



Pengembangan *Four-Tier Test* untuk Mengidentifikasi Miskonsepsi pada Materi Sistem Periodik Unsur Fase E SMA/MA

Umi Salamah¹, Andromeda^{2✉}

Universitas Negeri Padang, Indonesia^{1,2}

e-mail : umisalamah11223344@gmail.com¹, andromeda@fmipa.unp.ac.id²

Abstrak

Penelitian ini dilatar belakangi oleh banyaknya permasalahan yang dihadapi oleh peserta didik di SMA/MA pada proses pembelajaran kimia yang berhubungan dengan kesalahpahaman peserta didik memahami konsep kimia sehingga diperlukan sebuah alat ukur untuk mengetahui permasalahan tersebut. Tujuan penelitian ini adalah untuk mengembangkan instrumen tes diagnostik *four-tier* untuk mengidentifikasi miskonsepsi peserta didik pada materi sistem periodik unsur yang memenuhi kriteria instrumen tes yang baik. Penelitian ini merupakan penelitian *Research and Development* dengan menggunakan model pengembangan Treagust (1988). Berdasarkan hasil validasi terhadap lima orang validator, instrumen tes diagnostik dinyatakan valid secara logis dengan nilai rata-rata 0,88. Berdasarkan hasil uji coba terhadap peserta didik, didapatkan nilai reliabilitas instrumen tes diagnostik pada tingkat pertama (*first-tier*) dan tingkat ketiga (*third-tier*) sebesar 0,88 dan 0,90 dengan kategori reliabilitas sangat tinggi, serta memiliki indeks kesukaran, daya pembeda serta fungsi distraktor yang baik. Implikasi penelitian ini adalah dapat dijadikan sebagai alat ukur bagi guru untuk mengetahui miskonsepsi peserta didik SMA/MA pada materi sistem periodik unsur.

Kata Kunci: Tes Diagnostik, Pilihan Ganda Four-Tier, Miskonsepsi, Materi Sistem Periodik Unsur.

Abstract

This research is motivated by the many problems faced by high school/MA students in the chemistry learning process which are related to students' misunderstanding of understanding chemical concepts, so a measuring tool is needed to find out these problems. The aim of this research is to develop a four-tier diagnostic test instrument to identify students' misconceptions about material on the periodic system of elements that meets the criteria for a good test instrument. This research is a Research and Development study using the Treagust (1988) development model. Based on the validation results of five validators, the diagnostic test instrument was declared logically valid with an average value of 0.88. Based on the results of trials on students, the reliability values of the diagnostic test instruments at the first level and third-tier were obtained at 0.88 and 0.90 with very high reliability categories, and had indexes of difficulty and strength. differentiator and good distractor function. The implication of this research is that it can be used as a measuring tool for teachers to find out misconceptions of SMA/MA students regarding the material on the periodic system of elements.

Keywords: Diagnostic Test, Misconceptions, Four-Tier Multiple Choice, Periodic System Elements Material.

Copyright (c) 2024 Umi Salamah, Andromeda

✉ Corresponding author :

Email : andromeda@fmipa.unp.ac.id

DOI : <https://doi.org/10.31004/edukatif.v6i4.7397>

ISSN 2656-8063 (Media Cetak)

ISSN 2656-8071 (Media Online)

PENDAHULUAN

Sains adalah pengetahuan yang merupakan kumpulan fakta-fakta, konsep-konsep, prinsip-prinsip. Kimia merupakan ilmu yang mempelajari tentang materi, sifat-sifatnya, perubahan yang terjadi pada materi, dan energi yang terkait perubahan materi tersebut (Muderawan *et al.*, 2019). Kimia mempelajari bagaimana komposisi suatu zat terhadap sifat-sifatnya, bagaimana zat itu berubah, dan bagaimana materi berinteraksi dalam reaksi kimia. Selain itu, juga mempelajari struktur utama materi dan kekuatan yang menentukan sifat materi yang kita amati melalui indra kita (Sariati *et al.*, 2020). Pembelajaran kimia memuat konsep-konsep yang abstrak, hal ini sering membuat peserta didik merasa kesulitan dan gagal untuk memahami dan menghubungkan konsep-konsep yang ada dalam pembelajaran kimia sehingga belum mencapai pembelajaran yang bermakna (Rozi & Romadhoni, 2023). Tingkat kesulitan materi kimia bervariasi berdasarkan pengetahuan awal peserta didik. Perbedaan pengetahuan awal peserta didik selama proses pembelajaran sangat mempengaruhi pemahaman peserta didik terhadap konsep yang dipelajari. Perkembangan konsep bisa lambat atau cepat. Jika perkembangan peserta didik lambat, mereka mungkin merasa pelajaran sangat sulit. Akibatnya, sehingga menghambat peserta didik untuk menyelaraskan pengetahuan awal peserta didik dengan pengetahuan baru. Jika peserta didik tidak bisa menyelaraskan pemahaman baru dengan pemahaman sebelumnya, peserta didik akan menafsirkan konsep-konsep secara berbeda dari apa yang diketahui oleh para ilmuwan, yang disebut dengan miskonsepsi (Diannisa, 2023).

