



## Pengembangan Instrumen Evaluasi untuk Mengukur Keterampilan Berpikir Tingkat Tinggi pada Materi Kesetimbangan Kimia

Nur' Athiyah Muyassar Siregar<sup>\*1</sup>, Zainuddin Muchtar<sup>2</sup>, Ani Sutiani<sup>3</sup>, Ratu Evina Dibyantini<sup>4</sup>,  
Marudut Sinaga<sup>5</sup>

<sup>1,2,3,4,5</sup>Program Studi Pendidikan Kimia, Universitas Negeri Medan, Indonesia  
E-mail: [nurathiyah1004@gmail.com](mailto:nurathiyah1004@gmail.com)

Article Info	Abstract
<b>Article History</b> Received: 2023-05-22 Revised: 2023-06-15 Published: 2023-07-01	This study aims to develop an evaluation instrument that can measure high-order thinking skills in chemical equilibrium material, determine the feasibility of evaluation instruments and determine the high-order thinking skills of class XI C students of SMAN 2 Percut Sei Tuan in solving HOTS questions. This study uses a 4D development model which includes stages define, design, develop, and disseminate. This research was conducted at SMAN 2 Percut Sei Tuan with research subjects namely 36 students of class XI C. The results of this study showed that the value of validation by experts was 87% with a very valid category. Based on the analysis of the validity of the test items, there are 20 valid questions, the reliability coefficient is 0.885, the difficulty level is 0.31-1.00, the differential power is 0.21-0.70. The results of students' high-level thinking skills in the chemical equilibrium material are included in the high category with an overall average student score of 72. The distribution of HOTS questions for the cognitive level C4 is 80%, C5 is 10%, and C6 is 10%. Based on these criteria, 20 items have been declared good and fit to be used as an evaluation instrument to measure students' high-level thinking skills in chemical equilibrium material.
<b>Keywords:</b> HOTS; Evaluation Instruments; Chemical Equilibrium.	

Artikel Info	Abstrak
<b>Sejarah Artikel</b> Diterima: 2023-05-22 Direvisi: 2023-06-15 Dipublikasi: 2023-07-01	Penelitian ini bertujuan untuk mengembangkan instrumen evaluasi yang dapat mengukur keterampilan berpikir tingkat tinggi pada materi kesetimbangan kimia, mengetahui kelayakan instrumen evaluasi dan mengetahui keterampilan berpikir tingkat tinggi siswa kelas XI C SMAN 2 Percut Sei Tuan dalam menyelesaikan soal HOTS. Penelitian ini menggunakan model pengembangan 4D yang meliputi tahap <i>define, design, develop, dan disseminate</i> . Penelitian ini dilakukan di SMAN 2 Percut Sei Tuan dengan subjek penelitian yaitu 36 siswa kelas XI C. Hasil penelitian ini menunjukkan nilai validasi oleh ahli sebesar 87% dengan kategori sangat valid. Berdasarkan analisis validitas butir soal hasil uji coba soal yang valid berjumlah 20 butir soal, koefisien reabilitas 0,885, tingkat kesukaran 0,31-1,00, daya beda 0,21-0,70. Hasil keterampilan berpikir tingkat tinggi siswa pada materi kesetimbangan kimia termasuk pada kategori tinggi dengan nilai rata-rata siswa secara keseluruhan sebesar 72. Adapun sebaran soal HOTS untuk jenjang kognitif C4 sebesar 80%, C5 sebesar 10%, dan C6 sebesar 10%. Berdasarkan kriteria tersebut, maka 20 butir soal telah dinyatakan baik dan layak digunakan sebagai instrumen evaluasi untuk mengukur keterampilan berpikir tingkat tinggi siswa pada materi kesetimbangan kimia.
<b>Kata kunci:</b> HOTS; Instrumen Evaluasi; Kesetimbangan Kimia.	

### I. PENDAHULUAN

Mutu sumber daya manusia (SDM) merupakan tanggung jawab pendidikan di masa society 5.0 saat ini, yakni guna menjamin dan memajukan eksistensinya dalam mengalami perkembangan ilmu pengetahuan serta teknologi yang semakin pesat. Pendidikan mempunyai kedudukan berarti sebagai bekal peserta didik dalam menyikapi bermacam tantangan kehidupan yang bakal dihadapinya. Manusia selaku pusat perkembangan di masa society 5.0 yang berfungsi sebagai bagian dari teknologi itu sendiri. Maksudnya manusia sebagai komponen utama dari sebuah teknologi sekaligus menggunakannya dalam kehidupan sehari-hari. Manusia tidak bisa ter-

pisahkan dari teknologi, begitupula teknologi tidak bisa dipisahkan dari manusia selaku pengguna utamanya dalam menempuh kehidupan sehari-hari (Amaliya & Fathurohman, 2022).

Perkembangan teknologi serta informasi pada abad 21 sudah memberikan dampak yang signifikan terhadap dunia pendidikan. Pola pendidikan abad 21 mengedepankan peserta didik berpikir kritis, sanggup mengintegrasikan seluruh ilmu dalam kehidupan nyata, menguasai teknologi serta informasi dan cakap dalam berbicara serta bekerjasama, guna menghadapi pendidikan abad 21 tiap siswa SMA paling tidak wajib mempunyai 4C yakni kemampuan berpikir kritis (*critical thinking*), kemampuan berpikir

kreatif (*creative thinking*), kemampuan komunikasi (*communication*) serta kemampuan kerja sama (*collaboration*) ataupun yang dikenal juga dengan sebutan kecakapan hidup abad 21 (Malikah & Wafroturrohmah, 2022).

