

PENGEMBANGAN INSTRUMEN TES BERPIKIR KRITIS BERBASIS SINTESIS WATSON-GLASER DAN ENNIS

Tariza Fairuz^{1*}, Nadia Mubarakah², Sri Masnita Pardosi³, Shofia Lubis⁴

Pendidikan IPA, Universitas Negeri Medan^{1,2,3}

Pendidikan Biologi, Universitas Negeri Medan⁴

Email: tarizafairuz@unimed.ac.id*

Abstrak

Keterbatasan instrumen asesmen IPA yang mampu mengukur kemampuan berpikir kritis secara komprehensif, terutama dalam mengintegrasikan dimensi kognitif dan indikator operasional ke dalam satu konstruk yang utuh masih menjadi permasalahan dalam praktik evaluasi pembelajaran. Penelitian ini bertujuan mengembangkan instrumen tes kemampuan berpikir kritis berbasis sintesis kerangka Watson–Glaser dan Ennis dalam konteks pembelajaran IPA. Metode penelitian menggunakan pendekatan *Research and Development* dengan model ADDIE yang meliputi tahap analisis, desain, pengembangan, implementasi, dan evaluasi. Instrumen dikembangkan melalui penyusunan indikator sintesis Watson–Glaser dan Ennis, kemudian dioperasionalkan ke dalam butir soal pilihan ganda berbasis penalaran pada materi Gelombang dan Optik. Validasi dilakukan oleh ahli asesmen, pendidikan IPA, dan berpikir kritis untuk menilai aspek isi, konstruk, dan bahasa. Uji coba instrumen dilakukan pada 38 mahasiswa Pendidikan IPA Universitas Negeri Medan. Hasil penelitian menunjukkan bahwa instrumen berada pada kategori valid berdasarkan penilaian ahli dan memiliki validitas konstruk yang memadai. Uji reliabilitas menggunakan koefisien Cronbach's Alpha menghasilkan nilai 0,876 yang termasuk kategori tinggi. Temuan ini menunjukkan bahwa instrumen yang dikembangkan mampu mengukur kemampuan berpikir kritis secara konsisten dan sesuai dengan konstruk sintesis Watson–Glaser dan Ennis. Instrumen ini berpotensi digunakan sebagai alternatif asesmen pembelajaran IPA yang lebih komprehensif dan berorientasi pada penalaran kritis.

Kata Kunci : instrumen tes, berpikir kritis, Ennis, Watson-Glaser

Abstract

The limitations of science assessment instruments that are able to measure critical thinking skills comprehensively, particularly in integrating cognitive dimensions and operational indicators into one complete construct, is still a problem in learning evaluation practices. This study aimed to develop a critical thinking test instrument based on the synthesis of the Watson–Glaser and Ennis frameworks in the context of science education. The research employed a Research and Development approach using the ADDIE model, consisting of analysis, design, development, implementation, and evaluation stages. The instrument was developed by synthesizing Watson–Glaser critical thinking components and Ennis' operational indicators, which were then translated into multiple-choice reasoning-based items on wave and optics topics. Expert validation was conducted to evaluate content, construct, and language aspects. The instrument was tested on 38 undergraduate students of science education in Universitas Negeri Medan. The results indicated that the instrument was valid based on expert judgment and demonstrated adequate construct validity. Reliability analysis using Cronbach's Alpha yielded a coefficient of 0.876, indicating high internal consistency. These findings suggest that the developed instrument reliably measures critical thinking skills in accordance with the synthesized Watson–Glaser and Ennis framework. The instrument can serve as an alternative assessment tool for science learning that emphasizes critical reasoning.

Key Words : test instrument, critical thinking, Ennis, Watson-Glaser

PENDAHULUAN

Kemampuan berpikir kritis sudah menjadi fondasi utama dalam pembelajaran IPA modern. Dalam pembelajaran abad ke-21,

siswa dituntut mampu menilai bukti, mengevaluasi argumen, dan membangun penjelasan ilmiah yang berbasis data [1]. IPA tidak lagi sekadar transfer konsep,

tetapi latihan bernalar. Tanpa keterampilan ini, proses belajar akan berakhir pada hafalan, bukan pemahaman konseptual yang bermakna.

Masalahnya, asesmen IPA di sekolah masih belum mampu menangkap proses berpikir tingkat tinggi. Banyak guru masih mengandalkan soal pilihan ganda yang berfokus pada pengetahuan faktual saja, bukan analisis [2]. Padahal, penelitian terbaru menunjukkan bahwa asesmen yang hanya menilai ingatan konseptual tidak mencerminkan kemampuan siswa dalam merespons dan menyelesaikan masalah fenomena ilmiah yang kompleks [3].

