

Pengembangan Instrumen Tes Diagnostik *Four-Tier* untuk Melihat Gambaran Model Mental Peserta Didik pada Materi Asam Basa

Novi A. Devi¹ and Fajriah Azra^{1*}

¹Program Studi Pendidikan Kimia, Fakultas Matematika dan IPA, Universitas Negeri Padang, Sumatera Barat, Indonesia.

*Email: bunda_syasfa@yahoo.com

ABSTRACT

Mental models is a student's personal mental representation of an idea or concept. Mental models represent ideas in an individual's mind that are used to describe and explain a phenomenon. The mental models of students is very important for the teacher to know because it can determine the level of understanding, difficulties, and misconceptions experienced by students. Mental models of students can be known by giving a diagnostic test. Therefore, this study aims to develop a four-tier diagnostic test instrument to see of students' mental models in acid-base material that meet the criteria for a good test instrument. This research is a research and development (R&D) study using the Treagust (1988) development model. Based on the results of the validation of five validators, the diagnostic test instrument was declared valid in terms of logic with an average value of 0.84. Based on the results of trials on students, the reliability value of the diagnostic test instrument was obtained at the first-tier and third-tier is 0.84 and 0.83 with the very high category. The results showed that the developed diagnostic test instrument was valid and has the criteria for a good test instrument.

Keywords: Diagnostic Test, Mental Models, Four-tier Multiple Choice, Acid Base

ABSTRAK

Model mental merupakan representasi mental pribadi peserta didik terhadap suatu ide atau konsep. Model mental mewakili ide-ide dalam pikiran individu yang digunakan untuk menggambarkan dan menjelaskan suatu fenomena. Model mental peserta didik sangat penting untuk diketahui oleh guru karena dapat mengetahui tingkat pemahaman, kesulitan, dan miskonsepsi yang dialami peserta didik. Model mental peserta didik dapat diketahui dengan memberikan suatu tes diagnostik. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan instrumen tes diagnostik *four-tier* untuk melihat gambaran model mental peserta didik pada materi asam basa yang memenuhi kriteria instrumen tes yang baik. Penelitian ini merupakan penelitian *research and development* (R&D) dengan menggunakan model pengembangan Treagust (1988). Berdasarkan hasil validasi terhadap lima orang validator, instrumen tes diagnostik dinyatakan valid secara logis dengan nilai rata-rata 0,84. Berdasarkan hasil uji coba terhadap peserta didik, didapatkan nilai reliabilitas instrumen tes diagnostik pada tingkat pertama (*first-tier*) dan tingkat ketiga (*third-tier*) berturut-turut adalah sebesar 0,84 dan 0,83 dengan kategori sangat tinggi. Hasil penelitian menunjukkan bahwa instrumen tes diagnostik yang dikembangkan sudah valid dan sudah memiliki kriteria instrumen tes yang baik.

Kata Kunci: Tes Diagnostik, Pilihan Ganda *Four-tier*, Model Mental, Materi Asam Basa

PENDAHULUAN

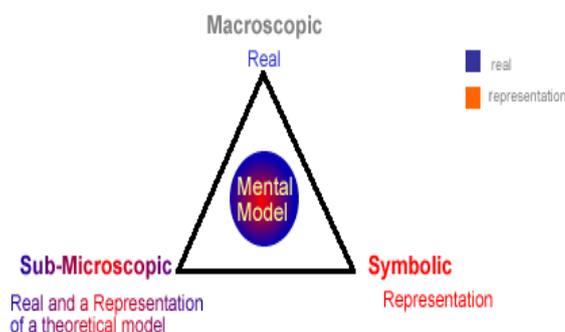
Kimia adalah ilmu yang mempelajari tentang materi dan sifat-sifatnya, perubahan yang dialami materi, serta energi yang menyertai perubahan materi tersebut (Silberberg, 2010). Selain mempelajari perubahan materi, kimia juga mempelajari energi yang menyertai perubahan tersebut. Dengan kata lain, kimia adalah ilmu yang sarat dengan konsep, mulai dari konsep yang sederhana hingga yang lebih kompleks, dan dari konsep yang konkrit hingga abstrak. Hal ini membuat sebagian besar peserta didik sulit untuk memahami konsep (Eliyawati dkk., 2020). Konsep kimia yang abstrak dapat lebih mudah untuk dipahami dengan melibatkan tiga level representasi kimia yaitu representasi makroskopik, submikroskopik, dan simbolik (Permatasari dkk., 2022). Pembelajaran kimia hendaknya ditekankan pada tiga level representasi kimia (makroskopik, submikroskopik, dan simbolik). Pemahaman seseorang terhadap konsep kimia ditentukan oleh kemampuannya dalam mentransfer dan menghubungkan ketiga level representasi tersebut (Sagita dkk., 2017)