Hasil riset *Trends in Mathematics and Science Study* (TIMSS) tahun 2015 menampilkan untuk literasi sains, peserta didik Indonesia menempati peringkat 45 dari 48 negara dengan pencapaian skor 397 (Siregar & Panggabean, 2020). Rendahnya kualitas pendidikan juga bisa dilihat dari laporan riset *Programme for International Student Assessment* (PISA) tahun 2018 yang diselenggarakan oleh *Organization for Economic Co-Operation and Development* (OECD) menampilkan Indonesia menduduki posisi 10 terbawah dari 79 negara yang berpartisipasi. Peserta didik Indonesia untuk kemampuan sains posisi di ranking 71 dengan skor rata-rata 396, sebaliknya untuk kemampuan matematika peserta didik Indonesia posisi di ranking 73 dengan nilai rata-rata 379, terlebih lagi untuk kemampuan membaca terletak di posisi 74 dengan skor rata-rata 371 (Kemendikbud, 2021).

Terdapatnya kesenjangan antara harapan dan realitas dalam melaksanakan menyesuaikan diri terhadap transformasi dan perkembangan dunia di era globalisasi merupakan kasus yang wajib diatasi, khususnya di dunia pendidikan. Oleh sebab itu butuh dikembangkan pendidikan yang bisa meningkatkan kemampuan dan keahlian siswa secara maksimal, sehingga sanggup menyesuaikan diri dengan kondisi dan transformasi yang berlangsung serta sanggup bekerja sama secara kolaboratif dalam memecahkan masalah kehidupan. Pendidikan mesti memusatkan pada pembentukan siswa guna bisa membangun kapasitas pengetahuannya dengan meningkatkan keterampilan berpikir tingkat tinggi, semacam berpikir kritis, membuat keputusan, serta memecahkan permasalahan (Sutiani *et al.*, 2022). Kurikulum 2013 yang sudah digunakan disaat ini sangat menekankan para pendidik untuk mempunyai keahlian dalam menyusun instrumen evaluasi HOTS. Instrumen evaluasi HOTS sebagai alat ukur yang digunakan untuk mengukur keterampilan berpikir tingkat tinggi peserta didik dalam meraih tujuan pembelajaran yang dirumuskan. HOTS sendiri ialah keterampilan berpikir tingkat tinggi yang mengacu pada ranah C4 (analisis), C5 (evaluasi), serta C6 (kreasi) (Dewi *et al.*, 2021).

Ilmu kimia merupakan bagian dari rumpun sains serta salah satu cabang ilmu pengetahuan alam yang mencakup konsep, aturan, hukum,

prinsip, dan teori (Panggabean, 2022). Ada 3 hal yang tidak bisa dipisahkan dalam belajar ilmu kimia, yakni kimia sebagai pengetahuan, sikap, dan proses (kinerja ilmiah). Konsep-konsep kimia pada dasarnya menggunakan dan mengaitkan 3 aspek kajian yang tidak bisa dipisahkan, yaitu aspek kajian makroskopis (sifat yang bisa dicermati), submikroskopis (partikel penyusun zat), dan simbolik (rumus kimia dan simbol kimia yang ada) (Gultom & Muchtar, 2022). Penerapan pembelajaran kimia di banyak sekolah cenderung kurang mencermati perkembangan keterampilan berpikir kritis siswa. Pembelajaran kimia di sekolah lebih bersifat teoritis, hafalan ataupun pengetahuan faktual. Perihal ini membuat pendidikan kimia tidak sesuai dengan tujuan pembelajaran nasional (Siahaan *et al.*, 2021).

Berdasarkan dari kajian literatur dikenal penggunaan instrumen evaluasi untuk mengukur keterampilan berpikir tingkat tinggi masih sangat jarang digunakan dalam proses evaluasi pembelajaran. Perihal ini diperoleh dari penelitian (Sholeh & Olenia, 2022), sebagian besar guru-guru kimia menyusun soal evaluasi mengambil dari buku paket dan juga jarang mengembangkan soal sendiri sebab memerlukan proses yang panjang dari penataan soal hingga tahap validasi soal. Dalam penelitian (Jayanti, 2020) pada proses penilaian masih jarang mengaplikasikan penilaian yang dapat mengukur keterampilan berpikir tingkat tinggi peserta didik karena masih kurang tersedianya soal-soal yang dirancang khusus untuk meningkatkan keterampilan berpikir tingkat tinggi yang menyebabkan peserta didik menjadi kurang terlatih dalam menyelesaikan soal-soal HOTS.

Penelitian yang berkaitan dengan pengembangan instrumen evaluasi untuk mengukur keterampilan berpikir tingkat tinggi lebih dahulu sudah dilakukan oleh (Risidiana *et al.*, 2022) yang menunjukkan hasil jika pengembangan soal HOTS pada materi Asam-Basa untuk Kelas XI SMA/ MA Sederajat dinyatakan sangat valid dan nilai reliabilitas soal secara keseluruhan pada kriteria sangat tinggi serta daya pembeda seluruh butir soal dapat diterima. Berikutnya menurut penelitian (Napitupulu & Yusuf, 2022) pada materi laju reaksi kimia, instrumen tes yang dikembangkan berupa tes pilihan berganda dapat mengukur ranah kognitif menganalisis (C4), mengevaluasi (C5), dan mencipta (C6). Hal ini didukung oleh penelitian (Dewi *et al.*, 2021) yang menyatakan bahwa pengembangan butir soal higher order thinking skill pada materi

kesetimbangan ion dalam larutan, distribusi butir soal HOTS yang dikembangkan bersumber pada jenjang kognitif taksonomi Bloom didominasi oleh soal dengan tingkat jenjang kognitif C4. Sebagaimana penelitian yang dilakukan oleh (Lubis *et al.*, 2022) melaporkan bahwa instrumen evaluasi HOTS yang dikembangkan untuk SMA/MA kelas XI pada materi Hidrolisis Garam sebanyak 40 pertanyaan sudah memenuhi validitas isi, dimana 37 pertanyaan dikatakan valid dan sesuai dengan kurikulum 2013.

