

Validitas Psikometrika Instrumen *Self Regulated Learning* dalam Pembelajaran Matematika SMA

Jihan Rofifah¹, Wardani Rahayu², Flavia Aurelia Hidajat³
^{1,2,3} Magister Pendidikan Matematika, Universitas Negeri Jakarta, Indonesia

Email: ✉ rofifahjihan31@gmail.com

Article Info	Abstract
Article History Submitted : 01-01-2025 Revised : 14-03-2025 Accepted : 25-03-2025	<i>Self Regulated Learning</i> adalah kemampuan seseorang untuk mengatur dirinya sendiri dalam proses belajar untuk membantu meningkatkan hasil belajar peserta didik. Dalam matematika, <i>Self Regulated Learning</i> merupakan strategi belajar yang dapat membantu peserta didik menjadi lebih mandiri dan meningkatkan prestasi akademik di mata pelajaran matematika. Penelitian ini bertujuan untuk menguji validitas psikometri <i>Self Regulated Learning</i> pada mata Pelajaran matematika untuk peserta didik Sekolah Menengah Atas (SMA). Responden penelitian ini adalah peserta didik kelas X pada SMA Negeri di Kota Jakarta Utara sebanyak 300 peserta didik. Hasil penelitian menggunakan winstep menunjukkan terdapat 40 pernyataan yang valid. Dengan nilai unidimensionality 39.4% (kategori cukup). Selain itu, Nilai <i>Cronbach alpha</i> (KR-20) person raw score sebesar 0,93 dan nilai item reliability sebesar 0.99 dapat dikatakan bahwa instrumen <i>Self Regulated Learning</i> reliabel.
Keywords: <i>Rasch Model; Validity; Reliability; Self Regulated Learning</i>	<i>Self Regulated Learning is a person's ability to regulate themselves in the learning process to help improve student learning outcomes. In mathematics, Self Regulated Learning is a learning strategy that can help students become more independent and improve academic achievement in mathematics. This study aims to test the psychometric validity of Self Regulated Learning in mathematics for Senior High School (SMA) students. The respondents of this study were 300 students in grade X at a State Senior High School in North Jakarta City. The results of the study using winstep showed that there were 40 valid statements. With a unidimensionality value of 39.4% (sufficient category). In addition, the Cronbach alpha (KR-20) person raw score value of 0.93 and the item reliability value of 0.99 can be said that the Self Regulated Learning instrument is reliable.</i>

PENDAHULUAN

Penting bagi setiap guru untuk mengajarkan peserta didik bagaimana cara mengatur strategi belajar sendiri atau yang bisa kita sebut dengan *Self Regulated Learning* (SLR) (Nugroho et al., 2018). SLR merupakan strategi belajar yang dapat membantu peserta didik menjadi lebih mandiri dan meningkatkan prestasi akademiknya (Cahyanto & Afifulloh, 2021). SRL dapat membantu peserta didik dalam menghadapi tantangan abad 21, di mana ilmu pengetahuan dan teknologi berkembang pesat. Dengan kemampuan SRL, peserta didik dapat mewujudkan keinginannya secara mandiri dan tidak bergantung pada orang lain (Fitriyah & Puspasari, 2021). SRL didasarkan pada paham konstruktivisme (Heri Retnawati, 2022). Zimmerman (1989) menyebutkan bahwa SRL terdiri dari tiga aspek, yaitu metakognisi, motivasi, dan perilaku.

Self Regulated Learning merupakan salah satu konsep penting dalam pendidikan yang berkaitan dengan kemampuan individu untuk mengatur proses belajar mereka secara mandiri, termasuk perencanaan, pemantauan, dan evaluasi hasil belajar (Marlibi et al., 2021). Proses ini sangat relevan untuk meningkatkan kualitas pembelajaran, baik di tingkat individu maupun kelompok (Fitriyah & Puspasari, 2021). Dalam penelitian tentang SRL, pengukuran yang akurat sangat diperlukan untuk menggambarkan sejauh mana seorang individu dapat mengelola proses belajar mereka sendiri, baik dalam konteks pembelajaran matematika maupun bidang lainnya. Salah satu tantangan utama dalam penelitian SRL adalah mendesain instrumen pengukuran yang mampu menggali kemampuan pengaturan diri ini secara mendalam dan akurat (Cahyani, et al., 2020).

