


Pengembangan Instrumen *Three Tier Diagnostic Test* Berbasis Multirepresentasi untuk Menganalisis Miskonsepsi Siswa pada Materi Asam Basa

Fina Ayu Lestari , Sigit Priatmoko, Endang Susilaningsih, dan Harjono

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

Info Artikel

Diterima: Juli 2023

Disetujui: September 2023

Dipublikasikan: Oktober 2023

Keywords:

Asam basa

Miskonsepsi

Rasch model

Three Tier Diagnostic Test

Abstrak

Penelitian ini bertujuan untuk mengembangkan instrumen *three tier diagnostic test* berbasis multirepresentasi untuk mengidentifikasi miskonsepsi siswa yang dianalisis menggunakan model Rasch. Penelitian dilakukan menggunakan desain 4D yang dikemukakan oleh Thiagarajan meliputi tahap: *define, design, development, dan disseminate*, namun dalam penelitian ini hanya sampai pada tahap *development*. Hasil validasi ahli menunjukkan instrumen tes masuk dalam kategori sangat valid dengan rata-rata skor 111,3 dari total skor 120. Hasil analisis reliabilitas menggunakan model Rasch menunjukkan bahwa reliabilitas instrumen belum cukup baik. Hasil analisis miskonsepsi secara klasikal menunjukkan sebanyak 47% siswa paham konsep, 16% siswa mengalami miskonsepsi positif, 12% siswa mengalami miskonsepsi negatif, 15% siswa mengalami miskonsepsi utuh, 4% siswa kurang paham, 1% siswa kurang yakin, dan 5% siswa tidak paham konsep. Siswa secara umum memberikan respon positif terhadap instrumen tes dibuktikan dengan angket. Persentase siswa yang memberikan respon setuju dan sangat setuju yakni sebanyak 45% dan 21% pada uji skala kecil, 63% dan 14% pada uji skala besar, serta 62% dan 14% pada uji implementasi.

Abstract

This study aims to develop a multirepresentation-based three-tier diagnostic test instrument to identify student misconceptions analyzed using the Rasch model. This research was conducted using the 4D design proposed by Thiagarajan covering the stages: define, design, development, and disseminate, but in this research only reached the development stage. The results of expert validation show that the test instrument is in the very valid category with an average score of 111.3 out of a total score of 120. The results of reliability analysis using the Rasch model show that the reliability of the instrument is not good enough. The results of the classical misconception analysis showed that as many as 47% of students understood concepts, 16% of students experienced positive misconceptions, 12% of students experienced negative misconceptions, 15% of students experienced complete misconceptions, 4% of students did not understand, 1% of students were not sure, and 5% of students did not understand concepts. Students generally give positive responses to test instruments evidenced by questionnaires. The percentage of students who gave affirmative and strongly affirmative responses was 45% and 21% in small-scale tests, 63% and 14% in large-scale tests, and 62% and 14% in implementation tests.

PENDAHULUAN

Kurikulum 2013 menekankan pada penilaian belajar siswa yang diperoleh melalui proses evaluasi pembelajaran. Pembelajaran pada kurikulum 2013 dikembangkan atas prinsip pembelajaran siswa aktif melalui serangkaian kegiatan mengamati (melihat, membaca, mendengar, menyimak), menanya (lisan, tulis), menganalisis (menghubungkan, menentukan keterkaitan, membangun cerita/konsep), dan mengkomunikasikan (lisan, tulis, gambar, grafik, tabel, chart). Evaluasi pembelajaran dalam kurikulum 2013 dinilai melalui aspek sikap spiritual, sosial, pengetahuan dan keterampilan. Siswa harus mampu menguasai aspek pengetahuan untuk mengaplikasikannya dalam aspek keterampilan. Keterampilan dalam diri siswa akan muncul jika siswa memahami konsep-konsep dalam pengetahuan yang diajarkan. Aspek pengetahuan dapat tercapai dengan baik apabila siswa memiliki pemahaman konsep yang benar (Wahyuningtyas *et al.*, 2020).

Salah satu ilmu yang konsep-konsepnya bersifat abstrak dan berurutan adalah ilmu kimia. Pembelajaran kimia harus dimulai dari konsep-konsep yang sederhana terlebih dahulu, kemudian dari konsep-konsep sederhana itu dibangun konsep-konsep yang lebih kompleks (Istiqomah *et al.*, 2021). Selama proses pembelajaran kimia seringkali siswa mengalami ketidakmampuan dalam menafsirkan ilmu kimia dengan konsep yang sebenarnya, hal ini disebut miskonsepsi. Miskonsepsi terjadi apabila siswa memiliki prasangka atau pemahaman konsep yang diyakini secara kuat namun konsep yang diyakini tersebut tidak sesuai dengan konsep ilmiah dari para ahli (Yuliati, 2017).