Keterampilan berpikir tingkat tinggi dapat dilatihkan pada materi kimia, salah satunya ialah kesetimbangan kimia. Alasan pemilihan materi kesetimbangan kimia karena materi tersebut dapat memberikan kesempatan kepada peserta didik untuk membangun pengetahuannya sendiri lewat kemampuan berpikir tingkat tinggi karena berhubungan erat dengan fenomena-fenomena yang terdapat di sekitar (Andriani & Yonata, 2018). Materi ini memerlukan keterampilan berpikir tingkat tinggi peserta didik karena terletak pada tingkat kognitif C4. Kesetimbangan kimia ialah materi yang kompleks sehingga diperlukan proses berpikir yang lebih dan tidak hanya semata-mata menghafal konsep-konsep (Effendi & Andromeda, 2022). Selain itu, ciri dari materi kesetimbangan kimia dapat dikembangkan soal dengan jenjang C5, sebab materi tersebut ada konsep teori dan hitungan yang cakupannya lebih luas serta pula berhubungan pada kehidupan sehari-hari, contohnya pada bidang industry (Andari *et al.*, 2021).

Berdasarkan latar belakang yang sudah diuraikan, maka tujuan dalam penelitian ini yaitu mengembangkan instrumen evaluasi yang dapat mengukur keterampilan berpikir tingkat tinggi pada materi kesetimbangan kimia, mengetahui kelayakan instrumen evaluasi serta mengetahui keterampilan berpikir tingkat tinggi siswa kelas XI C SMAN 2 Percut Sei Tuan dalam menyelesaikan soal HOTS.

## II. METODE PENELITIAN

Penelitian ini merupakan penelitian dan pengembangan (R&D). Model pengembangan yang digunakan yaitu model 4D. Model 4D terdiri dari empat tahapan yaitu *define* (pendefinisian), *design* (perancangan), *develop* (pengembangan), dan *disseminate* (penyebarluasan). Penelitian ini dilaksanakan di SMAN 2 Percut Sei Tuan, Kabupaten Deli Serdang, Provinsi Sumatera Utara. Subjek pada penelitian ini adalah siswa kelas XI C SMA Negeri 2 Percut Sei Tuan yang terdiri dari 36 siswa.

Teknik pengumpulan data yang digunakan dalam penelitian ini meliputi teknik tes dan angket. Instrumen penelitian yang digunakan terdiri dari instrumen tes dan instrumen non tes. Instrumen tes yang digunakan berupa soal HOTS pilihan ganda untuk mengukur keterampilan berpikir tingkat tinggi peserta didik, sedangkan instrumen non tes yang digunakan berupa lembar validasi. Teknik analisis data yang digunakan dalam penelitian ini adalah analisis kualitatif dan kuantitatif. Data kualitatif diwakili oleh saran dan masukan dari validator dan data kuantitatif diwakili oleh nilai tingkat kelayakan produk yang dinyatakan dalam angka dan perhitungan menggunakan skala *likert* berupa angka 1 sampai 5. Data kuantitatif juga menggunakan bantuan Microsoft Excel untuk menghitung validitas, reliabilitas, tingkat kesukaran, dan daya beda butir soal. Berikut penjelasan langkah-langkah yang dilakukan pada masing-masing tahapan:

Tahap pendefinisian (*define*) bertujuan untuk menetapkan dan mendefinisikan kebutuhan dan syarat-syarat yang sesuai untuk pengembangan instrumen evaluasi keterampilan berpikir tingkat tinggi dengan menganalisis tujuan pembelajaran dan materi berdasarkan silabus kimia kurikulum 2013. Dalam tahapan ini, peneliti melakukan beberapa kegiatan seperti analisis awal akhir, analisis peserta didik, analisis tugas, analisis konsep, dan perumusan tujuan pembelajaran.

Tahap perancangan (*design*) bertujuan untuk membuat rancangan awal instrumen evaluasi keterampilan berpikir tingkat tinggi pada materi kesetimbangan kimia yang akan dikembangkan berdasarkan hasil dari tahap pendefinisian. Langkah-langkah yang dilakukan dalam tahapan ini yaitu menetapkan bentuk instrumen, menyusun kisi-kisi soal HOTS, menyusun soal, membuat kunci jawaban, memilih jenjang kognitif, dan pedoman penskoran. Instrumen evaluasi yang dirancang pada tahap ini berupa soal pilihan ganda (*multiple choice*) yang berjumlah 30 soal yang terdiri jenjang kognitif C4, C5, dan C6. Instrumen evaluasi dirancang berdasarkan kisi-kisi soal dengan mengacu pada KI dan KD yang terdapat dalam kurikulum 2013 revisi 2017.

Tahap pengembangan (*develop*) bertujuan untuk menghasilkan produk instrumen evaluasi keterampilan berpikir tingkat tinggi. Instrumen evaluasi yang dihasilkan pada tahap pengembangan dilakukan pengujian produk yang meliputi validasi oleh validator ahli menggunakan angket validasi. Validasi ahli dilakukan oleh 2 dosen ahli

jurusan Kimia Universitas Negeri Medan dan 4 mahasiswa Program Profesi Guru (PPG) bidang studi kimia.

