

PENGEMBANGAN INSTRUMEN ASESMEN BERBASIS ENAM ASPEK PEMAHAMAN DALAM KERANGKA *UNDERSTANDING BY DESIGN* (UbD) PADA MATERI TERMOKIMIA KELAS XI SMA SEDERAJAT

Aufa Agustia^{1*}, Sri Haryati², Betty Holiwarni³

Program Studi Pendidikan Kimia, Universitas Riau, Indonesia

*Corresponding author email: aufa.agustia0898@student.unri.ac.id, sri.haryati@lecturer.unri.ac.id, betty.holiwarni@lecturer.unri.ac.id

Article History

Received: 12 June 2025

Revised: 24 October 2025

Published: 3 November 2025

ABSTRACT

This study aims to develop an assessment instrument based on the six facets of understanding within the Understanding by Design (UbD) framework, specifically for thermochemistry material. The research adopted the ADDIE model, which consists of five stages: Analysis, Design, Development, Implementation, and Evaluation. The study was conducted from February to May 2025 at the Faculty of Teacher Training and Education, University of Riau, with field testing carried out at SMA Negeri 4 Pekanbaru and SMA Negeri Plus Provinsi Riau. Data were collected through expert validation, one-on-one trials, small group trials, and item characteristic analysis. Expert validation results indicated that the developed instrument is valid in terms of content, construct, and language, with average percentage scores of 100%, 96.63%, and 100%, respectively. Student responses to the instrument fell into the "very good" category, with an average score percentage of 91.08%. Teacher responses also indicated a "very good" category, with an average score percentage of 95.83%. The item analysis revealed that the instrument has strong construct validity, high reliability, and appropriate levels of difficulty and discrimination. Therefore, this assessment instrument is deemed suitable for measuring students' understanding in alignment with the learning objectives outlined in the Merdeka Curriculum.

Keywords: *Assessment, Instrument Development, Six Facets of Understanding, Understanding by Design, Thermochemistry*

Copyright © 2025, The Author(s).

How to cite: Agustia, A., Haryati, S., & Holiwarni, B. (2025). Pengembangan Instrumen Asesmen Berbasis Enam Aspek Pemahaman dalam Kerangka Understanding by Design (UbD) Pada Materi Termokimia Kelas XI SMA Sederajat. *NUSRA : Jurnal Penelitian Dan Ilmu Pendidikan*, 6(4), 723–737. <https://doi.org/10.55681/nusra.v6i4.3970>



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

LATAR BELAKANG

Pembelajaran semakin berkembang seiring dengan perkembangan zaman dan perubahan kurikulum, begitu pula dengan tuntutan pembelajaran yang kian meningkat. Saat ini peserta didik dituntut untuk mampu memahami apa yang telah dipelajari secara mendalam. Bukan lagi sekadar mampu menyebutkan atau menjelaskan definisi, lebih dari itu pemahaman yang dimaksud memiliki kompleksitas yang lebih tinggi. Panduan Pembelajaran dan Asesmen yang disusun oleh Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi edisi revisi 2024 telah menunjukkan bahwa seluruh rangkaian proses pembelajaran memiliki tujuan akhir yang sama, yaitu peserta didik dapat memperoleh pemahaman terkait materi yang dipelajari. Sistem penilaian yang sesuai dengan karakteristik pembelajaran berbasis kompetensi adalah penggunaan instrumen penilaian yang dapat mengukur aspek psikomotor (Taufik & Mus'id, 2020).

Capaian pembelajaran yang telah dirumuskan didalam Panduan Pembelajaran dan Asesmen menitikberatkan pada pemahaman ditandai dengan tiap kalimatnya diawali dengan kata “memahami”, tidak terkecuali pada pembelajaran kimia dimana pendekatan yang dilakukan dalam pembelajaran saat ini yaitu metode pendekatan *Deep Learning* (Kemendikbudristek, 2024). Menurut pendekatan *Deep Learning*, pemahaman yang dimaksud mencakup berbagai aspek, seperti alasan suatu peristiwa dapat terjadi, pengaplikasian konsep dalam kehidupan nyata, pemecahan masalah di kehidupan nyata berdasarkan konsep yang telah dipelajari, serta aspek-aspek pemahaman yang mendalam lainnya. (Barlian, 2022).

Tujuan pembelajaran pada materi Termokimia misalnya, bukan sekadar mengetahui konsep Termokimia atau langkah menghitung nilai perubahan entalpi (ΔH), melainkan hasil apa yang akan diperoleh peserta didik jika mereka memiliki pemahaman terkait Termokimia. Untuk mengetahui tercapai atau tidaknya tujuan pembelajaran yang telah dirumuskan dapat ditinjau menggunakan alat ukur yang sesuai. Alat ukur ini sering dikenal dengan asesmen pembelajaran. Kemendikbudristek (2024) menyatakan ada tiga pendekatan yang dapat digunakan dalam penyusunan asesmen yaitu menggunakan Taksonomi Bloom, *Understanding by Design*, dan Marzano.

Ketika konsep pemahaman yang dimaksud dalam Capaian Pembelajaran memiliki makna yang lebih kompleks dan mendalam, maka asesmen berbasis taksonomi *bloom* dinilai tidak relevan digunakan untuk mengukur pemahaman yang ingin dicapai pada pembelajaran saat ini. Hal ini karena pada taksonomi *bloom* pemahaman dianggap bukan kata kerja operasional, sehingga tidak dapat diukur (Arda, 2020). Artinya menurut taksonomi *bloom* pemahaman adalah suatu hal yang tidak dapat diukur, sedangkan yang menjadi tujuan dari pembelajaran saat ini ialah pemahaman yang seharusnya dapat terukur untuk mengetahui ketercapaian tujuan pembelajaran.

Asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* akan lebih sesuai digunakan untuk mengukur pemahaman yang dimaksud dalam tujuan pembelajaran dimasa kini. Hal ini dikarenakan enam aspek pemahaman yang dikemukakan oleh Wiggins dalam kerangka *Understanding by Design* menilai pemahaman sebagai sesuatu yang dapat

diukur dengan melihat pada ketercapaian enam aspek pemahaman. Aspek tersebut diantaranya *explanation*, *perspective*, *interpretation*, *application*, *emphaty*, dan *self-knowledge* (Gloria, 2019). Enam aspek tersebut jauh lebih kompleks dan mendalam jika dibandingkan dengan aspek pemahaman yang tercantum dalam taksonomi *bloom* yang hanya menempati tingkat C2. Dengan demikian asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* lebih sesuai digunakan dalam pengukuran pemahaman peserta didik sesuai dengan tujuan pembelajaran saat ini.

