

**PENGGUNAAN RASCH MODEL UNTUK MENGEMBANGKAN
INSTRUMEN PENGUKURAN KEMAMPUAN BERIKIR KRITIS SISWA
PADA TOPIK IKATAN KIMIA**

**USING RASCH MEASUREMENT TO DEVELOP
AN INSTRUMENT TO ASSESS STUDENTS' CRITICAL THINKING ABILITY
IN CHEMICAL BONDING TOPIC**

Ardi Widhia Sabekti*, Fitriah Khoirunnisa

Program Studi Pendidikan Kimia, Universitas Maritim Raja Ali Haji
Jl. Politeknik Senggarang, Tanjungpinang, Kepulauan Riau. Kode Pos: 29115

*e-mail korespondensi: sabekti.ardi@umrah.ac.id

Abstrak

Salah satu kemampuan abad 21 yang harus dikuasai siswa adalah kemampuan berpikir kritis. Untuk mengetahui tingkat keberhasilan siswa dalam mengembangkan kemampuan berpikir kritis, diperlukan suatu instrumen yang dapat mengukur kemampuan tersebut. Artikel ini memperkenalkan penggunaan Rasch Model untuk mengembangkan instrumen pengukuran kemampuan berpikir kritis siswa pada topik kimia, terutama topik ikatan kimia. Berdasarkan analisis Kompetensi Dasar (KD) dan Indikator Berpikir Kritis, indikator soal dirancang. Berdasarkan indikator soal, butir soal beserta kunci jawaban dan pedoman penskoran dikembangkan. Instrumen terdiri dari 13 soal tipe uraian yang divalidasi oleh dua orang pakar dan direvisi berdasarkan masukan yang diberikan. Instrumen diujicobakan kepada 51 siswa SMA di Tanjungpinang. Model Rasch parsial-kredit diterapkan pada data hasil uji coba untuk memeriksa kualitas tiap butir soal dan kualitas instrumen secara keseluruhan. Hasil analisis menunjukkan bahwa 12 dari 13 item soal memenuhi kriteria soal yang baik menurut Rasch Model dengan reliabilitas instrumen 0,72 (cukup). Studi ini menunjukkan bahwa penggunaan Rasch model untuk mengembangkan instrumen pengukuran kemampuan berpikir kritis dalam bidang kimia cukup menjanjikan.

Kata kunci: berpikir kritis, rasch model, instrumen, ikatan kimia

Abstract

One of the 21st century abilities that students must master is critical thinking ability. To find out the level of success of students in developing critical thinking skills, an instrument is needed that can measure these abilities. This article introduces an approach to using Rasch measurement to develop instruments to assess students' critical thinking skills on chemical topics, especially chemical bonding topic. Based on the analysis of Basic Competencies and Critical Thinking Indicators, the question indicators are designed. Based on the question indicator, the item along with the key answers and scoring guidelines were developed. The instrument contained 13 essay that are validated by two experts and revised. The instrument was given to 51 students in Tanjungpinang. The partial-credit Rasch model was applied to the data to examine item and instrument properties. The research showed that 12 questions of 13 questions was valid through Rasch model analysis. The reliability of instrumen was 0,72, which is good enough category. This study demonstrates that using Rasch measurement to develop instruments for measuring critical thinking skills in chemistry is promising.

Keywords: critical thinking, rasch model, instrument, chemical bonding

PENDAHULUAN

Pembelajaran sains diharapkan dapat menghantarkan siswa memenuhi kemampuan abad 21. Salah satu kemampuan yang diperlukan pada abad 21 adalah kemampuan berpikir kritis. Kemampuan berpikir kritis penting bagi perkembangan siswa dan merupakan keterampilan yang berharga untuk hidup di tengah masyarakat (Danczak dkk., 2017). Perkembangan kemampuan berpikir kritis siswa perlu senantiasa dievaluasi sebagai tolak ukur keberhasilan proses pembelajaran dan sebagai bahan telaah reflektif untuk meningkatkan kualitas pembelajaran. Salah satu cara mengukur kemampuan berpikir kritis adalah dengan tes.

Beberapa tes telah dikembangkan untuk mengukur kemampuan berpikir kritis. Tes kemampuan berpikir kritis bebas materi yang telah dikembangkan antara lain *Cornell Critical Thinking Test* (Ennis dkk., 1964), *California Critical Thinking Skills Test* (Facione, 1990), dan *Ennis-Weir Critical Thinking Essay Test* (Ennis & Weir, 1985). Tes kemampuan berpikir kritis pada bidang studi kimia yang telah dikembangkan antara lain pada materi hidrokarbon (Kartimi dkk., 2012), termokimia (Kartimi & Liliyasi, 2012b), asam basa (Amalia & Susilaningih, 2014), dan Kesetimbangan Kimia (Kartimi & Liliyasi, 2012a). Adapun alat ukur kemampuan berpikir kritis terstandar pada materi ikatan kimia belum ada sehingga perlu dikembangkan.

