



PENGEMBANGAN INSTRUMEN TES LITERASI KIMIA PADA MATERI SISTEM PERIODIK UNSUR SMA/MA

Siti Nurfaizah Ariani¹, Yerimadesi^{2*}, Fauzana Gazali³, Faizah Qurrata Aini⁴

^{1,2,3,4}Program Studi Pendidikan Kimia, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Padang, Padang, Sumatera Barat, 25171, Indonesia

*E-mail: yeri@fmipa.unp.ac.id

Received: January 20, 2026; Accepted: February 28, 2026; Published: February 28, 2026

Abstract

Chemical literacy is an essential competency in chemistry learning as it enables students to understand concepts, interpret scientific phenomena, and make science-based decisions. Therefore, an appropriate evaluation instrument is needed to accurately measure chemical literacy, particularly in the topic of the Periodic Table of Elements (PTE) at the senior high school level (SMA/MA). This study aimed to develop a chemical literacy test instrument on the PTE topic and to analyze the quality of the developed instrument. This study employed a Research and Development (R&D) approach using the Rasch model. The instrument developed was a chemical literacy test. Content validity was analyzed using the Content Validity Ratio (CVR), while empirical validity, reliability, item difficulty, and item discrimination were analyzed using Rasch modeling with the assistance of Ministep software. The results showed that the CVR values ranged from 0.85 to 0.90, indicating that all items were valid in terms of content. Rasch analysis indicated that all test items fit the model and were acceptable. The person reliability value was 0.61, item reliability was 0.94, and Cronbach's Alpha was 0.61. Item difficulty indices ranged from 1.74 to -2.01 logits, with most items classified as moderate, and the item discrimination was categorized as high. Based on these findings, the chemical literacy test instrument on the PTE topic demonstrates good measurement quality and can be used as an assessment tool in chemistry learning at the SMA/MA level.

Keywords : Chemical literacy, Periodic table of elements, Rasch model, Test instrument

Abstrak

Literasi kimia merupakan kompetensi penting dalam pembelajaran kimia karena memungkinkan peserta didik memahami konsep, menafsirkan fenomena ilmiah, serta mengambil keputusan berbasis sains. Oleh karena itu, diperlukan instrumen evaluasi yang mampu mengukur literasi kimia secara tepat, khususnya pada materi Sistem Periodik Unsur (SPU) di tingkat SMA/MA. Penelitian ini bertujuan untuk mengembangkan instrumen tes literasi kimia pada materi SPU serta menganalisis kualitas instrumen yang dihasilkan. Penelitian ini menggunakan pendekatan Research and Development (R&D) dengan penerapan model Rasch. Instrumen yang dikembangkan berupa tes literasi kimia. Validitas isi dianalisis menggunakan Content Validity Ratio (CVR), sedangkan validitas empiris, reliabilitas, tingkat kesukaran, dan daya beda butir dianalisis menggunakan pemodelan Rasch dengan bantuan perangkat lunak Ministep. Hasil analisis menunjukkan nilai CVR berada pada rentang 0,85–0,90 sehingga seluruh butir dinyatakan valid secara isi. Analisis Rasch menunjukkan seluruh butir soal sesuai dengan

model dan dapat diterima. Nilai reliabilitas person sebesar 0,61, reliabilitas item sebesar 0,94, serta Cronbach's Alpha sebesar 0,61. Indeks kesukaran butir berada pada rentang 1,74 hingga -2,01 logit dengan mayoritas butir berkategori sedang, serta daya beda butir termasuk kategori tinggi. Berdasarkan hasil tersebut, instrumen tes literasi kimia pada materi SPU memiliki kualitas pengukuran yang baik dan dapat digunakan sebagai alat evaluasi pembelajaran kimia di SMA/MA.

Keywords : Instrumen tes, Literasi kimia, Pemodelan Rasch, Sistem periodik unsur

Pendahuluan

Pendidikan sains memiliki peran penting dalam membekali peserta didik agar dapat berpikir kritis, mengambil keputusan dengan berdasarkan bukti, serta berkontribusi aktif akan isu-isu sosial sebagai warga negara. Terutama dalam perkembangan abad ke-21, di mana Perkembangan ilmu pengetahuan dan teknologi semakin berlangsung dengan pesat, sehingga literasi sains menjadi kompetensi yang sangat penting dalam pendidikan sains. Menurut *Programme for International Student Assessment (PISA) 2022*, literasi sains merujuk pada kemampuan untuk menjelaskan fenomena, merencanakan serta menilai penyelidikan ilmiah, dan menganalisis data serta bukti dengan menggunakan penalaran ilmiah. Literasi ini tidak hanya menuntut pemahaman terhadap konten-konten sains, tetapi juga keterampilan berpikir tingkat tinggi, pemahaman epistemologis, serta kesadaran social (OECD, 2023). Literasi kimia merupakan salah satu cabang literasi sains. Literasi kimia mencakup kemampuan memahami konsep dasar, serta keterampilan dalam menggabungkan pengetahuan tersebut dengan konteks kehidupan nyata dan isu-isu sosiosaintifik. Literasi kimia tidak hanya sebatas pengetahuan konten, tetapi juga mencakup pemahaman terhadap bagaimana kimia diterapkan dalam masyarakat, serta sikap kritis terhadap isu-isu yang terkait dengan kimia (Kohen *et al.*, 2020). Kondisi literasi kimia peserta didik Indonesia di lapangan masih cukup jauh dari harapan. Berdasarkan hasil asesmen PISA 2022, skor literasi sains yang mencakup literasi kimia Indonesia berada pada tingkat yang lebih rendah daripada rata-rata negara OECD yang lain (OECD, 2023).