Hasil observasi yang dilakukan di SMA Negeri 4 Pekanbaru menunjukkan bahwa SMA tersebut menggunakan instrumen asesmen berbasis taksonomi *bloom*. Begitu pula dengan SMA Negeri Plus Provinsi Riau. Berdasarkan wawancara yang dilakukan kepada guru kimia, ditemukan informasi bahwa instrumen asesmen yang digunakan juga berbasis taksonomi *bloom*. Kedua sekolah juga belum pernah menggunakan asesmen berbasis *Understanding by Design* yang menilai pemahaman sebagai sesuatu yang dapat diukur untuk tujuan pembelajaran saat ini.

Berdasarkan permasalahan yang telah diuraikan, maka dikembangkan instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* (UbD) pada Materi Termokimia Kelas XI SMA sederajat yang diharapkan mampu menjadi solusi dan dapat digunakan sebagai asesmen yang relevan dengan tujuan pembelajaran saat ini, sehingga rumusan masalah dalam penelitian ini ialah bagaimana pengembangan instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* (UbD) pada materi termokimia yang valid berdasarkan aspek kelayakan konstruksi,

kelayakan materi, dan kelayakan bahasa; bagaimana respon guru dan peserta didik terhadap instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* (UbD) pada materi termokimia; serta bagaimana kualitas butir soal pada instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* (UbD) pada materi termokimia berdasarkan aspek validitas konstruk, reliabilitas, tingkat kesukaran, dan daya pembeda soal.

Tujuan dari penelitian ini meliputi untuk menghasilkan instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* (UbD) pada materi termokimia yang valid berdasarkan aspek kelayakan konstruksi, kelayakan materi, dan kelayakan bahasa; untuk mengetahui respon guru dan peserta didik terhadap instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* (UbD) pada materi termokimia; serta untuk mengetahui kualitas butir instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* (UbD) pada materi termokimia berdasarkan aspek validitas konstruk, reliabilitas, tingkat kesukaran, dan daya pembeda soal.

METODE PENELITIAN

Penelitian ini akan dilaksanakan di Fakultas Keguruan dan Ilmu Pendidikan Universitas Riau pada bulan Februari-Mei 2025, dengan uji coba yang dilakukan di dua sekolah, yaitu uji karakteristik butir soal di SMA Negeri 4 Pekanbaru dan uji satu-satu serta kelompok kecil dilakukan di SMA Negeri Plus Provinsi Riau. Penelitian ini merupakan penelitian *Research and Development* dengan model ADDIE yang terdiri dari tahap *Analysis*, *Design*,

Development, Implementation, dan Evaluations.

Pengembangan instrumen asesmen berbasis *Understanding by Design* (UbD) mengikuti prosedur ADDIE sebagai berikut:

Analysis: dilakukan analisis kurikulum, asesmen yang digunakan di sekolah, analisis materi termokimia, serta wawancara dengan guru kimia. Hasil analisis dievaluasi dan dijadikan landasan untuk melanjutkan ke tahap berikutnya.

Design: penyusunan kisi-kisi penyusunan asesmen dilakukan dengan mengacu pada indikator dari enam aspek pemahaman dalam kerangka UbD. Kisi-kisi ini kemudian dirancang menjadi soal esai yang disesuaikan dengan tujuan pembelajaran pada materi termokimia.

Development: tahap ini mencakup pembuatan soal, validasi ahli terhadap aspek konstruksi, materi dan bahasa. Selanjutnya uji coba satu-satu kepada tiga peserta didik dengan tingkat kemampuan berbeda untuk memperoleh saran deskriptif terhadap soal yang dikembangkan. Kemudian, uji coba kelompok kecil terhadap 10 peserta didik untuk memperoleh respon peserta didik dan respon guru sebagai pengguna. Terakhir, dilakukan uji karakteristik butir soal pada 30 peserta didik untuk menganalisis validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal.

Implementation: tahap ini belum dilakukan karena fokus penelitian hanya sampai tahap pengembangan produk yang valid dan layak digunakan.

Evaluation: evaluasi dilakukan dalam setiap tahap pengembangan. Evaluasi pada penelitian ini didasarkan pada analisis kurikulum, analisis asesmen, analisis kebutuhan, analisis materi, saran dari guru dan juga validator. Berdasarkan saran tersebut dilakukan perbaikan pada instrumen asesmen yang dihasilkan.

Subjek uji coba dalam penelitian terdiri dari 43 peserta didik XI IPA dari SMA Negeri 4 Pekanbaru dan SMA Negeri Plus Provinsi Riau tahun ajaran 2024/2025, yang terbagi sebagai 3 orang peserta uji satu-satu, 10 orang peserta uji kelompok kecil, dan 30 orang sebagai peserta uji dalam analisis karakteristik butir soal.

Jenis dan sumber data yang digunakan dalam penelitian ini yaitu data kuantitatif dan data kualitatif. Data kuantitatif diperoleh dari skor rata-rata lembar validasi oleh ahli dan indeks yang diperoleh dari uji karakteristik butir soal yang digunakan. Data kualitatif diperoleh dari & respon guru kimia SMA Negeri 4 Pekanbaru dan SMA Negeri Plus Provinsi Riau serta saran dari 3 orang validator.

Instrumen yang digunakan dalam pengumpulan data pada penelitian ini ialah lembar validasi ahli yang berisi pernyataan untuk meminta persetujuan validator meliputi 3 aspek, yaitu aspek materi, konstruksi dan bahasa. Selanjutnya, instrumen yang digunakan pada penelitian ini berupa angket atau kuesioner respon peserta didik yang meminta respon peserta didik terhadap instrumen asesmen yang diberikan meliputi aspek materi, bahasa, dan ketercukupan waktu.

Terakhir, instrumen yang digunakan ialah angket atau kuesioner respon guru yang meminta respon guru terhadap instrumen asesmen yang dikembangkan mencakup aspek keefektifan, bahasa, dan kemanfaatan.