Untuk memperoleh data kemampuan berpikir kritis yang terpercaya dalam materi ikatan kimia, dibutuhkan instrumen tes yang berkualitas. Untuk menentukan kualitas sebuah instrumen tes, seringkali masih digunakan teori tes klasik. Dalam teori tes klasik, tingkat kesukaran soal dan kemampuan siswa didasarkan pada persentase respon yang benar. Persentase statistik yang benar tidak sepenuhnya pada skala interval karena efek puncak dan efek dasar (yaitu, tertinggi adalah 100% dan terendah adalah 0%). Mereka juga tergantung pada tes dan kelompok siswa yang mengambilnya. Misalnya, tes lebih mudah bagi siswa yang lebih mampu, dan skor siswa lebih rendah pada tes yang lebih sulit. Oleh karena itu, dalam teori tes klasik, tingkat kesukaran soal dan kemampuan orang saling bergantung satu sama lain.

Salah satu konsekuensi dari ketergantungan ini adalah instrumen pengukuran yang dikembangkan menggunakan teori tes klasik perlu divalidasi ulang ketika sampel target berbeda dari sampel validasi awal. Rasch Model

mengatasi masalah ketergantungan ini, karena tingkat kesukaran soal tetap invariant, bagaimanapun karakter sampel yang terlibat dalam validasi awal (Wei dkk., 2012). Hal ini menyebabkan Rasch Model lebih direkomendasikan untuk digunakan dalam pengembangan instrumen tes. Keunggulan lain dari Rasch Model adalah kemampuannya dalam pengukuran *learning gain* (Pentecost & Barbera, 2013) Meskipun Rasch Model memiliki keunggulan dibandingkan dengan teori tes klasik, penggunaannya dalam penelitian masih minim.

Tujuan penelitian ini adalah untuk memperkenalkan penggunaan Rasch Model dalam pengembangan instrumen pengukuran kemampuan berpikir kritis pada materi ikatan kimia. Pertanyaan penelitian yang akan dijawab adalah: (1) bagaimana proses khas dari pengembangan instrumen pengukuran kemampuan berpikir kritis pada materi ikatan kimia menggunakan Rasch Model? dan (2) bukti validitas dan reliabilitas apa yang dapat disediakan Rasch Model untuk mendukung penggunaan instrumen pengukuran kemampuan berpikir kritis pada materi ikatan kimia?

METODE PENELITIAN

Proses penggunaan Rasch Model untuk mengembangkan instrumen tes adalah melalui siklus pengembangan dan revisi sehingga respon siswa terhadap pertanyaan dan data akan sesuai dengan Rasch Model. Prosesnya adalah proses sistematis di mana item tes secara sengaja disusun sesuai dengan teori yang dihipotesiskan dan diuji secara empiris dengan menerapkan model Rasch untuk menghasilkan serangkaian item yang memenuhi persyaratan model Rasch. Penelitian ini mengadopsi 10 langkah pengembangan instrumen tes menggunakan Rasch Model menurut Liu (2010), yaitu:

1. Mendefinisikan konstruk yang dapat dicirikan oleh sifat linier. Tahapan ini dilakukan dengan cara menganalisis kompetensi dasar, materi iktan kimia, dan indikator berpikir kritis.
2. Mengidentifikasi kemampuan yang sesuai dengan level yang berbeda dari konstruk yang ditentukan. Tahapan ini dilakukan dengan cara menyusun indikator soal dan memilih jenis tes yang akan digunakan.
3. Menentukan ruang hasil dari kemampuan siswa, yaitu dengan merancang butir soal beserta rubrik kunci jawaban dan pedoman penskoran. Selanjutnya, instrumen divalidasi

oleh dua orang ahli dan direvisi sesuai saran yang diberikan. Lembar validasi dengan skala sikap 1-4 digunakan untuk memperoleh respon dan masukan pakar terhadap instrumen tes yang dikembangkan. Lembar validasi mencakup tiga aspek penilaian, yaitu kesesuaian butir soal dengan indikator, kesesuaian kunci jawaban dan penskoran dengan konsep, serta kesesuaian kebahasaan. Validitas logis tiap butir soal dianalisis dengan formula Aiken's V.