Hasil penelitian memperlihatkan tren pergeseran asesmen IPA ke arah penilaian keterampilan berpikir kritis melalui pendekatan analisis data, argumentasi ilmiah, dan *problem-based learning* [4]. Meski begitu, instrumen yang benar-benar mampu mengukur aspek penalaran kritis secara komprehensif masih minim. Instrumen yang ada cenderung terpisah-pisah, mengikuti satu kerangka saja, sehingga tidak menangkap keragaman cara siswa bernalar.

Dua kerangka teori yang paling sering digunakan dalam asesmen berpikir kritis adalah Watson–Glaser dan Ennis. Watson–Glaser menekankan penarikan kesimpulan, asumsi, deduksi, interpretasi informasi, dan analisis argumen [5]. Sementara itu, Ennis menawarkan indikator operasional yaitu klarifikasi dasar, membangun keterampilan dasar, menyimpulkan, klarifikasi lanjut, dan mengatur strategi dan taktik [6]. Namun, karena keduanya digunakan secara terpisah, hasil pengukuran yang diperoleh menjadi kurang proporsional.

Instrumen berbasis Watson–Glaser saja sering terlalu psikometrik dan kurang kontekstual untuk IPA, sedangkan indikator Ennis yang sangat operasional terkadang

terlalu umum sehingga tidak mampu mengevaluasi kedalaman penalaran ilmiah siswa secara akurat [1]. Kekurangan inilah yang membuat sintesis dua kerangka menjadi solusi strategis: menggabungkan ketajaman logis Watson–Glaser dengan kejelasan operasional Ennis.

Pembelajaran IPA pada dasarnya menyediakan konteks yang sangat cocok untuk mengukur berpikir kritis. Setiap topik IPA, baik biologi, kimia, maupun fisika, selalu berhubungan dengan analisis bukti, penalaran sebab-akibat, dan interpretasi data [7]. Karena itu, asesmen IPA menuntut instrumen yang dapat menangkap cara siswa berpikir ketika menghadapi fenomena ilmiah, bukan hanya jawaban akhir.

Namun, instrumen yang secara eksplisit dirancang untuk mengukur berpikir kritis dalam IPA masih sedikit ditemukan. Banyak instrumen hanya mengambil indikator umum, tanpa keterkaitan kuat dengan karakteristik ilmiah pembelajaran IPA [8]. Hal ini berdampak pada asesmen yang kurang valid secara konseptual dan kurang sensitif mendeteksi kemampuan bernalar kritis siswa.

Sejumlah penelitian sebelumnya telah mengembangkan instrumen berpikir kritis berbasis kerangka Watson-Glaser maupun Ennis secara terpisah. Misalnya, penelitian oleh Lestari dkk. [9] menggunakan kerangka Watson-Glaser untuk mengukur kemampuan berpikir kritis siswa sekolah dasar, sementara Verawati dkk. [10] mengembangkan instrumen berbasis indikator Ennis dalam konteks pembelajaran sains. Penelitian lain yang sejenis juga menunjukkan pola yang sama, yaitu pengembangan instrumen berpikir kritis yang masih berfokus pada satu kerangka tertentu dan belum mengintegrasikan konstruk secara komprehensif, seperti pengembangan instrumen berbasis HOTS dan representasi

ganda dalam pembelajaran sains [11], [12], [13], [14], [15]. Persamaan penelitian-penelitian tersebut terletak pada fokus pengukuran kemampuan berpikir kritis, namun perbedaannya terletak pada pendekatan konstruk yang digunakan, yang masih parsial dan belum terintegrasi. Oleh karena itu, kebaruan (novelty) penelitian ini terletak pada sintesis dua kerangka utama berpikir kritis, yaitu Watson-Glaser dan Ennis, sehingga menghasilkan konstruk instrumen yang lebih komprehensif, sistematis, dan kontekstual dalam pembelajaran IPA.

Penelitian ini mengembangkan instrumen tes berpikir kritis berbasis sintesis Watson-Glaser dan Ennis yang dirancang khusus untuk materi IPA. Pengembangan dilakukan melalui analisis kebutuhan, penyusunan indikator sintesis, konstruksi butir soal, validasi ahli, dan uji coba terbatas. Pendekatan sintesis ini memungkinkan instrumen menangkap struktur berpikir kritis secara lebih komprehensif dan realistis sesuai tuntutan pembelajaran IPA. Sebagai konteks penerapan, instrumen diuji menggunakan materi gelombang dan optik, topik IPA yang kaya fenomena dan sangat menuntut analisis serta evaluasi bukti [5]. Tujuan utama penelitian adalah mengembangkan instrumen berpikir kritis berbasis sintesis Watson-Glaser dan Ennis. Dengan pendekatan ini, instrumen diharapkan memberikan kontribusi nyata bagi asesmen pembelajaran IPA yang lebih valid, modern, dan responsif terhadap kebutuhan pembelajaran.