Representasi makroskopik yaitu menggambarkan fenomena-fenomena nyata yang dapat dilihat secara kasat mata oleh peserta didik yang ditemui dalam kehidupan sehari-hari, misalnya ketika mengamati perubahan sifat materi (adanya perubahan warna, pH suatu larutan, pembentukan gas dan endapan dalam suatu reaksi kimia). Representasi submikroskopik merupakan penjelasan dari keadaan yang terjadi pada representasi makroskopik yang dijelaskan pada tingkat partikulat dimana materi digambarkan tersusun dari atom-atom, molekul, dan ion-ion. Representasi

simbolik yaitu representasi yang melibatkan penggunaan rumus senyawa, persamaan reaksi kimia, rumus dalam perhitungan kimia, serta gambar struktur molekul. (Chandrasegaran dkk., 2007).

Model mental merupakan representasi mental pribadi peserta didik terhadap suatu ide atau konsep (Chittleborough, 2004). Model mental mewakili ide-ide dalam pikiran individu yang digunakan untuk menggambarkan dan menjelaskan fenomena. Model mental dibangun dari persepsi, imajinasi, dan pemahaman wacana. (Jansoon dkk., 2009).

Pemahaman peserta didik terhadap representasi simbolik serta pengalaman peserta didik dengan level representasi makroskopik dan submikroskopik berkontribusi pada pengembangan model mental mereka terhadap suatu fenomena ilmiah. Hal ini digambarkan pada Gambar 1, dimana model mental merupakan produk dan pusat dari segitiga. Model mental peserta didik dapat dituangkan dan tercermin pada kemampuan interpretasi dari tiga level representasi kimia (Chittleborough, 2004).



Gambar 1. Hubungan tiga level representasi kimia dengan model mental (Chittleborough 2004).

Jika peserta didik memahami peran setiap tingkat representasi kimia, dan kemudian dapat mentransfer pengetahuan dari satu tingkat ke tingkat lainnya, artinya mereka mampu menghasilkan penjelasan yang dapat dipahami dan benar secara ilmiah, sehingga mengurangi konsepsi alternatif (Treagust dkk., 2003).

Model mental peserta didik penting untuk diketahui dan diteliti karena terdapat dua alasan. Pertama, model mental dapat berpengaruh terhadap fungsi kognitif peserta didik. Kedua, model mental dapat memberikan informasi yang berharga untuk para peneliti pendidikan sains tentang konsep-konsep yang dimiliki peserta didik (Laliyo, 2011). Selain itu, mengidentifikasi model mental peserta didik sangat penting bagi guru untuk mengetahui tingkat pemahaman, kesulitan, dan miskonsepsi yang dialami peserta didik (Amalia dkk., 2018) sehingga guru dapat memilih strategi, metode, serta media yang tepat (Darmiyanti dkk., 2017).

Menurut Wang (2007), model mental peserta didik dapat diketahui melalui pemberian suatu tes yang disebut sebagai tes diagnostik model mental. Tes diagnostik model mental adalah tes yang digunakan untuk melihat kemampuan peserta didik dalam menginterpretasikan data, memprediksi, dan menjelaskan fenomena kimia secara makroskopik dengan melibatkan penalaran pada tingkat submikroskopik serta menuliskan dalam bentuk simbolik atau persamaan reaksi kimia. Pemberian tes diagnostik model mental diharapkan dapat mengukur pemahaman peserta didik tentang konsep-konsep kimia (Coll & Taylor, 2002).