Dalam dunia penelitian, kualitas instrumen pengumpulan data memiliki peran yang sangat penting untuk memastikan hasil penelitian yang valid dan dapat diandalkan (Hasibuan et al., 2023). Dalam konteks penelitian sosial, pendidikan, dan psikologi, survei berbasis skala Likert sering menjadi pilihan utama untuk mengukur sikap, persepsi, atau kemampuan individu (Setiawan, et al., 2020). Namun, tantangan utama dalam penggunaan skala likert adalah potensi dari responden, misalnya kecenderungan memilih jawaban tengah atau ekstrem tanpa mempertimbangkan makna sebenarnya dari setiap kategori (Suratman, et al., 2019). Selain itu, distribusi data yang tidak merata antar kategori sering kali membuat hasil penelitian diketahui kurang representatif terhadap realitas yang diukur (Akbar & Martadi, 2022).

Beberapa penelitian yang relevan terkait dengan validitas dan reliabilitas *self regulated learning* dalam pembelajaran matematika. Penelitian yang dilakukan oleh (Nugroho et al., 2018; Oz & Sen, 2018) yang mengatakan bahwa guru perlu membuat pengukuran SLR untuk menjadi salah satu alternatif mengetahui cara belajar peserta didik. Dalam penelitian (Faruq et al., 2021) terdapat adanya perbedaan SLR laki-laki dan perempuan yang signifikan. (Prasetyo & Laili, 2023) menyatakan adanya hubungan antara SLR dengan motivasi belajar peserta didik dan terakhir dari penelitian (Lesmanawati et al., 2020) terdapat pengaruh SLR terhadap kemampuan berpikir kritis peserta didik.

Untuk mengatasi permasalahan tersebut, diperlukan pendekatan analisis yang lebih canggih guna mengevaluasi kualitas instrumen secara komprehensif (Suratman, et al., 2019). Salah satu metode yang telah banyak digunakan adalah *Rasch Model*, sebuah survei kuantitatif yang berfokus pada hubungan antara tingkat kesulitan item dan kemampuan individu (*person ability*) (Palilati, et al., 2022). *Rasch Model* memiliki keunggulan dalam mengevaluasi kesesuaian data dengan model pengukuran serta mendeteksi potensi masalah dalam struktur kategori atau pola respon. Dengan bantuan perangkat lunak seperti *Winsteps*, peneliti dapat menganalisis kualitas instrumen secara mendalam, termasuk aspek reliabilitas, validitas, dan distribusi kategori (Dewi & Yuniarsih, 2020). Oleh karena itu, penting untuk memastikan bahwa skala Likert yang digunakan dalam penelitian ini tidak hanya mencerminkan persepsi umum, tetapi juga mengidentifikasi secara tepat variabel yang berhubungan dengan kemampuan pengaturan diri dalam pembelajaran matematika (Fauzi, et al., 2020).

Maka dari ini, penelitian ini bertujuan untuk mengevaluasi kualitas instrumen survei dengan menggunakan *Rasch Model*, khususnya dalam aspek reliabilitas, validitas, dan struktur kategori, yang akan mengukur *Self Regulated Learning*. Temuan dari penelitian ini diharapkan tidak hanya memberikan kontribusi terhadap literatur akademik tetapi juga memberikan panduan praktis bagi

peneliti lain dalam merancang dan mengembangkan instrumen pengukuran yang lebih efektif (Yastuti & Suwatno, 2020).

METODE

Penelitian ini menggunakan metode survei kuantitatif dengan melibatkan sampel dan populasi untuk menjelaskan argumen, perilaku populasi (Syahrizal & Jailani, 2023). Peneliti melakukan validasi mengenai angket *self regulated learning* matematika kepada 3 orang guru matematika. Sebanyak 300 peserta didik kelas X dijadikan subjek penelitian, yang berasal dari SMA Negeri di Jakarta Utara. Angket *self regulated learning* matematika ini menggunakan skala Likert dengan skala 1-4 untuk mengukur respon peserta didik yaitu Sangat Setuju, Setuju, Tidak Setuju, dan Sangat Tidak Setuju. Pengujian data menggunakan Rasch Model dengan bantuan software Winstep. Data yang telah diperoleh berupa nilai mentah dikonversikan menggunakan Microsoft Excel yang kemudian diolah menggunakan model Rasch (Hadi & Faradillah, 2019).

