

Analisis Karakteristik Parameter Butir dan Keterampilan Berpikir Kritis Matematis Siswa dengan *Partial Credit Model (PCM)*

Rts. Ocha Putri Kunanti¹, Ilham Falani^{2✉}, Ade Kumalasari³

^{1,2} Program Studi Pendidikan Matematika, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Jambi, Jambi
Jl. Jambi- Muara Bulian No.KM 15, Jambi, Indonesia
rtsochaputri@email.com

Abstract

Measurement is one aspect that needs to be considered to determine the extent of the quality of learning that has been implemented. Education has a role in building the quality of human resources to face the development of the 21st century, precisely in mathematical critical thinking skills. However, most of the measurement process still uses a classical approach that has not provided in-depth measurement results to analyze the instruments and students' abilities. The purpose of this study was to analyze the instruments and measure critical thinking skills. This study uses a quantitative approach with a quantitative description method. The data used came from the results of student responses based on critical thinking indicators scored with the PCM model. The analysis was carried out with the help of the Winsteps application based on the Rasch Model. The results of this study showed that 88.89% of students matched the model. The level of difficulty of the instrument items is quite diverse and in accordance with the students' abilities. Based on the test information function, it was obtained that the test instrument is suitable for students with high and low abilities.

Keywords: Mathematical Critical Thinking, Partial Credit Model, Rasch Model

Abstrak

Pengukuran merupakan salah satu aspek yang perlu diperhatikan untuk mengetahui sejauh mana kualitas pembelajaran yang telah dilaksanakan. Pendidikan memiliki peran untuk membangun kualitas SDM guna menghadapi perkembangan abad 21 tepatnya pada keterampilan berpikir kritis matematis. Namun, sebagian besar proses pengukuran masih menggunakan pendekatan klasik yang belum memberikan hasil pengukuran mendalam untuk menganalisis instrumen dan kemampuan siswa. Tujuan penelitian ini untuk menganalisis instrumen serta mengukur keterampilan berpikir kritis. Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskripsi kuantitatif. Data yang digunakan berasal dari hasil respon siswa berdasarkan indikator berpikir kritis yang disekor dengan model PCM. Analisis dilakukan dengan berbantuan aplikasi Winsteps yang berbasis Rasch Model. Hasil penelitian ini didapatkan bahwa penggunaan PCM dapat memberikan hasil pengukuran yang akurat dan rinci terhadap variabel kemampuan dan butir instrumen. Analisis kecocokan model menunjukkan dari seluruh peserta didik 88,89% cocok dengan model. Tingkat kesukaran butir instrumen cukup beragam dan sesuai dengan kemampuan siswa. Berdasarkan fungsi informasi tes diperoleh bahwa instrumen tes cocok untuk siswa dengan kemampuan tinggi dan rendah.

Kata kunci: Berpikir Kritis Matematis, *Partial Credit Model*, *Rasch Model*

Copyright (c) 2025 Rts. Ocha Putri Kunanti, Ilham Falani, Ade Kumalasari

✉ Corresponding author: Ilham Falani

Email Address: ilhamfalani@unja.ac.id (Jl. Jambi- Muara Bulian KM 15, Jambi, Indonesia)

Received 21 June 2025, Accepted 04 August 2025, Published 27 August 2025

DoI: <https://doi.org/10.31004/cendekia.v9i3.4229>

PENDAHULUAN

Perkembangan dunia telah memasuki abad 21 dengan ditandai dengan adanya kemajuan ilmu pengetahuan dan teknologi yang begitu pesat. Partnership for 21st Century Skills mengidentifikasi keterampilan yang dibutuhkan pada abad 21 yaitu komunikasi, kolaborasi, berpikir kritis dan kreatifitas (Pratiwi et al., 2019). Salah satu kompetensi yang diprioritaskan pada dunia pendidikan abad 21 saat ini yaitu berpikir kritis (Halim, 2022).

Berpikir kritis adalah suatu proses yang melibatkan analisis, penilaian, pencarian solusi, dan penarikan kesimpulan terhadap suatu situasi atau masalah (Sofri et al., 2020). Keterampilan berpikir

kritis tidak hanya semata harus diasah, tetapi diperlukan proses pengukuran sebagai kegiatan evaluasi dari pengajar (Nuridha & Hardianti, 2022). Keakuratan hasil pengukuran keterampilan peserta didik sangat bergantung pada instrumen tes yang digunakan sehingga perlu dilakukan analisis butir soal. Hasil respon peserta didik perlu dilakukan analisis untuk membuktikan hasil pengukuran dalam pendidikan (Falani, 2023). Dalam pelaksanaannya, terdapat dua teori yang digunakan dalam pengukuran yaitu teori pengukuran klasik dan teori pengukuran modern.