Capaian literasi kimia peserta didik di Indonesia masih berada di bawah standar internasional, dengan skor rata-rata sebesar 383 poin, sedangkan rata-rata OECD mencapai 485 poin. Kesenjangan ini dipengaruhi oleh beberapa faktor, seperti pelaksanaan penilaian yang umumnya dilakukan setelah pembelajaran, penggunaan soal deskriptif sederhana yang belum memuat komponen literasi kimia, persepsi guru bahwa butir soal literasi kimia sulit untuk dikembangkan, serta dominasi soal yang berorientasi pada perhitungan numerik. (Nirmalasari *et al.*, 2022). Secara lebih jelas, kemampuan peserta didik dalam memahami konsep-konsep kimia kemudian mengaitkannya dengan fenomena kontekstual masih rendah. Materi kimia seperti sistem periodik unsur yang bersifat abstrak seringkali tidak dipahami secara mendalam oleh peserta didik sehingga pembelajaran menjadi kurang bermakna. Hal ini

menunjukkan perlunya upaya sistematis untuk meningkatkan kualitas pembelajaran dalam bidang sains, khususnya kimia.

Instrumen tes yang dikembangkan secara khusus untuk mengukur literasi kimia dalam kenyataannya masih sangat sedikit, terutama untuk materi sistem periodik unsur, sementara kurikulum menuntut agar pembelajaran tidak lagi berfokus pada hafalan semata, melainkan pada pencapaian kompetensi esensial seperti penguatan literasi (Wahyudin *et al.*, 2024). Bukti adanya instrumen yang digunakan untuk menilai kemampuan berpikir kritis serta literasi kimia peserta didik nyatanya masih Sangat sedikit (Ariefiani & Laksono, 2024). Kebanyakan soal yang digunakan dalam pembelajaran hanya terfokus pada aspek kognitif tingkat rendah seperti mengingat dan menjelaskan tanpa mempertimbangkan kemampuan peserta didik dalam mengintegrasikan konsep tersebut secara kontekstual. Instrumen penilaian yang biasa digunakan guru dalam pembelajaran peserta didik umumnya hanya terbatas untuk mengukur satu kemampuan saja, padahal dalam proses pembelajaran, terdapat beberapa kemampuan yang seharusnya diukur secara bersamaan (Ariefiani & Laksono, 2024). Hal ini mengakibatkan kemampuan berpikir kritis, analitis, dan reflektif peserta didik tidak terukur secara optimal, padahal aspek tersebut sangat penting dalam pembentukan karakter ilmiah peserta didik.

Beberapa solusi telah dikembangkan seperti penerapan soal *high order thinking skill* (HOTS) atau penggunaan asesmen otentik. Misalnya, penelitian oleh Sepe & Aquan (2023) yang menunjukkan bahwa soal-soal yang menekankan HOTS telah memenuhi kriteria validitas dan reliabilitas yang tinggi, namun tetap menghadapi kendala seperti rendahnya pemahaman siswa terhadap konteks soal. Hal ini memang dapat membantu pemahaman konseptual peserta didik, namun solusi tersebut belum sepenuhnya menjangkau cakupan literasi kimia yang utuh sebagaimana yang telah dirumuskan oleh Shwartz *et al* (2006), yang mencakup aspek konten, konteks, keterampilan belajar tingkat tinggi, dan sikap ilmiah. Beberapa temuan dari studi sebelumnya telah membuktikan bahwa instrumen tes literasi kimia dapat dijadikan solusi dalam mengatasi permasalahan ini. Penelitian yang dilakukan oleh Putri & Yusmaita (2024) yang mengembangkan instrumen tes literasi kimia pada materi Struktur Atom dan Nanoteknologi, serta penelitian Eliza & Yusmaita (2021) yang mengembangkan instrumen tes pada materi Sistem Koloid, menunjukkan soal yang dikembangkan memiliki item butir soal yang valid dan reliabel. Namun pengembangan instrumen tes literasi kimia yang terfokus pada materi Sistem Periodik Unsur nyatanya masih sangat terbatas. Ini artinya salah satu solusi untuk mengatasi permasalahan ini dapat berupa pengembangan instrumen tes literasi kimia, khususnya pada materi Sistem Periodik Unsur.

Instrumen yang dirancang dan dikembangkan berdasarkan kerangka literasi kimia dari Shwartz *et al* (2006) menjadikan instrumen lebih padu dibandingkan dengan soal konvensional yang hanya menilai dimensi kognitif rendah. Pendekatan ini tidak hanya mendorong pemahaman konseptual, tetapi juga mengembangkan kemampuan berpikir kritis melalui keterkaitan antara konsep kimia dan fenomena dunia nyata. Hal ini menegaskan bahwa instrumen atau soal yang disusun berbasis literasi kimia kontekstual memiliki potensi besar dalam menjangkau dimensi literasi kimia yang lebih menyeluruh. Kohen *et al* (2020) menyampaikan bahwa instrumen yang dirancang untuk mengukur literasi kimia harus

mencerminkan kompleksitas dunia nyata agar peserta didik mampu berpartisipasi aktif dalam isu-isu sains dan sosial yang aktual.

Penelitian ini difokuskan pada pengembangan instrumen penilaian literasi kimia pada materi Sistem Periodik Unsur. Instrumen yang dikembangkan bertujuan guna menghasilkan instrumen pengukuran yang memiliki validitas dan reliabilitas yang baik dalam menilai literasi kimia peserta didik yang mencakup pemahaman konsep, kemampuan mengaplikasikan pengetahuan kimia dalam kehidupan sehari-hari, dan keterampilan menalar fenomena kimia secara ilmiah. Instrumen ini diharapkan dapat digunakan sebagai rujukan dalam pelaksanaan penilaian pembelajaran kimia.

METODOLOGI

Penelitian ini merupakan penelitian *Educational Design Research* (EDR) yang bertujuan untuk menghasilkan instrumen tes literasi kimia pada materi Sistem Periodik Unsur (Liu, 2020). Penelitian pengembangan ini dilakukan melalui tahapan perancangan instrumen, validasi, uji coba, serta analisis kualitas butir soal hingga diperoleh instrumen tes literasi kimia yang layak digunakan. Tahap awal penelitian diawali dengan perancangan instrumen tes literasi kimia yang disusun berdasarkan indikator literasi kimia dan disesuaikan dengan materi Sistem Periodik Unsur serta karakteristik peserta didik. Instrumen yang telah dirancang selanjutnya divalidasi melalui uji validitas isi (validitas logis) oleh para ahli untuk menilai kesesuaian butir soal dengan indikator yang ditetapkan. Pengujian validitas isi dilaksanakan menggunakan metode *Content Validity Ratio* (CVR) (Wilson *et al.*, 2012).