Peninjauan dan pengukuran terhadap miskonsepsi siswa sangat penting untuk dilakukan guna mengetahui seberapa besar pemahaman konsep yang tidak akurat yang masih dialami oleh siswa. Jika tidak dilakukan pengukuran maka akan berpotensi terjadi miskonsepsi yang lebih kompleks pada materi-materi yang dipelajari selanjutnya. Terdapat berbagai metode untuk menganalisis miskonsepsi salah satunya adalah menggunakan tes diagnostik. Tes diagnostik tiga tingkat atau *three tier diagnostic test* merupakan salah satu tes yang dapat digunakan untuk menganalisis miskonsepsi. Tes diagnostik ini terdiri dari tiga *tier*, *tier* pertama terdiri atas pilihan jawaban, *tier* kedua terdiri atas pilihan alasan yang mendasari jawaban pada *tier* pertama, dan *tier* ketiga berisi tingkat keyakinan terhadap pilihan jawaban dan alasan (Mellyzar *et al.*, 2022).

Hasil identifikasi masalah di SMA Negeri 1 Tahunan Jepara diketahui bahwa hasil belajar siswa pada mata pelajaran kimia menurun pasca pembelajaran daring. Guru kimia di SMA Negeri 1 Tahunan juga menyatakan belum pernah diadakan tes yang secara khusus bertujuan untuk menganalisis profil pemahaman konsep siswa. Salah satu materi dalam pembelajaran kimia yang rawan terjadi miskonsepsi adalah materi asam basa. Materi asam basa saling berkesinambungan antara tiga level multirepresentasi yakni mikroskopik, makroskopik, dan simbolik. Penelitian ini bertujuan untuk mengembangkan instrumen *Three Tier Diagnostic Test* pendeteksi miskonsepsi siswa pada materi asam basa. Instrumen tes yang dikembangkan bermuatan multirepresentasi untuk mengukur kemampuan siswa pada ketiga level representasi. Uji kelayakan instrumen tes dilakukan dengan validasi ahli dan analisis menggunakan model Rasch, serta analisis angket respon siswa terhadap instrumen tes.

Instrumen *Three Tier Diagnostic Test* terbagi menjadi tiga tingkat (*tier*). Tingkat pertama berisi pilihan jawaban dari pertanyaan, tingkat kedua berisi pilihan alasan memilih jawaban pada tingkat pertama, dan tingkat ketiga berisi pilihan yakin atau tidak yakin atas jawaban yang telah dipilih pada tingkat pertama dan kedua (Nurhujaimah *et al.*, 2016). Instrumen *Three Tier Diagnostic Test* memiliki kelebihan jika dibandingkan dengan instrumen diagnostik pilihan ganda biasa dan *two-tier*. Kelebihan tersebut karena pada instrumen *Three Tier Diagnostic Test* ini memiliki tingkatan ketiga yang berisi keyakinan atas respon siswa. Peneliti akan mendapatkan informasi yang lebih lengkap mengenai miskonsepsi siswa sehingga dapat membedakan antara siswa yang mengalami miskonsepsi, menebak, paham konsep, dan tidak paham konsep (Qodriyah *et al.*, 2020). Identifikasi miskonsepsi menggunakan *Three Tier Diagnostic Test* memiliki keuntungan lebih dibandingkan dengan *two tier test*, karena pada *Three Tier Diagnostic Test* mampu membedakan antara siswa yang kurang pengetahuan berdasarkan keyakinan siswa saat menjawab pertanyaan pada tingkatan pertama dan tingkatan kedua (Nurulwati & Rahmadani, 2020).

Analisis data hasil uji coba untuk mengetahui validitas butir dan reliabilitas instrumen tes dilakukan dengan pemodelan Rasch. Alat bantu komputasi yang digunakan dalam pengolahan data dengan model Rasch adalah *software* Ministep. Ministep merupakan *software* khusus yang ditujukan untuk pengolahan data model Rasch yang dapat bekerja di dalam sistem Microsoft Windows. Ministep merupakan versi terbatas dari program Winsteps (Sumintono & Widhiarso, 2015). Alfarisa & Purnama, (2019) telah menggunakan program Winsteps untuk menganalisis butir soal ulangan dalam penelitian yang dilakukan. Program Winsteps juga telah digunakan oleh (Nuryanti *et al.*, 2018) dalam penelitiannya tentang pengembangan instrumen kemampuan berpikir kritis.