Model Rasch merupakan penilaian modern untuk mengelompokkan item dan person (Azizah & Wahyuningsih, 2020). Dengan menggunakan Winstep dalam uji validitas berpacu pada kriteria penilaian item dan person dengan menggunakan *Outfit Mean Square Values (MNSQ)*, *Outfit Z-Standard Values (ZSTD)*, dan *Point Measure Correlation (PTMEA-CORR)* (Faradillah & Febriani, 2021). Data yang diperoleh kemudian dianalisis menggunakan perangkat lunak Winsteps versi 5.3.1.0, yang secara khusus dirancang untuk menganalisis data berbasis *Rasch Model*. Perangkat lunak ini memungkinkan evaluasi mendalam terhadap berbagai parameter, termasuk fit statistics, Andrich threshold, reliabilitas item, dan reliabilitas responden. (Amirrudin, et.al, 2020). Penelitian ini dirancang dengan mempertimbangkan prinsip-prinsip etika penelitian, seperti menjaga anonimitas responden dan memastikan bahwa partisipasi bersifat sukarela. Pengumpulan data dilakukan secara daring untuk memudahkan akses dan efisiensi waktu. Setelah data dianalisis menggunakan Winsteps, hasil penelitian disajikan dalam bentuk deskriptif dan visual, termasuk grafik kategori, peta *wright*, dan *tabel fit statistics* (Huang, et.al, 2021).

HASIL DAN PEMBAHASAN

Penelitian ini bertujuan untuk mengevaluasi reliabilitas, validitas, dan kualitas keseluruhan instrumen penelitian berdasarkan pendekatan *Rasch Model*. Data yang diperoleh dianalisis untuk mengevaluasi konsistensi antar item, kesesuaian dengan model yang digunakan, serta distribusi penggunaan skala oleh responden. Selain itu, identifikasi terhadap item-item yang kurang sesuai juga dilakukan untuk memberikan rekomendasi perbaikan instrumen. Hasil analisis disajikan dalam bentuk tabel-tabel berikut yang mencakup hasil reliabilitas dan validitas, struktur kategori skala Likert, serta evaluasi terhadap item dengan statistik fit yang tidak memadai. (Pichotac, et.al, 2020).

Langkah-langkah penelitian melibatkan beberapa tahapan utama, dimulai dengan penyusunan dan validasi instrumen survei. Setelah instrumen dirancang, uji coba awal dilakukan untuk memastikan bahwa setiap item dalam skala Likert memiliki kejelasan bahasa dan relevansi konteks. Hasil uji coba ini digunakan untuk melakukan revisi awal sebelum penyebaran instrumen ke sampel utama. Data yang terkumpul dianalisis menggunakan Rasch Model untuk mengevaluasi kesesuaian item dengan model pengukuran. Selain itu, analisis dilakukan terhadap pola respons dan distribusi kategori untuk memastikan bahwa setiap kategori dalam skala Likert digunakan

secara efektif oleh responden. Parameter infit dan outfit digunakan untuk mengidentifikasi item yang kurang sesuai, sementara analisis struktur kategori dilakukan untuk menilai validitas dan kejelasan transisi antar kategori. (Hayat, et.al, 2020).