Teori pengukuran klasik atau *Classical True-Score Theory* merupakan teori yang memiliki prinsip bahwa skor pengukuran diperoleh dari penjumlahan skor sebenarnya dengan kesalahan (Sumaryanta, 2021). Aspek yang menentukan kualitas butir soal pada pendekatan ini adalah tingkat kesukaran soal namun memiliki karakteristik yang inkonsisten (Perdana, 2018). Tingkat kesukaran sangat bergantung pada kemampuan peserta tes sehingga memberikan hasil yang berubah-ubah (Erfan et al., 2020; Hardianti et al., 2023; Perdana, 2018). Maka dari itu diperlukan pendekatan baru yang mampu meminimalisir kekurangan dari pendekatan teori klasik yaitu Teori Pengukuran Modern.

Teori Pengukuran Modern (*Items Response Theory*) adalah suatu pendekatan dalam pengukuran yang menggali hubungan antara kemampuan individu dan kemungkinan mereka menjawab butir dengan benar (Amelia & Kriswantoro, 2017). Rasch Model merupakan model dari IRT yang dapat menganalisis bentuk soal salah satunya politomi dengan aplikasi Winsteps. Model yang umumnya digunakan dari Rasch Model adalah *Partial Credit Model (PCM)*.

Partial Credit Model (PCM) merupakan bagian dari model Rasch pengembangan dari Model IRT berbentuk politomi yang mempunyai *1 parameter of logistic (1-PL)* yaitu tingkat kesulitan (Purwanti & Sumandya, 2019). Asumsi pada PCM yakni setiap butir mempunyai daya beda yang sama. PCM dapat menyekor butir dalam kategori berjenjang, namun indeks kesukaran dalam setiap langkah tidak perlu terurut (WD Arif, T Rijanto, R Harimurti, 2023). PCM mampu mengevaluasi tes berlandaskan parameter butir dan kemampuan siswa sehingga kesesuaian instrumen tes terhadap tingkat kemampuan siswa dapat diketahui (Purwanti & Sumandya, 2019).

Penelitian ini memiliki keterkaitan dengan beberapa penelitian terdahulu. WD Arif, et al (2023) meneliti tentang penerapan PCM untuk meningkatkan kemampuan berpikir kritis siswa, dan menunjukkan hasil bahwa instrumen soal sesuai dengan model PCM dengan aspek unidimensional 40,4% serta hasil peta wright maps berada dalam rentang 2 sampai -2. Penelitian lain yang dilakukan oleh Fernanda & Wulan (2024) juga menemukan bahwa nilai infit pada setiap soal berada pada rentang 0,5 sampai 1,5 serta didapatkan informasi butir soal yang sesuai dan tidak sesuai dengan kemampuan peserta didik. Kesamaan penelitian ini dengan penelitian sebelumnya terletak pada model dan aspek yang dianalisis, sedangkan perbedaannya terdapat pada analisis butir yang dirincikan berdasarkan indikator keterampilan berpikir kritis matematis. Oleh karena itu, penelitian ini diharapkan dapat memperkuat dan melengkapi temuan-temuan yang telah ada.

Penelitian ini bertujuan untuk menganalisis instrumen serta mengukur keterampilan berpikir kritis siswa. Analisis menggunakan Winsteps berbasis *Rasch Model* dengan model penyekoran *Partial*

Credit Model (PCM). Hasil analisis diperoleh dari kesesuaian model dengan responden, tingkat kesukaran butir, estimasi kemampuan siswa, *wright maps* dan fungsi informasi tes.

METODE

Penelitian ini menggunakan pendekatan deskripsi kuantitatif. Tujuan dilakukan penelitian ini untuk mendeskripsikan hasil analisis instrumen dan kemampuan siswa dalam berpikir kritis matematis. Subjek penelitian terdiri atas seluruh siswa kelas VII semester genap tahun 2024/2025 di SMP Negeri Kecamatan Bahar Utara Kabupaten Muaro Jambi. Teknik pengumpulan data pada penelitian ini melalui pemberian tes keterampilan berpikir kritis matematis. Instrumen tes terdiri dari 2 butir soal berbentuk uraian. Proses validasi instrumen dilakukan oleh 5 pakar/ahli matematika. Hasil validasi dinilai dengan metode Aiken dan diperoleh hasil perhitungan 0.9 dan 0.95 yang mana lebih besar dari 0.8 sehingga instrumen dikatakan valid. Pengambilan sampel menggunakan teknik total sampling dimana seluruh populasi dijadikan sampel. Data yang digunakan berupa hasil respon siswa dari instrumen berpikir kritis matematis. Model penyekoran dengan format (0,1,2,3) menggunakan Partial Credit Model (PCM). Data dianalisis dengan Rasch Model berbantuan software Winsteps. Hasil analisis Winsteps berupa kesesuaian model dengan responden, tingkat kesukaran butir, estimasi kemampuan siswa, *wright maps* dan fungsi informasi.