Pernyataan sikap pada angket atau kuesioner respon peserta didik dan angket respon guru kimia sebagai pengguna, serta lembar validasi ahli menggunakan skala pengukuran sikap 1-4 seperti yang tertera pada Tabel 1.

Tabel 1. Skor Alternatif Pernyataan Positif

Pernyataan Sikap	Skor
Sangat Setuju (SS)	4
Setuju (S)	3
Tidak Setuju (KS)	2
Sangat Tidak Setuju (KS)	1

Teknik analisis data kuantitatif dilakukan terhadap data berupa skor yang diperoleh pada penelitian ini. Skor yang diperoleh dari validasi ahli kemudian dihitung persentasenya dengan menggunakan persamaan berikut:

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

Keterangan:

P = Persentase skor

n = Jumlah skor yang diperoleh

N = Jumlah skor maksimum

Hasil persentase dari perhitungan skor validasi yang diperoleh kemudian dikonversi menjadi nilai kualitatif sesuai kriteria penilaian skala pengukuran sikap dengan skor 1-4. Kriteria hasil validasi yang diperoleh seperti yang tertera pada Tabel 2.

Tabel 2. Kriteria Validitas

Persentase Skor	Kriteria Validitas
76% – 100%	Valid
51% – 75%	Cukup Valid
26% – 50%	Tidak Valid
0% – 25%	Sangat Tidak Valid

Skor yang diperoleh dari respon peserta didik dan guru juga dihitung persentasenya menggunakan persamaan yang serupa pada perhitungan persentase skor validasi. Hasil presentase yang diperoleh dikonversi menjadi nilai kualitatif dengan kriteria seperti yang tertera pada Tabel 3.

Tabel 3. Kriteria Respon Peserta Didik dan Guru

Persentase Skor	Kriteria Respon Pengguna
81% – 100%	Sangat Baik
61% – 80%	Baik
41% – 60%	Cukup Baik
21% – 40%	Tidak Baik
0% – 20%	Sangat Tidak Baik

Analisis karakteristik butir soal meliputi validitas konstruk, reliabilitas, tingkat kesukaran, dan daya pembeda soal. Validitas konstruk dilakukan dengan menggunakan data nilai yang diperoleh peserta didik uji. Metode yang digunakan ialah metode *Pearson Correlation* berbantuan *SPSS for Windows* versi 27.0. Nilai r yang diperoleh dari hasil *Pearson Correlation* berbantuan *SPSS for Windows* versi 27.0 atau disebut dengan r hitung kemudian akan dibandingkan dengan harga r tabel dengan $\alpha = 0,05$ (signifikansi 5%) dan $df = N-2$ (N = banyaknya siswa). Bila $r_{hitung} > r_{tabel}$ maka instrumen dinyatakan valid. Karena jumlah peserta uji pada penelitian ini berjumlah 30 peserta didik sehingga r tabel yang digunakan ialah 0,361 (Sudijono, 2015).

Reliabilitas diperoleh menggunakan koefisien *Alpha Cronboach*. Pada penelitian

ini nilai koefisien *Alpha Cronboach* dihitung menggunakan bantuan *SPSS for Windows* versi 27.0. Nilai r yang diperoleh dari hasil perhitungan *Alpha Cronbach* kemudian akan dibandingkan dengan harga r tabel dengan $\alpha = 0,05$ dan $df = N-2$ (N = banyaknya siswa). Bila $r_{hitung} > r_{tabel}$ maka instrumen dinyatakan reliabel. Karena jumlah sampel pada uji ini berjumlah 30 orang peserta didik, maka r tabel yang digunakan bernilai 0,377 (Sudijono, 2015). Sedangkan untuk mengetahui tinggi rendahnya reliabilitas suatu instrumen asesmen digunakan kategori pada Tabel 4.

Tabel 4. Kriteria Indeks Reliabilitas Soal

Indeks	Keterangan
0,000 – 0,199	Sangat rendah
0,200 – 0,399	Rendah
0,400 – 0,599	Cukup
0,600 – 0,799	Tinggi
0,800 – 1,000	Sangat tinggi

Tingkat kesukaran Tingkat kesukaran pada soal uraian dapat diketahui dengan menggunakan rumus sebagai berikut (Nitko, 1996):

$$TK = \frac{\text{Mean}}{\text{Skor maksimum yang ditetapkan}}$$

Mean dapat dihitung dengan cara:

$$\text{Mean} = \frac{\text{Jumlah skor peserta tes pada suatu soal}}{\text{Jumlah peserta didik yang mengikuti tes}}$$

Menentukan kategori tingkat kesukaran berdasarkan kriteria tingkat kesukaran dalam Arikunto (2013) seperti Tabel 5.

Tabel 5. Kriteria Indeks Kesukaran

Indeks	Tingkat Kesukaran
0.00–0.30	Sukar
0.31–0.70	Sedang
0.71–1.00	Mudah

Daya pembeda untuk soal bentuk uraian dihitung dengan rumus berikut.

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B}$$

Keterangan:

D : Daya beda soal

B_A : jumlah peserta didik kelompok atas yang menjawab soal dengan benar

B_B : jumlah peserta didik kelompok bawah yang menjawab soal dengan benar

J_A : jumlah peserta didik kelompok atas

J_B : jumlah peserta didik kelompok bawah

Hasil perhitungan dengan menggunakan rumus di atas dapat menggambarkan tingkat kemampuan soal dalam membedakan antar peserta didik yang sudah memahami materi yang diujikan dengan peserta didik yang belum/tidak memahami materi yang diujikan. Adapun klasifikasinya adalah seperti pada Tabel 6 berikut ini (Crocker dan Algina, 1986).

Tabel 6. Kriteria Indeks Daya Pembeda Soal

Indeks	Keterangan
0,19 - 0,00	Sangat buruk
0,20 - 0,29	Buruk
0,30 - 0,39	Cukup
0,40 - 1,00	Baik

HASIL DAN PEMBAHASAN

Produk yang dihasilkan dari penelitian pengembangan berupa instrumen asesmen berbentuk soal esai berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* yang sudah divalidasi dan diuji karakteristik butir soal. Hasil dari tahapan penelitian pengembangan dengan model ADDIE dapat dijelaskan seperti dibawah ini:

1. Tahap *Analyze*

Analisis kurikulum diperoleh dari hasil observasi serta wawancara guru kimia SMAN 4 Pekanbaru dan SMAN Plus Provinsi Riau. Dari kegiatan tersebut diperoleh informasi bahwa pembelajaran di kedua SMA tersebut sudah menggunakan Kurikulum Merdeka. Pembelajaran dilakukan sesuai dengan Kurikulum Merdeka, mulai dari capaian pembelajaran hingga tujuan pembelajaran yang disesuaikan dengan panduan pembelajaran dari Kemendikbudristek (2024).

