaktif dan sesuai dengan karakteristik generasi digital-native seperti peserta didik saat ini. Peserta didik merasa lebih tertarik dan termotivasi untuk mengerjakan soal karena tampilan yang modern dan tidak monoton. Stimulus kontekstual dan berbasis fenomena nyata dalam setiap butir soal mendorong peserta didik untuk mengaitkan materi ekonomi dengan kehidupan sehari-hari. Hal ini selaras dengan teori contextual learning yang menekankan pada keterkaitan antara pengalaman belajar dan lingkungan peserta didik. Kesiapan guru dalam mendampingi dan memberikan pembekalan tentang bagaimana mengerjakan soal HOTS sebelum ujian berlangsung juga sangat berpengaruh. Diberikannya kisi-kisi, pelatihan, dan tips menjawab soal kepada peserta didik terbukti meningkatkan kesiapan mental dan pemahaman mereka. Penggunaan aplikasi CAT yang sederhana dan mudah diakses, mempermudah guru dalam mengelola asesmen sekaligus mempercepat proses koreksi hasil ujian. Hal ini juga memberikan umpan balik yang lebih cepat kepada peserta didik, sehingga proses refleksi belajar menjadi lebih optimal.

Meski demikian, pengembangan dan implementasi instrumen ini tidak terlepas dari tantangan. Beberapa faktor penghambat yang teridentifikasi selama pelaksanaan antara lain Keterbatasan akses internet di lingkungan sekolah atau rumah peserta didik menyebabkan gangguan dalam proses ujian berbasis komputer. Ketika jaringan tidak stabil, sistem mengalami loading atau hang, sehingga mengganggu konsentrasi dan alokasi waktu peserta didik. Durasi pengerjaan soal yang kurang proporsional, terutama karena panjangnya stimulus yang membutuhkan waktu baca dan pemahaman yang lebih lama. Hal ini menjadi dilema dalam penentuan waktu ujian yang efektif, karena terlalu singkat akan mengurangi kualitas jawaban, namun terlalu lama akan mengganggu jadwal pembelajaran lainnya. Keterbatasan materi yang diuji hanya pada topik pertumbuhan

dan pembangunan ekonomi juga membatasi generalisasi hasil instrumen ini ke topik ekonomi lain. Keterbatasan jumlah item soal dan cakupan indikator HOTS yang digunakan juga menjadi catatan. Meskipun validasi ahli menunjukkan bahwa instrumen sangat layak digunakan, instrumen ini masih perlu dikembangkan untuk mencakup variasi topik ekonomi dan lebih banyak dimensi HOTS.

Secara keseluruhan, hasil validasi oleh ahli materi, media, dan desain pembelajaran menunjukkan bahwa instrumen ini masuk dalam kategori “sangat layak” dengan skor di atas 84%. Uji kepraktisan dari guru dan peserta didik juga menunjukkan bahwa instrumen ini mudah digunakan dan sesuai dengan tujuan pembelajaran. Hal ini memperkuat bahwa pengembangan instrumen asesmen HOTS berbasis CAT ini tidak hanya valid secara konten, tetapi juga praktis dan dapat diimplementasikan secara luas.

Temuan ini diperkuat oleh hasil penelitian sebelumnya seperti oleh Supranoto (2021), Sa'idah & Isnaini (2020), dan Masitoh & Aedi (2020), meskipun penelitian-penelitian tersebut tidak berbasis CAT. Dengan integrasi teknologi, instrumen dalam penelitian ini memiliki kelebihan dari segi efisiensi, daya tarik peserta didik, serta efektivitas penilaian.

Dengan demikian, instrumen asesmen HOTS berbasis CAT yang dikembangkan dalam penelitian ini terbukti efektif, layak, dan praktis untuk digunakan dalam pembelajaran ekonomi SMA. Diharapkan hasil penelitian ini dapat menjadi referensi dan acuan dalam pengembangan instrumen sejenis di mata pelajaran lain dan jenjang pendidikan lainnya, sekaligus menjadi solusi atas rendahnya kemampuan berpikir tingkat tinggi peserta didik di Indonesia sebagaimana dilaporkan oleh TIMSS dan PISA.