HASIL DAN DISKUSI

Uji Prasyarat

Langkah awal untuk menganalisis menggunakan pendekatan teori respon butir adalah uji prasyarat yang sering dikenal sebagai uji asumsi teori. Jika hal ini dapat dipenuhi maka proses analisis dapat dijalankan. Uji asumsi teori terdiri atas uji unidimensional, independensi lokal dan invariansi parameter (Hambleton, Swaminathan, & Rogers; 1991). Unidimensional adalah komponen utama untuk memastikan instrumen dapat mengukur objek yang diukur, yang mana pada penelitian ini aspek keterampilan berpikir kritis matematis. Kriteria yang ditinjau untuk memastikan aspek unidimensional terlihat pada komponen Raw variance explained by measure..

Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)			
		-- Empirical --	Modeled
Total raw variance in observations	=	36.7 100.0%	100.0%
Raw variance explained by measures	=	24.7 67.3%	67.3%
Raw variance explained by persons	=	11.6 31.7%	31.7%
Raw Variance explained by items	=	13.1 35.6%	35.6%
Raw unexplained variance (total)	=	12.0 32.7%	100.0% 32.7%
Unexplned variance in 1st contrast	=	3.8 10.4%	31.9%
Unexplned variance in 2nd contrast	=	1.9 5.3%	16.2%
Unexplned variance in 3rd contrast	=	1.4 3.8%	11.6%
Unexplned variance in 4th contrast	=	1.0 2.8%	8.6%
Unexplned variance in 5th contrast	=	.9 2.4%	7.3%

Gambar 1. *Butir Dimensionaly*

Aspek unidimensional minimal yang dipenuhi sebesar 20%, jika lebih dari 40% dikategorikan lebih bagus, lebih dari 60% dikategorikan istimewa (Kumalasari & Mahmudi, 2024). Berdasarkan

hasil Gambar 1 didapatkan hasil *raw variance* data sebesar 67,3% termasuk kategori sangat istimewa. Adapun nilai *Unexplned Variance* terdiri dari 10,4% dan selebihnya dibawah 5,3% yang mana tidak melebihi 15%. Hal ini sejalan dengan penelitian Prayoga et al., (2024) bahwa instrumen yang digunakan sudah memenuhi syarat unidimensi yang mampu mengukur satu dimensi yaitu keterampilan berpikir kritis matematis.

Selanjutnya adalah independensi lokal merupakan aspek yang meninjau hubungan antara respon terhadap butir tes. Independensi lokal dapat terpenuhi jika tidak ada keterkaitan antara respon siswa terhadap satu butir dengan butir yang lainnya (Alfarisa & Purnama, 2019). Asumsi yang terakhir adalah invariansi parameter dimana karakteristik parameter butir tidak dipengaruhi oleh respon siswa yang beragam (Suwanto, 2016). Suatu butir memiliki tingkat kesukaran yang tetap meskipun dikerjakan oleh siswa dengan kemampuan tinggi atau rendah. Independensi lokal dan invariansi parameter dapat terpenuhi jika aspek unidimensi terpenuhi (Hambleton et al., 1992)

Estimasi Parameter Keterampilan Berpikir Kritis Matematis

Tingkat kecocokan atau kesesuaian responden (*person fit*)

Analisis ini menunjukkan apakah terdapat kecocokan antara responden dengan model analisis yang digunakan yaitu *Rasch Model*. Model ini dapat mengidentifikasi pola jawaban responden yang berbeda atau tidak sesuai dengan model (Fitria et al., 2024). Dalam hal ini responden dikatakan cocok dengna model jika memberikan hasil yang konsisten dan sesuai dengan model. Hasil analisisnya sangat berguna bagi guru dalam menilai konsistensi pemikiran siswa serta mendeteksi potensi kecurangan saat menjawab soal atau butir (Kumalasari & Mahmudi, 2024).

Person STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	TOTAL MEASURE	MODEL S. E.		INFIT		OUTFIT		PT-MEASURE		EXACT MATCH		Person
				MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%			
15	27	12	1.15	.36	1.69	1.4	9.90	4.5	A	.12	.66	58.3	55.8	A019
61	12	12	-.77	.35	4.85	4.8	8.17	4.2	B	-.29	.71	16.7	50.7	A078
12	32	12	1.93	.45	1.04	.3	6.38	2.3	C	.07	.47	58.3	72.7	A014
19	26	12	1.02	.36	.63	-.8	5.53	2.5	D	.59	.69	83.3	56.1	A023
74	25	12	.89	.35	2.33	2.3	5.41	2.6	E	.27	.71	33.3	59.7	A093
31	2	12	-2.43	.64	.75	.1	4.94	2.0	F	-.03	.31	83.3	87.0	A037