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 dkk., 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 serta masih sedikit menerapkan multirepresentasi kimia terutama pada materi asam basa.

Berdasarkan uraian di atas, maka perlu dikembangkan instrumen tes diagnostik *four-tier* untuk melihat gambaran model mental peserta didik pada materi asam basa.

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 yang dapat dilihat pada Lampiran 1 dan diuraikan sebagai berikut:

1. Menentukan Isi

Pada tahap ini, terdapat empat langkah, diantaranya a) mendefinisikan pernyataan pengetahuan proposisional

yang dilakukan dengan cara menganalisis capaian pembelajaran dan tujuan pembelajaran pada materi asam basa 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 konten pernyataan proposisional dan peta konsep oleh ahli.

2. Memperoleh Informasi Tentang Model Mental Peserta Didik

Pada langkah ini dilakukan dua kegiatan, yaitu a) melakukan wawancara dengan salah satu guru kimia di SMA Negeri 1 Tilatang Kamang; b) melakukan studi literatur terkait model mental peserta didik.

3. 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. Pada penelitian ini, tahap penyempurnaan lanjutan dilakukan analisis kualitatif dan kuantitatif. Analisis kualitatif dilakukan untuk melihat validitas logis dari instrumen tes yang telah dikembangkan dengan melakukan validasi terhadap ahli. Sedangkan analisis kuantitatif, dilakukan uji coba skala kecil yang dilakukan terhadap 23 orang peserta didik di SMA Negeri 1 Tilatang Kamang untuk melihat validitas empiris, reliabilitas tes, indeks, kesukaran, daya pembeda soal, dan fungsi distraktor.

Subjek dari penelitian ini adalah peserta didik yang telah mempelajari materi asam basa, guru kimia, dan dosen kimia

FMIPA UNP. Objek penelitian pada penelitian ini yaitu Instrumen Tes Diagnostik *Four-Tier* Untuk Melihat Gambaran Model Mental Peserta Didik Pada Materi Asam Basa.

HASIL DAN DISKUSI

Menentukan Isi

Berdasarkan analisis yang telah dilakukan terhadap capaian pembelajaran dan tujuan pembelajaran, terdapat beberapa materi prasyarat untuk materi asam basa, diantaranya reaksi kimia, konsep mol, larutan elektrolit dan non elektrolit, serta kesetimbangan kimia. Terdapat 16 konsep pada materi asam basa, 2 konsep terkait reaksi kimia, 2 konsep terkait konsep mol, 3 konsep terkait larutan elektrolit dan non elektrolit, serta 1 konsep terkait kesetimbangan kimia. Konsep beserta definisi konsep pada materi asam basa dan materi prasyarat dibuatkan ke dalam tabel pernyataan proposisional.

Konsep-konsep yang terdapat pada tabel pernyataan proposisional, disusun ke dalam sebuah peta konsep. Peta konsep ini dihubungkan dengan pernyataan proposisional, hal ini bertujuan untuk memastikan peta konsep yang dibuat sudah memuat konsep-konsep yang terdapat pada pernyataan pengetahuan proposisional sehingga instrumen tes diagnostik yang dikembangkan selalu relevan dan sesuai dengan kurikulum merdeka yang digunakan yang ini.

Selanjutnya, pernyataan proposisional dan peta konsep yang telah dirancang, divalidasi oleh empat orang dosen kimia FMIPA UNP dan satu orang guru kimia SMA. Validasi dilakukan dengan mengisi angket validasi, dimana angket ini berisi 11 pertanyaan yang berkaitan dengan kesesuaian terhadap capaian pembelajaran dan tujuan pembelajaran, kesesuaian antara fakta/konsep/prinsip berdasarkan keilmuan kimia, dan dapat membantu peserta didik

dalam menyelesaikan tes diagnostik model mental. Data hasil validasi dianalisis dengan menggunakan rumus Aiken's dan didapatkan nilai rata-rata validitas sebesar 0,83, yang menunjukkan bahwa pernyataan proposisional dan peta konsep dikembangkan telah valid. Jika validator berjumlah 5 orang dengan skala 5 maka nilai dari koefisien validitas yaitu 0,8 (Aiken, 1985).