Instrumen yang telah direvisi berdasarkan hasil validasi kemudian diuji cobakan secara terbatas kepada 35 peserta didik SMA N 7 Padang untuk memperoleh data empiris. Data hasil uji coba selanjutnya dianalisis menggunakan pendekatan model Rasch untuk analisis kualitas instrumen penilaian literasi sains (Liu, 2020). Analisis model Rasch berbantuan perangkat lunak Ministep digunakan untuk mengkaji validitas empiris butir soal, reliabilitas instrumen, daya beda, serta indeks kesukaran soal (Sari & Mahmudi, 2024). Butir soal dapat dikatakan valid atau mempunyai kualitas yang baik adalah butir soal yang memiliki kualifikasi sesuai dengan kriteria *Outfit Mean Square* (MNSQ), *Outfit Z-Standard* (ZSTD), dan *Poin Measure Correlation* (PT Mean Corr). Kriteria minimum yang terpenuhi adalah sebanyak 2 dari 3 kriteria fit pada setiap butir soal untuk soal dapat dikatakan valid (Putri & Yusmaita, 2024). Kriteria nilai *outfit* MNSQ, *outfit* ZSTD, dan PT Mean Corr dapat ditentukan dari: 1) nilai *outfit* MNSQ dinyatakan dapat diterima apabila berada pada rentang 0,5–1,5; 2) nilai *outfit* ZSTD dianggap memenuhi kriteria jika berada pada rentang –2,0 hingga +2,0; dan 3) nilai PT Mean Corr berada dalam kisaran 0,4–0,85. (Sumintono & Widhiarso, 2015).

Nilai *Person Reliability* dan *Item Reliability* dapat dilihat dari nilai reliabilitas yang diperoleh dan disesuaikan dengan kategori sebagai berikut: 1) <0,67 berada pada kategori lemah; 2) 0,67-0,8 berada pada kategori cukup; 3) 0,81-0,9 berada pada kategori bagus; 4) 0,91-0,94 berada pada kategori bagus sekali; dan 5) >0,94 berada pada kategori istimewa (Sumintono & Widhiarso, 2015). Setelahnya nilai alpha Cronbach juga dapat diukur dari

kriteria: 1) $<0,5$ berada pada kategori sangat buruk; 2) $0,5-0,6$ berada pada kategori buruk; 3) $0,6-0,7$ berada pada kategori bagus; 4) $0,7-0,8$ berada pada kategori cukup; dan 5) $>0,8$ berada pada kategori sangat bagus (Ramadhan *et al.*, 2024). Sementara kriteria tingkat kesukaran yaitu; 1) $Measure\ logit > SD\ logit$ berarti item sangat sulit; 2) $0 = Measure\ logit = SD\ logit$ berarti item sulit; 3) $-SD\ logit = Measure\ logit = 0$ berarti item mudah; dan 4) $Measure\ logit < -SD\ logit$ berarti item sangat mudah. Hasil analisis Rasch digunakan sebagai dasar untuk merevisi dan menyempurnakan instrumen tes literasi kimia yang dikembangkan. Dengan demikian, instrumen yang dihasilkan diharapkan memenuhi kriteria valid dan reliabel serta mampu mengukur literasi kimia peserta didik secara akurat.

TEMUAN DAN PEMBAHASAN

Penelitian ini dilakukan untuk menghasilkan instrumen tes literasi kimia pada materi Sistem Periodik Unsur yang berfungsi mengukur kemampuan literasi kimia peserta didik. Penyusunan instrumen tes berupa butir soal literasi kimia dilakukan dengan mengacu pada tahapan pengembangan instrumen yang dikemukakan oleh Xiufeng Liu (2020). Langkah awal yang dilakukan ialah menetapkan tujuan instrumen. Penelitian ini bertujuan menghasilkan instrumen tes sumatif yang dirancang agar selaras dan berkaitan secara langsung dengan tujuan pembelajaran yang telah ditetapkan (Liu, 2020). Penelitian ini mengkaji pemahaman peserta didik terhadap konsep sistem periodik unsur, sehingga hirarki pemahaman dirancang secara bertahap, mulai dari pemahaman tentang dasar tabel periodik modern hingga kemampuan mengaplikasikan konsep sistem periodik unsur pada ranah sosio-saintifik.

Tahap selanjutnya yaitu menyusun indikator soal berdasarkan materi sistem periodik unsur untuk menghasilkan instrumen tes literasi kimia yang berfungsi menilai kemampuan literasi kimia peserta didik. Indikator soal dikembangkan dan disusun berdasarkan kisi-kisi instrumen tes dari hasil analisis Capaian Pembelajaran (CP) dan Tujuan Pembelajaran (TP). Tipe tes yang diterapkan dalam penelitian ini adalah tes uraian terbatas di mana Batasan tersebut meliputi format, isi, serta ruang lingkup jawaban. Meskipun jawaban yang diberikan peserta didik dapat bervariasi, jawaban tersebut tetap harus memuat pokok-pokok penting yang tersusun secara sistematis sesuai dengan ketentuan dan batasan yang telah ditetapkan dalam soal. (Susanto, 2023). Pengembangan instrumen ini memiliki 4 wacana dengan 4 soal. Wacana soal memuat 4 aspek literasi kimia (konten, konteks, *high order learning skill* (HOLS), dan aspek afektif). Contoh salah satu wacana yang dikembangkan terlihat pada Tabel 1. Adapun rubrik penilaian yang memuat kunci jawaban soal, level literasi, serta pedoman penskoran. Rubrik penilaian yang digunakan didasarkan pada penilaian literasi sains oleh Bybee dalam Schwartz, Ben-zvi, *et al.* (2006) yang terdiri dari 5 level literasi (*scientific illiteracy, nominal scientific literacy, functional scientific literacy, conceptual scientific literacy, dan multidimensional scientific literacy*). Namun pada penelitian ini terbatas sampai *conceptual scientific literacy*. Contoh rubrik penilaian disajikan pada Tabel 2.