84	19	12	.14	.36	.34	-1.6	.28	-1.2	g	.94	.79	75.0	61.6	A107
30	22	12	.52	.35	.31	-1.9	.26	-1.0	f	.91	.76	66.7	58.7	A036
44	22	12	.52	.35	.31	-1.9	.26	-1.0	e	.91	.76	66.7	58.7	A051
82	22	12	.52	.35	.25	-2.3	.26	-1.0	d	.92	.76	83.3	58.7	A105
79	3	12	-2.10	.52	.21	-1.2	.14	-.7	c	.70	.37	91.7	77.2	A099
87	4	12	-1.87	.45	.21	-1.7	.16	-.8	b	.76	.42	83.3	69.3	A113
23	29	12	1.43	.38	.19	-2.5	.15	-.6	a	.91	.60	75.0	66.9	A028

MEAN	18.6	12.0	.09	.38	.96	-.2	1.20	.1				61.9	61.7	
S. D.	8.8	.0	1.18	.05	.73	1.3	1.76	1.2				13.8	7.0	

Gambar 2. *Person Fit*

Untuk memeriksa responden yang fit dan tidak fit dapat ditinjau dari nilai INFIT MNSQ dari setiap *person*, yang mana nilai tersebut tidak melebihi penjumlahan nilai rata-rata dan standar deviasi (Sumintono & Widhiarso, 2014). Jika nilai logit yang diperoleh lebih besar dari akumulasi skor

tersebut maka dikatakan *misfit*.

Tabel 1. Klasifikasi *Person Fit*

No	Kriteria	Jumlah responden (Person)	
		Cocok	Tidak cocok
1	INFIT MNSQ > MEAN INFIT MNSQ + S.D INFIT MNSQ > 0,96 + 0,73	80 Responden	10 Responden
Persentase		88,89%	11,11%

Berdasarkan analisis tersebut diperoleh akumulasi logit butir dari rata-rata dan standar deviasi : $0,96 + 0,73 = 1,69$. Dari 90 responden terdapat 80 responden (88,89%) yang termasuk kategori *person fit* dan 10 responden(11,11%) yang tidak fit (*misfit*). Sehingga dapat disimpulkan sebagian besar responden cocok dengan model yang digunakan.

Tingkat Keterampilan Berpikir Kritis Matematis

Person Measure berfungsi untuk menampilkan data secara merinci dan menyeluruh mengenai abilitas peserta didik berdasarkan rata-rata jawaban pada setiap butir butir. Kolom *measure* menunjukkan nilai kemampuan siswa yang dinyatakan dalam satuan logit. Terdapat 12 butir pada instrumen tes dengan skor maksimal adalah 3, jumlah skor tertinggi untuk soal secara keseluruhan adalah 36.

Person STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	TOTAL MEASURE	MODEL S.E.	INFIT		OUTFIT		PT-MEASURE		EXACT MATCH		Person
					MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%	
7	34	12	2.43	.59	1.01	.3	.33	-.2	.41	.34	83.3	81.0	A007
11	34	12	2.43	.59	.71	.0	.21	-.4	.47	.34	83.3	81.0	A012
2	33	12	2.14	.50	.81	-.1	.32	-.2	.49	.41	75.0	79.3	A002
1	32	12	1.93	.45	.37	-1.2	.17	-.5	.65	.47	75.0	72.7	A001
3	32	12	1.93	.45	.56	-.7	.68	-.2	.53	.47	75.0	72.7	A003
12	32	12	1.93	.45	1.04	.3	6.38	2.3	.07	.47	58.3	72.7	A014
57	32	12	1.93	.45	.37	-1.2	.17	-.5	.65	.47	75.0	72.7	A068
.....													
77	6	12	-1.53	.38	.47	-1.3	.28	-.7	.70	.51	58.3	64.3	A096
78	5	12	-1.69	.41	.34	-1.5	.23	-.7	.72	.47	83.3	65.0	A097
64	4	12	-1.87	.45	.28	-1.4	.43	-.2	.65	.42	83.3	69.3	A082
66	4	12	-1.87	.45	.28	-1.4	.43	-.2	.65	.42	83.3	69.3	A084
87	4	12	-1.87	.45	.21	-1.7	.16	-.8	.76	.42	83.3	69.3	A113
79	3	12	-2.10	.52	.21	-1.2	.14	-.7	.70	.37	91.7	77.2	A099
31	2	12	-2.43	.64	.75	.1	4.94	2.0	-.03	.31	83.3	87.0	A037
.....													
MEAN	18.6	12.0	.09	.38	.96	-.2	1.20	.1			61.9	61.7	
S.D.	8.8	.0	1.18	.05	.73	1.3	1.76	1.2			13.8	7.0	

Gambar 3. *Person Measure*

Gambar 3 menunjukkan urutan tertinggi ke terendah berdasarkan perolehan skor benar terbanyak. Nilai *person* memiliki nilai rata-rata tertinggi dan terendah. Rata-rata nilai responden adalah 0,09 yang mana nilai tersebut lebih dari logit 0,0. Hal ini menunjukkan bahwa responden memiliki kecenderungan menjawab soal dengan benar. Menurut Kumalasari & Mahmudi (2024) kategori tingkat kemampuan siswa dapat diklasifikasi berdasarkan *mean measure dan standar deviation measure* yang disajikan sebagai berikut.