METODE

Penelitian dilakukan di SMA Negeri 1 Tahunan Jepara pada tahun ajaran 2022/2023. Subjek penelitian ini terdiri dari siswa kelas XI MIPA 4, XI MIPA 5, XII MIPA 4, XII MIPA 5, dan XII MIPA 6 yang keseluruhan berjumlah 125 siswa. Penelitian pengembangan ini menggunakan desain 4D yang dikembangkan oleh Thiagarajan *et al.*, (1974). Pengembangan model 4D terdiri dari 4 tahapan yaitu *Define* (Pendefinisian), *Design* (Perancangan), *Development* (Pengembangan), dan *Disseminate* (Penyebaran). Penelitian ini mereduksi model 4D menjadi 3D sehingga tahapan yang dilakukan hanya sampai pada tahap *development* (pengembangan). Analisis butir soal dilakukan dengan model Rasch meliputi analisis *item fit*, *reliability*, *wright map*, *item measure*, *person measure*, dan *person fit*.

Tahap *define* terdiri dari dua langkah yang dilakukan yakni studi literatur dan studi lapangan. Tahap *design* terdiri dari beberapa tahapan yaitu (1) merumuskan tujuan pengembangan; (2) membuat penggalan silabus materi asam basa; (3) menyusun kisi-kisi butir soal; (4) menyusun butir soal *three tier diagnostic test*; (5) merumuskan pedoman penilaian soal; (6) menyusun kisi-kisi angket respon siswa; (7) menyusun angket respon siswa; (8) menyusun lembar validasi instrumen tes dan angket respon siswa. Tahap *development* terdiri dari uji validitas isi oleh ahli, uji coba skala kecil, uji coba skala besar, dan uji implementasi. Instrumen tes pada setiap tahapan diperbaiki sesuai dengan hasil analisis untuk mendapatkan instrumen yang layak.

Profil miskonsepsi siswa diukur berdasarkan kategori kombinasi jawaban yang diberikan. Interpretasi kombinasi jawaban siswa terhadap kategori miskonsepsi disajikan pada Tabel 1.

Tabel 1. Interpretasi Kombinasi Jawaban Siswa

Kombinasi Jawaban			Klasifikasi Jawaban
Tingkat 1	Tingkat 2	Tingkat 3	
Benar	Benar	Yakin	Pemahaman Utuh
Benar	Salah	Yakin	Miskonsepsi (+)
Salah	Benar	Yakin	Miskonsepsi (-)
Salah	Salah	Yakin	Miskonsepsi
Benar	Benar	Tidak Yakin	Beruntung / Kurang Yakin
Benar	Salah	Tidak Yakin	Kurang Paham
Salah	Benar	Tidak Yakin	Kurang Paham
Salah	Salah	Tidak Yakin	Tidak Paham

(Sumber: Mubarak *et al.*, 2016)

HASIL DAN PEMBAHASAN

Hasil penelitian terdiri dari analisis kualitas instrumen tes dan analisis miskonsepsi siswa. Hasil penelitian diuraikan sebagai berikut.

Hasil Validasi Ahli

Uji validitas isi dilakukan oleh ahli yang berasal dari dua dosen jurusan kimia dan satu guru kimia. Hasil validasi menunjukkan instrumen tes termasuk dalam kategori sangat valid dengan rata-rata skor sebesar 111,3 dari 120. Instrumen tes mengalami sedikit revisi berdasarkan saran dari ahli untuk meningkatkan kualitas instrumen. Rekapitulasi hasil validasi instrumen tes oleh ahli disajikan pada Tabel 2.

Tabel 2. Rekapitulasi Hasil Validasi Ahli

Validator	Total Skor	Kategori	Kriteria
Validator 1	96	Valid	Dapat digunakan dengan sedikit revisi
Validator 2	120	Sangat Valid	Dapat digunakan tanpa revisi
Validator 3	118	Sangat Valid	Dapat digunakan tanpa revisi

Analisis *Item Fit* (Tingkat Kesesuaian Butir Soal)

Tingkat kesesuaian butir soal (*item fit*) memberikan informasi tentang validitas tiap butir soal pada instrumen tes. Kriteria butir soal yang masuk dalam kategori *fit* ditentukan berdasarkan nilai-nilai berikut: (1) Nilai *Outfit mean square* (MNSQ) yang diterima: $0,5 < MNSQ < 1,5$; (2) Nilai *Outfit Z-standar* (ZSTD) yang diterima: $-2,0 < ZSTD < +2,0$; (3) Nilai *Point measure correlation* (Pt Measure Corr) yang diterima: $0,4 < Pt Measure Corr < 0,85$ (Boone *et al.*, 2014). Butir soal yang tidak memenuhi ketiga kriteria tersebut maka dikategorikan sebagai soal *misfit* atau tidak valid sehingga perlu diperbaiki atau dihilangkan.