Tabel 1. Wacana dan soal literasi kimia

Wacana 3: Pertambangan Batubara

Bacaalah wacana di bawah ini dengan cermat!

Salah satu pertambangan besar di Indonesia terletak di daerah Sawahlunto, Sumatera Barat, yang berlangsung pada masa kolonial. Aktivitas ini memberikan manfaat ekonomi bagi masyarakat, tetapi juga dapat menimbulkan pencemaran lingkungan. Secara umum salah satu dampak yang perlu di waspadai adalah terbentuknya limbah logam berat, seperti timbal (Pb) dan merkuri (Hg), yang dapat mencemari tanah maupun air di sekitarnya. Dalam sistem periodik unsur, timbal (Pb) berada pada golongan 14 periode 6, sedangkan merkuri (Hg) berada pada golongan 12 periode 6. Keduanya termasuk unsur logam yang terletak di bagian bawah tabel periodik, sehingga memiliki jumlah kulit elektron yang lebih banyak dibandingkan unsur-unsur yang berada di bagian atas golongannya.

Konteks

Unsur yang terletak semakin ke bawah dalam satu golongan memiliki jari-jari atom yang lebih besar karena jumlah kulit elektronnya semakin banyak. Akibatnya, gaya tarik inti terhadap elektron terluar menjadi lebih lemah. Kondisi ini menyebabkan energi ionisasi (energi yang dibutuhkan untuk melepaskan elektron) menjadi lebih rendah, sehingga atom-atom logam berat seperti Pb dan Hg lebih mudah melepaskan elektron dan bereaksi membentuk senyawa di lingkungan. Selain itu, keelektronegatifan dan afinitas elektron unsur-unsur di bagian bawah golongan cenderung lebih kecil, yang menunjukkan kemampuan atom untuk menarik atau menerima elektron semakin lemah. Hal ini menyebabkan sifat logamnya semakin kuat dan senyawanya relatif stabil serta sulit terurai secara alami.

Konten

Sifat kestabilan inilah yang membuat logam berat seperti timbal dan merkuri dapat bertahan lama di alam, menumpuk di dalam tanah air, dan tubuh makhluk hidup melalui rantai makanan. Jika jumlahnya terus meningkat, logam berat tersebut dapat mengganggu sistem saraf, hati, dan ginjal pada manusia maupun hewan, serta menurunkan kualitas tanah dan air di sekitarnya seperti yang terlihat pada Gambar 3. Oleh karena itu, pengelolaan limbah pertambangan perlu dilakukan dengan hati-hati agar tidak merusak keseimbangan ekosistem di daerah tersebut.

HOLS

Afektif



Gambar 1. Pencemaran Tanah dan Air

Sumber: Tambang.co.id. (2014). Greenpeace: Sumber air di Kalsel tercemar limbah batu bara. <https://www.tambang.co.id/greenpeace-sumber-air-di-kalsel-tercemar-limbah-batu-bara>

Jawablah pertanyaan di bawah ini dengan tepat!

Berdasarkan wacana di atas, jelaskan mengapa logam berat seperti merkuri (Hg) dan timbal (Pb) dapat bersifat berbahaya dan bertahan lama di lingkungan jika dikaitkan dengan letaknya dalam tabel periodik unsur dan sifat-sifat periodik yang dimilikinya!

Aspek konten mencakup pemahaman terhadap konsep ilmiah umum serta hakikat kimia sebagai ilmu eksperimental yang memberikan pengetahuan untuk menjelaskan fenomena (Sjöström *et al.*, 2024). Pada wacana 3 aspek konten merujuk pada pertambangan yang dapat mencemari lingkungan. Aspek konteks merujuk pada penerapan konsep kimia kemudian mengaitkan pengetahuan kimia untuk menjelaskan fenomena. Aspek *High Order Learning Skills* (HOLS) mencakup keterampilan untuk merumuskan pertanyaan, menelusuri serta mengolah informasi, dan menganalisis berbagai sudut pandang secara kritis terhadap wacana yang disajikan. (Marfuatun *et al.*, 2024). Terakhir adalah aspek afektif yang berkaitan dengan sikap objektif, realistis terhadap kimia serta penerapannya dalam kehidupan (Shwartz, Ben-Zvi, *et al.*, 2006).

Rubrik penilaian literasi sains (Tabel 2) yang pertama yaitu *Scientific Illiteracy* di mana peserta didik tidak mampu mengaitkan atau menanggapi pertanyaan yang wajar terkait sains. Mereka belum memiliki kapasitas kognitif untuk mengenali pertanyaan tersebut sebagai pertanyaan ilmiah. Pada tingkatan *Nominal scientific literacy*, peserta didik telah mulai mengenali suatu konsep sebagai bagian dari sains, namun tingkat pemahamannya masih menunjukkan adanya miskonsepsi yang jelas. Pada tingkatan *Functional Scientific Literacy*, peserta didik mampu mendeskripsikan suatu konsep dengan tepat, namun memiliki pengetahuan yang masih terbatas terhadap konsep tersebut. Dan pada tingkatan *Conceptual Scientific Literacy*, peserta didik mulai membangun pemahaman terhadap skema konseptual utama dalam suatu disiplin ilmu serta mengaitkannya dengan pemahaman umum tentang sains (Shwartz, Ben-zvi, *et al.*, 2006).