Memperoleh Informasi Tentang Model Mental Peserta Didik

Berdasarkan wawancara yang telah dilakukan terhadap beberapa orang guru kimia SMA, diperoleh bahwa banyaknya peserta didik yang mengalami ketidaktuntasan pada materi asam basa. Menurut guru, kesulitan yang dihadapi oleh peserta didik paling besar yaitu menerapkan stoikiometri dalam menentukan pH larutan asam basa. Hal ini dapat disimpulkan bahwa peserta didik masih belum mampu untuk merepresentasikan level simbolik. Siswa masih terpaku pada rumus-rumus yang tersedia pada bahan ajar sehingga belum terjadi pembelajaran yang bermakna (Sagita dkk., 2017). Ketidakmampuan peserta didik dalam merepresentasikan salah satu dari tiga level representasi kimia, maka dapat menghambat kemampuan untuk memecahkan masalah kimia untuk level representasi lainnya (Eliyawati dkk., 2020). Dari hasil penelitian yang telah dilakukan oleh Atika dan Latisma (2022) juga ditemukan bahwa sebanyak 72% peserta didik mengalami kesulitan dalam menganalisis sifat larutan berdasarkan teori asam basa menurut Arrhenius, Bronsted-Lowry, dan Lewis.

Hasil Pengembangan Instrumen Tes Diagnostik

Instrumen tes 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 alasan terhadap pilihan jawaban, dan tingkat keempat (*four-tier*) merupakan keyakinan terhadap pilihan alasan. Soal yang dikembangkan terdiri dari 27 butir soal berbasis multirepresentasi kimia yaitu adanya level makroskopik, submikroskopik, dan simbolik. 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.

Instrumen tes 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 (Putri dkk., 2022). 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 dan Reliabilitas

Validitas terbagi menjadi dua, yaitu validitas logis dan validitas empiris. Validitas logis merupakan validitas yang secara hasil penalaran menyatakan instrumen penilaian mempunyai validitas logis apabila sudah direncanakan dengan teratur dan mengikuti teori maupun ketentuan yang berlaku. Validitas logis ada dua macam, yaitu validitas isi dan validitas konstruk.

Suatu instrumen tes disebut memiliki validitas isi apabila memiliki kesesuaian antara materi atau isi pembelajaran dengan tujuan pembelajaran yang diharapkan. validitas isi suatu instrumen penilaian memiliki tujuan untuk mengukur taraf penguasaan terhadap suatu materi berdasarkan tujuan pembelajaran. Validitas

konstruk merupakan soal tes yang diukur validitasnya berdasarkan konstruk pada perangkat pembelajaran yang membangun aspek kognitif seperti pengetahuan dan pemahaman (Kurniawan dkk., 2022).

Validitas logis diperoleh dari hasil validasi oleh lima orang validator yang terdiri dari empat orang dosen kimia FMIPA UNP dan satu orang guru kimia SMA dengan tiga aspek yang diukur yaitu kesesuaian isi (konten), kesesuaian penyajian (konstruk), dan kesesuaian bahasa. Menurut Yusrizal dan Rahmati (2020), untuk mengetahui validitas yang didasarkan pada hasil penilaian dari ahli sebanyak n orang terhadap suatu item dari sejauh mana item tersebut mewakili konstruk yang diukur, dapat menggunakan formula rumus *Aiken's*.