Hasil analisis *item fit* pada uji coba skala kecil menunjukkan terdapat 11 butir soal *misfit*, pada uji coba skala besar terdapat 6 butir soal *misfit*, dan pada uji implementasi terdapat 6 butir soal *misfit*. Beberapa butir soal tidak muncul dalam *output* hasil analisis *item fit* karena butir tersebut memiliki nilai *maximum measure*. Butir soal dengan nilai *maximum measure* berarti tidak terdapat siswa yang menjawab dengan benar butir soal tersebut. Cuplikan hasil analisis *item fit* untuk butir soal *misfit* disajikan pada Tabel 3.

Tabel 3. Cuplikan Hasil Analisis *Item Fit*

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-CORR.	AL-EXP.	EXACT MATCH OBS%	EXP%	Item
4	2	15	1.77	.81	1.38	.84	9.90	5.00	A-.32	.27	86.7	86.7	S4
24	1	15	2.61	1.07	1.28	.59	9.90	3.25	B-.29	.19	93.3	93.4	S24
17	5	15	.37	.61	1.53	2.10	1.28	.60	C-.17	.42	40.0	72.2	S17

Analisis Reliabilitas

Analisis reliabilitas pada pemodelan Rasch meliputi analisis *item reliability*, *person reliability*, dan *alpha Cronbach*. Nilai *alpha Cronbach* memberikan informasi tentang interaksi antara siswa dengan butir soal secara keseluruhan. Reliabilitas menunjukkan seberapa besar konsistensi instrumen tes dalam mengukur kemampuan siswa (Sumintono & Widhiarso, 2015b). Suatu instrumen tes dikatakan reliabel apabila menunjukkan hasil yang relatif sama ketika tes dilakukan secara berulang. Kategori nilai *alpha Cronbach* disajikan pada Tabel 4.

Tabel 4. Kategori Nilai *alpha Cronbach*

Nilai <i>Alpha Cronbach</i>	Kategori
< 0,50	Buruk
0,50 – 0,60	Jelek
0,60 – 0,70	Cukup
0,70 – 0,80	Bagus
> 0,80	Bagus sekali

(Sumber: Sumintono & Widhiarso, 2015)

Hasil analisis reliabilitas pada uji coba skala kecil diperoleh sebesar 0,71 dengan kategori bagus, uji coba skala besar sebesar 0,52 dengan kategori jelek, dan uji implementasi sebesar 0,69 dengan kategori cukup. Rendahnya nilai reliabilitas menunjukkan interaksi antara siswa dengan butir soal belum cukup baik. Dugaan penyebab rendahnya reliabilitas pada uji skala besar dan uji implementasi dibandingkan pada uji skala kecil adalah karena perbedaan durasi pelaksanaan tes. Uji skala besar dan uji implementasi dilaksanakan lebih singkat dibandingkan estimasi waktu yang sebenarnya karena menyesuaikan durasi KBM di sekolah tempat penelitian.

Analisis *Item Measure* (Tingkat Kesulitan Butir Soal)

Tingkat kesulitan butir soal pada pemodelan Rasch ditunjukkan pada tabel *item measure*. Hasil yang ditampilkan berupa nilai *logit* tiap butir soal beserta nilai rata-rata *logit* dan standar deviasi. Nilai rata-rata *logit* dan standar deviasi digunakan sebagai dasar untuk pengelompokan kategori kesulitan butir soal. Cuplikan hasil analisis *item measure* disajikan pada Tabel 5.

Tabel 5. Cuplikan Hasil Analisis *Item Measure*

12	57	60	-3.77	.64	.70	-.61	.15	-1.01	.57	.36	95.0	94.9	S12
10	59	60	-5.03	1.03	.90	.18	.14	-.54	.32	.22	98.3	98.3	S10
MEAN	27.2	60.0	.46	.56	.95	.08	1.40	.27			83.5	82.9	
P.SD	20.6	.0	2.82	.41	.18	1.21	1.89	1.66			14.6	11.7	

Tabel 5 merupakan hasil analisis *item measure* pada uji implementasi. Tingkat kesulitan butir soal dibedakan menjadi empat kategori yaitu sangat sulit, sulit, sedang, dan mudah. Pengelompokan kategori tingkat kesulitan butir soal pada uji implementasi berdasarkan Tabel 5 disajikan pada Tabel 6.