4. Uji coba dengan sampel representatif dari populasi target. Sampel adalah 51 siswa SMA di Kota Tanjungpinang Provinsi Kepulauan Riau yang dipilih secara acak dari dua sekolah dengan akreditasi yang berbeda (Akreditasi A= 26 siswa; Akreditasi B= 25 siswa). Tes diujicobakan pada bulan Agustus tahun 2018.
5. Menerapkan model Rasch. Model Rasch parsial-kredit dipilih karena tipe soal berupa uraian dengan skor maksimum berbeda pada beberapa butir soal.
6. Meninjau item sesuai statistik dan merevisi item jika perlu.
7. Meninjau peta Wright dan menambahkan atau menghapus item jika perlu.
8. Mengulangi langkah 4-7 hingga sekumpulan item sesuai dengan model Rasch.
9. Menetapkan validitas dan klaim reliabilitas untuk instrumen tes.
10. Mengembangkan dokumentasi untuk instrumen tes.

Adapun pada penelitian ini, uji coba hanya dilakukan satu kali sehingga langkah nomor 8 tidak dilaksanakan.

HASIL DAN PEMBAHASAN

Mendefinisikan Konstruk

Dalam mendefinisikan konstruk yang diukur, hal penting yang harus dipertimbangkan adalah bahwa konstruk secara teoritis harus bersifat unidimensional dari level rendah ke level tinggi (Wei dkk., 2012). Konstruk yang akan diukur pada penelitian ini adalah kemampuan berpikir kritis siswa pada materi ikatan kimia.

Materi ikatan kimia dipilih karena materi ini diajarkan pada kelas X, sehingga kelas XI dan XII juga telah menerima materi tersebut. Syarat sebuah tes diberikan adalah siswa telah pernah belajar materi tersebut. Tes materi ikatan kimia dapat diberikan pada tiga level kelas sehingga dapat memberikan gambaran

kemampuan berpikir kritis secara lebih komprehensif. Kompetensi Dasar (KD) materi ikatan kimia adalah sebagai berikut:

- 3.5 Membandingkan ikatan ion, ikatan kovalen, ikatan kovalen koordinasi, dan ikatan logam serta kaitannya dengan sifat zat
- 4.5 Merancang dan melakukan percobaan untuk menunjukkan karakteristik senyawa ion atau senyawa kovalen berdasarkan beberapa sifat fisika (titik leleh, titik didih, daya hantar listrik, atau sifat lainnya)
- 3.6 Menerapkan teori Tolakan Pasangan Elektron Kulit Valensi (VSEPR) dan Teori Domain Elektron dalam menentukan bentuk molekul
- 4.6 Membuat model bentuk molekul dengan menggunakan bahan-bahan yang ada di lingkungan sekitar atau perangkat lunak komputer
- 3.7 Menghubungkan interaksi antar ion, atom, dan molekul dengan sifat fisika zat
- 4.7 Menerapkan prinsip interaksi antar ion, atom, dan molekul dalam menjelaskan sifat-sifat fisik zat di sekitarnya.

Semua KD tersebut digunakan dalam pengembangan tes kecuali KD 4.6, karena penilaian pada KD tersebut lebih tepat jika berbasis produk. Cakupan keluasan dan kedalaman materi sebagai bahan pengembangan tes ditentukan berdasarkan cakupan kompetensi dasar yang telah dipilih, antara lain: Teori Lewis, ikatan ion, ikatan kovalen, ikatan kovalen koordinasi, ikatan logam, Teori VSEPR, bentuk molekul, kepolaran molekul, interaksi antar partikel, dan sifat/karakteristik fisika zat.

Indikator berpikir kritis yang dipilih pada penelitian ini adalah indikator Ennis karena lebih komprehensif dan mudah diaplikasikan. Indikator berpikir kritis Ennis (2011) meliputi lima aspek dengan perincian indikator masing-masing adalah sebagai berikut:

1. Memberikan penjelasan sederhana, meliputi: (1) memfokuskan pertanyaan, (2) menganalisis pernyataan, (3) bertanya dan menjawab pertanyaan tentang suatu penjelasan,
2. Membangun keterampilan dasar, meliputi: (4) mempertimbangkan apakah sumber dapat dipercaya/tidak, dan (5) mengamati dan mempertimbangkan suatu laporan hasil observasi,
3. Menyimpulkan, meliputi: (6) mendeduksi dan mempertimbangkan hasil deduksi, (7) menginduksi dan mempertimbangkan hasil induksi, (8) membuat dan menentukan nilai pertimbangan,

4. Mengatur strategi dan taktik, meliputi: (9) menentukan tindakan, (10) berinteraksi dengan orang lain.
5. Memberikan penjelasan lanjut, meliputi: (11) mendefinisikan istilah dan pertimbangan dalam tiga dimensi, (12) mengidentifikasi asumsi

Dari kedua belas indikator berpikir kritis Ennis, indikator nomor 10 dan 11 tidak dikembangkan dalam tes karena lebih cocok dievaluasi dengan teknik observasi.