Analisis asesmen dari wawancara guru kimia SMA diperoleh informasi bahwa instrumen asesmen yang digunakan berbasis taksonomi *bloom*, yang melihat pemahaman sebagai sesuatu yang tidak dapat diukur.

Analisis materi pembelajaran kimia di sekolah, diperoleh hasil berupa informasi bahwa seluruh materi kimia yang dipelajari di sekolah mengacu pada capaian pembelajaran pada pedoman dari kemendikbud dengan asesmen yang digunakan berbasis taksonomi *bloom*, termasuk materi termokimia yang sudah dipelajari. Analisis materi Termokimia untuk mengetahui tujuan pembelajaran yang sesuai dengan capaian pembelajaran saat ini. Analisis materi dilakukan dengan menganalisis capaian pembelajaran kemudian mengkonversikan ke dalam tujuan

pembelajaran yang disesuaikan dengan enam aspek pemahaman dalam kerangka *Understanding by Design*.

Studi kepustakaan dilakukan dengan membaca dan mempelajari jurnal, artikel penelitian, dan buku mengenai instrumen tes, pengembangan instrumen tes, serta *Understanding by Design*. Berdasarkan studi kepustakaan, diperoleh informasi bahwa instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* dinilai sesuai untuk mengukur pemahaman yang menjadi capaian pembelajaran peserta didik dalam kurikulum saat ini. Instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* mampu mengukur pemahaman yang dinilai tidak terukur atau bukan kata kerja operasional pada taksonomi bloom (Tim Penyusun, 2013).

Wiggins dan Tighe (2005) mengemukakan bahwa pemahaman (*understanding*) adalah proses berpikir tingkat tinggi, bukan sekadar menggunakan informasi untuk menjawab pertanyaan sederhana atau menjelaskan defenisi sahaja, lebih dari itu, pemahaman merupakan suatu kemampuan yang holistik yang diperoleh dari proses yang sistematis. Pemahaman dibagi menjadi enam aspek yaitu *explanation*, *interpretation*, *application*, *perspective*, *emphatize*, dan *self-knowledge*. Keenam aspek ini dapat diukur untuk menjadi acuan ketercapaian tujuan dalam pembelajaran saat ini.

Evaluasi kemudian dilakukan terhadap hasil analisis yang telah diperoleh, sehingga didapatkan kesimpulan bahwa perlu adanya penyesuaian instrumen asesmen yang digunakan untuk mengukur ketercapaian tujuan pembelajaran saat ini.

2. Tahap Design

Bentuk instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* ialah tes uraian atau esai yang terdiri dari 12 soal dengan pokok bahasan Termokimia.

Kisi-kisi penyusunan instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* disusun berdasarkan analisis materi yang telah dilakukan pada tahap analisis. Kisi-kisi berisikan indikator enam aspek pemahaman dalam kerangka *Understanding by Design*, tujuan pembelajaran, kisi-kisi penulisan soal serta nomor butir soal.

3. Tahap Development

Hasil yang diperoleh pada tahap pengembangan ialah instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* pada materi Termokimia yang disusun berdasarkan kisi-kisi yang telah dibuat pada tahap sebelumnya. Instrumen asesmen telah mencakup seluruh sub materi pada materi Termokimia dan keenam aspek pemahaman dalam kerangka *Understanding by Design* yang dikemukakan oleh Wiggins dan Tighe. Instrumen asesmen yang dikembangkan kemudian divalidasi oleh Ahli dan diuji untuk mendapatkan respon pengguna, serta dianalisis karakteristik butir soalnya.

Hasil validasi diperoleh dari proses validator menilai dan memberikan saran serta masukan terhadap soal yang dikembangkan dengan memberikan skor pada lembar validasi. Lembar validasi terdiri dari 14 butir pernyataan yang terdiri dari 3 aspek, dengan aspek konstruksi sebanyak 7 soal, aspek materi sebanyak 4 soal, dan aspek bahasa sebanyak 3 soal.

Aspek konstruksi yang dinilai pada validasi ini meliputi kejelasan dan ketegasan

rumusan soal dan efektivitas stimulus yang digunakan. Terdapat 7 pernyataan pada lembar validasi terkait aspek konstruksi yaitu menanyakan kesetujuan ahli terkait kesesuaian rumusan kalimat soal dengan kata tanya atau perintah yang digunakan, kata tanya atau perintah menuntut jawaban teruai dari peserta didik, rumusan jelas tidak menimbulkan penafsiran ganda, keluasaan pertanyaan dengan perkembangan peserta didik di Sekolah Menengah Atas, kesesuaian soal dengan stimulus yang digunakan, stimulus yang digunakan bebas pernyataan yang tidak relevan dan bebas pernyataan negatif ganda, serta kemampuan soal dalam menggunakan informasi untuk menyelesaikan masalah.

Aspek materi yang dinilai pada validasi ini ialah kesesuaian soal dengan materi yang diajarkan disekolah serta dengan capaian pembelajaran, tujuan pembelajaran, dan indikator ketercapaian enam aspek pemahaman pada kisi-kisi yang telah disusun. Terdapat 4 pernyataan pada aspek materi yaitu kesesuaian butir soal dengan capaian pembelajaran, kesesuaian butir soal dengan tujuan pembelajaran yang dirumuskan, kesesuaian butir soal dengan indikator enam aspek pemahaman dalam kerangka *Understanding by Design*, dan butir soal benar sesuai konsep materi termokimia.

Aspek bahasa yang dinilai pada validasi ini ialah kesesuaian rumusan soal dengan Pedoman Umum Ejaan Bahasa Indonesia (PUEBI). Terdapat tiga pernyataan pada aspek ini, yaitu kesesuaian butir soal dengan kaidah bahasa Indonesia, penyajian rumusan/kalimat soal tidak mengandung ungkapan yang bermakna tidak pasti (misal sebaiknya, kadang-kadang, pada umumnya), dan butir soal tidak menggunakan idiom-idiom lokal (bahasa yang berlaku setempat).