Tahap penyebaran (*disseminate*) bertujuan untuk dapat melakukan uji coba kepada siswa. Penyebarluasan dilakukan dengan memberikan instrumen evaluasi keterampilan berpikir tingkat tinggi pada sekolah tempat penelitian di SMA Negeri 2 Percut Sei Tuan. Dalam penelitian ini yang dilakukan adalah penyebaran terbatas karena dilakukan dalam lingkup instansi saja disebabkan adanya keterbatasan waktu.

### III. HASIL DAN PEMBAHASAN

#### A. Hasil Penelitian

Penelitian ini menghasilkan produk berupa instrumen evaluasi keterampilan berpikir tingkat tinggi pada materi kesetimbangan kimia dalam bentuk pilihan ganda. Berdasarkan penelitian yang telah dilakukan diperoleh analisis data untuk menentukan kelayakan instrumen berdasarkan aspek validitas, reliabilitas, tingkat kesukaran dan daya beda butir soal.

Validasi soal pada penelitian ini dilakukan dengan 2 tahapan, yaitu validasi oleh ahli dan validitas butir soal hasil uji coba. Berdasarkan hasil validasi ahli melalui dua orang dosen kimia Universitas Negeri Medan dan empat orang mahasiswa PPG bidang studi kimia didapatkan bahwa instrumen evaluasi yang telah dikembangkan dinyatakan valid dengan rata-rata persentase hasil validitas instrumen masing-masing adalah 87% dengan kategori sangat valid. Hal ini menunjukkan bahwa keenam validator setuju instrumen evaluasi yang telah dikembangkan dilanjutkan ke tahap penyebaran. Untuk hasil perhitungan validasi instrumen evaluasi dari para validator ahli dapat dilihat pada Tabel 1.

**Tabel 1.** Hasil validasi instrumen evaluasi oleh validator ahli

No.	Aspek Penilaian	Skor Penilaian (%)						Rata-Rata (%)
		V1	V2	V3	V4	V5	V6	
1	Materi	85	85	90	70	73	90	82
2	Konstruksi	88	88	90	63	70	83	80
3	Bahasa	100	80	100	80	90	100	92
4	Aturan Tambahan	100	80	100	80	100	100	93
Rata-Rata Keseluruhan Validasi (%)								87
Tafsiran Persentase								Sangat Tinggi
Kriteria Validasi Analisis Persentase								Layak

Keterangan: (V1) Dosen ahli 1; (V2) Dosen ahli 2; (V3) Mahasiswa PPG 1; (V4) Mahasiswa PPG 2; (V5) Mahasiswa PPG 3; (V6) Mahasiswa PPG 4

Sedangkan berdasarkan analisis validitas butir soal hasil uji coba diketahui bahwa soal yang valid berjumlah 20, sedangkan soal yang tidak valid berjumlah 10 soal. Hasil analisis validitas instrumen evaluasi dapat dilihat pada Tabel 2.

**Tabel 2.** Validitas instrumen evaluasi

Nomor Soal	Rentang Nilai	Kriteria	Persentase
1, 2, 4, 5, 8, 9, 10, 15, 16, 17, 18, 19, 20, 24, 25, 26, 27, 28, 29, 30	$r > 0,329$	Valid	67%
3, 6, 7, 11, 12, 13, 14, 21, 22, 23	$r < 0,329$	Tidak Valid	33%

Berdasarkan Tabel 2, diketahui bahwa dari 30 butir soal yang telah dilakukan uji validitas diperoleh 20 butir soal yang dinyatakan valid karena memiliki nilai  $r$  hitung  $>$   $r$  tabel. Berdasarkan hasil perhitungan nilai  $r$  tabel memiliki nilai 0,329. Jika kurang dari nilai  $r$  tabel maka butir soal tidak valid. Analisis data selanjutnya adalah menghitung reabilitas tes menggunakan rumus Kuder & Richardson (K-R.21). Diketahui reabilitas instrumen evaluasi dalam penelitian ini adalah 0,885 dan berada pada kategori reabilitas sangat tinggi.

Aspek lain yang menentukan kelayakan soal lainnya adalah tingkat kesukaran masing-masing butir soal. Butir soal dapat dikatakan baik apabila butir soal tersebut tidak terlalu mudah dan tidak terlalu sulit. Suatu butir soal dikatakan memenuhi syarat apabila harga tingkat kesukaran berkisar antara: 0,20-0,80. Hasil analisis tingkat kesukaran pada instrumen evaluasi dapat dilihat pada Tabel 3.

**Tabel 3.** Tingkat kesukaran instrumen evaluasi

Nomor Soal	Rentang Nilai	Kriteria	Persentase
1, 3, 4, 6, 8, 18, 19, 20	0,31-0,70	Sedang	40%
2, 5, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17	0,71-1,00	Mudah	60%

Berikutnya adalah daya beda pada instrumen evaluasi. Butir soal dapat dikatakan baik apabila butir soal memiliki indeks daya beda

(diskriminasi) antara 0,20-1,0. Sehingga diperoleh 14 soal berkategori daya beda cukup, dan 6 soal berkategori daya beda baik. Hasil analisis daya beda untuk instrumen evaluasi dapat dilihat pada Tabel 4.

**Tabel 4.** Daya beda instrumen evaluasi

Nomor Soal	Rentang Nilai	Kriteria	Persentase
2, 3, 5, 6, 7, 9, 10, 13, 14, 15, 16, 17, 18, 19	0,21-0,40	Cukup	70%
1, 4, 8, 11, 12, 20	0,41-0,70	Baik	30%

Berdasarkan Tabel 4, semua soal telah memenuhi kriteria daya beda yang baik. Butir soal yang memiliki kriteria sangat baik dan baik memiliki arti dapat digunakan untuk membedakan antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah (Dewi *et al.*, 2021).