Mengidentifikasi Kemampuan dari Konstruk yang Telah Didefinisikan

Berdasarkan konstruk yang telah didefinisikan, 13 indikator soal disusun. Jenis instrumen tes yang dikembangkan pada penelitian ini adalah tes jenis uraian. Tes uraian dipilih karena dapat mengevaluasi kemampuan berpikir kritis secara lebih mendalam dibandingkan pilihan ganda. Tes jenis uraian juga meminimalisir pemilihan jawaban benar secara acak yang tidak dilandasi konsep yang benar. Hasil penelitian juga menunjukkan bahwa secara empirik dan simulasi, tes yang disajikan dalam bentuk uraian cenderung memiliki nilai fungsi informasi item yang lebih tinggi (Susongko, 2010).

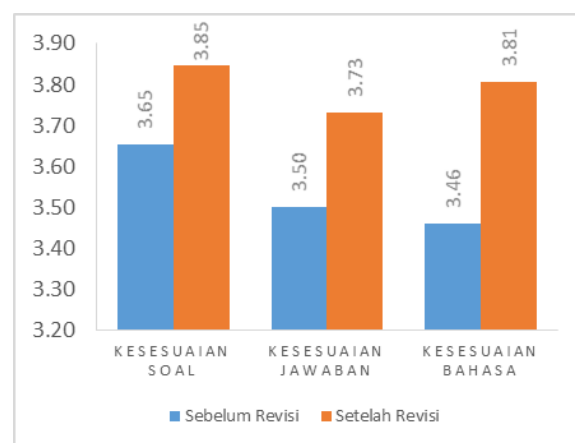
Menentukan Ruang Hasil dari Kemampuan Siswa

Berdasarkan indikator soal yang telah disusun, dikembangkan 13 butir soal esai beserta rubrik kunci jawaban dan pedoman penskoran. Instrumen yang telah dirancang selanjutnya divalidasi oleh dua orang dosen pakar evaluasi pembelajaran di Program Studi Pendidikan Kimia Universitas Maritim Raja Ali Haji dan direvisi sesuai saran perbaikan dari pakar. Saran perbaikan dari pakar untuk butir soal dalam instrumen dapat dilihat pada Tabel 1.

Perbandingan rata-rata respon pakar sebelum dan setelah revisi pada tiap aspeknya ditunjukkan pada Gambar 1. Perubahan validitas logis butir soal sebelum dan setelah revisi yang dihitung dengan formula Aiken's V dapat dilihat pada Tabel 2. Hasil analisis menunjukkan bahwa semua butir soal memiliki nilai validitas logis yang sangat tinggi. Meskipun demikian, instrumen tetap perlu diuji coba untuk mengetahui respon siswa sehingga dapat dianalisis validitas empiris dan nilai reliabilitas instrumen.

Tabel 1. Saran Perbaikan Instrumen

No	Saran Perbaikan
1	- Pada akhir soal, siswa perlu menyimpulkan - Pada kunci jawaban, daftar pertanyaan benar yang mungkin muncul perlu ditambah
4	- Memerlukan kasus yang lebih realistis - Lebih menyesuaikan indikator dengan cara merubah pertimbangan siswa dari membeli/tidak menjadi setuju/tidak
7	Kalimat soal dan penskoran perlu diperbaiki
9	Siswa perlu diminta untuk menggambarkan kutub positif dan negatif
10	Tambahkan alasan: ikatan kovalen lebih sulit untuk diputuskan
11	Kalimat soal ambigu, perlu diperbaiki



Gambar 1. Respon Pakar (Skala 1-4)

Tabel 2. Validitas Logis Butir Soal

No Soal	Sebelum Revisi		Setelah Revisi	
	Nilai V	Validitas	Nilai V	Validitas
1	0.72	tinggi	0.94	sangat tinggi
2	0.89	sangat tinggi	0.94	sangat tinggi
3	0.94	sangat tinggi	1.00	sangat tinggi
4	0.61	tinggi	0.83	sangat tinggi
5	1.00	sangat tinggi	1.00	sangat tinggi
6	1.00	sangat tinggi	1.00	sangat tinggi
7	0.72	tinggi	1.00	sangat tinggi
8	0.94	sangat tinggi	0.94	sangat tinggi
9	0.78	tinggi	0.83	sangat tinggi
10	0.83	sangat tinggi	0.94	sangat tinggi
11	0.78	tinggi	0.83	sangat tinggi
12	0.94	sangat tinggi	0.89	sangat tinggi
13	0.89	sangat tinggi	1.00	sangat tinggi

