

Pengembangan Instrumen Tes Diagnostik Bermuatan Literasi Numerasi Untuk Analisis Pemahaman Konsep Siswa Pada Materi Termokimia

Uswannas Ghaniyul Ilmi ✉, Endang Susilaningsih, Sri Nurhayati, dan Sri Wardani

Jurusan Kimia, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Semarang
Gedung D6 Kampus Sekaran Gunungpati Telp. (024)8508112 Semarang 50229

Info Artikel

Diterima: Nov 2024

Disetujui: Des 2024

Dipublikasikan: Jan 2025

Keywords:

PBL

Pembelajaran Berferensiasi

Kemampuan Berpikir Kritis

Stoikiometri

Abstrak

Tes diagnostik yang tersedia berkonteks sederhana dan belum menyesuaikan dengan kategori soal dalam tes PISA dan tes AKM. Penelitian bertujuan mengembangkan instrumen tes diagnostik bermuatan literasi numerasi untuk analisis pemahaman konsep siswa pada materi Termokimia. Metode penelitian yang digunakan adalah Penelitian Pengembangan dengan model Plomp mencakup lima tahapan. Subjek penelitian dilakukan pada siswa kelas XI SMA Negeri 1 Jepara. Teknik analisis data meliputi kelayakan, kualitas, angket respon serta pemahaman konsep. Instrumen tes layak dari sisi validitas isi dengan skor 60 dari 72 dengan bukti empirik Aiken's V skor 0,89 termasuk valid tinggi, validitas butir soal 22 dari 25 butir valid dan reliabilitas model Rasch dengan skor person reliability 0,84, item reliability 0,89, alpha cronbach 0,87, person separation 3, dan item separation 4. Instrumen tes mengungkap profil pemahaman konsep siswa pada materi Termokimia dengan persentase sebesar 42,5% siswa memahami konsep, 14,9% tidak memahami konsep, 9,1% miskonsepsi positif, 9,4% miskonsepsi negatif, serta 24,10% miskonsepsi. Simpulan diperoleh produk instrumen tes diagnostik bermuatan literasi numerasi yang layak, berkualitas, dan memberi respon positif.

Abstract

The diagnostic tests are simple in context and have not been adjusted to question categories in PISA and AKM tests. The study aims to develop diagnostic test instrument containing numeracy literacy for analysis of students' conceptual understanding in Thermochemistry. The research method used is Development Research with Plomp model covering five stages. The research subjects were grade XI students of SMA Negeri 1 Jepara. Data analysis techniques include validation, quality, response questionnaire and students' conceptual understanding. The research results showed test instrument was declared feasible in terms of aspects content validity score 60 from 72 with empirical evidence from Aiken's V score 0.89 including high valid, validity of test items 22 of 25 are valid and reliability Rasch model with person reliability 0.84, item reliability 0.89, alpha cronbach 0.87, person separation 3, and item separation 4. The test instrument can reveal profile of students' conceptual understanding in Thermochemistry with percentage of 42.5% understanding concept, 14.9% not understanding concept, 9.1% positive misconceptions, 9.4% negative misconceptions and 24.10% high misconceptions. The conclusion obtained is diagnostic test instrument product containing numeracy literacy that is feasible, high quality, and provides positive response..

©2022 Universitas Negeri Semarang

✉ Alamat korespondensi:

Gedung D6 Lantai 2 Kampus Sekaran, Gunungpati, Semarang 50229

E-mail: uswannasilmi27@students.unnes.ac.id

p-ISSN 1979-0503

e-ISSN 2503-1244

PENDAHULUAN

Ali (2019) menjelaskan bahwa masalah dunia pendidikan dapat bersumber dari berbagai faktor, salah satunya yaitu rendahnya pemahaman konsep bahkan terdapat miskonsepsi siswa terhadap suatu materi pelajaran. Termokimia sebagai salah satu materi esensial pada pembelajaran kimia siswa SMA tidak luput pula dari masalah pemahaman konsep. Materi Termokimia sangat penting untuk dipelajari dan dikuasai sebab berkesinambungan terhadap materi lain. Skema peristiwa kimia, transformasi wujud, dan pembuatan larutan merupakan beberapa materi yang menjadikan termokimia sebagai konsep prasyarat karena melibatkan data perubahan energi (Roghdah *et al.*, 2021).