Jika peserta didik salah memahami konsep dasar, kemungkinan akan salah memahami konsep yang saling berhubungan (Savira *et al.*, 2019). Pendapat tersebut didukung oleh (Winarni, S. & Syahrial, 2016) bahwa konsep dalam kimia bersifat hirarkis, sehingga kesalahpahaman satu konsep awal akan menimbulkan kesalahpahaman konsep lainnya. Salah satu materi kimia yang paling sering mengalami miskonsepsi adalah materi sistem periodik unsur. Karena materi sistem periodik unsur merupakan materi yang memiliki banyak istilah dan kosakata baru, Pokok bahasannya mengenai perkembangan sistem periodik unsur, letak unsur dalam tabel periodik unsur, sifat-sifat keperiodikan unsur diantaranya yaitu jari-jari atom, energi ionisasi, keelektronegatifan, dan afinitas elektron (Halimatusya *et al.*, 2021). Materi sistem periodik unsur merupakan materi yang paling dasar dalam ilmu kimia sehingga harus menguasainya secara menyeluruh agar materi kimia selanjutnya tidak menimbulkan kesulitan. Materinya juga bersifat abstrak dan mengandung banyak teori. Karena sistem periodik unsur merupakan materi dasar maka harus dikenali dan diperbaiki agar miskonsepsi tersebut tidak berlanjut pada materi kimia lainnya (Ramadhan *et al.*, 2020).

Untuk mengetahui siswa yang mengalami miskonsepsi atau tidak, guru dan peneliti sering menggunakan berbagai macam teknik untuk mengidentifikasi miskonsepsi, yaitu diskusi pemecahan masalah, wawancara diagnostik, peta konsep, dan tes diagnostik. Metode wawancara, terutama dengan banyak peserta didik, memerlukan waktu yang lebih lama, sehingga kurang efektif. Cara ini hanya digunakan pada beberapa siswa karena memerlukan waktu lebih lama dan hasilnya bersifat umum. Akibatnya, jawaban yang diperoleh memiliki waktu yang lebih lama maka menjadi kurang efektif. Dengan demikian, miskonsepsi yang terjadi pada peserta didik tidak dapat terdeteksi (Nurhafsari & Rismaningsih, 2023).

Tes diagnostik adalah cara yang tepat untuk mengidentifikasi kesalahpahaman peserta didik dalam proses pembelajaran. Cara ini tidak membutuhkan waktu lama seperti wawancara. Hal ini memudahkan guru dalam memberikan penilaian, yang akan lebih bermanfaat bagi pendidik di kelas (Treagust, 1988). Tes diagnostik pilihan ganda merupakan salah satu tes diagnostik. Tes diagnostik pilihan ganda terbagi atas beberapa jenis, di antaranya tes diagnostik pilihan ganda satu tingkat (*one-tier*), dua tingkat (*two-tier*), tiga tingkat (*three-tier*), dan empat tingkat (*four-tier*). Tes diagnostik pilihan ganda satu tingkat terdiri dari beberapa pilihan jawaban yang harus dipilih peserta didik. Tes diagnostik pilihan ganda satu tingkat tidak memuat alasan terhadap soal dan jawaban, sehingga tes ini tidak dapat membedakan antara kurangnya pengetahuannya, adanya konsep alternatif, atau hanya memang tebakan saja (Rusilowati, 2015). Tes diagnostik

dua dan tiga tingkat dapat mendeteksi miskonsepsi dan tingkat pemahaman peserta didik tetapi memiliki keterbatasan yaitu tidak dapat mengetahui tingkat keyakinan pada tingkat pertama dan kedua. Berdasarkan hal itu, saat ini telah dikembangkan tes diagnostik empat tingkat (*Four-tier*) dengan tingkat keyakinan jawaban dan alasan berjenjang (Tumanggor *et al.*, 2020). Namun, berdasarkan hasil wawancara yang telah dilakukan terhadap beberapa orang guru kimia SMA di Sumatera Barat, menyebutkan bahwa soal evaluasi yang digunakan oleh guru masih menggunakan soal pilihan ganda biasa.

Instrumen tes *four-tier* berupa soal pilihan ganda yang terdiri dari empat tingkat yaitu pertanyaan soal (*first-tier*), keyakinan terhadap tingkat satu (*two-tier*), alasan soal (*three-tier*), keyakinan terhadap tingkat ketiga (*four-tier*). Instrumen tes *four-tier* dapat melihat tentang pemahaman peserta didik. Penelitian mengenai pengembangan instrumen tes diagnostik sudah banyak dikembangkan, seperti pada penelitian (Agustin *et al.*, 2022) mengenai pengembangan instrumen tes diagnostik *four-tier multiple choice* untuk identifikasi miskonsepsi siswa pada materi kesetimbangan kimia. Selanjutnya penelitian (Ritonga & Yasthophi, 2019) mengenai pengembangan instrumen tes diagnostik *four-tier multiple choice* pada materi ikatan kimia, dan pada penelitian (Devi & Azra, 2023) mengenai pengembangan instrumen tes diagnostik *four-tier* untuk melihat gambaran model mental peserta didik pada materi asam basa. Namun, berdasarkan penelusuran referensi dan observasi di sekolah, belum ditemukan adanya instrumen tes diagnostik *four-tier* untuk mengidentifikasi miskonsepsi peserta didik pada materi sistem periodik unsur Fase E SMA/MA.