METODE

Penelitian ini dilaksanakan pada semester ganjil tahun akademik 2024/2025 di Program Studi Pendidikan IPA, Universitas Negeri Medan. Pendekatan *Research and Development* (R&D) dengan model ADDIE sebagai kerangka utama pengembangan instrumen digunakan sebagai jenis

penelitian. Model ini dipilih karena memberikan alur yang sistematis namun tetap fleksibel untuk menghasilkan produk instrumen asesmen yang dapat dipertanggungjawabkan secara teoretis maupun empiris. Fokus penelitian adalah mengembangkan instrumen tes kemampuan berpikir kritis berbasis sintesis indikator Watson-Glaser dan Ennis, kemudian mengujinya pada pembelajaran IPA. Proses pengembangan dilakukan secara bertahap mulai dari analisis kebutuhan hingga evaluasi kualitas instrumen berdasarkan data ahli dan data uji coba lapangan.

Tahap pertama adalah *Analysis*, yaitu fase untuk menganalisis kebutuhan instrumen yang dapat mengukur kemampuan berpikir kritis secara lebih komprehensif dalam pembelajaran IPA. Pada tahap ini dilakukan kajian teoretis terhadap komponen-komponen berpikir kritis menurut Watson-Glaser dan indikator berpikir kritis menurut Ennis, kemudian dianalisis kesesuaiannya dengan tuntutan asesmen pembelajaran IPA terkini. Analisis kebutuhan juga mencakup peninjauan ruang lingkup materi IPA yang relevan sebagai konteks soal, serta evaluasi instrumen yang sudah ada untuk memastikan urgensi pengembangan instrumen baru berbasis sintesis dua teori tersebut.

Tahap berikutnya adalah *Design*, yaitu merancang struktur instrumen berdasarkan hasil analisis sebelumnya. Pada tahap ini dilakukan penyusunan indikator sintesis yang menggabungkan dimensi kognitif Watson-Glaser dengan indikator operasional Ennis. Indikator tersebut kemudian dijabarkan ke dalam kisi-kisi instrumen. Melalui desain ini, rancangan awal instrumen diperkuat sehingga landasan pengembangannya jelas dan konsisten dengan konstruksi teoretis yang dirumuskan.

Tahap *Development* dilakukan untuk menyusun butir soal berdasarkan kisi-kisi

yang telah dibuat. Pada tahap ini, setiap butir ditulis dalam format pilihan ganda. Draft instrumen kemudian dinilai oleh para ahli yang memiliki kompetensi dalam asesmen, konten, dan kajian berpikir kritis. Validasi ahli digunakan untuk menilai kesesuaian konstruk, ketepatan konten, dan kejelasan bahasa. Berdasarkan masukan para ahli, instrumen direvisi hingga diperoleh versi yang dianggap layak untuk diuji secara empiris.

Tahap *Implementation* dilakukan dengan menguji instrumen yang telah direvisi pada subjek penelitian, yaitu mahasiswa sarjana Pendidikan IPA di Universitas Negeri Medan. Mahasiswa dipilih sebagai subjek uji coba karena mereka memiliki kompetensi konseptual IPA yang lebih stabil dibandingkan siswa sekolah, sehingga respons mereka dapat memberikan gambaran yang lebih akurat tentang kualitas butir instrumen pada tahap awal pengembangan. Pelaksanaan uji coba dilakukan secara terkontrol menggunakan lembar soal dan lembar jawaban yang telah disiapkan.

Tahap terakhir adalah *Evaluation*, merupakan fase untuk menganalisis data hasil uji coba guna menentukan kelayakan instrumen. Setiap butir dianalisis validitas konstraknya dihitung dengan korelasi item-total dan reliabilitas internal instrumen dihitung menggunakan koefisien Cronbach's Alpha untuk memastikan konsistensi antarbutir dalam mengukur konstruk berpikir kritis. Data validasi ahli dan hasil analisis empiris digunakan secara bersamaan untuk menyempurnakan instrumen hingga diperoleh versi final yang konsisten dan layak dipakai dalam asesmen pembelajaran IPA atau penelitian lanjutan.