Uji Coba Instrumen dan Penerapan Rasch untuk Meninjau Kesesuaian Butir Soal

Instrumen yang telah divalidasi pakar diujicobakan kepada 51 siswa dan telah memenuhi jumlah sampel minimum yang dianjurkan untuk uji coba instrumen tes jenis

uraian pada Rasch Model, yaitu 50 orang (Linacre, 1994). Sebelum diberikan tes, siswa diminta untuk mempelajari kembali materi ikatan kimia. Siswa diberikan waktu 60 menit untuk mengerjakan 13 butir soal uraian secara mandiri. Setelah siswa mengerjakan tes, jawaban dikoreksi dan diberikan skor sesuai pedoman penskoran yang telah disusun.

Data skor dianalisis dengan Rasch Model berbantuan program Winsteps Versi 3.73. Berdasarkan analisis *Item Measure* (Tabel 3), tingkat kesulitan soal dapat diklasifikasikan. Pengklasifikasian tingkat kesulitan dapat dilakukan dengan membandingkan nilai *Measure* tiap butir soal dengan nilai S.D Measure. Pengelompokan butir soal berdasarkan tingkat kesulitannya dapat dilihat pada Tabel 4. Analisis *Item Fit Order* (Tabel 5) menunjukkan bahwa terdapat satu butir soal yang tidak fit, yaitu nomor 13, sehingga perlu dibuang atau diganti. Hal ini karena soal nomor 13 tidak memenuhi satupun kriteria item fit. Dua belas soal yang lain dapat dipertahankan karena memenuhi salah satu kriteria item fit berikut (Sumintono & Widhiarso, 2015):

- a. $0,5 < \text{Outfit MNSQ} < 1,5$;
- b. $-2,0 < \text{Outfit ZSTD} < +2,0$;
- c. $0,4 < \text{Pt-Measure Corr.} < 0,85$.

Tabel 3. *Item Measure*

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.
9	1	5	1.95	1.05
4	12	39	1.26	.31
7	4	6	1.03	.59
2	34	46	.93	.22
10	17	18	.58	.34
12	5	5	-.17	.56
8	14	11	-.33	.36
6	35	33	-.33	.22
3	54	38	-.43	.25
5	34	28	-.66	.23
11	19	15	-.86	.31
1	60	39	-1.05	.20
13	44	22	-1.92	.26
MEAN	25.6	23.5	.00	.38
S.D.	18.6	14.0	1.04	.23

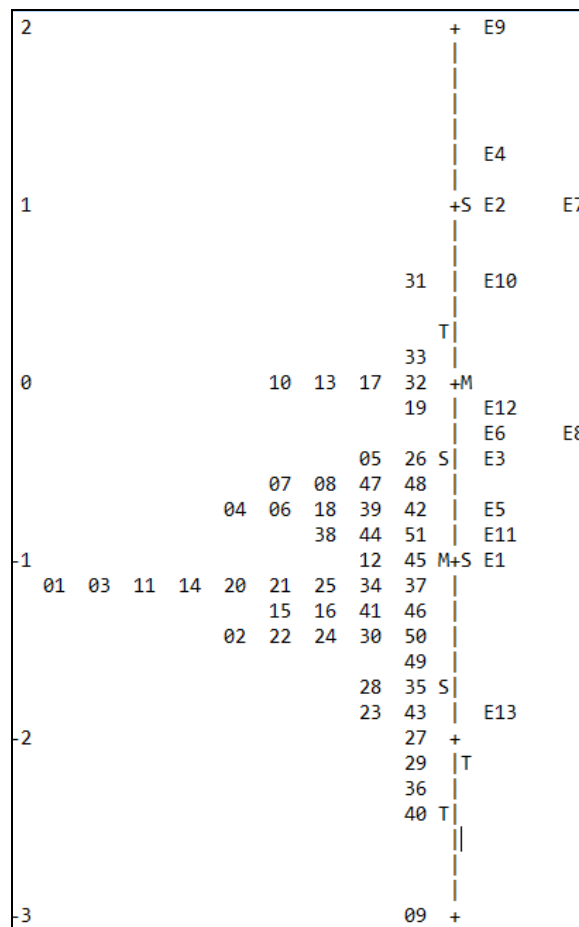
Tabel 4. Klasifikasi Soal Berdasar Tingkat Kesulitan