Tabel 2. Klasifikasi *Person Measure*

Nilai Logit Abilitas Siswa	Kriteria	Jumlah Siswa
< 0,09	Rendah	46 responden
0,1 – 1,17	Sedang	25 responden
>1,18	Tinggi	19 responden

Tabel 2 menyajikan klasifikasi tingkat kemampuan siswa pada aspek keterampilan berpikir kritis matematis. Berdasarkan kategori tersebut didapatkan siswa yang memiliki kemampuan rendah berjumlah 46 siswa, berkemampuan sedang 25 siswa dan berkemampuan tinggi 19 siswa. Abilitas yang paling tinggi dimiliki oleh A007 dan A012 (measure = +2,43 logit) dengan perolehan skor 34 poin yang mengindikasikan bahwa siswa tersebut hampir menjawab seluruhnya dengan benar. Begitupun sebaliknya abilitas yang paling rendah yang dimiliki oleh A037 (measure = -2,43 logit) dengan total skor 2 artinya siswa tersebut sebagian besar tidak menjawab soal dengan benar. Sejalan dengan (Kumalasari & Mahmudi, 2024) yang menyatakan bahwa siswa yang memiliki nilai logit tinggi mengindikasikan bahwa siswa tersebut memiliki probabilitas menjawab soal dengan benar begitupun sebaliknya. Berdasarkan hasil *person fit* terlihat bahwa sebagian besar peserta didik berada dalam kategori sedang dan rendah sehingga diperlukan peningkatan kualitas pembelajaran guna meningkatkan keterampilan berpikir kritis matematis siswa.

Estimasi Parameter Butir

Tingkat kesukaran butir dapat dilihat berdasarkan Output Tabel *Items Measure Order*. Kategori tingkat kesukaran butir dapat dikelompokkan berdasarkan *mean measure* dan standar deviasi (SD) Sumintono & Widhiarso (2015). Hasil tabel tersebut diketahui nilai Standar Deviasi (SD) sebesar 1,14. Berikut pengelompokan tingkat kesukaran butir disajikan sebagai berikut.

Tabel 6. Klasifikasi *Kesukaran Butir*

<i>Tabel Kesukaran Butir</i>	
Kategori	Interval
Sangat Sukar	$TK > 1,14$
Sukar	$1,14 \geq TK > 0,00$
Mudah	$0,00 \geq TK \geq -1,14$
Sangat Mudah	$< -1,14$

Kategori tingkat kesukaran butir dibagi menjadi 4 berdasarkan interval logit, kemungkinan besar berbasis mean = 0 dan standar deviasi 1,03. Interval dibagi berdasarkan nilai logit dari mean = 0 dengan batas ± 1 SD (yaitu $\pm 1,14$) digunakan sebagai batas kategorinya. Interval kategori tersebut digunakan untuk mengklasifikasikan tingkat kesukaran butir dalam model *Rasch*.

Item STATISTICS: MEASURE ORDER

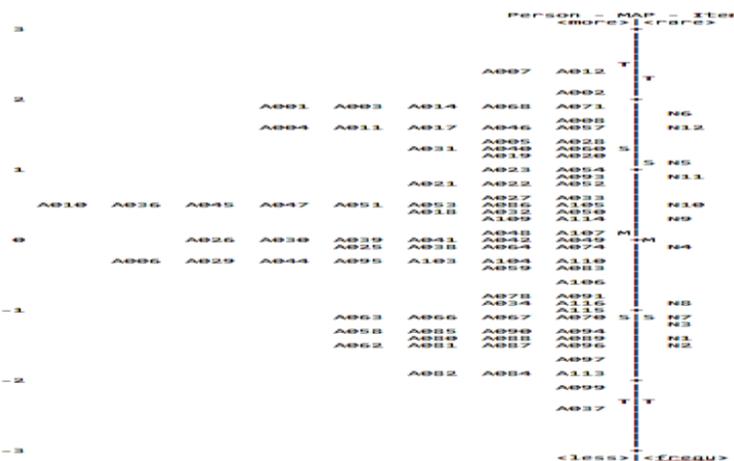
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFINIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PT-MEASURE CORR.	EXP.	EXACT OBS%	MATCH EXP%	Item
6	39	90	1.79	.16	1.08	.4	.96	.0	.52	.54	68.9	71.0	N6
12	49	90	1.55	.15	.72	-1.5	1.34	1.0	.63	.59	72.2	66.9	N12
5	74	90	1.06	.13	1.22	1.2	.79	-.6	.67	.67	56.7	56.9	N5
11	81	90	.93	.13	.70	-1.8	.87	-.4	.76	.69	65.6	58.2	N11
10	106	90	.52	.12	1.02	.2	1.02	.2	.73	.72	52.2	53.4	N10
9	119	90	.32	.12	1.26	1.6	1.07	.3	.70	.73	53.3	53.1	N9
4	145	90	-.07	.12	1.02	-.2	1.02	-.2	.71	.74	46.7	52.2	N4
8	193	90	-.85	.13	.70	-1.7	.43	-1.8	.78	.69	56.7	57.2	N8
7	207	90	-1.12	.14	.96	-.2	.87	-.2	.68	.66	62.2	63.4	N7
3	212	90	-1.22	.14	.96	-.1	.84	-.2	.65	.64	71.1	68.0	N3
1	220	90	-1.38	.15	1.10	-.6	2.64	2.5	.50	.61	67.8	69.8	N1
2	227	90	-1.54	.15	1.26	1.3	3.43	3.0	.48	.58	70.0	70.5	N2
MEAN	139.3	90.0	.00	.14	1.00	.0	1.27	.3			61.9	61.7	
S.D.	67.3	.0	1.14	.01	.20	1.1	.83	1.3			8.2	7.0	