Pemahaman konsep siswa terhadap suatu materi dapat diukur dan diketahui melalui tes diagnostik. Tes diagnostik dapat digunakan untuk mengukur penilaian berbagai ranah kognitif yang mengetes kemampuan siswa (Ramadhan *et al.*, 2022). Dalam penelitian ini digunakan bentuk *three tier* atau tiga tingkat. *Three tier* umumnya berupa tingkat pertama yang mengandung pilihan jawaban dari pertanyaan inti terkait kompetensi yang diukur, level kedua mengandung alasan saat pemilihan jawaban pada level pertama, serta level ketiga mengandung pilihan atas kepercayaan siswa saat memberikan pilihan jawaban pada level pertama dan kedua (Lestari *et al.*, 2021). Ada beberapa keunggulan atas penggunaan tes *three tier*, diantaranya menganalisis kesalahpahaman konsep yang dialami siswa lebih luas, memastikan poin - poin materi yang membutuhkan penekanan tertentu selama pembelajaran, dan membantu menurunkan kesalahpahaman konsep siswa melalui rancangan pembelajaran yang berkualitas (Elvia *et al.*, 2021).

Tes diagnostik yang selama ini beredar hanya berkonteks sederhana dan belum menyesuaikan dengan kategori soal - soal yang termuat dalam tes PISA. Muatan soal dalam tes PISA tidak sekadar memberi kepastian siswa memperoleh ilmu, melainkan pula memeriksa sejauh mana siswa mampu mengeksplorasi dari sesuatu yang sudah dipelajari (Rastuti & Prahmana, 2021). Pemerintah Indonesia telah mengadaptasi soal tes PISA dengan nama AKM (Asesmen Kompetensi Minimum) yaitu salah satu bentuk penilaian kemampuan minimum atau paling pokok yang wajib dikuasai oleh siswa pada jenjang terkhusus meliputi lingkup literasi membaca dan numerasi (Pusat Pembelajaran dan Asesmen, 2020). Literasi membaca tidak sekadar mengukur kecakapan membaca secara harfiah tanpa mengerti kandungan atau makna pada bacaan tersebut, tetapi kecakapan dalam pemahaman konsep bacaan (Sani, 2021). Penalaran matematis berperan penting untuk mendukung penilaian dalam aspek kontribusi pada pengembangan keterampilan abad-21 (OECD, 2023). Unsur literasi dapat ditilik dari keberadaan teks atau bacaan yang akan dipakai sebagai stimulus dalam pengembangan soal. Beberapa materi pembelajaran kimia seperti termokimia mengandung penerapan numerasi karena mengandung banyak konsep perhitungan dan penerapan rumus yang menjadi dasar peningkatan kemampuan numerasi (Rokhim *et al.*, 2022). Penerapan numerasi diperkuat oleh pendapat Nirmalasari *et al.* (2022) mengenai adanya unsur numerasi dapat ditandai melalui penerapan angka, simbol dan informasi matematis yang disajikan ke dalam teks maupun butir soal.

Hasil observasi dan diskusi bersama guru yang dikerjakan di SMA Negeri 1 Jepara pada topik mapel kimia kelas XI memperlihatkan pengerjaan tes dengan kategori kompleks mencakup tes diagnostik, literasi dan numerasi masih jarang dilakukan. Padahal soal tes berkategori kompleks sangat bermanfaat dalam membantu menemukan kelemahan – kelemahan siswa pada suatu konsep dalam suatu materi pelajaran. Ketercapaian siswa dalam tujuan pembelajaran umumnya diukur memakai suatu tes yang biasanya dilaksanakan di akhir sesi pembelajaran dalam satu keseluruhan bab materi tertentu maka dibutuhkan sesuatu seperti asesmen diagnostik bermuatan literasi numerasi yang dapat mengukur pemahaman konsep siswa.

METODE

Model penelitian pengembangan merujuk pada pengembangan Plomp dengan tujuan guna mengembangkan instrumen tes diagnostik bermuatan literasi numerasi untuk menganalisis pemahaman konsep siswa pada materi Termokimia. Terdapat lima tahapan yang dilalui dalam model pengembangan Plomp yaitu (1) investigasi awal, (2) desain, (3) realisasi/konstruksi, (4) tes, evaluasi dan revisi, serta (5) implementasi (Plomp, 2013). Langkah - langkah pada tahap investigasi awal mencakup (1) mengumpulkan data, (2) menganalisis data, (3) mendefinisikan masalah, dan (4) merencanakan tindak lanjut. Tahap desain mengandung dua langkah yaitu perancangan instrumen tes diagnostik dan angket respon. Tahap realisasi terdiri atas dua langkah yakni pembuatan draft instrumen tes dan pembuatan draft angket respon. Tahap tes, evaluasi dan revisi terdiri atas dua langkah utama yaitu uji validasi ahli dan uji skala kecil. Setiap masing - masing dari dua langkah mengandung pengujian atau tes, evaluasi berdasarkan hasil pengujian, dan revisi berdasarkan data hasil evaluasi. Terdapat dua langkah dalam tahap implementasi yakni pelaksanaan uji skala besar dan analisis data hasil uji skala besar.