HASIL DAN PEMBAHASAN

Pengembangan instrumen tes berpikir kritis mengikuti alur ADDIE. Tahap analisis menghasilkan identifikasi jelas mengenai

kebutuhan instrumen evaluasi yang mampu menangkap kemampuan berpikir kritis dalam konteks pembelajaran IPA. Kajian teoretis menunjukkan bahwa perangkat evaluasi yang umum digunakan selama ini masih cenderung mengukur aspek pengetahuan faktual dan belum menyentuh dimensi penalaran argumentatif secara memadai. Berdasarkan analisis tersebut, disusunlah indikator sintesis berbasis Watson-Glaser dan Ennis, lalu dijabarkan ke dalam kisi-kisi instrumen yang memuat nomor soal, aspek Watson-Glaser *Critical Thinking Assessment* (WGCTA), indikator Ennis, dan indikator soal yang sudah memuat konten pembelajaran IPA berupa materi Gelombang dan Optik. Struktur ini disajikan dalam Tabel 1, yang menjadi fondasi penulisan seluruh butir soal.

Tabel 1. Kisi-kisi instrumen

No Soal	Aspek (WGCTA)	Indikator (Ennis)	Indikator Soal
1	Interferensi	Mencari fakta relevan	Mengidentifikasi data relevan dari fenomena interferensi
2		Menguji kesimpulan awal	Menarik inferensi logis dari data gelombang
3		Menilai kekuatan kesimpulan	Membedakan kesimpulan kuat-lemah pada pembiasan
4		Mengambil kemungkinan paling rasional	Menentukan kemungkinan paling logis dari fenomena resonansi
5		Menilai kecukupan bukti	Menilai apakah data sudah cukup untuk mendukung inferensi gelombang
6	Mengenal asumsi	Mengidentifikasi asumsi tersembunyi	Menemukan asumsi implisit pada argumen pembiasan
7		Menguji asumsi	Menilai apakah klaim lensa bergantung pada

No Soal	Aspek (WGCTA)	Indikator (Ennis)	Indikator Soal	No Soal	Aspek (WGCTA)	Indikator (Ennis)	Indikator Soal
8		Menganalisis koherensi asumsi	asumsi tersembunyi Menguji konsistensi asumsi dengan teori pemantulan	19		Menginterpretasi situasi kompleks	Menafsirkan fenomena polarisasi menjadi makna ilmiah tepat
9		Membedakan asumsi vs opini	Mengidentifikasi pernyataan mana yang merupakan asumsi pada gelombang mekanik	20		Konsistensi interpretasi dengan bukti	Menilai konsistensi interpretasi dengan data pembiasan total
10		Menguji asumsi dapat diuji	Menentukan apakah asumsi tentang medium gelombang terfalsifikasi	21	Evaluasi argumen	Menilai kekuatan bukti	Menilai dukungan bukti pada argumen interferensi
11	Deduksi	Menilai validitas premis-kesimpulan	Menentukan apakah kesimpulan sesuai premis pada hukum Snellius	22		Mendeteksi bias	Menemukan bias dalam klaim tentang alat optik
12		Mengkritisi kecukupan premis	Mengidentifikasi premis yang belum cukup pada interferensi cahaya	23		Mengkritisi struktur argumen	Menilai kelemahan argumen pada kasus fiber optik
13		Menganalisis struktur argumen	Menguji struktur logis pada argumen difraksi	24		Menilai relevansi bukti	Menguji relevansi bukti dalam argumen gelombang longitudinal
14		Menilai koherensi logis	Menguji konsistensi klaim eksperimen gelombang stasioner	25		Memilih argumen terkuat	Menentukan argumen paling kuat terkait sistem kamera
15		Menolak kesimpulan yang tidak valid	Mengidentifikasi kesimpulan salah pada optik geometri	<p>Tahap desain mengembangkan indikator tersebut menjadi rancangan teknis instrumen. Penyusunan butir soal dalam bentuk pilihan ganda diarahkan agar mampu merepresentasikan indikator sintesis secara konsisten. Rancangan ini juga memastikan bahwa setiap butir memiliki stimulus dan alasan jawaban yang selaras dengan karakter penalaran kritis yang diacu dalam dua teori tersebut. Desain menjadi acuan kerja utama penulis butir sebelum masuk ke tahap pengembangan. Pada tahap pengembangan, butir yang telah disusun kemudian divalidasi oleh para ahli yang memiliki keahlian pada asesmen, pendidikan IPA, dan kajian berpikir kritis. Validitas isi dan konstruk dikaji melalui penilaian kesesuaian indikator, ketepatan konten, kejelasan bahasa, serta kualitas opsi</p>			
16	Interpretasi	Menafsirkan representasi kuantitatif	Menafsirkan grafik intensitas interferensi				
17		Menilai interpretasi sah/tidak sah	Membedakan interpretasi valid dari pola difraksi celah tunggal				
18		Menghubungkan multi-representasi	Menghubungkan model matematis, grafis, verbal pada lensa cembung				

jawaban. Hasil validasi ahli secara umum menunjukkan bahwa instrumen berada dalam kategori layak, dengan revisi minor terutama pada efisiensi redaksi dan penyelarasan stimulus. Ringkasan penilaian ahli ditampilkan dalam Tabel 2.