Tabel 6. Pengelompokan Tingkat Kesulitan Butir Soal

Kategori	Kriteria
Logit $\geq +2,82$	Sangat Sulit
$+0,46 < \text{logit} < +2,82$	Sulit
$-2,82 < \text{logit} < +0,46$	Sedang
Logit $\leq -2,82$	Mudah

Hasil pengelompokan tingkat kesulitan butir soal pada uji implementasi diperoleh sebanyak 7 butir soal termasuk dalam kategori sangat sulit, 5 butir soal sulit, 9 butir soal sedang, dan 4 butir soal mudah. Hasil menunjukkan bahwa instrumen tes didominasi oleh butir soal dengan tingkat kesulitan sedang. Butir soal yang baik adalah butir soal dengan tingkat kesulitan sedang karena dapat menjangkau siswa dengan kemampuan tinggi maupun rendah (Rahmasari & Ismiyati, 2016).

Analisis *Person Measure* (Tingkat Kemampuan Siswa)

Pemodelan Rasch memberikan informasi mengenai tingkat kemampuan siswa yang ditampilkan pada tabel *person measure*. Hasil yang ditampilkan berupa nilai *logit* tiap siswa beserta nilai rata-rata *logit* dan standar deviasi. Nilai rata-rata *logit* dan standar deviasi digunakan sebagai acuan pengelompokan tingkat kemampuan siswa. Cuplikan hasil yang ditampilkan pada tabel *person measure* disajikan pada Tabel 7.

Tabel 7. Cuplikan Hasil Analisis *Person Measure*

40	4	25	-2.99	.72	2.27	2.74	9.90	4.08	.15	.54	73.9	87.5	40P
51	4	25	-2.99	.72	.56	-1.35	.22	-.37	.63	.54	91.3	87.5	51L
55	4	25	-2.99	.72	1.28	.84	1.46	.76	.47	.54	82.6	87.5	55P
MEAN	11.3	25.0	-0.01	.61	1.00	-.12	1.23	.15			83.5	82.9	
P.SD	3.1	.0	1.21	.04	.42	1.30	1.72	.96			9.6	2.2	

Tabel 7 menunjukkan cuplikan hasil analisis *person measure* pada uji implementasi. Tabel tersebut menunjukkan nilai rata-rata *logit* siswa sebesar -0,01 dan standar deviasi sebesar 1,21 yang digunakan sebagai acuan pengelompokan tingkat kemampuan siswa. Tingkat kemampuan siswa dikelompokkan menjadi empat kategori yakni sangat tinggi, tinggi, sedang, dan rendah. Pengelompokan tingkat kemampuan siswa berdasarkan data pada Tabel 7 disajikan dalam Tabel 8.

Tabel 8. Pengelompokan Tingkat Kemampuan Siswa

Kategori	Kriteria
Logit $\geq +1,21$	Sangat Tinggi
$-0,01 < \text{logit} < +1,21$	Tinggi
$-1,21 < \text{logit} < -0,01$	Sedang
Logit $\leq -1,21$	Rendah

Tabel 8 memberikan informasi tentang pengelompokan kemampuan siswa pada uji implementasi. Hasil analisis diperoleh sebanyak 5 siswa memiliki kemampuan sangat tinggi, 37 siswa memiliki kemampuan tinggi, 12 siswa memiliki kemampuan sedang, dan 6 siswa memiliki kemampuan rendah.

Analisis *Person Fit* (Tingkat Kesesuaian Peserta Didik)

Analisis *person fit* bertujuan untuk mengetahui adanya pola respon siswa yang tidak sesuai (*misfit*) dalam mengerjakan tes. Kriteria tingkat kesesuaian siswa pada analisis *person fit* sebagai berikut: (1) Nilai *Outfit mean square* (MNSQ) yang diterima: $0,5 < \text{MNSQ} < 1,5$; (2) Nilai *Outfit Z-standar* (ZSTD) yang diterima: $-2,0 < \text{ZSTD} < +2,0$; (3) Nilai *Point measure correlation* (*Pt Measure Corr*) yang diterima: $0,4 < \text{Pt Measure Corr} < 0,85$ (Boone *et al.*, 2014). Siswa yang tidak memenuhi ketiga kriteria tersebut maka termasuk dalam kategori *misfit*. Cuplikan hasil analisis *person fit* pada uji implementasi disajikan pada Tabel 9.