**Tabel 5.** Hasil analisis keterampilan berpikir tingkat tinggi siswa

No	Nilai Siswa	Kriteria HOTS	Jumlah	Persentase
1	10 < Nilai ≤ 20	Sangat Kurang	3	8%
2	20 < Nilai ≤ 40	Kurang	3	8%
3	40 < Nilai ≤ 60	Cukup	6	17%
4	60 < Nilai ≤ 80	Baik	6	17%
5	80 < Nilai ≤ 100	Sangat Baik	18	50%

Berdasarkan Tabel 5, terdapat 18 siswa berkriteria sangat baik dengan rentang nilai 80-100, 6 siswa berkriteria baik dengan rentang nilai 60-80, 6 siswa berkriteria cukup dengan rentang nilai 40-60, 3 siswa berkriteria kurang dengan rentang nilai 20-40, dan 3 siswa berkriteria sangat kurang dengan rentang nilai 0-20 dalam menyelesaikan soal keterampilan berpikir tingkat tinggi pada materi kesetimbangan kimia.

**Tabel 6.** Persentase HOTS siswa berdasarkan ranah kognitif

Indikator	Jumlah	Persentase
Menganalisis (C4)	16	80%
Mengevaluasi (C5)	2	10%
Mengkreasi (C6)	2	10%

Berdasarkan Tabel 6, persentase terbesar siswa dalam menjawab instrumen evaluasi keterampilan berpikir tingkat tinggi pada

materi kesetimbangan kimia adalah pada level kognitif menganalisis (C4) yang berjumlah 16 soal dengan persentase sebesar 80%, level kognitif mengevaluasi (C5) yang berjumlah 2 butir soal dengan persentase sebesar 10%, dan level kognitif mengkreasi (C6) yang berjumlah 2 butir soal dengan persentase sebesar 10%.

## B. Pembahasan

Penelitian pengembangan merupakan penelitian yang menghasilkan suatu produk. Adapun penelitian pengembangan yang dimaksud dalam penelitian ini yaitu mengembangkan instrumen evaluasi keterampilan berpikir tingkat tinggi pada materi kesetimbangan kimia. Instrumen evaluasi yang dihasilkan berupa 20 butir soal dengan pembagian 16 soal menganalisis (C4), 2 soal mengevaluasi (C5) dan 2 soal mengkreasi (C6). Untuk menghasilkan instrumen evaluasi ini, maka peneliti menggunakan model 4D. Model ini terdiri dari 4 tahap yaitu *define* (pendefinisian), *design* (perancangan), *develop* (pengembangan), dan *disseminate* (penyebaran).

Pengembangan instrumen evaluasi memerlukan sebuah kriteria untuk menentukan kelayakan instrumen evaluasi yang telah dikembangkan apakah sudah layak atau tidak. Kriteria tersebut diperlukan sebagai acuan dalam menentukan sejauh mana proses pengembangan dilakukan. Berdasarkan penelitian yang telah dilakukan diperoleh analisis data untuk menentukan kelayakan instrumen evaluasi berdasarkan aspek validitas, reliabilitas, tingkat kesukaran dan daya beda.

Validitas instrumen evaluasi pada penelitian ini diukur melalui 2 tahap, yaitu validasi oleh ahli dan validitas butir soal hasil uji coba. Validasi tahap pertama dilakukan validasi oleh ahli untuk menilai kevaliditasan soal berdasarkan aspek materi, konstruksi, bahasa, dan aturan tambahan. Berdasarkan hasil analisis data diketahui bahwa rata-rata validasi oleh ahli masing-masing sebesar 87% dengan kategori sangat valid. Akan tetapi masih ada beberapa soal yang masih harus disempurnakan untuk dapat memperbaiki penulisan, kalimat pada soal yang masih bermakna ganda, pengkategorian jenjang kognitif, dan penyesuaian dengan karakteristik soal berpikir tingkat tinggi. Hasil tersebut hampir sama dengan penelitian Arum *et al.*, (2022), nilai persentase rata-rata validitas

sebesar 86,7% dan termasuk dalam kategori sangat valid. Dengan demikian instrumen soal HOTS dapat dianggap sangat layak untuk digunakan berdasarkan kriteria penilaian yang telah ditetapkan.

Tahap selanjutnya yaitu validasi butir soal hasil uji coba, untuk menentukan validitas butir soal dilakukan dengan menghitung koefisien validitas (koefisien korelasi) antara butir soal dengan skor total dengan menggunakan rumus *Product Moment*. Berdasarkan hasil analisis dari 30 butir soal, diketahui bahwa soal yang valid ( $r > 0,329$ ) berjumlah 20 soal, sedangkan butir soal yang tidak valid berjumlah 10 soal. Masing-masing pada kriteria valid 63% dan pada kriteria tidak valid 37%. Dengan demikian soal pada instrumen evaluasi ini dapat digunakan untuk mengukur keterampilan berpikir tingkat tinggi peserta didik pada materi kesetimbangan kimia.

Kemudian untuk menentukan koefisien reabilitas dalam penelitian ini, digunakan rumus Kuder & Richardson (K-R.21). Reabilitas soal dalam penelitian ini adalah 0,885 dan berada pada kategori sangat tinggi. Hasil tersebut hampir sama dengan penelitian (Aprilia & Panggabean, 2021), nilai reabilitas soal sebesar 0,897 dan berada pada kategori sangat tinggi. Hal ini berarti instrumen tes yang dikembangkan memiliki konsistensi tinggi dalam memberikan hasil penilaian. Instrumen yang reliabel berarti instrumen tersebut dapat digunakan untuk mengukur objek yang lain dan dapat dipercaya sebagai alat pengumpul data (Netri *et al.*, 2018).