Tabel 2. Ringkasan validasi ahli

Aspek	Rata-rata	Kategori
Isi	85,33	Valid
Konstruk	82,92	Valid
Bahasa	86,17	Sangat Valid
Rata-rata	84,81	Valid

Instrumen yang telah direvisi kemudian diimplementasikan pada 38 mahasiswa Pendidikan IPA sebagai subjek uji coba. Data hasil uji coba dianalisis untuk memeriksa kualitas konstruk instrumen terhadap pengukuran kemampuan berpikir kritis berbasis sintesis Watson-Glaser dan Ennis. Analisis validitas konstruk dilakukan menggunakan korelasi item-total, dan hasilnya menunjukkan bahwa sebagian besar butir soal berada pada kategori valid. Dua butir soal dengan nilai korelasi mendekati batas minimal tetap direkomendasikan untuk revisi minor, tetapi tidak mengganggu struktur keseluruhan instrumen. Rekapitulasi nilai validitas konstruk disajikan dalam Tabel 3.

Tabel 3. Ringkasan validasi konstruk

<i>Item-Total Statistics</i>		
	<i>Corrected Item-Total Correlation</i>	<i>Cronbach's Alpha if Item Deleted</i>
soal1	.387	.873
soal2	.298	.875
soal3	.545	.868
soal4	.350	.874
soal5	.525	.869
soal6	.499	.869
soal7	.550	.868
soal8	.430	.871
soal9	.549	.868
soal10	.329	.874
soal11	.330	.874
soal12	.303	.875
soal13	.364	.873
soal14	.320	.874
soal15	.500	.869
soal16	.465	.870

<i>Item-Total Statistics</i>		
	<i>Corrected Item-Total Correlation</i>	<i>Cronbach's Alpha if Item Deleted</i>
soal17	.461	.871
soal18	.436	.871
soal19	.557	.868
soal20	.344	.874
soal21	.437	.871
soal22	.492	.870
soal23	.528	.869
soal24	.464	.870
soal25	.563	.868

Selanjutnya, reliabilitas instrumen dihitung menggunakan koefisien *Cronbach's Alpha* untuk memastikan konsistensi internal antarbutir. Nilai reliabilitas yang diperoleh berada pada kategori tinggi, yang menguatkan bahwa instrumen memiliki kestabilan pengukuran yang baik dan mampu merepresentasikan konstruk berpikir kritis secara konsisten. Hasil reliabilitas dapat dilihat dalam Tabel 4.

Tabel 4. Hasil Reliabilitas

<i>Reliability Statistics</i>		
<i>Cronbach's Alpha</i>	<i>Cronbach's Alpha Based on Standardized Items</i>	<i>N of Items</i>
.876	.878	25

Tahap evaluasi dilakukan dengan meninjau kembali kesesuaian antara tujuan pengembangan instrumen, hasil validasi ahli, dan temuan uji empiris. Secara keseluruhan, instrumen menunjukkan konsistensi antara indikator sintesis WG-Ennis dan bukti empiris yang mendukung. Hasil validitas dan reliabilitas menunjukkan bahwa instrumen ini bekerja dengan konsisten dan mengukur apa yang seharusnya diukur. Dengan kata lain, instrumen ini sudah cukup kuat untuk digunakan dalam asesmen pembelajaran IPA maupun penelitian lanjutan yang membutuhkan pengukuran kemampuan berpikir kritis secara lebih terarah.

Pengembangan instrumen berpikir kritis dalam penelitian ini diformulasikan melalui sintesis antara kerangka Watson-Glaser

Critical Thinking Appraisal (WGCTA) dan indikator dari kerangka *Ennis Critical Thinking Framework*, yakni sebuah pendekatan yang jarang dieksplorasi dalam konteks pembelajaran IPA di Indonesia. WGCTA menawarkan struktur kognitif inti: inferensi, deduksi, interpretasi, pengenalan asumsi, dan evaluasi argumentasi, yang telah sejak lama diakui sebagai komponen utama berpikir kritis [9]. Namun, WGCTA sendiri pada praktik luas cenderung bersifat umum dan kurang beradaptasi terhadap karakteristik khas penalaran ilmiah pada materi sains, seperti analisis data, interpretasi fenomena, atau evaluasi bukti konkret.