Subyek utama dilakukan terhadap siswa kelas XI SMA Negeri 1 Jepara yang telah mempelajari materi Termokimia dengan dengan jumlah dan karakteristik siswa yang berbeda untuk setiap pengujian. Uji coba skala kecil dilaksanakan pada kelas XI 10 berjumlah 20 siswa sementara uji skala besar pada kelas XI 1 dan XI 2 masing – masing 20 siswa sehingga total berjumlah 40 siswa. Teknik pengumpulan data meliputi studi literatur, studi lapangan, metode tes, lembar validasi dan metode angket. Teknik analisis data yang diterapkan mencakup (1) analisis kelayakan instrumen tes, (2) analisis kualitas instrumen tes menggunakan permodelan Rasch berbantuan software Ministep, (3) analisis pemahaman konsep siswa pada materi Termokimia, serta (4) analisis angket respon siswa terhadap instrumen tes diagnostik bermuatan literasi numerasi. Analisis kelayakan instrumen tes ditinjau dari keempat aspek yaitu (1) validitas isi dari para validator ahli, (2) analisis kelayakan secara bukti empirik (skor Aiken's V), (3) analisis validitas butir soal menggunakan model Rasch, serta (4) analisis reliabilitas instrumen tes menggunakan model Rasch. Analisis kualitas instrumen tes menggunakan model Rasch terdiri atas (1) analisis reliabilitas instrumen, (2) analisis kesesuaian butir soal, (3) analisis tingkat kesulitan butir soal, (4) analisis tingkat kemampuan siswa, (5) analisis peta wright, (6) analisis tingkat kesesuaian siswa, dan (7) analisis scalogram. Analisis pemahaman konsep siswa pada materi termokimia dilakukan terhadap siswa yang telah mengikuti uji skala besar. Analisis pemahaman konsep siswa pada materi Termokimia mencakup (1) analisis kategori pemahaman konsep siswa secara klasikal, (2) analisis kategori pemahaman konsep siswa pada setiap indikator pemahaman konsep, serta (3) analisis kategori pemahaman konsep siswa pada setiap butir soal.

PEMBAHASAN

Perancangan Instrumen Tes Diagnostik Bermuatan Literasi Numerasi

Instrumen tes mengandung 25 butir soal bertingkat tiga yakni level pertama berisi jawaban pertanyaan inti, level kedua berisi sebab pemilihan jawaban pertanyaan inti, serta level ketiga berisi tingkat kepercayaan memilih jawaban pada tingkat pertama dan kedua menggunakan Certainty of Responses Index (CRI) dengan skala 0 sampai 5. Setiap butir soal mengandung stimulus yang dapat berbeda atau sama tetapi pertanyaan yang diberikan berbeda - beda sesuai tujuan pembelajaran yang ingin dicapai. Pernyataan jawaban dalam tiap butir soal disesuaikan dengan bentuk soal AKM baik tingkat pertama maupun kedua. Bentuk soal yang dapat digunakan antara lain pilihan ganda biasa dengan hanya satu jawaban benar, pilihan ganda kompleks dengan lebih dari satu jawaban tepat, menjodohkan, jawaban atau isian singkat, dan esai atau uraian. Soal juga disesuaikan dengan indikator pemahaman konsep dan indikator muatan literasi numerasi.

Validasi Isi Menurut Validator Ahli

Kesimpulan yang didapatkan menurut hasil pengolahan data skor penilaian validasi adalah instrumen tes diagnostik bermuatan literasi numerasi yang dikembangkan masuk dalam kategori "sangat valid" bagi validator 1 dan 2 serta "valid" bagi validator 3. Apabila dihitung rata - rata skornya maka diperoleh skor sebanyak 60 dari ketercapaian keseluruhan skor 72 sehingga instrumen tes diagnostik bermuatan literasi numerasi memenuhi kategori "sangat valid" dari sisi validitas menurut para ahli.

Instrumen tes juga perlu diuji dari sisi bukti empirik, salah satunya menggunakan rumus Aiken's V. Perhitungan nilai Aiken's V dilakukan terhadap data hasil jawaban siswa uji coba skala kecil. Setelah dilakukan perhitungan diperoleh skor keseluruhan Aiken's V sebesar 0,89 dan masuk kategori validitas tinggi. Kesimpulan yang didapatkan menurut hasil penilaian validasi dan pengolahan data nilai Aiken's V adalah instrumen tes dinyatakan layak dari sisi validitas isi dengan skor 60 dari skor keseluruhan 72 dengan bukti empirik Aiken's V skor 0,89 termasuk kategori valid tinggi.