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

$$s = r - l_0$$

Keterangan:

V = indeks validitas isi *Aiken* (*Aiken's V*)

n = banyaknya penilai (rater)

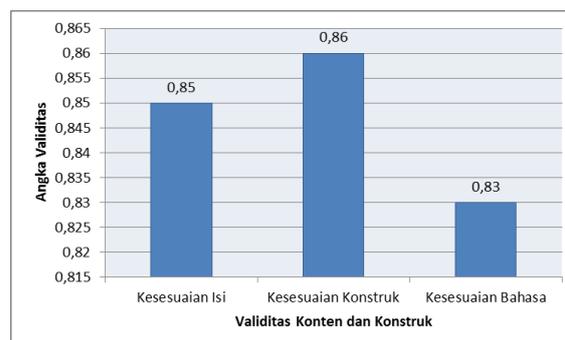
l_0 = angka penilaian terendah (misal 1)

c = angka penilaian tertinggi (misal 5)

r = angka yang diberikan oleh penilai

Validasi dilakukan sebanyak dua kali, karena terdapat beberapa item soal yang tidak valid dengan angka validitas $< 0,8$. Butir soal yang tidak valid ini dikarenakan beberapa hal, antara lain terdapat kesalahan dalam menuliskan redaksi soal serta adanya pilihan jawaban yang belum sesuai dengan konsep keilmuan kimia, gambar submikroskopik yang kurang tepat, 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 didapatkan dengan melakukan uji coba instrumen tes diagnostik pada peserta didik yang telah mempelajari materi asam basa. Untuk menentukan nilai validitas empiris dapat digunakan rumus korelasi point biserial (Supardi, 2017). 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 paling tinggi terdapat pada soal nomor 14 yaitu sebesar 0,66. 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 2 yang memiliki nilai validitas sebesar -0,56. 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 pertanyaan nomor 2 belum mewakili pemahaman mereka yang sebenarnya. Rekap validitas empiris pada *first-tier* dan *third-tier* dapat dilihat pada Tabel 1 dan Tabel 2.

Tabel 1. Validitas *Firs-tier*

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

Tabel 2. Validitas *Third-tier*

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

Relibilitas merupakan tingkat atau derajat konsistensi dari suatu instrumen. Suatu instrumen dapat dipercaya jika memberikan hasil yang tetap apabila tes tersebut dilakukan secara berulang-ulang (Amri & Syahriani, 2018). Untuk menentukan nilai reliabilitas dapat digunakan rumus K-R 20 (Supardi, 2017). Nilai reliabilitas untuk tingkat pertama (*first-tier*) dan tingkat ketiga (*third-tier*) sebesar 0,84 dan 0,83 dengan kategori reliabilitas sangat tinggi.

Tingkat Kesukaran, Daya Pembeda, dan Fungsi Distraktor

Tingkat kesukaran merupakan besarnya peluang bagi peserta didik untuk dapat menjawab benar suatu butir soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk angka indeks. Semakin besar angka indeks tingkat kesukaran yang diperoleh, maka semakin mudah soal itu dan begitu sebaliknya, semakin kecil angka indeks yang diperoleh maka semakin sukar soal tersebut (Supriyadi, 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 0,09 yaitu soal nomor 16. 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 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
1,2,3,4,5,10,22	Mudah	25,9%
6,7,8,9,11,12,13,14,15,17,18,19,23,25,26,27	Sedang	59,3%
16,20,21,24	Sukar	14,8%

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

No. Butir Soal	Kategori	Persentase
5,8,10,13,17,19,20,22,23,24,26	Mudah	40,7%
1,2,3,4,6,7,9,11,12,14,15,18,25	Sedang	48,2%
16,21,27	Sukar	11,1%

Analisis daya pembeda dilakukan untuk melihat apakah suatu butir soal mampu membedakan antara peserta didik didik yang memiliki kemampuan tinggi dengan peserta didik yang memiliki kemampuan rendah (Faradilla dkk., 2020). Daya pembeda dapat dikelompokkan menjadi empat, yaitu jelek, sedang, baik, dan sangat baik. Soal nomor 6 merupakan salah satu soal dengan daya pembeda sangat baik, nilai daya pembeda pada *first-tier* sebesar 0,73 sedangkan pada *third-tier* sebesar 0,64. Hal ini menunjukkan bahwa soal ini mampu untuk membedakan antara peserta didik yang memiliki kemampuan tinggi dengan kemampuan rendah. Sebaliknya soal nomor 2 memiliki daya beda negatif yang menunjukkan menunjukkan bahwa soal tersebut belum mampu membedakan antara peserta didik yang memiliki kemampuan tinggi dengan kemampuan rendah. Berikut ini disajikan

rekap daya pembeda yang dapat dilihat pada Tabel 5.