Validasi instrumen asesmen berbasis enam aspek pemahaman dalam kerangka *Understanding by Design* pada materi termokimia yang dikembangkan, menghasilkan data kualitatif dan data kuantitatif. Data kualitatif diperoleh melalui saran dan komentar validator terhadap instrumen asesmen yang dikembangkan. Saran dan komentar dari masing-masing validator kemudian dijadikan dasar evaluasi instrumen asesmen yang dikembangkan.

Evaluasi dilakukan terhadap instrumen asesmen yang telah mendapatkan saran dan komentar dari validator dengan melakukan tindak lanjut berupa revisi terhadap butir soal yang mendapatkan saran atau komentar dari validator.

Setelah dilakukan revisi, instrumen asesmen yang dikembangkan dinyatakan valid oleh validator pada validasi kedua, baik pada aspek konstruksi, materi, dan bahasa. Skor dari ketiga aspek yang diperoleh dari ketiga validator kemudian dikonversi sehingga diperoleh kriteria validitas masing-masing soal sebagai hasil akhir dari tahap validasi. Skor dan kriteria validasi pada aspek konstruksi dapat dilihat pada Tabel 7.

Tabel 7. Hasil Validasi Ahli pada Aspek Kontruksi.

No Soal	Persentase Skor Validator (%)			Rata-rata Presentase (%)	Kategori
	I	II	III		
1	100	96,43	92,85	96,43	Valid
2	100	96,43	100	98,81	Valid
3	100	92,85	100	97,62	Valid
4	100	92,85	92,85	95,23	Valid
5	100	100	96,43	98,81	Valid
6	100	92,85	96,43	96,43	Valid
7	100	96,43	92,85	96,43	Valid
8	100	92,85	92,85	95,23	Valid
9	100	92,85	92,85	95,23	Valid
10	100	92,85	96,43	96,43	Valid
11	100	92,85	100	97,62	Valid
12	96,43	92,85	96,43	95,24	Valid

Hasil pada Tabel 7 menunjukkan bahwa instrumen asesmen yang dikembangkan aspek konstruksi, dengan

rata-rata persentase skor dari validator 1 sebesar 99,7%, validator 2 sebesar 94,34%, dan validator 3 sebesar 95,83%. Artinya 12 soal esai yang dikembangkan sudah sesuai rumusan kalimatnya dengan kata tanya atau perintah yang digunakan, kata tanya yang digunakan pada setiap soal sudah menuntut jawaban terurai dari peserta didik.

Rumusan soal yang dikembangkan pada instrumen asesmen yang dikembangkan sudah jelas dan tidak menimbulkan penafsiran ganda. Sesuai dengan yang disampaikan Nitko dan Brookhart (2011) mengemukakan bahwa bahasa yang digunakan pada soal harus jelas, sederhana, dan tidak ambigu agar semua peserta didik dapat memahami maksud soal secara konsisten.

Soal yang baik harus dirumuskan dengan jelas dan tidak ambigu sehingga peserta didik dapat memahami maksud soal secara seragam. Penggunaan bahasa yang tidak tepat dapat menurunkan validitas instrumen yang dikembangkan, karena peserta didik berpotensi salah dalam memahami maksud soal bukan karena tidak menguasai materi, tetapi karena soal ditulis secara tidak jelas. Keluasan soal yang dikembangkan juga sudah sesuai dengan perkembangan peserta didik di tingkat SMA. Artinya segala istilah dan peristiwa yang disebutkan pada soal sudah diketahui oleh peserta didik di tingkat SMA.

Soal yang dikembangkan sudah sesuai dengan stimulus yang digunakan, dan stimulus yang digunakan bebas dari pernyataan tidak relevan. Menurut Zimmaro (2010) bahwa stimulus harus berkaitan langsung dengan soal dan tidak membingungkan peserta didik. Berdasarkan hal tersebut instrumen asesmen yang dikembangkan sudah fokus dalam mengukur pemahaman peserta didik. Selanjutnya hasil perhitungan validasi ahli terhadap instrumen

asesmen yang dikembangkan pada aspek materi dapat dilihat pada Tabel 8.

Tabel 8. Hasil Validasi Ahli pada Aspek Materi

No Soal	Persentase Skor Validator (%)			Rata-rata Presentase (%)	Kategori
	I	II	III		
1	100	100	100	100	Valid
2	100	100	100	100	Valid
3	100	100	100	100	Valid
4	100	100	100	100	Valid
5	100	100	100	100	Valid
6	100	100	100	100	Valid
7	100	100	100	100	Valid
8	100	100	100	100	Valid
9	100	100	100	100	Valid
10	100	100	100	100	Valid
11	100	100	100	100	Valid
12	100	100	100	100	Valid

Hasil pada Tabel 8 menunjukkan bahwa instrumen asesmen yang dikembangkan valid pada aspek materi. Seluruh soal yang dikembangkan telah dinyatakan valid pada aspek materi dengan rata-rata presentase skor validator 1 sebesar 100%, validator 2 sebesar 100%, dan validator 3 sebesar 100%.

Hasil validasi berarti bahwa 12 soal esai yang dikembangkan sudah sesuai dengan capaian pembelajaran saat ini. 12 soal esai yang dikembangkan sudah sesuai dengan tujuan pembelajaran dan indikator yang disusun berdasarkan enam aspek pemahaman dalam kerangka *Understanding by Design* yang dikemukakan oleh Wiggins (2005) meliputi aspek *explanation*, *interpretation*, *perspective*, *application*, *empathize*, dan *self-knowledge*. Berdasarkan hal tersebut, instrumen asesmen yang dikembangkan sudah mampu mengukur pemahaman peserta didik.

Selanjutnya hasil perhitungan validasi ahli terhadap instrumen asesmen yang dikembangkan pada aspek bahasa dapat dilihat pada Tabel 9.