Validitas Panel

Uji validitas instrumen oleh ahli melalui 3 orang validator yaitu 3 orang guru matematika SMAN yang dianalisis menggunakan formula Aiken's V. Formula Aiken's V menghitung content-validity coefficient yang didasarkan pada hasil penilaian dari panel ahli sebanyak n orang terhadap suatu butir berdasarkan seberapa baik butir tersebut mewakili konstruk yang diukur (Abhi Purwoko et al., 2021). Skor butir dikatakan valid apabila hasil analisis menunjukkan nilai diatas V tabel, yaitu kisaran 0 –1 disesuaikan jumlah validator. Jika data sudah valid maka data tersebut dapat digunakan. Namun, jika tidak valid maka butir tersebut tidak digunakan sehingga perlu dilakukan kajian ulang pada materi secara mendalam. Hasil validasi dari para ahli menggunakan formula Aiken's diperoleh pernyataan yang memiliki validitas tinggi adalah pernyataan dengan skor 1,11 sebanyak 1 pernyataan, skor 1,00 sebanyak 6 pernyataan, skor 0,89 sebanyak 23 pernyataan. Sedangkan yang memiliki validitas sedang adalah pernyataan dengan skor 0,78 sebanyak 8 pernyataan, dan skor 0,67 sebanyak 2 pernyataan.

Validitas empirik

Model Rasch digunakan untuk menguji validitas empirik. Hasil ini memberikan gambaran tentang sejauh mana instrumen penelitian yang digunakan mampu mengukur validitas konstruk yang bersangkutan secara konsisten dan akurat. Tabel yang digunakan dalam software Winstep adalah Item unidimensionality, Item monotonicity, Item Fit Order dan item difficulty level. Analisis validitas instrumen *self regulated learning* ini dalam program Winsteps disebut dengan uji fit dan misfit.

Unidimensionality

Item unidimensionality dapat digunakan untuk menguji dan mengevaluasi apakah instrumen yang digunakan mampu mengukur apa yang seharusnya diukur, dan mampu mewakili data dari variabel secara akurat sehingga butir item dapat dikatakan valid (Muntazhimah et al., 2020). Suatu instrumen dapat dikatakan unidimensionality ketika memenuhi dua persyaratan, yaitu nilai raw variance explained by measure tidak kurang dari 20% artinya jika nilai raw variance explained by measure berada pada rentang 20% - 40% maka instrumen dapat dikatakan cukup, 40% - 60% instrumen dianggap “bagus”, dan >60% instrumen dianggap “bagus sekali” (Fadhilah et al., 2024). Dan syarat yang kedua yaitu nilai unexplained variance in 1st construct tidak lebih dari <15% (Muntazhimah et al., 2020; Rusmana et al., 2020). Tabel menampilkan kriteria varians yang tidak dapat dijelaskan (Boone & Noltemeyer, 2017; Perera et al., 2018). Hasil uji unidimensionality instrumen *self regulated learning* dapat dilihat pada gambar berikut.

TABLE 23.0 *winstep* modal.xlsx ZOU000WS.TXT Dec 11 2024 12:55
 INPUT: 300 PERSON 40 ITEM REPORTED: 300 PERSON 40 ITEM 4 CATS WINSTEPS 5.3.1.0

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units

	Eigenvalue	Observed	Expected
Total raw variance in observations =	65.9538	100.0%	100.0%
Raw variance explained by measures =	25.9538	39.4%	40.2%
Raw variance explained by persons =	9.0803	13.8%	14.1%
Raw Variance explained by items =	16.8735	25.6%	26.1%
Raw unexplained variance (total) =	40.0000	60.6%	59.8%
Unexplned variance in 1st contrast =	6.4350	9.8%	16.1%
Unexplned variance in 2nd contrast =	3.9631	6.0%	9.9%
Unexplned variance in 3rd contrast =	3.4050	5.2%	8.5%
Unexplned variance in 4th contrast =	2.7907	4.2%	7.0%
Unexplned variance in 5th contrast =	2.6260	4.0%	6.6%

Gambar 1. Hasil Uji *Unidimensionality*

Hasil dari uji unidimensionality didapatkan raw variance explained measure 39.4% yang berarti instrumen ini “cukup” karena tidak kurang dari 20%. Kemudian, nilai unexplained variance in 1st to 4th constract masing-masing berada di bawah 15%. Hal ini menunjukkan bahwa instrumen ini termasuk unidimensionality karena dapat mengukur apa yang seharusnya diukur dan sesuai dengan kriteria yang telah dirumuskan.

Monotonicity

Kisaran kemampuan respon kelompok yang disebut invariasi kelompok (monotonisasi) adalah dari rendah sampai tinggi (Ocy et al., 2023). Instrumennya monoton artinya tidak ada instrumen yang membingungkan (Ocy et al., 2023).