Tabel 9. Cuplikan Hasil Analisis *Person Fit*

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%	Person	
40	4	25	-2.99	.72	2.27	2.74	9.90	4.08	A	.15	.54	73.9	87.5	40P
16	7	25	-1.63	.65	1.47	1.22	5.41	2.23	B	.50	.66	82.6	85.9	16P
7	7	25	-1.63	.65	1.40	1.07	5.36	2.21	C	.51	.66	82.6	85.9	07L
18	7	25	-1.63	.65	1.40	1.07	5.36	2.21	D	.51	.66	82.6	85.9	18P
31	7	25	-1.63	.65	1.40	1.07	5.36	2.21	E	.51	.66	82.6	85.9	31L

Tabel 9 menunjukkan bahwa siswa dengan kode 40P tidak memenuhi tiga kriteria yang diterima Rasch sehingga dikategorikan *misfit*. Siswa 40P hanya mampu menjawab sebanyak 4 dari 25 butir soal dengan benar. Siswa 16P, 07L, 18P, dan 31L tidak memenuhi dua kriteria yaitu *outfit* MNSQ dan *outfit* ZSTD. *Outfit* ZSTD merupakan tolok ukur utama tingkat kesesuaian siswa pada model Rasch sehingga siswa yang tidak memenuhi dikategorikan *misfit*. Siswa yang *misfit* berarti mampu mengerjakan butir soal yang memiliki nilai *logit* lebih besar dari nilai *logit* siswa. Siswa *misfit* terindikasi melakukan kecurangan atau menebak jawaban. Data yang digunakan untuk mendukung hasil analisis *person fit* adalah *scalogram*. *Scalogram* menampilkan data pola respon siswa sehingga dapat diketahui siswa tersebut melakukan kecurangan atau tidak. Cuplikan *scalogram* pada uji implementasi disajikan pada Gambar 1.

```

28 +111010111101100000000000 28L
36 +111111110000010100000000 36L
41 +111111110000011000000000 41L
46 +111111101000001100000000 46L
42 +111111110000010000000000 42L
45 +111111110000010000000000 45L
7 +111101000000010000010000 07L
15 +1101101001001000000000000 15P
16 +1110101000000100000010000 16P
18 +1111010000000100000010000 18P
31 +1111010000000100000010000 31L
50 +1111000010000001000000000 50L
14 +1100001000010000000000000 14P
35 +1000100000110000000000000 35L
40 +0010010001000000001000000 40P
51 +1111000000000000000000000 51L
55 +1001100000010000000000000 55P
-----
|11 111 212112222 11
|0219453892836312450451767

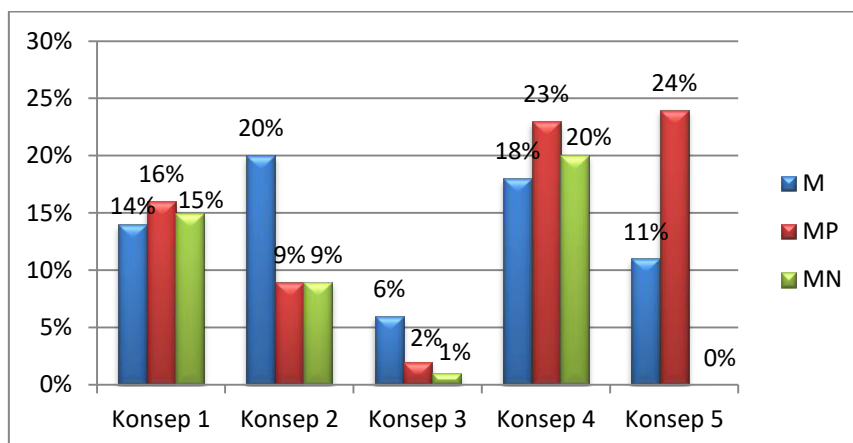
```

Gambar 1. Cuplikan *Scalogram* pada Uji Implementasi

Hasil analisis *person fit* menunjukkan bahwa siswa 16P, 07L, 18P, dan 31L terindikasi melakukan kecurangan karena memiliki nilai *outfit* MNSQ yang sama. *Scalogram* menunjukkan bahwa siswa 07L, 18P, dan 31L memiliki pola respon yang sama sehingga kemungkinan penyebab siswa tersebut *misfit* adalah kecurangan. Sementara siswa 16P *misfit* kemungkinan karena menebak jawaban atau kurang cermat dalam mengerjakan tes.

Analisis Profil Miskonsepsi Siswa

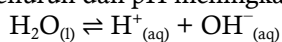
Kategori miskonsepsi siswa pada penelitian ini diukur berdasarkan kombinasi jawaban pada *three tier diagnostic test*. Butir soal disusun berkesinambungan antara satu *tier* dengan *tier* lainnya sehingga hasil tes dapat diinterpretasikan berdasarkan kategori miskonsepsi. Soal disusun berdasarkan konsep-konsep yang terdapat pada materi asam basa meliputi: (1) teori asam basa; (2) kesetimbangan ion dalam larutan; (3) peranan senyawa asam basa; (4) derajat keasaman; dan (5) indikator asam basa. Profil miskonsepsi siswa berdasarkan peta konsep pada uji implementasi disajikan pada Gambar 2.