Berikutnya berdasarkan hasil analisis dari 20 butir soal, diketahui bahwa keseluruhan soal berkisar pada tingkat kesukaran sedang dan mudah. Masing-masing pada kategori sedang dengan interval  $0,31 < P \leq 0,70$  adalah 40%, pada kategori mudah dengan interval  $0,71 < P \leq 1,00$  adalah 60%. Soal yang memiliki kriteria sedang dan sukar dibutuhkan lebih banyak daripada soal yang mudah dalam pengembangan soal HOTS. Soal yang mudah dijawab peserta didik tidak akan mengembangkan berpikir tinggi dalam menganalisis (Wantoro *et al.*, 2019). Menurut (Sriyanti, 2022), dalam upaya menentukan apakah suatu soal dikelompokkan dalam soal HOTS atau tidak, maka tidak cukup hanya dengan melihat dari soal tetapi perlu dilakukan triangulasi metode dalam bentuk kajian lebih lanjut dengan menganalisis bagaimana proses

dalam menyelesaikan persoalan tersebut. Dengan demikian, soal-soal keterampilan berpikir tingkat tinggi belum tentu soal-soal yang memiliki tingkat kesukaran yang tinggi.

Aspek yang mendukung penentuan kelayakan instrumen evaluasi selanjutnya adalah daya beda soal. Daya pembeda soal adalah kemampuan suatu soal untuk membedakan antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah (Jayanti, 2020). Berdasarkan hasil analisis data, diketahui bahwa 70% soal berkategori daya beda cukup dan 40% soal berkategori daya beda baik. 14 soal yang diujikan berada antara 0,21-0,40 pada kategori cukup dan 6 soal yang diujikan berada antara 0,41-0,70 pada kategori baik. Soal yang sebaiknya digunakan adalah soal yang memiliki daya pembeda sangat baik dan baik, atau sekurang-kurangnya memiliki daya pembeda cukup (Yustiana & Ramadhoni, 2017).

Berdasarkan hasil uji coba yang telah dilakukan menunjukkan bahwa butir soal tersebut memenuhi kriteria kualitas soal yang baik karena pada setiap butir soal yang dikembangkan memenuhi tiga dari empat kriteria kualitas soal yang baik. Instrumen evaluasi yang baik dan layak digunakan harus memiliki dua syarat penting yaitu valid dan reliabel. Suatu validitas adalah tingkatan yang menunjukkan kebenaran dari sebuah instrumen. Sedangkan reliabilitas merupakan suatu keakuratan ataupun ketepatan dari sebuah instrumen sehingga dapat digunakan berkali-kali.

Berdasarkan instrumen evaluasi keterampilan dalam berpikir tingkat tinggi yang dikembangkan ini, karakteristik soal HOTS yang dimiliki meliputi jenjang kognitif C4, C5, dan C6. Akan tetapi, dalam pembuatan butir soal terdapat beberapa soal yang memiliki kemampuan untuk berpikir tingkat tinggi tetapi termasuk dalam kategori MOTS yakni jenjang kognitif C3 (menerapkan atau mengaplikasikan). Butir soal yang telah dikembangkan selanjutnya diklasifikasikan berdasarkan jenjang kognitif sesuai dengan taksonomi Bloom revisi yang meliputi C3 (mengaplikasikan), C4 (menganalisis), C5 (mengevaluasi), dan C6 (menciptakan).

Berdasarkan dari hasil penelitian terkait distribusi butir soal HOTS pada saat uji coba didapatkan bahwa pada sub pokok bahasan konsep kesetimbangan dinamis memiliki sebaran persentase jenjang kognitif C3

sebesar 3,3%, dan jenjang kognitif C4 sebesar 13,3%. Kemudian pada sub pokok bahasan tetapan kesetimbangan kimia (Kc dan Kp) memiliki sebaran persentase jenjang kognitif C4 sebesar 36,7%, dan jenjang kognitif C5 sebesar 10%. Selanjutnya pada sub pokok bahasan pergeseran arah kesetimbangan memiliki sebaran persentase jenjang kognitif C4 sebesar 20% dan jenjang kognitif C6 sebesar 3,3%. Terakhir pada sub pokok bahasan penerapan konsep kesetimbangan kimia memiliki sebaran persentase jenjang kognitif C4 sebesar 6,7% dan jenjang kognitif C6 sebesar 6,7%.

Berdasarkan distribusi butir soal HOTS yang telah diuji cobakan tersebut dihasilkan 10 butir soal yang harus dibuang sehingga tersisa 20 butir soal. Dari 20 butir soal yang dinyatakan valid ini memiliki persentase distribusi jenjang kognitif yang berbeda pada setiap sub pokok bahasannya. Pada sub pokok bahasan konsep kesetimbangan dinamis, persentase sebaran jenjang kognitif C4 sebesar 20%. Pada sub pokok bahasan tetapan kesetimbangan kimia (Kc dan Kp) memiliki persentase sebaran jenjang kognitif C4 sebesar 30% dan jenjang kognitif C5 sebesar 10%. Pada pokok bahasan pergeseran arah kesetimbangan persentase sebaran jenjang kognitif C4 sebesar 20%. Sedangkan pada sub pokok bahasan penerapan konsep kesetimbangan kimia memiliki persentase sebaran jenjang kognitif C4 sebesar 10% dan jenjang kognitif C6 sebesar 10%.