Sebaliknya, kerangka Ennis memberikan indikator yang lebih operasional dan kontekstual terhadap penalaran kritis dalam setting pendidikan dan ilmiah, termasuk kemampuan mengenali asumsi (eksplisit maupun implisit), menilai kredibilitas bukti, membuat inferensi logis berdasarkan data, dan menilai relevansi serta kekuatan argumen dalam konteks nyata. Pendekatan ini telah banyak dipakai dalam riset instrumen berpikir kritis IPA di berbagai jenjang pendidikan [16], [17].

Dengan mensintesis kedua kerangka, yaitu mengkombinasikan struktur kognitif WGCTA dengan indikator operasional Ennis, penelitian ini merumuskan konstruk berpikir kritis yang kuat, relevan dengan konten IPA, seperti materi gelombang & optik, dan mampu menjembatani antara teori penalaran umum dan karakteristik penalaran ilmiah. Hal ini menjadi inovasi dalam ranah pengembangan instrumen IPA, karena sebagian besar penelitian sebelumnya hanya mengadopsi satu kerangka saja. Misalnya, penelitian terdahulu menggunakan WGCTA saja untuk menilai berpikir kritis siswa [9], [15], [18], dan ada pula yang memakai kerangka Ennis, tetapi tanpa struktur konstruk yang jelas dan sistematis [10], [12], [14].

Hasil validasi ahli mendukung kekokohan konstruk sintesis ini. Aspek isi, konstruk, dan bahasa semuanya memperoleh skor “valid” dengan rerata tinggi, menunjukkan bahwa ahli sepakat indikator dan butir soal sudah mewakili konsep berpikir kritis dan sesuai konteks IPA. Selain itu, uji empiris pada mahasiswa Pendidikan IPA menunjukkan bahwa korelasi item-total pada mayoritas butir memenuhi kriteria validitas konstruk empiris yang membuktikan bahwa indikator sintesis bukan hanya valid di atas kertas, tetapi relevan dan sensitif terhadap variasi kemampuan berpikir kritis nyata.

Nilai reliabilitas yang tinggi memperkuat bahwa instrumen bekerja konsisten sebagai satu kesatuan konstruk. Konsistensi ini penting karena sintesis antara dua teori tidak menimbulkan fragmentasi konstruk, artinya indikator berbeda tidak saling bertabrakan, melainkan saling mendukung dalam mengukur kemampuan berpikir kritis secara holistik. Ini mengindikasikan bahwa struktur sintesis WGCTA-Ennis berhasil diterjemahkan ke dalam butir soal yang kohesif dan stabil secara psikometrik.

Secara praktis, sintesis ini memberi keuntungan besar bagi pengembangan asesmen IPA. Guru atau peneliti tidak perlu memilih antara tes penalaran abstrak ala WGCTA atau tes kontekstual ala Ennis, tetapi mendapat alat ukur hybrid yang mampu menangkap keduanya, yaitu penalaran logis dan proses berpikir ilmiah. Ini sesuai dengan tuntutan abad ke-21 bahwa siswa tidak hanya harus bisa memahami konsep IPA, tetapi juga mampu berpikir kritis ketika dihadapkan pada fenomena, data, atau eksperimen.

Namun, sintesis ini juga membawa tantangan. Karena indikator meliputi dua kerangka, beban kognitif butir soal cenderung tinggi, seperti memerlukan

stimulus yang kompleks, redaksi yang presisi, dan butir yang dirancang dengan sangat hati-hati agar tidak membingungkan siswa. Oleh karena itu validasi ahli dan uji empiris menjadi krusial untuk memastikan bahwa butir soal tetap dapat diakses dan dapat dipahami dengan benar.

Selain itu, hasil penelitian ini hanya diuji pada mahasiswa Pendidikan IPA. Untuk memperkuat generalisasi instrumen, perlu replikasi pada siswa sekolah menengah dan pada topik IPA lain di luar topik gelombang dan optik. Ini penting agar kekuatan sintesis WGCTA-Ennis tetap berlaku dalam konteks yang lebih luas. Studi pengembangan instrumen kritis IPA sebelumnya juga merekomendasikan validasi lebih lanjut pada populasi berbeda.