Validitas Butir Soal Secara Model Rasch (*Item Fit Order*)

Tingkat kesesuaian butir soal (*item fit order*) dipakai sebagai landasan guna penentuan validitas pada tiap - tiap butir soal. Analisis validitas item soal memperlihatkan adanya 22 soal fit dan 3 soal berkategori misfit (soal nomor 6,9,dan 11) pada pengujian skala kecil dan 25 soal berkategori fit pada pengujian skala besar. Contoh cuplikan analisis validitas butir soal misalnya pada pengujian skala kecil dapat dilihat melalui Tabel 1.

Tabel 1. Cuplikan Hasil Validitas Butir Soal

Butir Soal	<i>Outfit</i> MNSQ	<i>Outfit</i> ZSTD	<i>PT Measure Corr</i>	Keterangan
Soal 5	1,03	0,2	0,49	Sangat Valid
Soal 21	0,76	0,08	0,40	Valid
Soal 3	0,49	-0,26	0,46	Valid
Soal 1	0,46	-0,2	0,37	Valid
Soal 6	4,71	2,82	-0,25	Tidak Valid

Tabel 1 menunjukkan bahwa soal 5 dinyatakan fit dan masuk kategori sangat valid sebab nilai *Outfit* MNSQ, *Outfit* ZSTD, dan *PT Measure Corr* telah memenuhi ketentuan. Soal 21 masih dianggap fit dan masuk kategori valid meskipun nilai *PT Measure Corr* tidak sesuai ketentuan karena nilai *Outfit* ZSTD dan *Outfit* MNSQ masih sesuai ketentuan. Soal 3 masih dianggap fit dan masuk kategori valid meskipun nilai *Outfit* MNSQ tidak sesuai ketentuan karena nilai *Outfit* ZSTD dan *PT Measure Corr* masih sesuai ketentuan. Soal 1 masih dianggap fit dan masuk kategori valid meskipun nilai *Outfit* MNSQ dan *PT Measure Corr* tidak sesuai ketentuan karena nilai *Outfit* ZSTD masih sesuai ketentuan. Soal 6 dinyatakan misfit dan masuk kategori tidak valid sebab tidak memenuhi ketentuan nilai *Outfit* MNSQ, *Outfit* ZSTD, dan *PT Measure Corr*.

Berdasarkan output reliabilitas tes melalui aplikasi Ministep diperoleh hasil ringkas pengolahan data skor dari analisis reliabilitas instrumen tes dalam dua tahapan pengujian ditunjukkan melalui Tabel 2.

Tabel 2. Hasil Reliabilitas Instrumen Tes Model Rasch

Analisis Reliabilitas	Skor Uji Coba Skala Kecil	Skor Uji Skala Besar	Kategori
<i>Person Reliability</i>	0,83 dan 0,85	0,84 dan 0,85	Bagus
<i>Item Reliability</i>	0,85 dan 0,86	0,89 dan 0,90	Bagus
<i>Alpha Cronbach</i>	0,85	0,87	Bagus Sekali
<i>Person Separation</i>	3	3	Baik
<i>Item Separation</i>	4	4	Baik

Kesimpulan yang didapatkan menurut hasil skor reliabilitas tersebut adalah konsistensi dari jawaban siswa masuk kategori bagus dan kualitas butir-butir pernyataan pada instrumen tes juga masuk kategori bagus berdasarkan kriteria Sanjaya *et al* (2020). Kelima kriteria reliabilitas instrumen mencakup *item reliability*, *person reliability*, *alpha cronbach*, nilai separasi individu (*person separation*) dan separasi butir (*item separation*) masuk kategori bagus sehingga dapat dinyatakan bahwa instrumen tes diagnostik bermuatan literasi numerasi reliabel.

Tingkat Kesukaran Butir Soal (*Item Measure*)

Ketentuan pengukuran level kesukaran butir soal berpatokan dari nilai deviasi standar (SD) yang termuat dalam output tabel item measure bagian terbawah dengan nilai angka sebesar +1,67 pada pengujian skala kecil dan +1,32 pada pengujian skala besar. Setelah nilai standar deviasi diketahui dilanjutkan dengan melakukan pengelompokkan soal Berdasarkan hasil pengelompokkan soal diperoleh jumlah soal masing - masing kategori sesuai tingkat kesulitan soal pada uji coba skala kecil yakni sangat sulit terdapat 7 butir soal, sulit 4 soal, sedang 10 soal dan mudah 4 soal. Sementara, jumlah soal masing - masing kategori sesuai tingkat kesulitan soal pada uji skala besar yakni sangat sulit terdapat 4 butir soal, sulit 11 soal, sedang 7 soal dan mudah 3 soal.