Tingkat Kesulitan	Nilai Measure	Butir Soal
Sangat Sulit	Lebih besar dari 1,04	9,4
Sulit	0,00 – 1,04	7, 2, 10
Mudah	-1,04 – 0,00	12, 8, 6, 3, 5, 11
Sangat Mudah	Lebih kecil dari 1,04	1, 13

(S.D.=1,04)

Tabel 5. *Item Fit Order*

MNSQ	ZSTD	PT-MEASURE		EXACT MATCH		Item	G
		CORR.	EXP.	OBS%	EXP%		
2.00	3.0	A .14	.38	22.7	41.2	E13	C
1.54	1.0	B-.01	.44	40.0	45.4	E12	C
1.28	.8	C .31	.34	79.5	72.7	E4	C
1.10	.6	D .51	.41	39.1	47.8	E2	B
1.08	.5	E .05	.30	80.0	81.8	E9	C
1.05	.3	F .66	.49	30.3	43.4	E6	C
.73	-.1	G .49	.46	66.7	56.3	E7	C
.96	-.1	f .52	.43	47.4	52.1	E3	B
.86	-.5	e .10	.48	50.0	42.6	E5	C
.81	-.9	d .61	.49	56.4	44.6	E1	A
.68	-.8	c .42	.47	63.6	41.9	E8	C
.59	-1.4	b .40	.47	60.0	41.7	E11	C
.39	-2.8	a .61	.47	72.2	46.4	E10	B
1.01	.0			54.5	50.6		
.41	1.3			17.5	12.2		



Gambar 2. Peta Wright

Tabel 6. *Summary Statistic*

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	24.1	23.6	.00	.41	1.05	.1	1.07	.2
S.D.	18.5	14.6	1.02	.24	.33	.8	.33	.7
MAX.	60.0	46.0	1.99	1.08	1.97	1.6	2.00	1.6
MIN.	1.0	5.0	-1.48	.21	.67	-1.1	.66	-1.1
REAL RMSE	.54	TRUE SD	.87	SEPARATION	1.62	Item	RELIABILITY	.72
MODEL RMSE	.48	TRUE SD	.90	SEPARATION	1.88	Item	RELIABILITY	.78
S.E. OF Item MEAN = .31								

Meninjau Peta Wright

Sebaran kemampuan responden dan sebaran tingkat kesulitan soal dengan skala yang sama dapan ditinjau dari Peta Wright. Peta Wright hasil uji coba instrumen dapat dilihat pada Gambar 2. Peta Wright sebelah kiri menggambarkan kemampuan siswa, dan terlihat bahwa siswa dengan kode 31 (S31) memiliki kemampuan tertinggi dibandingkan siswa yang lain. Meskipun memiliki kemampuan tertinggi, nilai *logit* dari siswa ini kurang dari +1 *logit*. Siswa S31 berada di luar dari batas dua deviasi standar (T) yang menunjukkan kecerdasan tinggi yang berbeda (*outlier*). Adapun siswa yang paling rendah kemampuannya adalah S09, dengan nilai *logit* -3, yang juga menunjukkan kemampuan yang sangat rendah (*outlier*), karena di luar batas T.

Peta wright sebelah kanan menjelaskan sebaran nilai *logit* butir soal. Butir soal esai nomor 9 (E9) merupakan soal dengan tingkat kesukaran tertinggi (+2 *logit*), yang bermakna probabilitas semua siswa untuk mengerjakan soal ini dengan benar kecil sekali. Adapun soal E13 adalah soal dengan nilai *logit* terendah (mendekati -2 *logit*). Dalam hal ini, lebih banyak siswa mampu mengerjakan soal ini dengan benar. Soal E6 dan E8 memiliki tingkat kesulitan yang setara karena nilai *logit*-nya sama, begitu juga soal E2 dan E7.

Nilai rata-rata *logit* butir selalu ditetapkan dalam 0,0 *logit* yang menunjukkan titik acuan awal skala. Dari Peta Wright, rata-rata *logit person* didapati -1 *logit* (di bawah 0,0 *logit*). Hal ini menunjukkan bahwa kemampuan rata-rata siswa di bawah rata-rata tingkat kesukaran standar soal. Dari peta wright juga terlihat bahwa 28 siswa memiliki nilai *logit* di bawah -1 *logit*, mencerminkan bahwa lebih dari separuh sampel memiliki kemampuan berpikir kritis yang rendah.