Gambar 4. *Items Measure*

Nilai kesukaran butir dapat dilihat dari kolom *measure* yang dinyatakan dalam satuan logit. Nilai *measure* diurutkan dari urutan tertinggi sampai terendah dimana nilai terbesar menyatakan butir tersukar, dan begitupun sebaliknya. Berdasarkan Gambar 4 didapatkan bahwa butir 6 dan 12 termasuk dalam kategori sangat sukar. Kategori sukar pada butir 5,11,10. Adapun kategori mudah terdapat pada butir 4,8, dan 7 sementara pada kategori sangat mudah terdapat tiga butir, yakni butir nomor 3, 1 dan 2.

Berdasarkan indikator keterampilan berpikir kritis matematis ditunjukkan bahwa butir 6 dan 12 merupakan indikator *overview*. Hasil analisis menunjukkan *overview* termasuk kategori sangat sukar dimana aspek ini sebagian besar siswa masih mengalami kendala dalam membuat interpretasi pada hasil penyelesaian soal. Indikator *clarity, situation & inference* termasuk dalam kategori sukar dimana siswa menyajikan penjelasan dari hasil jawaban, siswa mengkaitkan hal yang memiliki hubungan dalam penyelesaian masalah, dan siswa membuat kesimpulan. Indikator *focus & reason* termasuk dalam kategori mudah dan sangat mudah dimana siswa dapat menentukan apa yang diketahui dan ditanya sebelum menyelesaikan permasalahan.

Wright Map

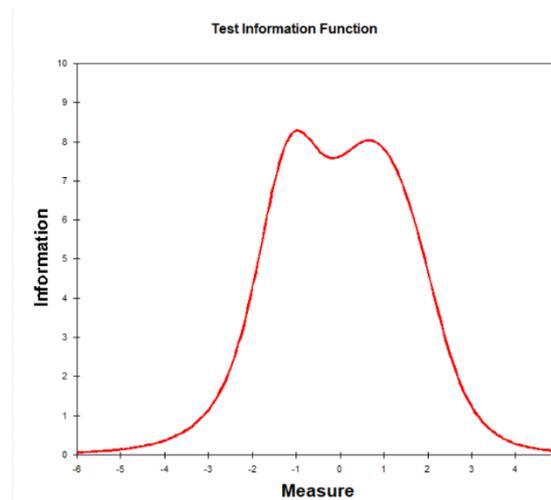


Gambar 5. *Wright Maps*

Hasil dari Wright Map menggambarkan persebaran kemampuan subjek dan sebaran tingkat

kesulitan butir yang dinyatakan dalam skala yang sama (Falani et al., 2022). Pada sisi sebelah kiri menunjukkan persebaran kemampuan subjek sementara sisi sebelah kanan adalah persebaran butir. Dari peta tersebut dapat diketahui bahwa terdapat beberapa siswa yang memiliki kemampuan lebih tinggi dibandingkan dengan tingkat kesukaran butir. Adapun siswa A007 dan A012 dengan nilai logit 2,43 dan mampu menjawab butir N6 dengan tingkat kesukaran 1,79 kategori sangat sukar. Siswa yang menempatkan posisi paling atas menunjukkan paling banyak menjawab butir. Kemudian terdapat satu siswa yang dengan kategori tertinggi kedua sebesar 2,14 yaitu A102. Sementara terdapat siswa dengan kemampuan paling rendah yaitu A037 dengan perolehan nilai sebesar -2,43 mengindikasikan bahwa siswa tersebut paling sedikit menjawab butir.