Dengan demikian, sintesis WGCTA-Ennis serta hasil penelitian ini memberi kontribusi signifikan pada literatur pengembangan asesmen berpikir kritis di IPA. Secara konseptual memberi landasan konstruk yang lebih komprehensif, secara empiris menghasilkan instrumen yang valid dan reliabel, dan secara praktis memberi alat ukur yang relevan dan responsif dengan kebutuhan pembelajaran sains yang modern.

SIMPULAN

Penelitian ini berhasil mengembangkan instrumen tes berpikir kritis berbasis sintesis Watson-Glaser dan Ennis sebagai upaya menghadirkan asesmen pembelajaran IPA yang lebih valid dan relevan dengan tuntutan pembelajaran modern. Melalui tahapan ADDIE, instrumen dirancang dengan mengintegrasikan struktur kognitif Watson-Glaser dan indikator operasional Ennis ke dalam konteks pembelajaran IPA. Hasil validasi ahli serta uji empiris menunjukkan bahwa instrumen memiliki kualitas konstruk yang baik dan reliabilitas yang tinggi, sehingga mampu mengukur

kemampuan berpikir kritis secara konsisten dan sesuai dengan tujuan pengembangan. Implikasi penelitian ini terletak pada penyediaan alternatif instrumen asesmen yang tidak hanya menilai penguasaan konsep IPA, tetapi juga menekankan penalaran kritis berbasis bukti dan argumentasi ilmiah. Instrumen ini dapat dimanfaatkan dalam asesmen pembelajaran maupun penelitian pendidikan IPA, sekaligus menjadi rujukan bagi pengembangan instrumen serupa pada materi dan jenjang pendidikan yang berbeda guna memperkuat praktik evaluasi pembelajaran yang lebih modern dan bermakna.

DAFTAR PUSTAKA

- [1] S. Mahany *dkk.*, “Peran Problem-Based Learning dalam Pengembangan Berpikir Kritis pada Pembelajaran Fisika: Analisis Bibliometrik,” *Epistemic: Scientific Thinking and Literacy*, vol. 1, no. 1, hlm. 1–11, Agu 2025, doi: 10.58706/epistemic.v1n1.p1-11.
- [2] Sulbiana, A. Haris, dan A. Yani, “Kemampuan Berpikir Kritis Dalam Pembelajaran Fisika Peserta Didik Kelas XI MIPA SMA Negeri 3 Makassar,” *Jurnal Sains dan Pendidikan Fisika*, vol. 20, no. 2, hlm. 174–189, Agu 2024, doi: 10.35580/jspf.v20i2.3690.
- [3] R. Setiawan, N. Lestari, dan N. Apsari, “Berpikir Kritis dalam Fisika Menggunakan Aplikasi Mobile,” *Jurnal Pembelajaran IPA dan Aplikasinya (QUANTUM)*, vol. 1, no. 2, hlm. 33–43, 2021.
- [4] F. Naswar, M. H. Rahman, dan N. A. Rahman, “Pengaruh Kemampuan Berpikir Kritis Terhadap Kemampuan Literasi Sains Siswa Kelas VIII SMP Negeri 7 Kota Ternate Pada Konsep Getaran Dan Gelombang Dengan Menggunakan Model Pembelajaran Discovery