Tingkat Kemampuan Siswa (*Person Measure*)

Ketentuan pengukuran tingkat kemampuan siswa berpatokan dari nilai deviasi standar (SD) yang termuat dalam output tabel person measure bagian terbawah dengan nilai angka sebesar +1,38 pada pengujian skala kecil dan +1,37 pada pengujian skala besar. Setelah nilai standar deviasi diketahui dilanjutkan dengan melakukan pengelompokkan siswa. Berdasarkan hasil pengelompokkan siswa diperoleh jumlah siswa masing - masing kategori sesuai tingkat abilitas siswa pada uji coba skala kecil yakni sangat tinggi terdapat 3 siswa, tinggi 6 siswa, sedang 8 siswa dan rendah 3 siswa. Sementara, jumlah siswa masing - masing kategori sesuai tingkat abilitas siswa pada uji skala besar yakni sangat tinggi terdapat 2 siswa, tinggi 12 siswa, sedang 14 siswa dan rendah 12 siswa.

Analisis Peta Wright

Peta Wright memberikan gambaran mengenai penyebaran abilitas siswa selaku person dan penyebaran tingkat kesukaran soal selaku item pada ukuran skala yang identik (Susilaningasih *et al.*, 2021). Berdasarkan output Variable (Wright) Map sisi sebelah kanan pada uji skala besar, butir S23 dan S9 masuk dalam kelompok paling sulit dijawab benar sebab butir tersebut mempunyai skor logit terbesar yaitu lebih dari +2 logit (outlier/data ekstrim). Meskipun begitu, butir S23 dan S9 masih tergolong normal sebab skor logitnya masih terletak dalam area diantara dua sisi batas deviasi standar (T). Berbeda dengan butir S4 dan S7 yang masuk dalam kelompok paling mudah dijawab benar oleh siswa sebab butir tersebut mempunyai skor logit terkecil yakni kurang dari -2 logit dan butir ini tampak berada di luar area dua standar deviasi (T). Maksudnya, kemungkinan siswa menjawab dengan benar pada butir S23 besar sekali.

Tingkat Kesesuaian Siswa (*Person Fit Order*)

Hasil analisis kesesuaian respon siswa (*person fit order*) dalam dua tahapan uji menunjukkan terdapat satu siswa pada pengujian skala kecil yang memiliki pola respon tidak sesuai (misfit) yaitu siswa berkode 10LK sedangkan pengujian skala besar terdapat dua siswa yang memiliki pola respon tidak sesuai (misfit)

yaitu siswa berkode 14PK dan 05PK. Pola respon misfit yang ditemukan hendaknya dianalisis lebih mendalam pada bagian analisis scalogram.

Scalogram

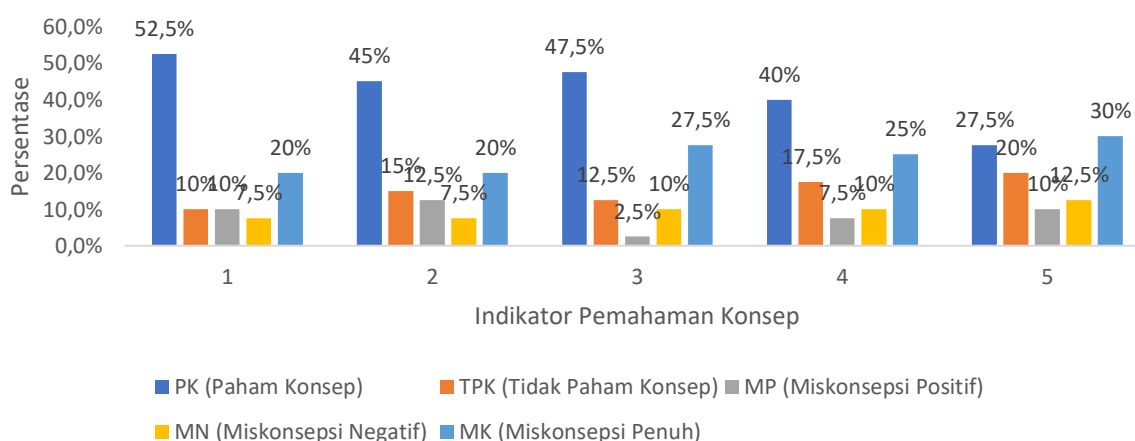
Scalogram memperlihatkan bahwa siswa berkode 10LK bisa memberikan jawaban benar untuk soal nomor 5, 9, 6, 14, 16, dan 17. Ternyata skor logit keenam soal tersebut lebih tinggi dibandingkan skor logit kemampuan siswa berkode 10LK sehingga dapat mengindikasikan terkait adanya jawaban yang mengasal atau hasil tebakan. Kemudian, siswa berkode 10LK juga terindikasi kurang teliti saat memberikan jawaban. Keadaan tersebut dapat diidentifikasi dari kesalahan jawaban pada beberapa soal yang tergolong kategori mudah yakni soal nomor 3, 7 dan 24.