Tabel 9. Hasil Validasi Ahli pada Aspek Bahasa

No Soal	Persentase Skor Validator (%)			Rata-rata Presentase (%)	Kategori
	I	II	III		
1	100	100	100	100	Valid
2	100	100	100	100	Valid
3	100	100	100	100	Valid
4	100	100	100	100	Valid
5	100	100	100	100	Valid
6	100	100	100	100	Valid
7	100	100	100	100	Valid
8	100	100	100	100	Valid
9	100	100	100	100	Valid
10	100	100	100	100	Valid
11	100	100	100	100	Valid
12	100	100	100	100	Valid

Hasil validasi menunjukkan bahwa soal yang dirumuskan dengan bahasa yang sesuai dengan Pedoman Umum Ejaan Bahasa Indonesia (PUEBI), tidak mengandung ungkapan bermakna ganda atau tidak pasti, dan tidak menggunakan idiom lokal yang tidak dipahami secara luas, akan membantu peserta didik memahami soal dengan benar. Hal ini penting agar asesmen mengukur kompetensi yang dimaksud, bukan kemampuan menafsirkan soal yang ambigu. Nitko dan Brookhart (2011) mengemukakan bahwa penggunaan bahasa yang tidak tepat dapat menurunkan validitas instrumen yang dikembangkan, karena peserta didik berpotensi salah dalam memahami maksud soal bukan karena tidak menguasai materi, tetapi karena soal ditulis secara tidak jelas.

Hasil validasi yang menunjukkan skor maksimal dari seluruh validator merupakan indikator bahwa instrumen yang dikembangkan sudah memenuhi kriteria kebahasaan yang sangat baik dengan kategori valid. Dengan demikian, 12 soal esai yang dikembangkan dapat dianggap layak digunakan sebagai instrumen asesmen dalam pembelajaran kimia khususnya pada materi termokimia.

Uji satu-satu pada tiga orang peserta didik memperoleh hasil berupa durasi pengerjaan dan hasil wawancara bersama

peserta didik. Durasi pengerjaan peserta didik dapat dilihat pada Tabel 10.

Tabel 10. Durasi Pengerjaan Peserta Didik

Kode Peserta Didik	Kemampuan	Durasi
PD-01	Tinggi	60
PD-02	Sedang	75
PD-03	Rendah	95

Hasil wawancara bersama peserta didik memperoleh hasil berupa informasi bahwa narasi soal yang dikembangkan secara keseluruhan sudah bisa dimengerti, hal ini karena butir soal sudah memenuhi aspek validitas dan layak uji coba. Hanya saja perlu waktu tambahan dalam pengerjaannya.

Evaluasi dilakukan berdasarkan hasil yang ditemukan pada uji satu-satu, sehingga dipertimbangkan waktu pengerjaan soal diubah dari semula 75 menit menjadi 90 menit atau setara dengan 2 jam pelajaran. Hal ini untuk memastikan bahwa waktu yang diberikan cukup untuk uji berikutnya sehingga tidak mempengaruhi hasil uji karakteristik butir soal nantinya. Sejalan dengan yang disampaikan oleh Nitko dan Brookhart (2011) yang menekankan bahwa waktu yang diberikan dalam asesmen harus cukup agar peserta didik dapat menunjukkan kompetensinya secara optimal.

Uji kelompok kecil terhadap 10 orang peserta didik dengan meminta peserta didik mengerjakan instrumen asesmen yang dikembangkan, kemudian mengisi angket repon peserta didik, diperoleh hasil berupa respon peserta didik terhadap instrumen yang dikembangkan seperti pada tabel 11.

Tabel 11. Rekapitulasi Skor Respon Peserta Didik

No	Aspek penilaian	Rata-rata skor respon peserta didik (%)	Kriteria
1	Aspek Materi	97,5	Sangat baik
2	Aspek Bahasa	92	Sangat baik
3	Aspek Waktu	83,75	Sangat baik
Rata-rata persentase		91,08%	Sangat baik

Hasil angket respon peserta didik menunjukkan bahwa secara keseluruhan instrumen asesmen diterima dengan sangat baik oleh peserta didik dari segi materi, keterbacaan, dan ketercukupan waktu. Respon positif ini menjadi indikator bahwa asesmen yang dikembangkan tidak hanya valid secara teoretis (validasi ahli), tetapi juga praktis dan layak digunakan di lapangan. Pada uji kelompok kecil juga diberikan angket kepada guru, dimana hasil respon guru dapat dilihat pada Tabel 12.

Tabel 12. Rekapitulasi Respon Guru

No.	Aspek	Persentase skor penilaian (%)		Kriteria
		Guru 1	Guru 2	
1.	Keefektifan	100	100	Sangat baik
2.	Bahasa	87,5	87,5	Sangat baik
3.	Kemanfaatan	100	100	Sangat baik
Rata-rata		95,83	95,83	Sangat baik

Hasil angket respons guru secara keseluruhan menunjukkan bahwa instrumen asesmen yang dikembangkan sangat layak untuk digunakan dalam proses

pembelajaran, karena telah memenuhi aspek keefektifan, kebahasaan, dan kemanfaatan. Penilaian ini juga menunjukkan bahwa dari sudut pandang praktisi, instrumen tidak hanya valid secara teoritis, tetapi juga praktis dan diaplikasikan dalam pembelajaran kimia.

Analisis karakteristik butir soal yang dilakukan terhadap instrumen asesmen yang dikembangkan memperoleh hasil kualitas butir soal dari segi validitas konstruk, reliabilitas, tingkat kesukaran, dan daya pembeda soal. Hasil uji karakteristik butir soal pada validitas konstruk dapat dilihat pada Tabel 13.

Tabel 13. Hasil Validitas Konstruk

No Soal	r Hitung	r Tabel	Kriteria
1	0,703	0,361	Valid
2	0,720	0,361	Valid
3	0,422	0,361	Valid
4	0,727	0,361	Valid
5	0,794	0,361	Valid
6	0,856	0,361	Valid
7	0,626	0,361	Valid
8	0,774	0,361	Valid
9	0,806	0,361	Valid
10	0,816	0,361	Valid
11	0,823	0,361	Valid
12	0,505	0,361	Valid

Hasil yang diperoleh dengan membandingkan r hitung butir soal pada Tabel 13 dengan r tabel menunjukkan bahwa seluruh soal yang dikembangkan valid. Hal ini dapat dilihat dari r hitung pada Tabel 13 yang menunjukkan seluruh butir soal memiliki r hitung $> 0,361$ (lebih besar dari 0,361). Artinya r hitung butir soal $> r$ tabel, sehingga seluruh soal yang dikembangkan valid secara konstruk. Tidak ada soal yang gugur sehingga dapat dilakukan analisis butir soal berikutnya.

Hasil uji reliabilitas dapat dilihat pada Tabel 14.