Tujuan dari penelitian ini adalah untuk menghasilkan instrumen tes diagnostik *four-tier* untuk mengidentifikasi miskonsepsi peserta didik pada materi sistem periodik unsur Fase E SMA/MA yang valid dan reliabel agar ketersediaan instrumen tes pada materi terkait memenuhi tujuan pembelajaran.

METODE

Penelitian ini merupakan jenis penelitian pengembangan atau *Research and Development* (R&D). penelitian ini menggunakan model pengembangan yang telah dikembangkan oleh Treagust (1988), dimana model pengembangan ini terdiri dari tiga tahap yaitu sebagai berikut:

Pada tahap pertama menentukan isi. Terdapat empat, langkah diantaranya a) mendefinisikan pernyataan pengetahuan proposisional yang dilakukan dengan cara menganalisis capaian pembelajaran pada materi sistem periodik unsur yang sesuai dengan kurikulum merdeka; b) mengembangkan peta konsep, konsep-konsep yang terdapat pada tabel pernyataan proposisional disusun ke dalam sebuah peta konsep; c) menghubungkan pengetahuan proposisional dengan peta konsep; d) validasi konsep pernyataan proposisional dengan peta konsep oleh ahli.

Pada tahap kedua, memperoleh informasi tentang miskonsepsi peserta didik. langkah ini dilakukan dua kegiatan, yaitu a) melakukan wawancara dengan salah satu guru kimia di SMA Negeri 12 Padang; b) melakukan studi literatur terkait model mental peserta didik.

Tahap ketiga, mengembangkan instrumen tes diagnostik. Dalam mengembangkan instrumen tes diagnostik dilakukan tiga langkah, yaitu a) mengembangkan tes diagnostik, b) merancang kisi-kisi spesifik instrumen tes diagnostik, c) penyempurnaan lanjutan.

Penelitian ini dilakukan di SMA Negeri 12 Padang, Kota Padang pada semester Januari-Juni tahun ajaran 2023/2024. Subjek dari penelitian ini adalah peserta didik yang telah mempelajari materi sistem periodik unsur guru kimia, dan dosen kimia FMIPA UNP. Objek penelitian pada penelitian ini yaitu Instrumen Tes Diagnostik *Four-Tier* untuk mengidentifikasi miskonsepsi peserta didik fase E SMA/MA.

Dalam penelitian pengembangan tes diagnostik *four-tier* pada materi sistem periodik unsur digunakan instrumen untuk mengumpulkan data pada penelitian ini yaitu: lembar wawancara, angket validasi dan tes. Pada lembar angket validasi ini berisi penilaian terhadap kelayakan isi, kesesuaian konstruk, dan kesesuaian Bahasa.

Oleh karena itu, dilakukan validasi ke 5 validator. Validator akan diberikan lembar validasi soal dan mengisi dengan skala likert rentang 1 sampai 5. Hasil validasi dari para validator akan dihitung menggunakan Aiken'V, dengan rumus sebagai berikut:

$$V = \frac{\sum S}{n(c - 1)}$$
$$s = r - l_0$$

Keterangan :

lo = Angka penilaian validitas terendah

c = Angka penilaian validitas tertinggi

r = angka yang diberikan oleh penilai

n = jumlah penilai

Setelah dilakukan validasi selanjutnya instrument tes diujikan dalam kelompok kecil (*small group*) kepada 20 orang peserta didik di SMA Negeri 12 Padang. Kemudian hasil dari uji coba yang dilakukan secara langsung ke sekolah, selanjutnya dilakukan analisis butir soal yang terdiri dari validitas empiris, reliabilitas tes, indeks, kesukaran, daya pembeda soal, dan fungsi distraktor. Analisis butir soal menggunakan *Microsoft excel*.

HASIL DAN PEMBAHASAN

Hasil

Menentukan Isi

Berdasarkan analisis yang telah dilakukan terhadap capaian pembelajaran dan tujuan pembelajaran, terdapat materi prasyarat untuk materi sistem periodik unsur, yaitu struktur atom. Terdapat 11 konsep pada materi sistem periodik unsur, 13 konsep pada materi struktur atom. Konsep beserta definisi konsep pada materi sistem periodik unsur dan materi prasyarat dibuatkan ke dalam tabel pernyataan proposisional. Didapatkan angka validitas sebesar 0,88 dan ini dinyatakan valid.

Memperoleh Informasi Tentang Mengidentifikasi Miskonsepsi Pada Peserta Didik

Berdasarkan wawancara yang dilakukan terhadap salah satu guru kimia SMA, diperoleh bahwa banyaknya peserta didik yang mengalami ketidaktuntasan pada materi sistem periodik unsur. Menurut guru, kesulitan yang dihadapi peserta didik disebabkan adanya miskonsepsi pada peserta didik. Miskonsepsi yang paling besar yaitu pada menentukan letak golongan dan periodik didalam tabel periodik unsur dan juga pada materi sifat-sifat keperiodikan unsur. Dari hasil ujian harian disekolah didapatkan sebanyak 70,3% peserta didik mengalami kesulitan pada materi sistem periodik unsur.