Instrumen yang telah dikembangkan kemudian divalidasi oleh 5 orang validator, yaitu 5 orang dosen Departemen Kimia, Fakultas Matematika dan ilmu Pengetahuan Alam, Universitas Negeri Padang. Pengumpulan data hasil validasi dilakukan menggunakan angket yang disusun oleh tiga komponen yaitu wacana soal, pertanyaan dan jawaban, serta level literasi kimia. Untuk penilaian dari ketiga komponen tersebut diberikan skor dengan skala -1 s/d 1 di mana -1 menyatakan tidak relevan/tidak sesuai, 0 menyatakan kurang relevan/kurang sesuai, dan 1 menyatakan relevan atau sesuai (Obilor & Miwari, 2022). Hasil validitas isi disajikan pada Tabel 3.

Tabel 2. Rubrik penilaian wacana 3

Jawaban	Level Literasi Kimia	Skor
Jawaban salah atau tidak terjawab	<i>Scientific Illiteracy</i>	0
Semakin ke bawah suatu golongan: <ul style="list-style-type: none"> • Jari-jari atom bertambah besar, sehingga gaya tarik inti terhadap elektron berkurang. • Energi ionisasi, keelektronegatifan, dan afinitas elektron menurun, membuat atom lebih mudah melepaskan elektron. 	<i>Nominal Scientific Literacy</i>	1
Merkuri (Hg) dan timbal (Pb) berada di bagian bawah golongan logam dalam sistem periodik. Semakin ke bawah suatu golongan: <ul style="list-style-type: none"> • Jari-jari atom bertambah besar, sehingga gaya tarik inti terhadap elektron berkurang. • Energi ionisasi, keelektronegatifan, dan afinitas elektron menurun, membuat atom lebih mudah melepaskan elektron. 	<i>Functional Scientific Literacy</i>	2
Merkuri (Hg) dan timbal (Pb) berada di bagian bawah golongan logam dalam sistem periodik. Semakin ke bawah suatu golongan: <ul style="list-style-type: none"> • Jari-jari atom bertambah besar, sehingga gaya tarik inti terhadap elektron berkurang. • Energi ionisasi, keelektronegatifan, dan afinitas elektron menurun, membuat atom lebih mudah melepaskan elektron. <p>Akibatnya, logam berat seperti Hg dan Pb mudah membentuk senyawa yang stabil, tetapi sulit diuraikan secara alami. Sifat stabil inilah yang menyebabkan logam berat bertahan lama di lingkungan, menumpuk dalam tanah, air, dan tubuh makhluk hidup, sehingga bersifat racun dan berbahaya bagi ekosistem.</p>	<i>Conceptual Scientific Literacy</i>	3

Tabel 3. Hasil validitas isi

Butir Soal	Nilai CVR	Keterangan
S1	0,85	Valid
S2	0,9	Valid
S3	0,9	Valid
S4	0,88	Valid

Validitas isi instrumen dianalisis dengan metode *Content Validity Ratio* (CVR) yang dikembangkan oleh Lawshe untuk mengetahui tingkat ketepatan setiap butir soal dengan indikator yang telah ditetapkan berdasarkan penilaian para validator ahli. CVR digunakan untuk menilai sejauh mana suatu butir soal dianggap esensial dalam mengukur konstruk yang diinginkan (Wilson *et al.*, 2012). Penilaian dilakukan oleh validator dengan memberikan keputusan terhadap setiap butir aspek penilaian, yaitu apakah butir tersebut tergolong sesuai, kurang sesuai, dan tidak sesuai dengan instrumen yang dikembangkan. Nilai CVR kemudian dihitung untuk masing-masing butir berdasarkan jumlah validator yang menyatakan item tersebut sesuai atau tidak. Hasil akhir CVR yang didapatkan selanjutnya dibandingkan dengan nilai minimum CVR sesuai jumlah validator berdasarkan kriteria Lawshe. Berdasarkan kriteria minimum Lawshe, untuk batas kriteria 5 validator, sebesar 0,736 (Wilson *et al.*, 2012). Temuan ini menunjukkan bahwa soal-soal yang dibuat telah disesuaikan dengan indikator literasi kimia serta tujuan pengukuran yang ditetapkan. Hasil analisis ini menjadi dasar dalam menentukan kelayakan setiap butir soal sebelum instrumen digunakan pada tahap selanjutnya (G *et al.*, 2024).

Instrumen tes yang telah disusun diujicobakan kepada sampel yang telah ditentukan untuk mengumpulkan data. Instrumen tersebut diberikan kepada 35 peserta didik kelas X SMAN 7 Padang yang telah mempelajari materi sistem periodik unsur. Peserta didik diberi waktu selama 45 menit untuk menyelesaikan empat butir soal literasi kimia. Setelah uji coba, jawaban peserta didik terhadap butir soal dianalisis menggunakan model Rasch. Analisis model Rasch dilakukan dengan bantuan perangkat lunak Ministep berdasarkan data skor jawaban peserta didik yang diperoleh dari uji lapangan. Penilaian kualitas butir soal mencakup aspek validitas, reliabilitas, tingkat kesukaran, dan daya pembeda.

Pemodelan Rasch merupakan alternatif dalam pengembangan instrumen pengukuran pada penilaian pendidikan selain pendekatan teori tes klasik. (Sumintono & Widhiarso, 2015). Baik tidaknya suatu instrumen ditentukan oleh beberapa faktor, salah satunya validitas (Haq, 2022). Instrumen tes dianggap memiliki validitas tinggi apabila mampu menjalankan fungsi ukur secara tepat dan menyajikan data yang sesuai dengan tujuan pengukuran (Ramadhan *et al.*, 2024). Hasil validitas butir soal disajikan pada Gambar 2. Berdasarkan data yang diperoleh, seluruh butir soal dapat diterima sebab memenuhi semua kriteria.