- Learning,” *SAINTIFIK@: Jurnal Pendidikan MIPA*, vol. 8, no. 1, Mar 2023, doi: 10.33387/saintifik.v8i1.6218.
- [5] M. Makhrus, S. Ayub, N. Nyoman Putu Sri Verawati, A. Busyairi, dan P. Studi Pendidikan Fisika, “Kemampuan Berpikir Kritis Mahasiswa pada Materi Optik Saat Proses Pembelajaran dengan CCM-CCA Berbantuan Eksperimen Virtual,” *Kappa Journal*, vol. 4, no. 2, hlm. 2549–2950, Des 2020, [Daring]. Tersedia pada: <http://e-journal.hamzanwadi.ac.id/index.php/kpj/index>
- [6] H. Imtihani, F. Bakri, dan S. Sunaryo, “Pengembangan Video Pembelajaran Untuk Melatih Keterampilan Berpikir Kritis Siswa Sma Pada Materi Gelombang Bunyi Dan Cahaya,” dalam *Prosiding Seminar Nasional Fisika (E-Journal)*, Prodi Pendidikan Fisika dan Fisika UNJ, 2023. doi: 10.21009/03.1102.pf34.
- [7] E. Susilowati, D. Rusdiana, dan I. Kaniawati, “Efektivitas Perkuliahan Gelombang Dan Optika Berbasis Scaffolding Terhadap Peningkatan Kebiasaan Berpikir Kritis Mahasiswa,” *Jurnal Wahana Pendidikan Fisika*, vol. 2, no. 2, hlm. 68–70, 2020, [Daring]. Tersedia pada: <http://www.ccsnh.edu/documents/CSNH>
- [8] A. D. Pramudita dkk., “Analisis Pengaruh Kemampuan Berpikir Kritis Terhadap Pemahaman Konsep Gelombang Elektromagnetik Mahasiswa Pendidikan Fisika,” *Jurnal Ilmiah Wahana Pendidikan*, vol. 9, no. 14, hlm. 105–113, 2023, doi: 10.5281/zenodo.8170616.
- [9] F. P. Lestari, F. Ahmadi, dan R. Rochmad, “The Critical Thinking Ability in Watson-Glaser Framework in Fourth Grade Students,” *Educational Management*, vol. 9, no. 2, hlm. 233–240, 2020, [Daring]. Tersedia pada: <http://journal.unnes.ac.id/sju/index.php/eduman>
- [10] N. N. S. P. Verawati, S. Prayogi, M. Y. Yusup, dan H. Taha, “Development of the Test Instrument for Measuring Students’ Critical Thinking Abilities on Fluid Material,” *Prisma Sains : Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram*, vol. 8, no. 1, hlm. 46, Jul 2020, doi: 10.33394/j-ps.v8i1.2487.
- [11] T. Fairuz, “Assessment Instrument of Critical Thinking Skills for Student on Light Interaction with Organisms,” *Jurnal Studi Guru dan Pembelajaran*, vol. 4, no. 1, hlm. 67, 2021, doi: 10.30605/jsgp.4.1.2021.500.
- [12] I. G. E. Saputra, I. N. Jampel, dan I. G. L. A. Parwata, “Pengembangan Instrumen Penilaian Kemampuan Berpikir Kritis Peserta Didik Kelas VIII SMP pada Materi Getaran dan Gelombang,” *Jurnal Pendidikan dan Pembelajaran Sains Indonesia (JPPSI)*, vol. 5, no. 2, hlm. 154–164, Okt 2022, doi: 10.23887/jppsi.v5i2.51772.
- [13] H. Haryanti dan P. Susongko, “Pengembangan Instrumen Penilaian Berpikir Kritis Menurut FACIONE Pada Pembelajaran IPA Di Sekolah Menengah Pertama Berbasis Model Rasch,” *PSEJ (Pancasakti Science Education Journal)*, vol. 9, no. 1, hlm. 78–87, Nov 2024, doi: 10.24905/psej.v9i2.207.
- [14] Gusti Ngurah Arya Surya Wangsa, Nyoman Dantes, dan I Wayan Suastra, “Pengembangan Instrumen Kemampuan Berpikir Kritis Dan Hasil Belajar IPA Kelas V SD Gugus IV Kecamatan Gerokgak,”

- PENDASI: Jurnal Pendidikan Dasar Indonesia*, vol. 5, no. 1, hlm. 139–150, Mar 2021, doi: 10.23887/jurnal_pendas.v5i1.267.
- [15] P. A. Gunawan, H. Ansori, dan I. Budiarti, “Analisis Berpikir Kritis Siswa Berkategori Tinggi Menggunakan Watson-Glaser Critical Thinking Appraisal Ditinjau Dari Jenis Kelamin,” *EDU-MAT: Jurnal Pendidikan Matematika*, vol. 11, no. 2, hlm. 276, Okt 2023, doi: 10.20527/edumat.v11i2.16133.
- [16] N. U. Sa'idah, E. Handoyo, Supriyadi, dan W. Sumarni, “Development of a Critical Thinking Assessment Instrument for Earth Changes in Elementary School Science Lessons,” *Journal of Environmental and Science Education*, vol. 5, no. 1, hlm. 30–37, Apr 2025, doi: 10.15294/jese.v5i1.6297.
- [17] A. Asma, M. T. A, S. Potiua, dan W. Y. Anuli, “Critical Reasoning Assessment Instruments on Statistics Material for Phase D Students: Validity and Reliability,” *Jurnal Penelitian dan Pengembangan Pendidikan*, vol. 9, no. 2, hlm. 424–433, Jul 2025, doi: 10.23887/jppp.v9i2.92251.
- [18] M. Widiastuti dan N. Hamidi, “Analisis Keterampilan Berpikir Kritis Siswa Sekolah Menengah Kejuruan Menggunakan Tes W-GCTA,” *Jurnal Pendidikan Akuntansi Indonesia*, vol. 23, no. 2, hlm. 22–36, 2025.