Hasil Pengembangan Instrumen Tes Diagnostik

Instrumen es yang dikembangkan merupakan instrumen tes diagnostik pilihan ganda empat tingkat (*four-tier*). Tingkat pertama (*first-tier*) merupakan pertanyaan dengan lima pilihan jawaban, tingkat kedua (*second-tier*) merupakan keyakinan terhadap pilihan jawaban, tingkat ketiga (*third-tier*) merupakan lima alasan terhadap pilihan jawaban pada tingkat pertama, dan tingkat keempat (*fourth-tier*) merupakan keyakinan terhadap pilihan alasan. Soal yang dikembangkan terdiri dari 30 butir soal. Untuk memastikan bahwa instrumen tes diagnostik yang telah dikembangkan sudah mencakup semua konsep-konsep pada pernyataan proposisional dan peta konsep dikembangkan juga kisi-kisi spesifik.

Validitas konten dan konstruk dilakukan oleh ahli dan untuk melihat validitas empiris, reliabilitas, daya pembeda, indeks kesukaran, serta fungsi distraktor dilakukan uji coba pada peserta didik.

Validitas

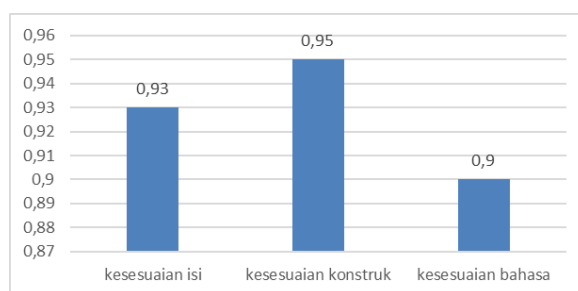
Validitas terbagi menjadi dua, yaitu validitas logis dan validitas empiris. Validitas logis merupakan validitas yang secara hasil penalaran menyatakan instrument penilaian mempunyai validitas logis apabila

sudah direncanakan dengan teratur dan mengikuti teori maupun ketentuan yang berlaku. Validitas isi dan validitas konstruk.

Validitas Logis

Validitas logis diperoleh dari hasil validitas oleh lima orang validator yang terdiri dari tiga orang dosen kimia FMIPA UNP dan dua orang guru kimia SMA dengan tiga aspek yang diukur yaitu kesesuaian isi (konten), kesesuaian penyajian (konstruk), dan kesesuaian Bahasa.

Validitas dilakukan sebanyak dua kali. Karena terdapat beberapa item soal yang belum sesuai. Dikarenakan beberapa hal, antara lain terdapat kesalahan dalam menuliskan redaksi soal serta adanya pilihan jawaban yang belum sesuai dengan konsep keilmuan kimia, serta pilihan jawaban dan alasan yang kurang logis. Kemudian dilakukan revisi terhadap soal yang tidak valid maupun soal yang valid namun terdapat beberapa masukan dari validator dan selanjutnya dilakukan validasi ulang oleh validator sehingga didapatkan soal yang valid secara keseluruhan. Nilai rata-rata validitas logis setelah revisi dapat dilihat pada gambar 2.



Gambar 2. Grafik Rata-Rata Validitas Logis Setelah Revisi

Validitas empiris

Validitas empiris didapatkan dengan melakukan uji coba instrumen tes diagnostik pada peserta didik yang telah mempelajari materi sistem periodik unsur. Untuk menentukan nilai validitas empiris dapat dilakukan rumus korelasi poin biserial (Supardi, 2021). Analisis validitas empiris dilakukan terhadap tingkat pertama (*first-tier*) yaitu pilihan jawaban dan tingkat ketiga (*third-tier*) pilihan alasan. Pada *first-tier* ditemukan soal dengan kategori valid yang memiliki nilai korelasi poin biserial yang tinggi terdapat pada soal nomor 6 yaitu 0,75. Hal ini menyatakan bahwa jawaban yang diberikan peserta didik pada soal ini mewakili pemahaman mereka yang sebenarnya (Habiddin & Page, 2019).

Selain itu ditemukan juga soal dengan kategori tidak valid salah satunya soal nomor 30 yang memiliki nilai validitas sebesar -0,24. Ketidakvalidan soal ini didukung dengan rendahnya nilai daya pembeda yang memperkuat dugaan bahwa banyaknya peserta didik yang menebak-nebak untuk menjawab soal ini, sehingga jawaban yang diberikan peserta didik terhadap pernyataan nomor 30 belum mewakili pemahaman mereka yang sebenarnya. Rekap empiris pada *first-tier* dan *third-tier* dapat dilihat pada tabel 1 dan tabel 2.