ITEM STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	ITEM
2	73	35	-.74	.35	1.24	.98	1.60	1.86	A .48	.68	60.0	70.1	S2
3	57	35	1.01	.32	.95	-.15	.92	-.27	B .75	.70	68.6	64.8	S3
4	50	35	1.74	.32	.92	-.29	.91	-.30	b .80	.70	54.3	64.9	S4
1	83	35	-2.01	.37	.75	-.99	.82	-.29	a .65	.62	82.9	73.2	S1
MEAN	65.8	35.0	.00	.34	.96	-.11	1.06	.25			66.4	68.2	
P.SD	13.0	.0	1.47	.02	.18	.71	.31	.93			10.8	3.6	

Gambar 2. Item fit

Tes dinyatakan reliabel apabila menghasilkan skor yang konsisten ketika diberikan kepada kelompok yang sama pada waktu atau kesempatan yang berbeda. Reliabilitas terdiri

atas dua jenis, yaitu reliabilitas konsistensi tanggapan dan reliabilitas konsistensi gabungan butir. Reliabilitas konsistensi tanggapan berkaitan dengan tingkat konsistensi respon peserta didik atau objek ukur terhadap tes atau instrumen yang digunakan. Sementara itu, reliabilitas konsistensi gabungan butir berhubungan dengan kestabilan dan keterpaduan antarbutir dalam suatu tes (Ramadhan et al., 2024). Hasil reliabilitas yang didapatkan menggunakan Ministeps terlihat seperti yang ada di Gambar 3. Dalam pemodelan Rasch, reliabilitas ditunjukkan melalui nilai separasi individu (*person separation*) dan separasi butir (*item separation*) (Sari & Mahmudi, 2024). Nilai reliabilitas butir soal sebesar 0,94 menunjukkan bahwa reliabilitas item berada pada kategori sangat baik. Sementara itu, nilai reliabilitas person sebesar 0,61 mengindikasikan bahwa reliabilitas peserta didik tergolong rendah. Nilai yang belum optimal ini mengindikasikan bahwa instrumen belum sepenuhnya mampu memisahkan kemampuan siswa secara stabil, sehingga terdapat kemungkinan variasi respons yang dipengaruhi oleh faktor eksternal, seperti tingkat konsentrasi siswa, pemahaman terhadap butir soal, atau homogenitas kemampuan responden (Sumintono & Widhiarso, 2015). Nilai reliabilitas person yang rendah seringkali terjadi ketika variasi kemampuan responden tidak cukup luas atau jumlah butir soal relatif sedikit sehingga kemampuan siswa yang berbeda tidak dapat ditangkap sepenuhnya oleh instrumen yang ada (Handayani et al., 2023). Untuk meningkatkan reliabilitas person pada penelitian lanjutan dapat dilakukan dengan menambah jumlah butir soal yang relevan dan memiliki rentang kesulitan yang lebih luas, sehingga instrumen dapat lebih membedakan kemampuan responden secara lebih akurat (Ibrahim, 2025). Sementara untuk nilai Alpha Cronbach yaitu sebesar 0.61 yang berada pada kategori cukup. Alpha Cronbach berfungsi untuk mengukur konsistensi keseluruhan interaksi antara individu dan butir soal (Ngadi, 2023).

SUMMARY OF 35 MEASURED PERSON								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	7.5	4.0	1.46	1.01	1.01	-.07	1.06	-.06
SEM	.3	.0	.33	.02	.14	.19	.17	.19
P.SD	1.9	.0	1.90	.11	.81	1.11	1.00	1.11
S.SD	1.9	.0	1.92	.11	.82	1.13	1.01	1.13
MAX.	11.0	4.0	5.29	1.28	3.38	2.16	4.37	2.36
MIN.	5.0	4.0	-.96	.95	.11	-1.72	.11	-1.72

REAL RMSE	1.18	TRUE SD	1.48	SEPARATION	1.26	PERSON RELIABILITY	.61	
MODEL RMSE	1.02	TRUE SD	1.60	SEPARATION	1.57	PERSON RELIABILITY	.71	
S.E. OF PERSON MEAN = .33								

PERSON RAW SCORE-TO-MEASURE CORRELATION = 1.00								
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .61 SEM = 1.16								
STANDARDIZED (50 ITEM) RELIABILITY = .97								
SUMMARY OF 4 MEASURED ITEM								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	65.8	35.0	.00	.34	.96	-.11	1.06	.25
SEM	7.5	.0	.85	.01	.10	.41	.18	.54
P.SD	13.0	.0	1.47	.02	.18	.71	.31	.93
S.SD	15.0	.0	1.70	.02	.21	.82	.36	1.07
MAX.	83.0	35.0	1.74	.37	1.24	.98	1.60	1.86
MIN.	50.0	35.0	-2.01	.32	.75	-.99	.82	-.30

REAL RMSE	.35	TRUE SD	1.43	SEPARATION	4.08	ITEM RELIABILITY	.94	
MODEL RMSE	.34	TRUE SD	1.43	SEPARATION	4.21	ITEM RELIABILITY	.95	
S.E. OF ITEM MEAN = .85								

ITEM RAW SCORE-TO-MEASURE CORRELATION = -1.00								
Global statistics: please see Table 44.								
UHEAN=.0000 USCALE=1.0000								

Gambar 3. Summary statistics

Tingkat kesukaran butir soal menunjukkan kategori tingkat kesukaran soal, apakah tergolong mudah, sedang, atau sulit. Informasi mengenai tingkat kesukaran diperoleh dari

Output Item Measure pada analisis Rasch. Indeks ini ditentukan dengan mempertimbangkan nilai rata-rata *logit* dan standar deviasi (SD). Hasil analisis menunjukkan nilai rata-rata *logit* adalah 0.00 dan standar deviasi sebesar 1.47. Nilai tersebut digunakan sebagai dasar dalam pengelompokan tingkat kesukaran butir soal, informasi tersebut dapat dilihat melalui Gambar 4. Penelitian ini menunjukkan nilai masing-masing item yakni S4 sebesar 1.74 yang berada pada kategori sangat sulit, S3 sebesar 1.01 yang berada pada kategori sulit, S2 sebesar -0.74 yang berada pada kategori mudah, serta S1 sebesar -2.01 yang berada pada kategori sangat mudah. Nilai *logit* yang lebih tinggi mengindikasikan soal dengan tingkat kesulitan yang lebih besar. Tingkat kesukaran setiap butir direpresentasikan oleh nilai *measure*, yaitu hasil transformasi data ordinal berupa skor jawaban ke dalam satuan *logit* yang memiliki interval setara (*equal interval*) (Nazharita et al., 2022).