Tabel 14. Hasil Uji Reliabilitas

No Soal	Cronbach's Alpha if Item Deleted	Cronbach's Alpha	Reliabilitas
1	0,907	0,913	Reliabel
2	0,906		Reliabel
3	0,921		Reliabel
4	0,905		Reliabel
5	0,903		Reliabel
6	0,899		Reliabel
7	0,910		Reliabel
8	0,903		Reliabel
9	0,902		Reliabel
10	0,900		Reliabel
11	0,901		Reliabel
12	0,915		Reliabel

Hasil uji pada Tabel 14 diketahui nilai koefisien *Cronbach's Alpha* sebesar 0,913, yang berarti soal reliabel secara keseluruhan karena $r \text{ hitung} > 0,377$ (r tabel) dengan kategori sangat tinggi. Hal ini sesuai dengan kriteria koefisien *Cronbach Alpha* menurut Hadi (1999), jika nilai koefisien *Cronbach Alpha* berkisar antara 0,8-1 maka dikategorikan sangat tinggi. Reliabilitas masing-masing butir soal juga menempati kategori sangat tinggi karena koefisien *Cronbach's Alpha if Item Deleted* (t hitung) pada Tabel 14 bernilai $> 0,8$.

Hasil uji tingkat kesukaran dapat dilihat pada Tabel 15.

Tabel 15. Hasil Uji Kriteria Kesukaran

No Soal	Indeks	Kriteria Kesukaran
1	0,9	Mudah
2	0,85	Mudah
3	0,48	Sedang
4	0,69	Sedang
5	0,65	Sedang
6	0,67	Sedang
7	0,65	Sedang
8	0,85	Mudah
9	0,7	Sedang
10	0,85	Mudah
11	0,62	Sedang
12	0,82	Mudah

Hasil uji coba pada Tabel 15 menunjukkan bahwa soal nomor 1, 2, 8, 10, dan 12 memperoleh kategori mudah. Hal ini

karena indeks kesukaran yang diperoleh $>0,71$. Menurut Arikunto (2013) soal yang memiliki indeks tingkat kesukaran $0,71-1,00$ dikategorikan sebagai soal yang mudah. Artinya soal tidak sulit dan mudah untuk diselesaikan, Sedangkan soal nomor 3, 4, 5, 6, 7, 9, dan 11 memperoleh kategori sedang dikarenakan indeks tingkat kesukarannya berkisar antara $0,31-0,7$. Tidak ada soal yang sukar, karena tidak ada butir soal yang memperoleh indeks kesukaran $<0,3$.

Hasil uji daya pembeda soal dapat dilihat pada Tabel 16.

Tabel 16. Hasil Uji Daya Pembeda Soal

No Soal	Indeks	Kriteria Indeks Daya Pembeda Soal
1	0,4	Cukup
2	0,53	Baik
3	0,33	Cukup
4	0,53	Baik
5	0,6	Baik
6	0,67	Baik
7	0,47	Baik
8	0,53	Baik
9	0,53	Baik
10	0,53	Baik
11	0,6	Baik
12	0,33	Cukup

Hasil uji coba pada Tabel 16 menunjukkan bahwa soal nomor 1, 3, dan 12 mendapatkan kategori cukup. Hal ini karena indeks daya pembeda pada ketiga soal tersebut berkisar $0,3-0,39$. Dimana menurut Crocker dan Algina (1986) soal dengan indeks daya pembeda berkisar $0,3-0,39$ berada pada kategori cukup. Artinya soal-soal tersebut cukup mampu dalam menunjukkan perbedaan pemahaman peserta didik dengan kemampuan yang berbeda-beda. Sedangkan soal nomor 2, 4, 5, 6, 7, 8, 9, 10 mendapatkan kategori baik karena memiliki indeks daya pembeda soal $>0,4$. Soal dengan indeks pembeda $0,4-1,00$

berada pada kategori baik. Artinya soal nomor 2,4, 5, 6, 7, 8, 9, dan 10 dapat menunjukkan dengan baik pemahaman peserta didik dengan kemampuan yang berbeda-beda.

KESIMPULAN

Pengembangan Instrumen Asesmen Berbasis Enam Aspek Pemahaman dalam Kerangka *Understanding by Design* (UbD) pada Materi Termokimia Kelas XI SMA Sederajat dinyatakan valid berdasarkan penilaian oleh validator. Hasil validasi pada aspek konstruksi memperoleh presentase rata-rata skor dari masing-masing validator sebesar 99,7%, 94,34%, dan 95,83% dengan kategori kelayakan valid. Aspek materi memperoleh presentase rata-rata skor dari masing-masing validator sebesar 100% dengan kategori kelayakan valid. Aspek bahasa memperoleh presentase rata-rata skor dari masing-masing validator sebesar 100% dengan kategori kelayakan valid.

Respon pengguna dari 2 orang guru kimia terhadap Instrumen Asesmen Berbasis Enam Aspek Pemahaman dalam Kerangka *Understanding by Design* (UbD) pada Materi Termokimia Kelas XI SMA Sederajat ialah sangat baik. Instrumen asesmen sudah memenuhi aspek keefektifan, bahasa, dan kemanfaatan dengan presentase rata-rata skor sebesar 95,83%. Respon peserta didik terhadap Instrumen Asesmen Berbasis Enam Aspek Pemahaman dalam Kerangka *Understanding by Design* (UbD) pada Materi Termokimia Kelas XI SMA Sederajat ialah sangat baik. Instrumen asesmen sudah memenuhi aspek materi, bahasa, dan ketercukupan waktu dengan presentase rata-rata skor sebesar 91,08%.

Uji karakteristik butir soal memperoleh hasil yang baik pada seluruh aspek yang diujikan. Seluruh soal

dinyatakan valid secara konstruk dengan memperoleh r hitung $>$ r tabel, sehingga tidak ada soal yang gugur. Seluruh soal reliabel dengan kategori sangat tinggi dengan nilai koefisien Cronbach Alpha sebesar 0,913. Tingkat kesukaran instrumen asesmen berada pada kategori mudah dan sedang, dengan 5 soal termasuk kategori mudah dan 7 soal termasuk kategori sedang. Daya pembeda soal berada pada kategori cukup dan baik, dengan 3 soal yang masuk ke kategori cukup dan 9 soal tergolong kedalam kategori baik.