Tabel 1. Validitas *first-tier*

No. Soal	Kategori
1,2,3,4,5,6,7,8,9,10,11,12,15,16,17,18,19,20,21,23,24,27,28,29	Valid
13,14,22,25,26,30	Tidak valid

Tabel 2. Validitas *third-tier*

No. Soal	Kategori
1,2,3,4,6,7,8,9,10,12,13,14,15,16,17,18,19,20,22,23,24,25,27,29	Valid

5,11,21,26,28,30

Tidak valid

Reliabilitas

Reliabilitas adalah tingkat atau derajat konsistensi dari suatu instrumen. Suatu tes dikatakan dapat dipercaya jika memberikan hasil yang tetap apabila diteskan secara berulang-ulang (Winarni, S. & Syahrial, 2016). Untuk menentukan nilai reliabilitas dapat digunakan rumus K-R 20 (Supardi, 2021). Nilai reliabilitas untuk tingkat pertama (*first-tier*) dan tingkat ketiga (*third-tier*) sebesar 0,88 dan 0,90 dengan kategori sangat tinggi.

Tingkat Kesukaran, Daya Pembeda dan Fungsi Distraktor

Tingkat kesukaran soal merupakan peluang peserta didik untuk dapat menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Semakin besar angka indeks tingkat kesukaran yang diperoleh, maka semakin mudah soal itu dan begitu sebaliknya. Soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sukar (Widodo, 2021).

Analisis tingkat kesukaran dilakukan pada tingkat pertama (*first-tier*) yaitu pilihan jawaban dan tingkat ketiga (*third-tier*) yaitu pilihan alasan. Pada tingkat pertama terdapat beberapa soal dengan kategori sangat sukar salah satunya dengan indeks kesukaran paling rendah sebesar 0,05 yaitu soal nomor 28. Soal ini tergolong sukar, yang dibuktikan dengan hanya dua orang peserta didik dari kelompok bagian atas yang mampu menjawab soal ini, akan tetapi kedua peserta didik ini belum mampu menjawab soal ini, akan tetapi kedua peserta didik ini belum mampu menjawab dengan alasan yang benar. Hal ini berarti mereka hanya mampu menerapkan pengetahuan konten mereka saja, tetapi belum mampu menerapkan penalaran konseptual dengan baik (Habiddin & Page, 2019). Rekap tingkat kesukaran dapat dilihat pada tabel 3 dan tabel 4.

Tabel 3. Tingkat kesukaran (*first-tier*)

No.Butir Soal	Kategori	Persentase
1,2,8,9,10	Mudah	16,7%
3,4,5,6,7,12,13,14,15,16,17,18,19,21,22,24,29,30	Sedang	60%
11,20,23,25,26,27,28	Sukar	23,3%

Tabel 4. Tingkat kesukaran (*third-tier*)

No.Butir Soal	Kategori	Persentase
1,2	Mudah	6,7%
3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,19,21,23,24,29	Sedang	66,7%
18,20,21,22,25,26,28,30	Sukar	26,6%

Analisis daya pembeda dilakukan untuk melihat apakah suatu butir soal mampu membedakan antara peserta didik yang memiliki kemampuan tinggi dengan peserta didik yang memiliki kemampuan rendah (Widodo, 2021). Daya pembeda dapat dikelompokkan menjadi empat, yaitu jelek, sedang, baik, dan sangat baik. Soal nomor 6 dan nomor 20 merupakan soal dengan daya pembeda sangat baik, nilai daya pembeda pada *first-tier* sebesar 0,70 sedangkan pada *third-tier* sebesar 0,70. Hal ini menunjukkan bahwa soal ini mampu untuk membedakan antara peserta didik yang memiliki kemampuan tinggi dengan kemampuan rendah. Sebaliknya soal nomor 13, 25 dan nomor 30 memiliki daya beda negatif. Ini menunjukkan bahwa soal tersebut belum mampu membedakan antara peserta didik yang memiliki kemampuan tinggi dengan kemampuan rendah (Widodo, 2021). Berikut ini disajikan rekap daya pembeda yang dapat dilihat pada tabel 5.

Tabel 5. Persentase daya pembeda

<i>Tier</i>	Persentase			
	Sangat baik	Baik	Sedang	Jelek
<i>First-tier</i>	56,7%	13,3%	3,3%	26,7%
<i>Third-tier</i>	53,3%	16,6%	10%	20%

Distraktor bertujuan untuk mengecoh mereka yang kurang mampu (tidak tahu) untuk dibedakan dengan yang mampu (lebih tahu) (Amri & Syahriani, 2018). Pada tingkat pertama (*first-tier*) dan tingkat ketiga (*third-tier*), distraktor pada semua soal berfungsi dengan baik, hal ini dapat dilihat bahwa pada setiap soal lebih dari 5% peserta didik memilih jawaban selain dari kunci jawaban.

Pembahasan

Tujuan dari penelitian ini adalah untuk menghasilkan instrumen tes diagnostik *four-tier* untuk mengidentifikasi miskonsepsi peserta didik pada materi sistem periodik unsur fase E SMA/MA yang valid dan reliabel. Penelitian ini merupakan penelitian pengembangan atau *Research and Development* (R&D) dan menggunakan model pengembangan yang dikembangkan oleh Treagust (1988) yang terdiri dari tiga tahap.