Gambar 2. Profil Miskonsepsi Siswa

Profil miskonsepsi utuh tertinggi ditemukan pada konsep kesetimbangan ion dalam larutan dengan persentase sebesar 20%. Miskonsepsi positif tertinggi terdapat pada konsep indikator asam basa dengan persentase sebesar 24%. Miskonsepsi negatif tertinggi yang terdeteksi terdapat pada konsep derajat keasaman dengan persentase sebesar 20%. Tingginya angka miskonsepsi menandakan kemampuan siswa dalam memahami konsep-konsep tersebut masih lemah.

Miskonsepsi utuh tertinggi terjadi pada konsep kesetimbangan ion dalam larutan khususnya butir soal nomor 16. Butir soal nomor 16 meminta siswa menganalisis pergeseran kesetimbangan pada reaksi ionisasi air yang ditambahkan OH^- . Tidak terdapat siswa yang mampu menjawab dengan benar butir soal tersebut yang menandakan tingkat pemahaman siswa masih rendah. Siswa yang mengalami miskonsepsi utuh beranggapan bahwa penambahan ion OH^- ke dalam kesetimbangan ionisasi air menyebabkan jumlah ion H^+ meningkat. Anggapan tersebut tidak tepat karena penambahan ion OH^- akan membuat jumlah ion H^+ menurun dan pH meningkat. Reaksi ionisasi air adalah sebagai berikut:



Penambahan ion OH^- dalam reaksi tersebut akan menyebabkan jumlah ion OH^- bertambah dan ion H^+ menurun sehingga pH meningkat. Butir soal nomor 16 termasuk ke dalam level representasi mikroskopik. Pemahaman siswa pada level representasi mikroskopis yang tertinggal akan menyebabkan siswa kesulitan mengembangkan pemahaman konseptual dan berpotensi menyebabkan kesalahan konsep (Nurhayati & Natasukma, 2019).

Miskonsepsi positif tertinggi terdapat pada konsep indikator asam basa khususnya pada butir soal nomor 23. Butir soal nomor 23 meminta siswa menganalisis sifat larutan yang telah diencerkan melalui pengujian menggunakan indikator PP dan kertas lakmus. Miskonsepsi yang terdeteksi pada butir soal ini adalah miskonsepsi positif yang berarti siswa memberikan jawaban tepat, namun alasan yang diberikan salah. Siswa yang mengalami miskonsepsi menganggap bahwa pengenceran dapat mempengaruhi sifat larutan yang mulanya asam menjadi basa. Anggapan siswa tersebut tidak tepat karena pengenceran hanya menyebabkan konsentrasi larutan semakin kecil dan kepekatannya berkurang. Perubahan konsentrasi larutan mempengaruhi pH larutan yang dihasilkan, namun tidak merubah sifat larutan.

Miskonsepsi negatif tertinggi terdapat pada konsep derajat keasaman (pH) khususnya pada butir soal nomor 17. Butir soal nomor 17 meminta siswa untuk menentukan konsentrasi larutan asam kuat yang diketahui nilai pH dan jumlah valensi asamnya. Penelusuran lebih lanjut menunjukkan tidak terdapat siswa menjawab dengan benar butir soal ini, artinya pemahaman siswa terhadap konsep perhitungan konsentrasi dan pH larutan masih rendah. Siswa yang mengalami miskonsepsi negatif berarti menjawab dengan benar *tier* pertama, namun salah pada *tier* kedua. Siswa kemungkinan belum dapat membedakan antara konsentrasi larutan asam dengan konsentrasi ion H^+ sehingga mengalami kesalahan dalam perhitungan.

Analisis Angket Respon Siswa

Angket respon siswa bertujuan untuk mengetahui respon siswa terhadap instrumen *three tier diagnostic test* yang dikembangkan. Secara keseluruhan siswa memberikan respon positif terhadap adanya instrumen *three tier diagnostic test* pada materi asam basa. Persentase siswa yang memberikan respon setuju

dan sangat setuju yakni sebanyak 45% dan 21% pada uji skala kecil, 63% dan 14% pada uji skala besar, serta 62% dan 14% pada uji implementasi.