Tabel 5. Persentase Daya Pembeda

Tier	Persentase			
	Sangat baik	Baik	Sedang	Jelek
First-tier	55,6%	3,7%	14,8%	25,9%
Third-tier	40,7%	18,6%	11,1%	29,6%

Distraktor atau pengecoh dibuat untuk mengecoh peserta didik yang kurang mampu (tidak tahu) dan untuk membedakan dengan peserta didik yang mampu (lebih tahu) (Amri & Syahriani, 2018). Hasil analisis distraktor pada instrumen tes menunjukkan pada tingkat pertama (*first-tier*) distraktor pada semua soal berfungsi dengan baik, hal ini dikarenakan pada setiap soal lebih dari 5% peserta didik memilih jawaban selain dari kunci jawaban. Sedangkan pada *third-tier*, terdapat satu soal yang memiliki fungsi distraktor yang tidak efektif yaitu soal nomor 10. Pada soal ini peserta didik yang memilih jawaban selain kunci jawaban kurang dari 5%, yang artinya sebagian besar peserta didik memilih jawaban yang benar.

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, dihasilkan instrumen tes diagnostik *four-tier* untuk melihat gambaran model mental peserta didik yang valid dan reliabel dengan rata-rata nilai Aiken's V sebesar 0,84 dengan kategori sangat tinggi dan reliabilitas instrumen tes diagnostik pada tingkat pertama (*first-tier*) dan tingkat ketiga (*third-tier*) sebesar 0,84 dan 0,83 serta memiliki daya pembeda dan indeks kesukaran yang baik.

REFERENSI

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of

ratings, Educational and Psychological Measurement. *Journal Articles; Reports - Research; Numerical/Quantitative Data*, 45(1), 131–142.

Amalia, F. R., Ibnu, S., Widarti, H. R., & Wuni, H. (2018). Students' mental models of acid and base concepts taught using the cognitive apprenticeship learning model. *Jurnal Pendidikan IPA Indonesia*, 7(2), 187–192.

<https://doi.org/10.15294/jpii.v7i2.14264>

Amri, & Syahriani. (2018). Pengembangan Tes Sumatif Mata Pelajaran Biologi. *Jurnal Biotek*, 6(01), 53–64.

Atika, M. D., & Latisma, L. (2022). Entalpi Pendidikan Kimia Deskripsi Kesulitan Belajar Kimia Siswa pada Materi Asam Basa Description of Student Chemistry Learning Difficulties in Acid Base Entalpi Pendidikan Kimia. *Entalpi Pendidikan Kimia*.

Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice*, 8(3), 293–307.

<https://doi.org/10.1039/B7RP90006F>

Chittleborough, G. D. (2004). The Role of Teaching Models and Chemical Representations in Developing Students' Mental Models of Chemical Phenomena. *Curtin University of Technology*, May, 1–494.

Coll, K. R., & Taylor, N. (2002). Mental Models In Chemistry: Senior Chemistry Students' Mental Models Of Chemical Bonding. *Chemistry Education: Research And Practice In Europe*, 3(2), 175–184.