DAFTAR PUSTAKA

- Aiken, Lewis R. (1976). *Psychological Testing And Assessment*. America: CIP
- Anas Sudijono. (2009). *Pengantar Evaluasi Pendidikan*. Jakarta: Rajawali Pers.
- Arda. (2020). Profil Kemampuan Kognitif Mahasiswa Tadris IPA IAIN Palu Pada Mata Kuliah Konsep Dasar IPA . *Jurnal Pembelajaran Matematika dan Sains* , 33-37.
- Ardian, A. (2016). Pengembangan Media Pembelajaran Berupa Buletin dalam Bentuk Buku Saku Untuk Pembelajaran IPA Terpadu. *Jurnal Ilmiah Pendidikan Fisika Albiruni*, 1.
- Arifin, Z., Rahmawati, R., & Nugroho, S. (2022). Analisis kualitas soal ditinjau dari aspek keterbacaan dan validitas isi. *Jurnal Evaluasi Pendidikan*, 13(1), 45–52.
- Arikunto, S. (2008). *Dasar-Dasar Evaluasi Pendidikan Edisi Revisi*. Jakarta: Bumi Aksara.
- Arikunto. (2013). *Dasar-Dasar Evaluasi Pendidikan Edisi II*. Jakarta: Bumi Aksara.
- Arikunto. (2021). *Dasar-Dasar Evaluasi Pendidikan Edisi 3*. Jakarta: Bumi Aksara.
- Arifin, Zainal. (2011). *Penelitian Pendidikan Metode dan Paradigma Baru*. Bandung: PT Remaja Rosdakarya.
- Arumsari, Lusiana Tiara. (2016). *Pengembangan Instrumen Asesmen Keterampilan Proses Sains Pada Materi Teori Tumbukan*. Skripsi. FKIP Universitas Lampung.
- Barlian, U. C. (2022). Implementasi Kurikulum Merdeka Dalam Meningkatkan Mutu Pendidikan. *Journal of Educational and Language Research*, 4-5.
- Basuki, Ismet dan Hariyanto. 2014. *Asesmen Pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Cahyadi, R. A. (2019). Pengembangan Bahan Ajar Berbasis Addie Model. *Halaqa Islamic Education Journal* , 35.
- Crocker dan Algina. (1986). *Introduction to Clasisical and Modern Test Theory*.
- Gloria, R. Y. (2019). Applying Formative Assessment through Understanding by Design (UbD) in the Lecture of Plant Physiology to Improve the Prospective Teacher Education Students' Understanding. *Journal of Turkish Science Education*.
- Hadi, Sutrisno. (1989). *Metodologi Research Jilid I & II*. Yogyakarta: Andi Offset.
- Hamruni. (2012). *Strategi Pembelajaran. Insan Madani*.
- Kemendikbudristek. (2024). *Panduan Pembelajaran dan Asesmen*. Badan Standar, Kurikulum, dan Asesmen Pendidikan (BSKAP).
- Khusna, A. H., & Annisa, A. K. (2021). *Pengembangan Instrumen Pembelajaran Berdasarkan RECCE-Model untuk Mengukur Kemampuan*

- Kognitif Mahasiswa Pada Materi Geometri Transformasi. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 10(1), 443-456.
- Kuntari, F. R. (2019). Understanding by Design (UbD) for the Physics Learning about Parabolic Motion. *Jurnal Penelitian Fisika dan Aplikasinya*, 32-43.
- Latief, M. A. (2020). *Research Methods on Language Learning and Education*. Malang: Universitas Negeri Malang Press.
- Mulyasa, E. 2009. Analisis, Validitas, Reliabilitas, dan Interpretasi Hasil Tes Implementasi Kurikulum 2004. Bandung: Remaja Rosdakarya.
- Nasution, S. (2021). Assesment kurikulum Merdeka belajar di Sekolah Dasar. *Jurnal Mahesa Center*, 135-145.
- Natala, V. E. (2003). Impelemtasi Understanding by Design dalam Kegiatan Pembelajaran: Literatur Review. Seminar Nasional Ilmu Pendidikan ke-2. Bandar Lampung: FKIP Universitas Lampung.
- Nitko, A. J. (1983). *Educational Test and Measurement an Introduction*. New York: Harcourt Brace Jovanovich, Inc.
- Nitko, A. J., & Brookhart, S. M. (2011). *Educational Assessment of Students* (6th ed.). Pearson Education.
- Presiden Republik Indonesia. (2016). Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 21 Tahun 2016 tentang Standar Isi Pendidikan dan Menengah.
- Putri, M. D., Arifin, M. B., & Syahrir, S. (2020). Analisis kualitas instrumen asesmen berdasarkan aspek kebahasaan dan keterbacaan. *Jurnal Pendidikan dan Evaluasi*, 8(2), 110–119.
- Rachmadtullah, R., Syaodih, E., & Sumantri, M. S. (2022). Validitas dan kepraktisan instrumen asesmen formatif berbasis keterampilan abad 21. *Jurnal Inovasi Pendidikan*, 18(1), 45–53.
- Rahmi, S., & Zulela, M. S. (2021). Pengembangan instrumen asesmen berbasis HOTS untuk meningkatkan kemampuan berpikir kritis siswa. *Jurnal Evaluasi Pendidikan*, 12(1), 23–31.
- Siagian, Dolok Putra., Yasthophi, Arif. (2021). Desain dan Uji Coba Media Pembelajaran Berorientasi Everyday Life Phenomena pada Materi Termokimia. *Jurnal Riset Pendidikan Kimia*.
- Sudijono, A. (2015). *Pengantar Statistik Pendidikan*. Jakarta: PT Rajagrafindo Persada.
- Sugiyono. (2017). *Metode Penelitian Kuantitatif Kualitatif Dan R&D*. Bandung: Alfabeta.
- Taufik, A., & Mus'id, R. (2020). Pengembangan Instrumen Penilaian Olahraga Panahan Pada Pemula untuk Guru Penjas Tingkat Sekolah Menengah Pertama, Kecamatan Praya Kabupaten Lombok Tengah. *JIGE: Jurnal Ilmiah Global Education*, 1(1), 58-67. doi:DOI: <https://doi.org/10.55681/jige.v1i1.87>
- Tim Penyusun. (2013). *Permendikbud Nomor 66 Tahun 2013 tentang Standar Penilaian*. Jakarta: Kemdikbud.