Berdasarkan distribusi butir soal HOTS yang dikembangkan tersebut sesuai dengan tingkatan jenjang kognitif dari taksonomi Bloom revisi, mayoritas butir soal didominasi oleh jenjang kognitif C4 yaitu menganalisis. Merujuk kepada hasil uji coba yang telah dilakukan terhadap butir soal yang dikembangkan tersebut diperoleh hasil bahwa siswa yang menyelesaikan instrumen evaluasi keterampilan berpikir tingkat tinggi pada materi kesetimbangan kimia, dengan kriteria sangat baik terdapat 18 siswa yang memiliki persentase 50% pada rentang nilai 80-100. Pada kriteria baik terdapat 6 siswa dengan persentase sebesar 17% dengan rentang nilai 60-80, untuk kriteria cukup terdapat 6 siswa dengan persentase sebesar 17% pada rentang nilai 40-60. Selanjutnya pada kriteria kurang terdapat 3 siswa dengan persentase 8% dalam rentang nilai 20-40, dan pada kriteria sangat kurang terdapat 3 siswa dengan persentase

8% dalam rentang nilai 0-20. Sehingga diperoleh nilai rata-rata siswa kelas XI C di SMAN 2 Percut Sei Tuan pada instrumen evaluasi keterampilan berpikir tingkat tinggi untuk materi kesetimbangan kimia adalah 72 dengan kriteria baik.

#### IV. SIMPULAN DAN SARAN

##### A. Simpulan

Pengembangan instrumen evaluasi untuk mengukur keterampilan berpikir tingkat tinggi pada materi kesetimbangan kimia dikembangkan dengan menggunakan model 4D, dimana model ini terdiri dari 4 tahapan yaitu pendefinisian, perencanaan, pengembangan, dan penyebaran. Instrumen evaluasi yang dikembangkan berupa soal pilihan ganda sebanyak 30 soal dengan masing-masing soal yang mengukur ranah kognitif menganalisis (C4), mengevaluasi (C5), dan mencipta (C6).

Instrumen evaluasi yang dikembangkan dinyatakan layak digunakan dengan kriteria telah memenuhi syarat validitas yang diperoleh dari penilaian validator ahli dengan presentase sebesar 87% dengan kategori sangat valid. Berdasarkan validitas butir soal hasil uji coba pada kategori valid ( $r_{hitung} \geq 0,329$ ), koefisien reliabilitas 0,885 pada kategori sangat tinggi karena berada pada interval  $0,80 < r_{11} < 1,00$ , tingkat kesukaran 0,31 sampai 1,00 pada kategori sedang dan mudah karena berada pada interval  $0,20 < P < 0,80$ , dan daya beda 0,21 sampai 0,70 pada kategori cukup hingga baik karena berada pada interval  $0,20 < D < 1,00$ .

Keterampilan berpikir tingkat tinggi pada materi kesetimbangan kimia di kelas XI C SMA Negeri 2 Percut Sei Tuan tahun ajaran 2022/2023 termasuk pada kategori tinggi. Hal ini berdasarkan nilai rata-rata yang diperoleh siswa secara keseluruhan adalah 72.

##### B. Saran

Pembahasan terkait penelitian ini masih sangat terbatas dan membutuhkan banyak masukan, saran untuk penulis selanjutnya adalah mengkaji lebih dalam dan secara komprehensif tentang Pengembangan Instrumen Evaluasi untuk Mengukur Keterampilan Berpikir Tingkat Tinggi pada Materi Kesetimbangan Kimia.

#### DAFTAR RUJUKAN

Amaliya, I., & Fathurohman, I. (2022). Analisis Kemampuan Literasi Matematika Ditinjau Dari Gaya Belajar Siswa Sekolah Dasar.