Fungsi Informasi Tes



Gambar 6. Fungsi Informasi Tes

Fungsi informasi tes (*Test Information Function*) bertujuan untuk memberikan informasi mengenai karakteristik kemampuan siswa terhadap parameter instrumen tes yang diujikan (Kumalasari & Mahmudi, 2024). Sumbu koordinat X menunjukkan nilai kemampuan siswa dan sumbu koordinat Y menunjukkan nilai informasi tes. Berdasarkan Gambar 6 didapatkan Nilai *TIF* (*Test Information Function*) maksimum dari instrumen keterampilan berpikir kritis matematis adalah 8,285 pada kemampuan siswa sekitar -1,1 dan 0,99. Dalam hal ini instrumen keterampilan berpikir kritis dapat memberikan informasi yang baik untuk siswa yang berada pada kemampuan sekitar -1,1 dan 0,99. Menurut (Kumalasari & Mahmudi, 2024) jika grafik menunjukkan kurva yang condong ke kiri dan kanan membentuk lembah di tengah mengindikasikan kemampuan siswa terhadap tes terbagi menjadi dua yaitu berkemampuan tinggi dan rendah. Sehingga instrumen ini cocok digunakan untuk mengukur kemampuan siswa yang beragam.

Penelitian ini memiliki keterbatasan yang perlu dilengkapi lebih lanjut untuk penelitian berikutnya. Adapun model yang digunakan pada penelitian ini menggunakan model PCM, diharapkan lebih lanjut dapat menggunakan model IRT lainnya agar hasil analisis dari pengukuran dapat lebih

akurat dengan mempertimbangkan aspek lainnya seperti daya beda dan peluang menebak benar. Selain itu, bentuk tes pada instrumen berupa uraian dimana untuk penelitian lebih lanjut dapat menggunakan bentuk tes kombinasi antara bentuk pilihan ganda dan uraian.

KESIMPULAN

Berdasarkan hasil analisis terhadap respon siswa diperoleh bahwa instrumen tes yang disajikan sudah memenuhi prasyarat. Hal ini ditunjukkan dari hasil pengukuran *raw variance* sebesar 65,2%. Kesesuaian siswa dengan model ditunjukkan dari hasil *output Person Fit* sebanyak 80 siswa. Adapun siswa dengan kemampuan tertinggi dengan nilai logit +2,43 sementara terendah yaitu -2,43. Tingkat kesukaran butir soal yang diujikan cukup beragam dari sangat mudah sampai sangat sukar. Terdapat kesetaraan antara kemampuan siswa dengan tingkat kesulitan butir terlihat dari peta *wright maps* dimana setiap butir dapat mewakili siswa. Hal ini didukung dengan hasil fungsi informasi tes yang menunjukkan instrumen tes yang diujikan cocok untuk mengukur siswa berkemampuan tinggi dan rendah. Adapun saran dari penelitian ini yaitu dapat dilakukan analisis lebih mendalam mengenai kemampuan dan instrumen dengan fitur lain dari software Winsteps dengan analisis PCM. Untuk penelitian selanjutnya, disarankan agar hasil analisis kemampuan dan kualitas butir soal yang diperoleh dari satuan logit dapat disajikan dalam skala nilai 0–100. Penyajian dalam skala ini diharapkan dapat memudahkan interpretasi hasil oleh pembaca maupun praktisi pendidikan, sehingga informasi yang diperoleh menjadi lebih jelas, aplikatif, dan dapat dibandingkan dengan standar penilaian yang umum digunakan di sekolah.

REFERENSI

- Alfarisa, F., & Purnama, D. N. (2019). *Analisis Butir Soal Ulangan Akhir Semester Mata Pelajaran Ekonomi SMA Menggunakan RASCH Model*. 11(2).
- Amelia, R. N., & Kriswantoro, K. (2017). Implementasi Item Response Theory sebagai Basis Analisis Kualitas Butir Soal dan Kemampuan Kimia Siswa Kota Yogyakarta. *JKPK (Jurnal Kimia Dan Pendidikan Kimia)*, 2(1), 1. <https://doi.org/10.20961/jkpk.v2i1.8512>
- Erfan, M., Mauliyda, M. A., Hidayati, V. R., Astria, F. P., & Ratu, T. (2020). Tes Klasik Dan Model Rasch. *Indonesian Journal of Educational Research and Review*, 3(1), 11–19. <https://ejournal.undiksha.ac.id/index.php/IJERR/article/view/24080/pdf>
- Falani, I. (2023). *Desain dan Validasi Aplikasi Tes Literasi Matematika Berbasis Komputer dengan Pendekatan Item Response Theory Design and Validation of Computer-Based Test of Mathematical Literacy with Item Response Theory Approach* (Vol. 12, Issue 2).
- Falani, I., Iriyadi, D., Wardani, I. Y., Susanti, H., & Nasution, R. A. (2022). A Rasch Analysis of Perceived Stigma of Covid-19 among Nurses in Indonesia Questionnaire. *Psychological Thought*, 15(1), 12–28. <https://doi.org/10.37708/psyc.v15i1.530>
- Fernanda, J. W., & Wulan, E. R. (2024). Analisis Polychotomous Rasch Model untuk Kalibrasi Soal