Berdasarkan hasil penelitian, tabel pernyataan proposisional yang berisi konsep-konsep pada materi sistem periodik unsur beserta materi prasyarat dan disusun ke dalam sebuah peta konsep, dimana keduanya saling berhubungan dinyatakan valid. Hal ini berarti tabel pernyataan proposisional dan peta konsep tersebut sudah dibuat sesuai dengan konsep/teori yang ada sehingga dapat dijadikan sebagai acuan dalam mengembangkan instrumen tes diagnostik.

Berdasarkan wawancara yang telah dilakukan terhadap beberapa orang guru kimia SMA, diperoleh bahwa banyaknya peserta didik yang mengalami ketidaktuntasan pada materi sistem periodik unsur. Menurut guru, miskonsepsi yang dihadapi oleh peserta didik paling besar yaitu menentukan letak golongan dan periodik didalam tabel periodik unsur dan juga pada materi sifat-sifat keperiodikan unsur. Materi sistem periodik unsur merupakan materi yang paling dasar dalam ilmu kimia sehingga harus menguasainya secara menyeluruh agar materi kimia selanjutnya tidak menimbulkan kesulitan. Karena materi sistem periodik unsur merupakan materi yang memiliki banyak istilah dan kosakata baru, Pokok bahasannya mengenai perkembangan sistem periodik unsur, letak unsur dalam tabel periodik unsur, sifat-sifat keperiodikan unsur diantaranya yaitu jari-jari atom, energi ionisasi, keelektronegatifan, dan afinitas elektron.

Pada penelitian ini telah dikembangkan instrumen tes diagnostik pilihan ganda empat tingkat. Tingkat pertama (*first-tier*) merupakan pertanyaan dengan lima pilihan ganda, tingkat kedua (*second-tier*) merupakan keyakinan terhadap pilihan jawaban, tingkat ketiga (*third-tier*) merupakan alasan terdapat pilihan jawaban pada tingkat pertama, tingkat keempat (*fourth-tier*) merupakan keyakinan terhadap pilihan alasan.

Instrumen yang baik adalah instrumen tes yang valid, reliabel, memiliki daya pembeda yang baik, tingkat kesukaran yang sedang (tidak terlalu mudah dan tidak terlalu sukar), dan memiliki fungsi distraktor yang baik. Untuk menghasilkan instrumen tes diagnostik yang baik, dilakukan penyempurnaan lanjutan yaitu analisis soal. Analisis pada umumnya dilakukan melalui dua acara, yaitu analisis kualitatif dan analisis kuantitatif (Ambiyar, 2012). Setelah dianalisis, masih terdapat item pernyataan belum sesuai dengan konsep ilmiah berdasarkan kesesuaian isi, konstruk, maupun Bahasa. Hal ini disebabkan karena terdapat beberapa pernyataan soal yang salah, pilihan jawaban dan alasan yang tidak logis, serta Bahasa yang digunakan kurang jelas dan ambigu. Oleh karena itu, dilakukan revisi, revisi ini tidak hanya dilakukan pada soal yang tidak valid tetapi juga pada soal yang sudah valid karena terdapat beberapa masukan dari validator yang dapat membuat soal tersebut menjadi lebih baik. Setelah soal direvisi, dilakukan validasi ulang kepada validator. Data hasil validasi ulang dianalisis kembali dan didapatkan hasil bahwa seluruh item soal memperoleh nilai validitas $\geq 0,8$ dan dinyatakan valid secara logis, hal ini menyatakan bahwa instrumen tes diagnostik yang dikembangkan telah sesuai dengan konsep/teori maupun ketentuan yang berlaku (Widodo, 2021).

Analisis kuantitatif menekankan pada analisis karakteristik internal tes melalui data yang diperoleh secara empiris. Karakteristik internal secara kuantitatif dimaksudkan meliputi parameter validitas empiris, reliabilitas, indeks kesukaran, daya beda, dan fungsi distraktor (Ambiyar, 2012). Berdasarkan hasil penelitian, validitas empiris yang terdapat pada lampiran 19. Pada *first-tier* ditemukan soal dengan kategori valid yang memiliki nilai korelasi poin biserial yang tinggi terdapat pada soal nomor 6 yaitu 0,75. Hal ini menyatakan bahwa jawaban yang diberikan peserta didik pada soal ini mewakili pemahaman mereka yang sebenarnya. Selain itu ditemukan juga soal dengan kategori tidak valid salah satunya soal nomor 30 yang memiliki nilai validitas sebesar -0,24. Ketidakvalidan soal ini didukung dengan rendahnya nilai daya pembeda yang memperkuat dugaan bahwa banyaknya peserta didik yang menebak-nebak untuk menjawab soal ini, sehingga jawaban yang diberikan peserta didik terhadap pernyataan nomor 30 belum mewakili pemahaman mereka yang sebenarnya.

Berdasarkan hasil penelitian, instrumen tes diagnostik yang telah dikembangkan memiliki reliabilitas sangat tinggi, hal ini menunjukkan bahwa instrumen tes tersebut memiliki tingkat konsistensi yang tinggi. Reliabilitas adalah tingkat atau derajat konsistensi dari suatu instrumen. Suatu tes dikatakan dapat dipercaya jika memberikan hasil yang tetap apabila diteskan secara berulang-ulang (Amri & Syahrani, 2018).