ITEM STATISTICS: MEASURE ORDER													
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT OBS%	MATCH EXP%	ITEM
4	50	35	1.74	.32	.92	-.29	.91	-.30	.80	.70	54.3	64.9	S4
3	57	35	1.01	.32	.95	-.15	.92	-.27	.75	.70	68.6	64.8	S3
2	73	35	-.74	.35	1.24	.98	1.60	1.86	.48	.68	60.0	70.1	S2
1	83	35	-2.01	.37	.75	-.99	.82	-.29	.65	.62	82.9	73.2	S1
MEAN	65.8	35.0	.00	.34	.96	-.11	1.06	.25			66.4	68.2	
P.SD	13.0	.0	1.47	.02	.18	.71	.31	.93			10.8	3.6	

Gambar 4. Item measure

Daya pembeda pada model Rasch yang menggunakan perangkat lunak Ministeps dapat dilihat pada menu *outfit tables* kemudian pada *summary statistic*. Daya pembeda berfungsi untuk mengidentifikasi perbedaan kemampuan peserta didik pada kategori tinggi, sedang, dan rendah. Nilai *separation* digunakan untuk mengetahui pengelompokan antara peserta didik dan butir soal. Nilai *separation* yang semakin besar mengindikasikan bahwa kualitas instrumen semakin baik karena mampu mengidentifikasi perbedaan kelompok kemampuan peserta didik maupun kelompok tingkat kesukaran butir secara lebih optimal. Nilai *separation* dapat dilihat pada Gambar 3. Setelah didapatkan nilai *separation*, selanjutnya dicari persamaan untuk mendapatkan daya beda person dan item (Sumintono & Widhiarso, 2015).

$$H = \frac{4(1.26)+1}{3} = 2.01 \tag{1}$$

Nilai *person separation* yang didapatkan adalah 2.01, kemudian dibulatkan menjadi 2. Artinya terdapat dua kelompok person, yang dapat diartikan dengan abilitas tinggi dan abilitas rendah.

$$H = \frac{4(4.08)+1}{3} = 5.7 \tag{2}$$

Nilai butir *separation* yang didapatkan adalah 5.7, kemudian dibulatkan menjadi 6. Artinya instrumen memiliki daya beda yang tinggi dan sangat baik. Nilai *separation* yang didapat digunakan untuk mengelompokkan daya beda soal, di mana semakin tinggi nilai *separation*, semakin baik kemampuan butir soal dalam membedakan tingkat kemampuan

peserta didik. Semakin besar nilai separation, semakin baik kualitas instrumen baik dari sisi responden maupun butir soal, karena instrumen tersebut mampu mengidentifikasi kelompok responden dan kelompok butir secara lebih jelas (Putri & Yusmaita, 2024). Temuan dari analisis memperlihatkan bahwa instrumen memiliki daya beda yang baik dan mampu mengelompokkan peserta didik ke dalam dua tingkat kemampuan yang artinya berada pada abilitas tinggi dan abailitas rendah.

Tahap akhir yaitu perlunya penyusunan pedoman untuk mendukung penggunaan serta pengembangan instrumen tes secara tepat, sehingga memberikan kejelasan dalam penerapannya (Liu, 2020). Pedoman yang disusun meliputi penjabaran Capaian Pembelajaran (CP) ke dalam Tujuan Pembelajaran (TP), penyusunan kisi-kisi instrumen tes, perumusan butir soal, serta pengembangan rubrik penilaian. Instrumen literasi kimia yang dikembangkan telah memenuhi kriteria kelayakan berdasarkan analisis empiris, sehingga dapat digunakan sebagai alat ukur literasi kimia peserta didik, yang dapat dimanfaatkan guru untuk mengidentifikasi kesenjangan pemahaman serta merancang tindak lanjut pembelajaran secara lebih tepat. Pengembangan lanjutan tetap diperlukan agar melalui penambahan butir soal, variasi kesukaran, hingga uji coba sampel yang lebih luas dapat memperoleh parameter yang lebih stabil.

SIMPULAN

Pengembangan butir soal literasi kimia pada materi sistem periodik unsur dilakukan dengan mengacu pada model pengembangan instrumen yang dikemukakan oleh Xiufeng Liu. Hasil analisis butir soal literasi kimia materi sistem periodik unsur menggunakan model Rasch mengindikasikan bahwa semua butir soal memenuhi kriteria validitas dan layak digunakan sebagai instrumen pengukuran. Instrumen tersebut memiliki reliabilitas item sebesar 0,94, reliabilitas person sebesar 0,61, dan nilai Alpha Cronbach sebesar 0,61. Instrumen yang dikembangkan memiliki tingkat kesulitan butir soal yang baik, sedangkan daya pembeda dapat mengelompokkan peserta didik menjadi dua kategori, yakni peserta didik berabilitas tinggi dan peserta didik berabilitas rendah.

PENGHARGAAN

Peneliti menyampaikan apresiasi dan ucapan terima kasih kepada Direktorat Penelitian dan Pengabdian Kepada Masyarakat Direktorat Jendral Riset dan Pengembangan Kementerian Pendidikan Tinggi, Sains dan Teknologi atas dukungan pendanaan terhadap penelitian ini melalui kontrak penelitian nomor 088/C3/DT.05.00/PL/2025.