## CONCLUSIONS

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa produk yang dikembangkan dalam bentuk instrumen asesmen *Higher Order Thinking Skills* (HOTS) berbasis *Computerized Assisted Test* (CAT) pada mata pelajaran Ekonomi, khususnya materi pertumbuhan dan pembangunan ekonomi, terbukti valid, reliabel, efektif, dan praktis. Validasi yang dilakukan oleh ahli materi, ahli media, ahli desain pembelajaran, guru mata pelajaran, serta peserta didik menunjukkan bahwa instrumen ini termasuk dalam kategori sangat layak dengan skor validasi rata-rata di atas 85%. Dari 60 butir soal yang dikembangkan, sebanyak 56 soal dinyatakan valid berdasarkan hasil uji validitas ( $\text{sig} < 0,05$ ), sementara 4 soal dinyatakan tidak valid dan telah direvisi sesuai masukan validator. Uji reliabilitas menghasilkan koefisien sebesar 0,987 yang menunjukkan reliabilitas tinggi. Selain itu, daya beda dan tingkat kesukaran soal termasuk dalam kategori baik setelah dilakukan perbaikan pada beberapa pengecoh.

Instrumen asesmen HOTS berbasis CAT ini juga terbukti efektif dalam meningkatkan hasil belajar peserta didik. Hal ini ditunjukkan oleh peningkatan signifikan nilai rata-rata dari pretest sebesar 29,95 menjadi 80,51 pada posttest, dengan selisih sebesar 50,56 poin. Hasil uji-t menunjukkan nilai signifikansi sebesar 0,000 ( $< 0,05$ ) dan nilai *effect size* Cohen's *d* sebesar 6,72 yang tergolong sangat besar, menunjukkan bahwa penggunaan instrumen ini memiliki pengaruh yang signifikan terhadap pencapaian belajar peserta didik. Selain itu, dari sisi kepraktisan, hasil validasi menunjukkan bahwa

instrumen ini dinilai sangat praktis digunakan baik oleh guru maupun peserta didik karena dapat diakses secara daring, memudahkan evaluasi, serta hemat waktu dan tenaga dalam pelaksanaan ujian.

Implikasi praktis dari temuan ini menunjukkan bahwa guru dapat memanfaatkan instrumen HOTS berbasis CAT ini sebagai alat asesmen yang adaptif dan sesuai dengan tuntutan pembelajaran abad ke-21. Selain itu, sekolah dapat mengintegrasikan instrumen ini ke dalam sistem evaluasi pembelajaran untuk mendukung transformasi digital di bidang pendidikan. Instrumen ini juga menjadi rujukan dalam menyusun soal yang mampu mengukur kemampuan berpikir tingkat tinggi secara sistematis dan kontekstual.

Namun demikian, penelitian ini memiliki beberapa keterbatasan, di antaranya pengembangan instrumen hanya mencakup satu materi, yaitu pertumbuhan dan pembangunan ekonomi, serta dilakukan pada satu sekolah dan satu jenjang kelas. Oleh karena itu, disarankan kepada peneliti selanjutnya untuk mengembangkan instrumen serupa pada materi ekonomi lain seperti inflasi, pasar, atau kebijakan fiskal. Penelitian lebih lanjut juga dapat dilakukan pada jenjang pendidikan lain seperti SMP atau SMA kelas X dan XII guna memperluas cakupan penggunaan instrumen. Selain itu, integrasi instrumen dengan fitur analitik otomatis dalam platform CAT juga perlu dikembangkan agar guru dapat secara langsung mengevaluasi kualitas soal dan pencapaian peserta didik. Dengan pengembangan dan penyebarluasan yang lebih luas, instrumen ini diharapkan mampu meningkatkan kualitas pembelajaran serta menumbuhkan kemampuan berpikir kritis dan reflektif peserta didik secara berkelanjutan.

## REFERENCES

- Budiman, A., & Jailani, J. (2014). Pengembangan Instrumen Asesmen Higher Order Thinking Skill (Hots) Pada Mata Pelajaran Matematika Smp Kelas Viii Semester 1. *Jurnal Riset Pendidikan Matematika*, 1(2), 139. <https://doi.org/10.21831/jrpm.v1i2.2671>
- Dalman, R. P., & Junaidi, J. (2022). *Penyebab Sulitnya Siswa Menjawab Soal HOTS dalam Pembelajaran Sosiologi di Kelas XI IPS SMAN 1 Batang Kapas Pesisir Selatan*. 1, 103–112.
- Ernawati, I., & Sukardiyono, T. (2017). Uji Kelayakan Media Pembelajaran Interaktif Pada Mata Pelajaran Administrasi Server. *Elinvo (Electronics, Informatics, and Vocational Education)*, 2(2), 204–210. <https://doi.org/10.21831/elinvo.v2i2.17315>
- Fauziyah MF, A., Hala, Y., & Asmawati Aziz, A. (2023). Uji Kevalidan Instrumen Tes Berorientasi Higher Order Thinking Skill (Hots) Pada Materi Sistem Ekskresi Kelas Xi Sma/Ma. *Jurnal Biogenerasi*, 8(1), 329–333. <https://doi.org/10.30605/biogenerasi.v8i1.2125>
- Fensham, P. J., & Bellocchi, A. (2013). Higher order thinking in chemistry curriculum and its assessment. *Thinking Skills and Creativity*, 10, 250–264. <https://doi.org/10.1016/j.tsc.2013.06.003>
- Ghani, I. B. A., Ibrahim, N. H., Yahaya, N. A., & Surif, J. (2017). Enhancing students' HOTS in laboratory educational activity by using concept map as an alternative assessment tool. *Chemistry Education Research and Practice*, 18(4), 849–874. <https://doi.org/10.1039/c7rp00120g>
- Hadi, S. dan N. (2022). Pengaruh Model Problem Based Learning Terhadap Keterampilan