- Jurnal Riset Pendidikan Dasar*, 5(1), 45–56.  
<https://doi.org/https://doi.org/10.26618/jrpd.v5i1.7294>
- Andari, D., Rohiat, S., & Nurhamidah. (2021). Analisis Soal Pada Buku Teks Kimia Sma Kelas Xi Berdasarkan Ranah Kognitif Taksonomi Bloom. *Jurnal Pendidikan Dan Ilmu Kimia*, 5(2), 175-182.  
<https://doi.org/10.33369/atp.v5i2.17139>
- Andriani, D. W., & Yonata, B. (2018). Melatihkan High Order Thinking Skills Peserta Didik Melalui Implementasi Model Pembelajaran Inkuiri Pada Materi Kesetimbangan Kimia. *Unesa Journal of Chemical Education*, 7(3), 333–339.  
<https://jurnalmahasiswa.unesa.ac.id/index.php/journal-of-chemical-education/article/view/25621>
- Arum, N. E., Sjaifuddin, & Taufik, A. N. (2022). Pengembangan Instrumen Soal HOTS (Higher Order Thinking Skill) Berbasis Quizizz pada Pembelajaran IPA Daring Tema Hipertensi untuk SMP Kelas VIII. *Eksakta: Jurnal Penelitian Dan Pembelajaran MIPA*, 7(2), 270–279.  
<https://doi.org/10.31604/eksakta.v7i2.270-279>
- Aprilia, T., & Panggabean, J.H. (2021). Pengembangan Instrumen Tes Berbasis Critical Thinking Skill pada Materi Pokok Fluida SMA. *Jurnal Inovasi Pembelajaran Fisika (INPAFI)*, 9(2)-65-72.  
<http://jurnal.unimed.ac.id/2012/index.php/inpafi>
- Dewi, P., Elvia, R., & Elvinawati. (2021). Pengembangan Butir Soal Hots Untuk Menguji Kemampuan Berpikir Tingkat Tinggi Siswa Di Ma Negeri 2 Kota Bengkulu. *Jurnal Pendidikan Dan Ilmu Kimia*, 5(2), 141–148.  
<https://doi.org/10.33369/atp.v5i2.17119>
- Effendi, Z., & Andromeda, A. (2022). Development of a Two-Tier Higher Pengembangan Instrumen Tes Two-Tier Higher Order Thinking Skill (HOTS) Berbasis IBT pada Materi Kesetimbangan Kimia untuk Siswa SMA/MA Development Order Thinking Skill (HOTS) Test Instrument Based on IBT on Chemical Equil. *Entalpi Pendidikan Kimia*, 6(1), 31–39.  
<https://doi.org/10.24036/epk.v3i2.264>
- Gultom, E. R., & Muchtar, Z. (2022). Penerapan Model Pembelajaran Berbasis Proyek Pada Pembelajaran E-Learning Materi Kesetimbangan Kimia. *Educenter: Jurnal Ilmiah Pendidikan*, 1(4), 327–334.  
<https://doi.org/10.55904/educenter.v1i4.91>
- Jayanti, E. (2020). Instrumen Tes Higher Order Thinking Skill Pada Materi Kimia Sma. *Orbital: Jurnal Pendidikan Kimia*, 4(2), 135–149.  
<https://doi.org/10.19109/ojpk.v4i2.4368>
- Pusat Penilaian Pendidikan. (2019). *Risalah Kebijakan Meningkatkan Kemampuan Literasi Dasar Siswa Indonesia Berdasarkan Analisis data PISA 2018*. Balitbang, Kemendikbud.
- Lubis, A. R., Sudrajat, A., & Nugraha, W. (2022). Pengembangan instrumen penilaian kemampuan berpikir tingkat tinggi (HOTS) hidrolisis garam. 1, 63–69.  
<https://doi.org/10.24114/jpkim.v14i1.31382>
- Malikah, S., & Wafroturrohmah, W. (2022). Konsep Pendidikan Abad 21: untuk Pengembangan Sumber Daya Manusia SMA. *JIIP - Jurnal Ilmiah Ilmu Pendidikan*, 5(7), 2609–2614.  
<https://doi.org/10.54371/jiip.v5i7.730>
- Napitupulu, N. M., & Yusuf, M. (2022). Pengembangan Instrumen Tes Berbasis Hots (Higher Order Thinking Skills) Pada Materi Laju Reaksi Kimia. *Jurnal Sekolah*, 6(3), 162–168.  
<https://doi.org/10.24114/js.v6i3.36179>
- Netri, N., Holiwarni, B., & Abdullah. (2018). Development Of Test Instruments Based Higher Order Thinking Skill (HOTS) On Chemical Equilibrium At Second Grade In Senior High School. *Jom FKIP*, 5(2), 1–11.
- Panggabean, F. T. M., Purba, J., Sutiani, A., & Panggabean, M. A. (2022). Analisis Hubungan Antara Kemampuan Matematika dan Analisis Kimia Terhadap Hasil Belajar Kimia Materi Kesetimbangan Kimia. *Jurnal Inovasi Pembelajaran Kimia*, 4(1), 18.  
<https://doi.org/10.24114/jipk.v4i1.32904>

- Risdiana, A., Erna, M., & Holiwarni, B. (2022). Pengembangan Soal HOTS (Higher Order Thinking Skills) pada Materi Asam-Basa untuk Kelas XI SMA/MA Sederajat. *Jurnal Inovasi Pendidikan Kimia*, 16(2), 111-117. <https://doi.org/10.15294/jipk.v16i2.32879>
- Sholeh, M. I., & Olenia, Y. (2022). Strategi dan Implementasi Penyusunan Soal HOTS Kimia Berbasis Lesson Study. *Orbital: Jurnal Pendidikan Kimia*, 6(1), 38-48. <https://doi.org/10.19109/ojpk.v6i1.12169>
- Siahaan, R., Sitorus, M., & Silaban, S. (2021). The development of teaching materials oriented to critical thinking skills for chemistry class XI high school. *Jurnal Pendidikan Kimia*, 13(1), 60-68. <https://doi.org/10.24114/jpkim.v13i1.24145>
- Siregar, S. L., & Panggabean, F. T. M. (2020). Analisis Pbl Dengan DI Menggunakan Macromedia Flash Terhadap Motivasi Dan Hasil Belajar Siswa Pada Materi Laju Reaksi Di Sma Negeri 10 Medan. *Jurnal Inovasi Pembelajaran Kimia*, 2(1), 21. <https://doi.org/10.24114/jipk.v2i1.17829>
- Sriyanti, A., Samdewi, A. R., Mania, S., & Yuliany, N. (2022). Analisis Soal Tipe Higher Order Thinking Skill (HOTS) pada Buku Ajar Matematika SMK Kelas XI. *Edukatif: Jurnal Ilmu Pendidikan*, 4(2), 2385-2394. <https://doi.org/10.31004/edukatif.v4i2.2400>
- Sutiani, A., Muchtar, Z., Dibyantini, R. E., Sinaga, M., & Purba, D. J. (2022). *JURNAL INOVASI PEMBELAJARAN KIMIA (Journal Of Innovation in Chemistry Education) Analisis Kemampuan Guru-Guru Kimia SMA Sumatera Utara Dalam Mengintegrasikan TPACK*. 3. <https://jurnal.unimed.ac.id/2012/index.php/jipk>
- Yustiana, S., & Romadhoni, D.R.M. (2017). Analisis Soal Pilihan Ganda pada Mata Pelajaran PKn Buatan Guru Sekolah Dasar. *Jurnal Inovasi Pendidikan Dasar*, 3(1), 9-14.