Tingkat kesukaran soal merupakan peluang peserta didik untuk dapat menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Semakin besar angka indeks tingkat kesukaran yang diperoleh, maka semakin mudah soal itu dan begitu sebaliknya. Soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sukar (Widodo, 2021). Hasil penelitian menunjukkan terdapat beberapa soal dengan kategori sangat sukar salah satunya dengan indeks kesukaran paling rendah sebesar 0,05 yaitu soal nomor 28. Soal ini tergolong sukar, yang dibuktikan dengan hanya dua orang peserta didik dari kelompok bagian atas yang mampu menjawab soal ini, akan tetapi kedua peserta didik ini belum mampu menjawab soal ini, akan tetapi kedua peserta didik ini belum mampu menjawab dengan alasan yang benar. Hal ini berarti mereka hanya mampu menerapkan pengetahuan konten mereka saja, tetapi belum mampu menerapkan penalaran konseptual dengan baik (Habiddin & Page, 2019). Berikut ini disajikan diagram persentase tingkat kesukaran instrumen tes diagnostik.

Berdasarkan hasil penelitian, terdapat soal dengan kategori daya pembeda sangat baik, salah satu nomor 6 dan nomor 20 dengan nilai daya pembeda pada *first-tier* yaitu sebesar 0,70 sedangkan pada *third-tier* sebesar 0,70. Hal ini menunjukkan bahwa soal ini mampu untuk membedakan antara peserta didik yang memiliki kemampuan tinggi dengan kemampuan rendah. Sebaliknya soal nomor 13, 25 dan nomor 30 memiliki daya beda negatif. Ini menunjukkan bahwa soal tersebut belum mampu membedakan antara peserta didik yang telah menguasai materi yang dinyatakan dengan peserta didik yang belum menguasai materi yang ditanyakan. Semakin besar nilai daya beda maka semakin baik daya beda suatu soal (Ambiyar, 2012).

Distraktor bertujuan untuk mengecoh mereka yang kurang mampu (tidak tahu) untuk dibedakan dengan yang mampu (lebih tahu). 1) apabila keempat jawaban pengecoh berguna maka soal dinyatakan efektivitas yang sangat baik, 2) apabila ditemukan tiga jawaban pengecoh yang berguna maka soal dinyatakan mempunyai efektivitas pengecoh yang baik, 3) apabila ditemukan dua jawaban pengecoh yang berguna maka soal dinyatakan mempunyai efektivitas pengecoh yang cukup baik, 4) apabila ditemukan 1 jawaban pengecoh yang berguna maka soal dinyatakan mempunyai efektivitas pengecoh yang tidak baik (Amri & Syahrani, 2018). Berdasarkan hasil penelitian, pada tingkat pertama (*first-tier*) dan tingkat ketiga (*third-tier*), distraktor pada semua soal berfungsi dengan baik, hal ini dapat dilihat bahwa pada setiap soal lebih dari 5% peserta didik memilih jawaban selain dari kunci jawaban.

Keterbatasan pada penelitian ini adalah jumlah sampel yang digunakan 20 siswa yang sampai pada tahap ujicoba skala kecil karena berdasarkan kepada tujuan utama penelitian ini adalah untuk mengembangkan suatu produk instrumen tes diagnostik *four-tier* untuk mengidentifikasi miskonsepsi pada materi sistem periodik unsur fase E SMA/MA. Pada uji coba skala kecil (*small group*) instrumen tes memiliki

reliabilitas yang sangat tinggi pada pertanyaan soal (*first tier*) dan alasan soal (*four-tier*), serta indeks kesukaran yang dapat berfungsi dengan baik, daya beda soal yang dapat membedakan siswa dengan kemampuan yang tinggi dan rendah, serta fungsi distraktor yang baik.

SIMPULAN

Berdasarkan penelitian yang telah dihasilkan instrumen tes diagnostik *four-tier* untuk mengidentifikasi miskonsepsi peserta didik pada materi sistem periodik unsur memiliki validitas logis dan empiris, serta reliabilitas soal yang sangat tinggi, serta menghasilkan analisis butir soal pada indeks kesukaran, daya pembeda yang efektif dan baiki digunakan serta fungsi distraktor yang baik.