- 
- Computational Thinking Matematis Siswa Kelas Iv Sdn Kebon Bawang 03 Jakarta. *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, 8(1), 375–385. <https://doi.org/10.36989/didaktik.v8i1.302>
- Haryati, M. (2020). *Analisis Soal Un Biologi Sma / Ma Berdasar Dimensi Proses Kognitif, Karakteristik Hots, Dan Bentuk Stimulus*. 8(2), 91–94.
- Istiyono, E., Marzuki, M., & Suparman, S. (2020). The development of computerized adaptive test to measure higher order thinking skills in physics. *European Journal of Educational Research*, 9(1), 91–101. <https://doi.org/10.12973/eu-jer.9.1.91>
- Koza, Y., Harso, A., & Doa, H. (2024). Pengembangan Instrumen Soal High Order Thinking Skill (Hots) Pada Materi Fluida Statis. *OPTIKA: Jurnal Pendidikan Fisika*, 8(1), 69–78. <https://doi.org/10.37478/optika.v8i1.3555>
- Masitoh, L. F., & Aedi, W. G. (2020). Pengembangan Instrumen Asesmen Higher Order Thinking Skills (HOTS) Matematika di SMP Kelas VII. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, 4(2), 886–897. <https://doi.org/10.31004/cendekia.v4i2.328>
- Muhson, A. (2017). Panduan Penggunaan AnBuso 2017. *Universitas Negeri Yogyakarta*, 1–19.
- Phito, V., Asmaiwaty, A., & ROza, M. (2019). Pengembangan Instrumen Asesmen Higher Order Thinking Skills (HOTS) Dalam Pembelajaran Fisika Pada Materi Hukum Newton Kelas X SMA/MA. *Prosiding Seminar Nasional Fisika PPs UNM*, 5, 787–799. <http://ojs.unm.ac.id/semnasfisika>
- Retnawati, H. (2016). *Analisis Kuantitatif Instrumen Penelitian (Pertama)*. Parama Publishing.
- Sa'idah, N., & Isnaini, I. K. (2020). Pengembangan Instrumen Tes Higher Order Thingking Skill (HOTS) Mata Pelajaran Fikih Kelas XI di MA Masalilik Huda Tahunan Jepara. *Al-Tadzkiyyah: Jurnal Pendidikan Islam*, 11(2), 209–220.
- Setiawati, W., Asmira, O., Ariyana, Y., Bestary, R., & Pudjiastuti, A. (2019). Buku Penilaian Berorientasi Higher Order Thinking Skills. *Direktorat Jenderal Guru dan Tenaga Kependidikan Kementerian Pendidikan dan Kebudayaan*, 1–82.
- Sugiyono. 2019. *Metode Penelitian Pendidikan (Pendekatan Kuantitatif, Kualitatif dan R&D)*. Bandung: Alfabeta.
- Supranoto, H. (2021). *Penggunaan Soal HOTS Ekonomi Berbasis CBT untuk Meningkatkan Hasil Belajar Siswa Kelas XII SMAN 2 Ulubelu*. 1(1), 1–9.
- Sutami, Y. Hudiyono, M. I. (2020). Pengembangan Instrumen Asesmen Higher Order Thinking Skills (Hots) Pada Mata Pelajaran Bahasa Indonesia Sma Dan Smk. *Diglosia: Jurnal Kajian, Sastra dan Pengajarannya*, 3(1), 102–113.
- Suwartini, S. H. & Y. P. (2017). *Wiyata dharma*. V(November), 162–171.
-