SIMPULAN

Pengembangan instrumen *three tier diagnostic test* berbasis multirepresentasi pada materi asam basa berhasil mengungkap miskonsepsi siswa pada materi asam basa. Instrumen yang dikembangkan dinyatakan valid namun belum cukup reliabel yang kemungkinan disebabkan karena faktor eksternal. Profil miskonsepsi utuh tertinggi ditemukan pada konsep kesetimbangan ion dalam larutan dengan persentase sebesar 20%. Miskonsepsi positif tertinggi terdapat pada konsep indikator asam basa dengan persentase sebesar 24%. Miskonsepsi negatif tertinggi yang terdeteksi terdapat pada konsep derajat keasaman dengan persentase sebesar 20%. Secara keseluruhan siswa memberikan respon positif terhadap instrumen *three tier diagnostic test* dibuktikan dengan angket. Persentase siswa yang memberikan respon setuju dan sangat setuju yakni sebanyak 45% dan 21% pada uji skala kecil, 63% dan 14% pada uji skala besar, serta 62% dan 14% pada uji implementasi.

DAFTAR PUSTAKA

- Alfarisa, F., & Purnama, D. N. 2019. Analisis Butir Soal Ulangan Akhir Semester Mata Pelajaran Ekonomi SMA Menggunakan Rasch Model. *Jurnal Pendidikan Ekonomi*. 11(2).
- Boone, W. J., Staver, J. ., & Yale, M. S. 2014. *Rasch Analysis in the Human Sciences*. Dordrecht: Springer.
- Istiqomah, W., Rahayu, S., & Muchson, M. 2021. Identifikasi Miskonsepsi Materi Ikatan Kovalen pada Mahasiswa Kimia Tahun Pertama Universitas Negeri Malang Menggunakan Tes Diagnostik Two-Tier. *Jurnal MIPA dan Pembelajarannya*. 1(6), 482–492.
- Mellyzar, Fakhrah, & Isnani. 2022. Analisis Miskonsepsi Siswa SMA Menggunakan Instrumen Three Tier Multiple Choice pada Materi Struktur Atom dengan Teknik Certanty of Response Index (CRI). *Edukatif: Jurnal Ilmu Pendidikan*. 4(2), 2556–2564.
- Mubarak, S., Susilningsih, E., & Cahyono, E. 2016. Pengembangan Tes Diagnostik Three Tier Multiple Choice untuk Mengidentifikasi Miskonsepsi Peserta Didik Kelas XI. *Journal of Innovative Science Education*. 5(2), 101–110.
- Nurhayati, S., & Natasukma, M. 2019. Profil Miskonsepsi Peserta Didik pada Pembelajaran Multirepresentasi Materi Asam Basa Melalui Model Blended Learning. *Chemistry in Education*. 8(2), 17–23.
- Nurhujaimah, R., Kartika, I. R., & Nurjaydi, M. 2016. Analisis Miskonsepsi Siswa Kelas XI SMA pada Materi Larutan Penyangga Menggunakan Instrumen Tes Three Tier Multiple Choice. *Jurnal Paedagogia*. 19(1), 15–28.
- Nurulwati, N., & Rahmadani, A. 2020. Perbandingan Hasil Diagnostik Miskonsepsi Menggunakan Threetier dan Fourtier Diagnostic Test pada Materi Gerak Lurus. *Jurnal Pendidikan Sains Indonesia*. 7(2), 101–110.
- Nuryanti, S., Masykuri, M., & Susilowati, E. 2018. Analisis Iteman dan Model Rasch pada Pengembangan Instrumen Kemampuan Berpikir Kritis Peserta Didik Sekolah Menengah Kejuruan. *Jurnal Inovasi Pendidikan IPA*. 4(2), 224–233.
- Qodriyah, N. R. L., Rokhim, D. A., Widarti, H. R., & Habiddin. 2020. Identifikasi Miskonsepsi Siswa Kelas XI SMA Negeri 4 Malang pada Materi Hidrokarbon Menggunakan Instrumen Diagnostik Three Tier. *Jurnal Inovasi Pendidikan Kimia*. 14(2), 2645.
- Rahmasari, D., & Ismiyati. 2016. Analisis Butir Soal Mata Pelajaran Pengantar Administrasi Perkantoran. *Economic Education Analysis Journal*. 5(1), 317–330.
- Sumintono, B., & Widhiarso, W. 2015. *Aplikasi Permodelan Rasch Pada Assessment Pendidikan*. Cimahi: Trimkomunikata.
- Thiagarajan, S., Semmel, D. S., & Semmel, M. I. 1974. *Instructional Development for Training Teachers of Exceptional Children*. National Center for Improvement Educational System.

Wahyuningtyas, W., Sumarti, S. S., Susilaningsih, E., & Wijayati, N. 2020. Analisis Miskonsepsi Asam Basa Menggunakan Instrumen Multirepresentasi Diagnostic Test Berbasis Web. *Chemistry in Education*. 9(1), 1–7.

Yuliati, Y. (2017). Miskonsepsi Siswa pada Pembelajaran IPA serta Remediasinya. *Jurnal Bio Education*. 2(2), 50–58.