Menetapkan Validitas dan Klaim Reliabilitas untuk Instrumen Tes

Setelah soal nomor 13 dibuang karena tidak fit (*misfit*), kualitas instrumen tes yang tersusun dari 12 soal dianalisis pada *Summary Statistic* (Tabel 6). Berdasarkan Tabel 6 diperoleh bahwa nilai Infit MNSQ = 1,05 dan Output MNSQ = 1,07. Keduanya memiliki nilai mendekati ideal (nilai ideal MNSQ = 1,00), sehingga instrumen secara keseluruhan baik digunakan untuk pengukuran. Tabel 6 juga menunjukkan bahwa nilai Infit ZSTD = 0,1 dan Outfit ZSTD = 0,2. Nilai ZSTD instrumen mendekati nilai ideal (0,00), sehingga mempunyai perkiraan yang logis. Hasil analisis juga menunjukkan bahwa nilai reliabilitas item adalah 0,72. Hal ini menunjukkan bahwa kualitas butir-butir soal dalam instrumen aspek reliabilitasnya cukup, sesuai acuan kriteria instrumen dalam Rasch Model (Sumintono & Widhiarso, 2015).

Mengembangkan Dokumentasi untuk Instrumen Tes

Dokumentasi memberikan informasi untuk membantu pengguna dalam mengaplikasikan instrumen yang telah dikembangkan. Dokumentasi instrumen tersedia pada website Jurnal Zarah. Sebaran butir soal berdasarkan aspek dan indikator berpikir kritis dapat dilihat pada Tabel 7. Adapun soal nomor 13 tidak dapat digunakan karena tidak fit.

Skor yang diperoleh dari penggunaan instrumen dapat diolah untuk mengetahui sebaran kemampuan berpikir kritis siswa pada tiap aspek berpikir kritis, indikator berpikir kritis, maupun kemampuan berpikir kritis secara keseluruhan. Instrumen juga dapat digunakan untuk membandingkan kemampuan berpikir kritis antara dua atau lebih kelompok siswa pada penerapan strategi pembelajarn yang berbeda.

Tabel 6. Sebaran Butir Soal

Aspek Berpikir Kritis	Indikator Berpikir Kritis	Butir Soal
Memberikan penjelasan sederhana	Memfokuskan pertanyaan	1
	Menganalisis pernyataan	2
	Menjawab pertanyaan tentang suatu penjelasan	3
Membangun keterampilan dasar	Mempertimbangkan apakah sumber dapat dipercaya	4
	Mengamati dan mempertimbangkan suatu laporan hasil observasi	5, 6
Menyimpulkan	Meneduksi dan mempertimbangkan hasil deduksi	7
	Menginduksi dan mempertimbangkan hasil induksi	8, 9
	Membuat dan menentukan nilai pertimbangan	10, 11
Mengatur strategi-taktik	Menentukan tindakan	12

Hasil penelitian ini menunjukkan bahwa Rasch Model sangat praktis dan informatif dalam proses pengembangan instrumen tes berbentuk uraian pada topik ikatan kimia. Hal ini sesuai dengan penelitian penggunaan Rasch Model yang telah dilakukan sebelumnya pada topik kimia yang lain maupun bentuk instrumen yang berbeda (Barbera, 2013; Glynn, 2012; Hadenfeldt dkk., 2013; He dkk., 2016; Herrmann-Abell & DeBoer, 2011; Park dkk., 2017; Pentecost & Barbera, 2013; Priyambodo & Marfuatun, 2016; Towns, 2014; Wei dkk., 2012).

KESIMPULAN

Dari penelitian yang telah dilakukan, diperoleh kesimpulan bahwa terdapat proses khas dari pengembangan instrumen pengukuran kemampuan berpikir kritis pada materi ikatan kimia menggunakan Rasch Model, yaitu pada pendefinisian konstruk, tipe analisis statistik, dan analisis Peta Wright. Instrumen yang telah dikembangkan memiliki reliabilitas cukup dan terdiri dari 12 butir soal yang valid. Analisis dengan Rasch Model memberikan informasi yang lebih komprehensif mengenai karakteristik instrumen yang telah dikembangkan.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Universitas Maritim Raja Ali Haji yang telah mendanai kegiatan penelitian ini melalui skema Penelitian Dosen Muda (PDM) Tahun 2018.