REFERENSI

Ariefiani, N. W., & Laksono, E. W. (2024). *Development of An Integrated Assessment Instrument to Measure Students ' Critical Thinking and Chemical Literacy Skills for Rate of*

- Reaction Topic*. 10(1), 7726–7734. <https://doi.org/10.29303/jppipa.v10i10.9069>
- Eliza, W., & Yusmaita, E. (2021). Pengembangan Butir Soal Literasi Kimia pada Materi Sistem Koloid Kelas XI IPA SMA / MA. *JEP (Jurnal Eksakta Pendidikan)*, 5(2), 197–204.
- Gymnatias, M. A., Noor, R. A. M., & Wiyono, A. Creation Of Assesment Instruments in Differential Learning System. Atikanoto: *Journal of Automotive Engineering Education*, 1(1). Handayani, Y., Rahmawati, & Widiasih. (2023). Using Rasch Model to Analyze Reliability and Validity of Concept Mastery Test on Electricity and Magnetism Topic. *JIPF (Jurnal Ilmu Pendidikan Fisika)*, 8(2), 226–239.
- Haq, V. A. (2022). Menguji Validitas dari Reliabilitas pada Mata Pelajaran Alquran Hadits Menggunakan Korelasi Produk Momen Spearman Brown. *An-Nawa: Jurnal Studi Islam*, 04(01), 11–24. <https://doi.org/https://doi.org/10.37758/annawa.v4i1.419>
- Ibrahim, N. L. (2025). Analisis Literatur Penggunaan Model Rasch untuk Validasi Instrumen Asesmen Pendidikan : Tren, Temuan, dan Implikasi. *Lencana: Jurnal Inovasi Ilmu Pendidikan*, 3(April).
- Kohen, Z., Herscovitz, O., & Dori, Y. J. (2020). How to promote chemical literacy? On-line question posing and communicating with scientists. *Chemistry Education Research and Practice*, 21(1), 250–266. <https://doi.org/10.1039/c9rp00134d>
- Liu, X. (2020). *Using and Developing Measurement Instruments in Science Education: A Rasch Modeling Approach 2nd edition* (C. S. Kalman (ed.)). Information Age Publishing, Inc.
- Marfuatun, Nahadi, Yuliani, G., & Hernani. (2024). The Framework and Types of Chemical Literacy Tests : A Systematic Review. *Jurnal Penelitian Pendidikan IPA*, 10(6), 269–276. <https://doi.org/10.29303/jppipa.v10i6.7641>
- Nazharita, E., Hidayat, R., & Rahmatudin, J. (2022). Analisis Butir Soal Tes Literasi Numerasi Domain geometri Siswa SMP Menggunakan Rasch Model Berbantuan Ministep. *Pedagogy: Jurnal Pendidikan Matematika*, 10(3), 989–1000.
- Ngadi. (2023). Analisis Model Rasch untuk Mengukur Kompetensi Pengetahuan Siswa SMKN 1 Kalianget pada Mata Pelajaran Perawatan Sistem Kelistrikan Sepeda motor. *Jurnal Pendidikan Vokasi Otomotif*, 6(November).
- Nirmalasari, M. . Y., Mago, O. Y. T., & Manuk, I. L. (2022). Validitas Instrumen Soal Literasi Numerasi Kimia Hidrokarbon dalam Integrasinya dengan Isu Sosiosaintifik Lokal Sikka M.A. *Jurnal Pendidikan MIPA*, 12(4), 1004–1011. <https://doi.org/https://doi.org/10.37630/jpm.v12i4.725>
- Obilor, E. I., & Miwari, G. U. (2022). Content Validity in Educational Assessment. *International Journal of Innovative Education Research*, 10(2), 57–69.
- OECD. (2023). Pisa 2022 Results. In *Factsheets: Vol. 1*. https://www.oecd-ilibrary.org/education/pisa-2022-results-volume-i_53f23881-en%0Ahttps://www.oecd.org/publication/pisa-2022-results/country-notes/germany-

1a2cf137/

- Putri, S. A. D., & Yusmaita, E. (2024). Pengembangan Instrumen Tes Berbasis Literasi Kimia pada Materi Struktur Atom dan Nanoteknologi Kurikulum Merdeka Fase E SMA. *Entalpi Pendidikan Kimia*.
<http://entalpipendidikan.ppj.unp.ac.id/index.php/epk/article/view/479>
- Ramadhan, M. F., Siroj, R. A., & Afgani, M. W. (2024). *Validitas and Reliabilitas*. 06(02), 10967–10975.
- Sari, E. D. K., & Mahmudi, I. (2024). *Analisis Pemodelan Rasch pada Assesment Pendidikan (Analisis dengan Menggunakan Aplikasi Winstep*. PT. Pena Persada Kerta Utama.
- Sepe, C. P. A., & Aquan, H. M. (2023). Developing and Testing HOTS-Based Evaluation Packages for Metabolism with Science Literacy Skills Aspects. *IJECA (International Journal of Education and Curriculum Application)*, 6(2), 123.
<https://doi.org/10.31764/ijeca.v6i2.15420>
- Shwartz, Y., Ben-zvi, R., & Hofstein, A. (2006). The use of scientific literacy taxonomy for assessing the development of chemical literacy among high-school students. *Chemistry Education Research and Practice*, 7(4), 203–225.
- Shwartz, Y., Ben-Zvi, R., & Hofstein, A. (2006). Chemical literacy: What does this mean to scientists and school teachers? *Journal of Chemical Education*, 83(10), 1557–1561.
<https://doi.org/10.1021/ed083p1557>
- Sjöström, J., Yavuzkaya, M., Guerrero, G., & Eilks, I. (2024). Critical Chemical Literacy as a Main Goal of Chemistry Education Aiming for Climate Empowerment and Agency. *Journal of Chemical Education*, 101, 4189–4195. <https://doi.org/10.1021/acs.jchemed.4c00452>
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch Pada Assesment Pendidikan* (B. Trim (ed.)). Trim Komunikata.
- Susanto, S. (2023). Pengembangan Alat dan Teknik Evaluasi Tes dalam Pendidikan. *Jurnal Tarbiyah Jamiat Kheir*, 1(1), 51–60.
- Wahyudin, D., Subkhan, E., Malik, A., Hakim, M. A., Sudiapermana, E., LeliAlhapip, M., Nur Rofika Ayu Shinta Amalia, L. S., Ali, N. B. V., & Krisna, F. N. (2024). Kajian Akademik Kurikulum Merdeka. *Kemendikbud*, 1–143.
- Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development*, 45(3), 197–210. <https://doi.org/10.1177/0748175612440286>