Profil Pemahaman Konsep Siswa

Hasil persentase rata - rata kategori pemahaman konsep keseluruhan siswa yang mengikuti uji skala besar mencakup 42,5% siswa berkategori paham konsep (PK), 14,9% berkategori tidak paham konsep (TPK), 9,1% berkategori miskonsepsi positif (MP), 9,4% berkategori miskonsepsi negatif (MN), serta 24,10% berkategori miskonsepsi (MK).

Apabila ditinjau dari hasil profil pemahaman konsep siswa secara klasikal terhadap pemahaman konsep siswa pada materi Termokimia maka pemahaman konsep siswa kelas XI SMA Negeri 1 Jepara berada pada taraf sedang, terdapat 42,5% siswa berkategori paham konsep. Namun saat dibandingkan dengan keseluruhan siswa tidak paham konsep dan miskonsepsi maka jumlahnya lebih rendah. Sebagian besar siswa masih belum memahami konsep bahkan ada yang mengalami miskonsepsi pada materi termokimia. Hasil yang diperoleh hampir serupa dengan penelitian dari Roghdah *et al* (2021) yang melakukan analisis pemahaman konsep siswa pada materi termokimia. Lestari *et al* (2021) mengungkapkan penyebab rendahnya pemahaman konsep dan miskonsepsi dapat dipengaruhi oleh aspek luar dan dalam. Aspek luar seperti kondisi yang minim ketenangan saat pembelajaran materi termokimia. Aspek dalam mencakup (1) sedikitnya afinitas siswa terhadap kegiatan belajar kimia, (2) proses pembelajaran yang tidak tepat, yang berakibat rendahnya daya ingat siswa terhadap materi yang telah diterima, (3) sedikitnya ketangguhan siswa dalam memberikan pertanyaan, (4) semangat menuntut ilmu yang sedikit, serta (5) kemampuan siswa yang kurang dalam menghubungkan konsep.

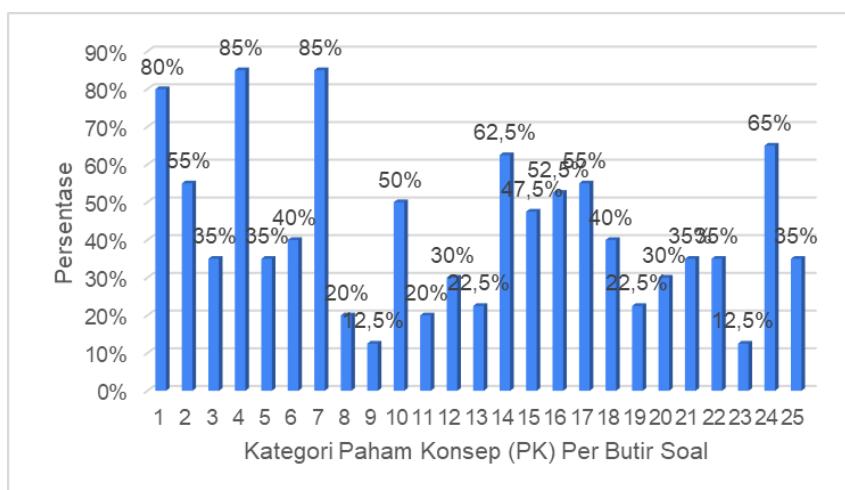
Terdapat lima indikator pemahaman konsep yang digunakan pada penelitian pengembangan instrumen tes diagnostik bermuatan literasi numerasi yaitu (1) mengklasifikasikan objek - objek menurut sifat tertentu/IPK-1, (2) memberi contoh dan non contoh dari konsep/IPK-2, (3) menyajikan konsep dalam beragam bentuk representasi matematis/IPK-3, (4) menggunakan, memanfaatkan, dan memilih prosedur atau operasi tertentu/IPK-4, serta (5) mengaplikasikan konsep atau algoritma pemecahan masalah/IPK-5. Penggambaran mengenai rekapitulasi kategori pemahaman konsep yang terjadi pada siswa uji skala besar dari setiap indikator pemahaman konsep siswa khususnya pada kategori Paham Konsep (PK) ditunjukkan melalui Gambar 1.



Gambar 1. Persentase Kategori Pemahaman Konsep Per Indikator Pemahaman Konsep

Indikator pemahaman konsep dengan mayoritas siswa tidak memahami konsep dan mengalami miskonsepsi pada indikator pemahaman konsep kelima. Banyaknya siswa yang tidak memenuhi indikator mengaplikasikan konsep menunjukkan soal - soal yang termuat berkategori tingkat kesulitan tinggi. Suleman *et al* (2023) menyatakan mayoritas siswa belum memahami topik bahasan Termokimia khususnya apabila siswa diberikan beberapa soal yang masuk kategori level pemahaman tinggi.