Darmiyanti, W., Rahmawati, Y., Kurniadewi, F., & Ridwan, A. (2017). Analisis Model Mental Siswa Dalam

- Penerapan Model Pembelajaran Learning Cycle 8e Pada Materi Hidrolisis Garam. *Jurnal Riset Pendidikan Kimia*, 1(1), 38–51.
- Eliyawati, E., Agustin, R. R., Sya'bandari, Y., & Putri, R. A. H. (2020). Smartchem: An Android Application for Learning Multiple Representations of Acid-Base Chemistry. *Journal of Science Learning*, 3(3), 196–204. <https://doi.org/10.17509/jsl.v3i3.23280>
- Faradilla, A., Hadi, W., & Soro, S. (2020). *Evaluasi Proses dan Hasil Belajar (EPHB) Matematika dengan Diskusi dan Simulasi (DiSi)*. Uhamka Press.
- Habiddin, & Page, mary E. (2019). *Development and Validation of a Four-Tier Diagnostic Instrument for Chemical Kinetics (FTDICK)*. 19(3), 720–736. <https://doi.org/10.22146/ijc.39218>
- Jansoon, N., Cooll, R. K., & Somsook, E. (2009). Understanding Mental Models of Dilution in Thai Students. *International Journal of Environmental & Science Education*. *International Journal of Environmental and Science Education*, 4(2), 147–168.
- Kurniawan, A., Febrianti Nandia, A., Hardianti, T., Ichsan, Desy, & Risan, R. (2022). *Evaluasi Pendidikan*. PT. Global Eksekutif Teknolog. https://books.google.co.id/books?id=nKeAEAAAQBAJ&pg=PA145&dq=Daya+beda+soal&hl=id&newbks=1&newbks_redir=0&source=gb_mobile_search&ovdme=1&sa=X&ved=2ahUKEwjm_ZbMibOAAxWG3jgGHR4aAP44ChDrAXoECAIQBQ#v=onepage&q&f=true
- Laliyo, A. L. R. (2011). Model Mental Siswa Dalam Memahami Perubahan Wujud Zat. *Jurnal Penelitian Dan Pendidikan*, c, 1–12.
- Permatasari, M. B., Rahayu, S., & Dasna, I. W. (2022). Chemistry Learning Using Multiple Representations: A Systematic Literature Review. *Journal of Science Learning*, 5(2), 334–341. <https://doi.org/10.17509/jsl.v5i2.42656>
- Putri, A. K., Azra, F., & Hayati, E. (2022). Effectiveness of Chemscrabb Media Games on Fundamental Laws of Chemistry Towards for Students' Learning Outcomes. *Journal of Educational Chemistry (JEC)*, 4(2), 91–98. <https://doi.org/10.21580/jec.2022.4.2.12583>
- Rusilowati, A. (2015). Development of Diagnostic Tests as an Evaluation Tool for Physics Learning Difficulties. *Prosiding Seminar Nasional Fisika Dan Pendidikan Fisika*, 6, 1–10. <https://jurnal.fkip.uns.ac.id/index.php/prosfis1/article/view/7684>
- Sagita, R., Azra, F., & Azhar, M. (2017). Pengembangan Modul Konsep Mol Berbasis Inkuiri Terstruktur Dengan Penekanan Pada Interkoneksi Tiga Level Representasi Kimia Untuk Kelas X Sma. *Jurnal Eksakta Pendidikan (Jep)*, 1(2), 25–32. <https://doi.org/10.24036/jep.v1i2.48>
- Silberberg, M. S. (2010). *Principles of general chemistry*. McGraw-Hill, a business unit of The McGraw-Hill Companies, Inc.
- Supardi. (2017). *Statistika Penelitian Pendidikan*. PT Rajagrafindo Persada.
- Supriyadi. (2021). *Evaluasi Pendidikan*. PT. Nasya Expanding Management.
- Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, July 2013, 37–41. <https://doi.org/10.1080/0950069880100204>
- Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2003). The role of submicroscopic and symbolic representations in chemical explanations. *International Journal of Science Education*, 25(11), 1353–1368. <https://doi.org/10.1080/095006903200070306>
- 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>

Wang, C.Y. (2007). The role of mental-modeling ability, content knowledge, and mental models in general chemistry students' understanding

about molecular polarity. *Disertasi*, 70(3-A), 835.

Yusrizal, & Rahmati. (2020). *Tes Hasil Belajar*. Jakarta: Bandar Publishing.

LAMPIRAN

Lampiran 1. Model Pengembangan Treagust (1988)