DAFTAR RUJUKAN

- Amalia, N. F., & Susilaningsih, E. (2014). Pengembangan Instrumen Penilaian Keterampilan Berpikir Kritis Siswa SMA pada Materi Asam Basa. *Jurnal Inovasi Pendidikan Kimia*, 8(2), 1380–1389.
- Barbera, J. (2013). A Psychometric Analysis of the Chemical Concepts Inventory. *Journal of Chemical Education*, 90, 546–553.
- Danczak, S. M., Thompson, C. D., & Overton, T. L. (2017). “What does the term Critical Thinking mean to you?” A qualitative analysis of chemistry undergraduate, teaching staff and employers’ views of critical thinking. *Chemistry Education Research and Practice*, 18(3), 420–434.
- Ennis, R. H. (2011). The Nature of Critical Thinking: An Outline of Critical Thinking Dispositions. *University of Illinois*, 1–8.
- Ennis, R. H., Gardiner, W. L., Morrow, R., Paulus, D., & Ringel, L. (1964). *The Cornell Class-Reasoning Test, Form X*. Urbana-Champaign: University of Illinois.
- Ennis, R. H., & Weir, E. (1985). *Ennis-Weir Critical Thinking Essay Test*. Pacific Grove: Midwest Publications.
- Facione, P. A. (1990). *The California Critical Thinking Skills Test - College Level. Experimental Validation and Content Validity*. California Academic Press. La Cruz Ave. Millbrae.
- Glynn, S. M. (2012). International assessment: A Rasch model and teachers’ evaluation of TIMSS science achievement items. *Journal of Research in Science Teaching*, 49(10), 1321–1344.
- Hadenfeldt, J. C., Bernholt, S., Liu, X., Neumann, K., & Parchmann, I. (2013). Using Ordered Multiple-Choice Items To Assess Students’ Understanding of the Structure and Composition of Matter BT - *Journal of Chemical Education*. *Journal of Chemical Education*, 90(12), 1602–1608.
- He, P., Liu, X., Zheng, C., & Jia, M. (2016). Using Rasch measurement to validate an instrument for measuring the quality of classroom teaching in secondary chemistry lessons. *Chem. Educ. Res. Pract.*, 17(2), 381–393.
- Herrmann-Abell, C. F., & DeBoer, G. E. (2011). Using distractor-driven standards-based multiple-choice assessments and Rasch modeling to investigate hierarchies of chemistry misconceptions and detect

- structural problems with individual items. *Chemistry Education Research and Practice*, 12(2), 184–192.
- Kartimi, K., Liliyasi, L., & Permanasari, A. (2012). Pengembangan alat ukur berpikir kritis pada konsep senyawa hidrokarbon untuk siswa SMA di Kabupaten Kuningan. *Jurnal Pendidikan MIPA*, 13(1), 18–25.
- Kartimi, & Liliyasi. (2012a). Implementasi Pengembangan Alat Ukur Berpikir Kritis pada Konsep Kestimbangan Kimia Untuk Siswa Sma. *Jurnal Scientiae Educatia*, 1(2), 1–12.
- Kartimi, & Liliyasi. (2012b). Pengembangan alat ukur berpikir kritis pada konsep termokimia untuk siswa sma peringkat atas dan menengah. *Jurnal Pendidikan IPA Indonesia*, 1(1), 21–26.
- Linacre, J. M. (1994). No Title. *Rasch Measurement Transactions*, 7(4), 328.
- Liu, X. (2010). *Using and Developing Measurement Instruments in Science Education: A Rasch Modeling Approach*. Charlotte, NC: IAP Press.
- Park, M., Liu, X., & Waight, N. (2017). Development of the Connected Chemistry as Formative Assessment Pedagogy for High School Chemistry Teaching. *Journal of Chemical Education*, 94(3), 273–281.
- Pentecost, T. C., & Barbera, J. (2013). Measuring learning gains in chemical education: A comparison of two methods. *Journal of Chemical Education*, 90(7).
- Priyambodo, E., & Marfuatun. (2016). Validity and Reliability of Chemistry Systemic Multiple Choices Questions (CSMCQs). *International Journal of Evaluation and Research in Education*, 5(4), 306–309.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch pada Assessment Pendidikan*. Cimahi: Trim Komunikata.
- Susongko, P. (2010). Perbandingan keefektifan bentuk tes uraian dan. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 14(2), 269–288.
- Towns, M. H. (2014). Guide to developing high-quality, reliable, and valid multiple-choice assessments. *Journal of Chemical Education*, 91(9), 1426–1431.
- Wei, S., Liu, X., Wang, Z., & Wang, X. (2012). Using rasch measurement to develop a computer modeling-based instrument to assess students' conceptual understanding of matter. *Journal of Chemical Education*, 89(3), 335–345.