Hasil rekapitulasi kategori paham konsep (PK) yang terjadi pada siswa uji skala besar dari setiap butir soal ditunjukkan melalui Gambar 2.



Gambar 2. Persentase Kategori Pemahaman Konsep Per Butir Soal

Butir soal yang rendah dipahami konsepnya oleh siswa adalah butir soal nomor 23 yang membahas materi energi pembakaran bahan bakar dengan persentase sebesar 12,5%. Soal nomor 23 meminta siswa untuk menentukan urutan bahan bakar dengan energi pembakaran satu gram dari yang paling besar. Salah satu siswa yang mengalami miskonsepsi pada butir soal nomor 23 menganggap semakin besar massa molar relatif suatu bahan bakar maka semakin besar energi yang dihasilkan dari pembakaran satu gram bahan bakar tersebut. Konsep energi pembakaran bahan bakar pada soal nomor 23 yang benar adalah semakin besar harga perubahan entalpi per massa molar relatif suatu bahan bakar maka semakin besar energi yang dihasilkan dari pembakaran satu gram bahan bakar tersebut. Salah satu jawaban siswa pada soal 23 yang telah diketahui terjadi miskonsepsi ditunjukkan melalui Gambar 3.

✗ Alasan dari jawaban soal nomor 23 *

0/1

- ☐ a. Semakin banyak atom karbon suatu bahan bakar maka semakin besar energi yang dihasilkan dari pembakaran satu gram bahan bakar tersebut
- ☐ b. Semakin kecil jumlah mol suatu bahan bakar maka semakin besar energi yang dihasilkan dari pembakaran satu gram bahan bakar tersebut
- ☐ c. Semakin besar perubahan entalpi suatu bahan bakar maka semakin besar energi yang dihasilkan dari pembakaran satu gram bahan bakar tersebut
- ☒ d. Semakin besar Mr suatu bahan bakar maka semakin besar energi yang dihasilkan dari pembakaran satu gram bahan bakar tersebut ✗
- ☐ e. Semakin besar harga $-\Delta H/Mr$ suatu bahan bakar maka semakin besar energi yang dihasilkan dari pembakaran satu gram bahan bakar tersebut

Jawaban yang benar

- ☒ e. Semakin besar harga $-\Delta H/Mr$ suatu bahan bakar maka semakin besar energi yang dihasilkan dari pembakaran satu gram bahan bakar tersebut

Gambar 3. Pola Jawaban Soal 23 yang Terjadi Miskonsepsi

Terdapat siswa lain yang menjawab pertanyaan tingkat kedua atau alasan dengan benar tetapi memilih pertanyaan yang kurang tepat serta yakin terhadap pilihannya. Hal tersebut dapat mengindikasikan bahwa siswa kurang cermat atau teliti terutama saat melakukan perhitungan sehingga memberikan jawaban yang keliru (Lestari *et al.*, 2021). Rendahnya pemahaman konsep pada butir soal yang membahas sub materi energi bahan bakar hampir identik dengan perolehan penelitian yang dikerjakan oleh Roghdah *et al* (2021) yaitu besar rerata pengetahuan siswa mengenai konsep menganalogikan energi panas pembakaran suatu

bahan bakar senilai 14% dan berada pada tingkat rendah. Berdasarkan analisis butir soal berdasarkan pemahaman konsepnya, maka guru perlu memberikan pembelajaran tambahan atau ulang mengenai sub materi energi bahan bakar karena butir soal yang berisi sub materi energi bahan bakar paling rendah dipahami konsepnya oleh siswa.

SIMPULAN

Instrumen tes diagnostik bermuatan literasi numerasi yang dikembangkan sudah disebut layak berlandaskan (1) Validitas isi dengan rerata skor sebesar 60 dari nilai keseluruhan 72 dan masuk kategori sangat valid dengan bukti empirik Aiken's V skor 0,89 termasuk kategori valid tinggi, (2) Validitas butir soal menggunakan model Rasch pada pengujian skala kecil dengan hasil 22 butir soal masuk kategori valid dari total 25 butir soal, (3) Reliabilitas instrumen tes menggunakan model Rasch pada pengujian skala kecil dengan skor 0,85 dan masuk kategori bagus sekali. Instrumen tes diagnostik masuk kategori berkualitas ditinjau dari sisi validitas dimana seluruh butir soal masuk kategori fit dan reliabilitas sebesar 0,87 menggunakan model Rasch dan masuk kategori bagus sekali. Instrumen tes dapat mengungkap profil pemahaman konsep siswa pada materi Termokimia dengan persentase sebesar 42,5% siswa memahami konsep (PK), 14,9% tidak memahami konsep (TPK), 9,1% miskonsepsi positif (MP), 9,4% miskonsepsi negatif (MN), serta 24,10% miskonsepsi (MK).