- Analisis Real pada Program Studi Tadris Matematika. *Jurnal Universitas Muhammadiyah Metro*, 5(1), 157–166. <https://scholar.ummetro.ac.id/index.php/emteka/article/download/4177/2405/>
- Fitria, A., Siburian, J., Falani, I., & Muhammad, D. (2024). Applying the Rasch Model to Assess Retention and Transfer Test Instruments in Science Education on Additive and Addictive Substances. *Integrated Science Education Journal*, 5(2), 101–109. <https://doi.org/10.37251/isej.v5i2.864>
- Halim, A. (2022). Signifikansi dan Implementasi Berpikir Kritis dalam Proyeksi Dunia Pendidikan Abad 21 Pada Tingkat Sekolah Dasar. *Jurnal Indonesia Sosial Teknologi*, 3(3), 404–418. <https://doi.org/10.36418/jist.v3i3.385>
- Hambleton, R. K., Swaminathan, H., & Rogers, R. J. (1992). Fundamentals of item response theory. In *Choice Reviews Online* (Vol. 29, Issue 07). <https://doi.org/10.5860/choice.29-4185>
- Hardianti, H., Liliawati, W., & Tayubi, Y. R. (2023). Karakteristik tes kemampuan berpikir kritis siswa SMA pada materi momentum dan impuls: Perbandingan classical theory test (CTT) dan model Rasch. *WaPFI (Wahana Pendidikan Fisika)*, 8(1), 21–28. <https://doi.org/10.17509/wapfi.v8i1.30958>
- Kumalasari, E. D., & Mahmudi, I. (2024). *Analisis Pemodelan Rasch Pada Asesmen Pendidikan* (Issue February).
- Nuridha, S., & Hardianti, R. D. (2022). Pengukuran Critical Thinking Skills Siswa Menggunakan Four-Tier Multiple Representation Test. *Proceeding Seminar Nasional IPA XII*, 234–238.
- Perdana, S. A. (2018). Analisis Kualitas Instrumen Pengukuran Pemahaman Konsep Persamaan Kuadrat Melalui Teori Tes Klasik Dan Rasch Model. *Jurnal Kiprah*, 6(1), 41–48. <https://doi.org/10.31629/kiprah.v6i1.574>
- Pratiwi, S. N., Cari, C., & Aminah, N. S. (2019). Pembelajaran IPA abad 21 dengan literasi sains siswa. *Jurnal Materi Dan Pembelajaran*, 9, 34–42. <https://jurnal.uns.ac.id/jmpf/article/view/31612%0Ahttps://jurnal.uns.ac.id/jmpf/article/download/31612/21184>
- Prayoga, K. P., Suryana, D., Supriatna, M., & Budiman, N. (2024). Penggunaan Rasch Model Untuk Menganalisis Konstruksi Instrumen Kontrol Diri Pada Siswa Sekolah Menengah. *G-Couns: Jurnal Bimbingan Dan Konseling*, 9(1), 367–381. <https://doi.org/10.31316/gcouns.v9i1.4459>
- Purwanti, N. K. R., & Sumandya, I. W. (2019). Penerapan Partial Credit Model (PCM) dalam Mengevaluasi Tes Uraian. *Jurnal EMASAINS*, VIII(1), 77–85.
- Sofri, D., Arif, F., & Nur, A. (2020). *Analisis Kemampuan Berpikir Kritis Matematis Pada Model Problem Based Learning (PBL) Berbantu Media Pembelajaran Interaktif dan Google Classroom. 2018*.
- Sumintono, B., & Widhiarso, W. (2014). *Aplikasi model rasch: Untuk penelitian ilmu-ilmu sosial*. Trim Komunikata Publishing House.
- Suwarto. (2016). Daya Beda , Tingkat Kesulitan , dan Tebaan Tes Biologi Kelas 8 Semester Gasal

Discrimination , Difficulty , and Guessing The Biology Test 8 th Grade By The Period Of The Odd Term. *Proceeding Biology Education Conference (ISSN: 2528-5742)*, 13(1), 151–158.

WD Arif, T Rijanto, R Harimurti, Y. F. (2023). Penerapan Partial Credit Model (PCM) Untuk Meningkatkan Kemampuan Berpikir Kritis Siswa Pada Mata Pelajaran Perbaikan Peralatan Listrik Pada Kelas XI Titl Di SMK Negeri 1 Kediri. *Jurnal Elektronika Dan Teknik Informatika Terapan (JENTIK)*, 1(3), 249–262. <https://doi.org/10.59061/jentik.v1i3.405>