DAFTAR PUSTAKA

- Agustin, U., Susilaningih, E., Nurhayati, S. N., & Wijayati, N. (2022). Pengembangan Instrumen Tes Diagnostik Four-Tier Multiple Choice Untuk Identifikasi Miskonsepsi Siswa Pada Materi Keseimbangan Kimia. *Journal Of Chemistry In Education*, 2(3), 133–139.
- Ambiyar. (2012). *Pengukuran Dan Tes Dalam Pendidikan*. Unp Press Padang.
- Amri, A., & Syahriani, S. (2018). Pengembangan Tes Sumatif Mata Pelajaran Biologi Semester Ganjil Kelas Ix. *Jurnal Biotek*, 6(1), 53. <https://doi.org/10.24252/Jb.V6i1.5072>
- Devi, N. A., & Azra, F. (2023). Pengembangan Instrumen Tes Diagnostik Untuk Melihat Gambaran Model Mental Peserta Didik Pada Materi Asam Basa. *Entalpi Pendidikan Kimia*, 16–26. <https://doi.org/10.24036/Epk.V4i3.345>
- Diannisa, N. F. Dk. (2023). Deskripsi Pemahaman Konsep Pada Materi Struktur Atom Di Kelas X Sma Negeri 01 Ngabang. *Kependidikan Kimia*, 11(August), 497–512.
- Habiddin, & Page, E. M. (2019). Development And Validation Of A Four-Tier Diagnostic Instrument For Chemical Kinetics (Ftdick). *Indonesian Journal Of Chemistry*, 19(3), 720–736. <https://doi.org/10.22146/Ijc.39218>
- Halimatusya, S., Helsy, I., & Sari. (2021). Pembuatan Media Pembelajaran Uno Spin Kimia Pada Materi Sistem Periodik Unsur Manufacture Of Chemical Uno Spin Learning Media On Elements Periodic System Materials. *Gunung Djati Conference Series*, 2, 234–343.
- Muderawan, I. W., Wiratma, I. G. L., & Nabila, M. Z. (2019). Analisis Faktor-Faktor Penyebab Kesulitan Belajar Siswa Pada Materi Kelarutan Dan Hasil Kali Kelarutan. *Jurnal Pendidikan Kimia Indonesia*, 3(1), 17. <https://doi.org/10.23887/Jpk.V3i1.20944>
- Nurhafsari, A., & Rismaningsih, F. (2023). *DEVELOPMENT OF A FOUR TIER DIAGNOSTIC TEST INSTRUMENT BASED ON ISPRING SUITE 9 TO IDENTIFY UNDERSTANDING OF CONCEPTS IN STUDENTS PENGEMBANGAN INSTRUMEN FOUR TIER DIAGNOSTIC TEST BERBASIS ISPRING SUITE 9 UNTUK MENG.* 5(2).
- Ramadhan, Y., Nisa, K. R., & Sunarwin, S. (2020). Analysis Of Students Misconception Using Certainly Of Response Index (Cri) In The Periodic System Of Elements Concept. *Educhemia (Jurnal Kimia Dan Pendidikan)*, 5(2), 210. <https://doi.org/10.30870/Educhemia.V5i2.8285>
- Ritonga, P. S., & Yasthophi, A.-. (2019). Pengembangan Instrumen Test Diagnostik Multiple Choice Four Tier Pada Materi Ikatan Kimia. *Konfigurasi : Jurnal Pendidikan Kimia Dan Terapan*, 3(1), 23. <https://doi.org/10.24014/Konfigurasi.V3i1.6797>
- Rozi, F., & Romadhoni, D. S. (2023). Edu Komputika Journal Media Pembelajaran Sistem Periodik Unsur Berbasis Android Dengan Metode Self Directed Learning. *Edu Komputika*, 10(1), 27–37. <http://journal.unnes.ac.id/sju/index.php/edukom>
- Rusilowati, A. (2015). Pengembangan Tes Diagnostik Sebagai Alat Evaluasi Kesulitan Belajar Fisika.

- 3553 *Pengembangan Four-Tier Test untuk Mengidentifikasi Miskonsepsi pada Materi Sistem Periodik Unsur Fase E SMA/MA - Umi Salamah, Andromeda*
DOI : <https://doi.org/10.31004/edukatif.v6i4.7397>
- Prosiding Seminar Nasional Fisika Dan Pendidikan Fisika*, 6(1), 1–10.
- Sariati, Kadek, N., Suardana, Nyoman, I., Wiratini, & Made, N. (2020). Analisis Kesulitan Belajar Kimia Siswa Kelas Xi Pada Materi Larutan Penyangga. *Jurnal Imiah Pendidikan Dan Pembelajaran P-Issn : 1858-4543 E-Issn : 2615-6091*, 4(1), 86–97. <https://doi.org/10.23887/jipp.v4i1.15469>.
- Savira, I., Wardani, S., Harjito, & Noorhayati, A. (2019). Desain Instrumen Tes Three Tiers Multiple Choice Untuk Analisis Miskonsepsi Siswa Terkait Larutan Penyangga. *Jurnal Inovasi Pendidikan Kimia*, 13(1), 2277–2286. <https://journal.unnes.ac.id/nju/index.php/jipk/article/view/15924>
- Supardi, G. (2021). *Statistik Penelitian Pendidikan*. Uny Press.
- Treagust, D. F. (1988). Development And Use Of Diagnostic Tests To Evaluate Students' Misconceptions In Science. *International Journal Of Science Education*, 10(2), 159–169. <https://doi.org/10.1080/0950069880100204>
- Tumanggor, A. M. R., Supahar, Kuswanto, H., & Ringo, E. S. (2020). Using Four-Tier Diagnostic Test Instruments To Detect Physics Teacher Candidates' Misconceptions: Case Of Mechanical Wave Concepts. *Journal Of Physics: Conference Series*, 1440(1). <https://doi.org/10.1088/1742-6596/1440/1/012059>
- Widodo, H. (2021). *Evaluasi Pendidikan*. Uad Press.
- Winarni, S. & Syahril. (2016). Miskonsepsi Kimia Yang Disebabkan Pernyataan Nonproposisi. *Jurnal Pendidikan Sains*, 4(4), 122–129.