DAFTAR PUSTAKA

- Ali, M. (2019). Analisis Miskonsepsi Siswa Berdasarkan Gender dalam Pembelajaran Fisika dengan Menggunakan Tes Diagnostik Two-Tier di Kotabaru. *CENDEKIA: Jurnal Ilmiah Pendidikan*, 7(1), 59–66. <https://doi.org/10.33659/cip.v7i1.120>
- Elvia, R., Rohiat, S., & Menda Ginting, S. (2021). Identifikasi Miskonsepsi Mahasiswa Pada Pembelajaran Daring Matematika Kimia Melalui Tes Diagnostik Three Tier Multiple Choice. *Hydrogen : Jurnal Kependidikan Kimia*, 8(2), 84–96.
- Lestari, A. E., Harjito, Susilaningih, E., & Wijayati, N. (2021). Analisis Miskonsepsi Menggunakan Tes Diagnostik Three-Tier Multiple Choice Pada Materi Stoikiometri. *Jurnal Inovasi Pendidikan Kimia*, 15(2), 2824–2830.
- Nirmalasari, M. A. Y., Oktavius Yoseph Tuta Mago, & Irene Lete Manuk. (2022). Validitas Instrumen Soal Literasi Numerasi Kimia Hidrokarbon dalam Integrasinya dengan Isu Sosiosaintifik Lokal Sikka. *JURNAL PENDIDIKAN MIPA*, 12(4), 1004–1011. <https://doi.org/10.37630/jpm.v12i4.725>
- OECD. (2023). PISA 2022 Assessment and Analytical Framework. *OECD Publishing*. <https://doi.org/10.1787/dfc0bf9c-en>
- Plomp, T. (2013). *Educational Design Research*. Enschede : University of Twente.
- Pusat Pembelajaran dan Asesmen. (2020). *Desain Pengembangan Soal AKM*. Kemendikbud.
- Ramadhan, A. H., Fadillah, H., Khaliza, R., & Nasution, I. (2022). Penerapan Kurikulum Darurat sebagai Strategi Pendidikan dalam Kondisi Pandemic Covid-19. *Edumaspul: Jurnal Pendidikan*, 6(1), 401–407. <https://doi.org/10.33487/edumaspul.v6i1.2383>
- Rastuti, M., & Prahmana, R. C. I. (2021). The Programme for International Student Assessment Research in Indonesia. *Jurnal Elemen*, 7(2), 232–253. <https://doi.org/10.29408/jel.v7i2.3289>
- Roghdah, S. J., Zammi, M., & Mardhiya, J. (2021). Development of Four-Tier Multiple Choice Diagnostic Test to Determine Students' Concept Understanding Level On Thermochemical Material Pengembangan Four-Tier Multiple Choice Diagnostic Test Untuk Mengetahui Tingkat Pemahaman Konsep Peserta Didik Pada Materi Termokimia. *Phenomenon*, 11(1), 57–74.
- Rokhim, D. A., Tyas, F. K., Rahayu, S., & Habiddin, H. (2022). Perspektif Siswa Dan Guru Dalam Pelaksanaan AKM (Asesmen Kompetensi Minimum) Pada Mata Pelajaran Kimia. *Administrasi Dan Manajemen Pendidikan*, 5(1), 46–52. <https://doi.org/10.17977/um027v5i12022p46>
- Sani, A. R. (2021). *Pembelajaran Berorientasi AKM*. Bumi Aksara.
- Sanjaya, M. R., Kurniawan, D., Saputra, A., Sari, N., & Destriani, R. (2020). Software Design Android (GIS) for Tourist Destination Seekers in Palembang City Using Rasch Model Measurements. *Proceedings of the Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019)*, 172, 505–511. <https://doi.org/10.2991/aisr.k.200424.078>

Suleman, N., Lukum, A., Rauf, N., Paputungan, M., Alio, L., & Sukamto, K. (2023). Identifikasi Pemahaman Konsep Siswa pada Materi Termokimia Menggunakan Tes Diagnostik Three-Tier Multiple Choice. *Jambura Journal of Educational Chemistry*, 5(2), 122–129.

<https://doi.org/10.34312/jjec.v5i2.13255>

Susilaningsih, E., Nuswowati, M., Aprilia, N., & Luthfiyah, A. (2021). Dissemination of test instruments as product of the development research to measure the problem-solving ability of class X students by online in the pandemic period. *Journal of Physics: Conference Series*, 1918(3).

<https://doi.org/10.1088/1742-6596/1918/3/032024>