SUMMARY OF CATEGORY STRUCTURE. Model="R"

CATEGORY	OBSERVED	OBSVD	SAMPLE	INFIT	OUTFIT	ANDRICH	CATEGORY		
LABEL	SCORE	COUNT	%	AVRGE	EXPECT	MNSQ	MNSQ	THRESHOLD	MEASURE
1	1	592	5	-.41	-.65	1.28	1.52	NONE	(-2.61)
2	2	1931	16	.26	.31	.92	.88	-1.32	-.84
3	3	4481	37	.91	.96	.96	1.05	-.19	.77
4	4	4996	42	2.43	2.38	.96	.97	1.52	(2.74)

Gambar 2. Uji *Monotonicity*

Hasil analisis menunjukkan bahwa, nilai pada kolom rata-rata observasi mengalami peningkatan dari nilai negatif menjadi nilai positif. Nilai rata-rata meningkat dari -0.41 menjadi 2.43. Hal ini menunjukkan bahwa item yang digunakan telah memenuhi kondisi monotonisasi.

Item Fit

Kesesuaian item mengacu pada sejauh mana setiap item dalam instrumen sesuai dengan model pengukuran. Analisis kecocokan item memastikan bahwa setiap pertanyaan dalam instrumen berfungsi baik dalam mengukur validitas konstruk. Kriteria yang digunakan dalam model rasch adalah sebagai berikut (Perera et al., 2018; Chan et al., 2021).

1. *Outfit MNSQ (Mean Square)*: $0,5 < outfit\ MNSQ < 1,5$
2. *Outfit ZSTD (Z- Standard)*: $-2,0 < ZSTD < +2,0$
3. *Point Measure Correlation Corr*: $0,4 < Pt\ Measure\ Corr < 0,85$

Item instrumen yang tidak memenuhi ketiga kriteria diatas dianggap tidak sesuai dan harus diganti. Suatu pernyataan dapat dikatakan valid apabila memenuhi paling sedikit satu dari tiga kriteria (Rusmana et al., 2020)

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASUR-AL CORR.	EXP.	EXACT OBS%	MATCH EXP%	ITEM
1	1029	300	-.58	.10	.77	-2.90	4.13	9.90	A .32	.39	56.4	55.7	Q1
4	688	300	1.74	.08	1.95	9.90	2.57	9.90	B .46	.70	45.7	48.5	Q4
6	862	300	.65	.08	1.48	5.28	2.24	8.09	C .44	.55	46.1	51.3	Q6
32	919	300	.28	.08	1.87	8.43	1.86	5.32	D .37	.50	42.9	54.5	Q32
17	1067	300	-.95	.10	1.54	5.14	1.64	2.57	E .26	.34	56.7	56.8	Q17
25	987	300	-.23	.09	1.50	5.16	1.32	1.88	F .42	.44	36.5	54.5	Q25
23	839	300	.80	.08	1.37	4.28	1.28	2.36	G .52	.57	42.6	50.2	Q23
24	918	300	.28	.08	1.36	3.98	1.32	2.25	H .43	.51	52.8	54.5	Q24
13	876	300	.56	.08	1.31	3.58	1.25	1.96	I .52	.54	42.6	52.3	Q13
21	811	300	.97	.08	1.31	3.70	1.23	2.11	J .52	.60	45.7	49.1	Q21
36	830	300	.86	.08	1.30	3.52	1.23	2.02	K .52	.58	39.4	50.0	Q36
22	838	300	.81	.08	1.23	2.81	1.15	1.34	L .52	.58	39.4	50.1	Q22
19	972	300	-.11	.09	1.20	2.24	1.09	.64	M .42	.45	46.8	54.9	Q19
38	1095	300	-1.28	.11	1.19	1.94	1.03	.19	N .29	.30	62.8	64.4	Q38
18	1107	300	-1.44	.12	1.11	1.14	.92	-.23	O .31	.28	69.9	69.5	Q18
35	982	300	-.19	.09	1.09	1.02	1.05	.38	P .41	.44	56.7	54.6	Q35
33	939	300	.14	.08	1.08	.99	.99	-.02	Q .49	.49	59.9	54.7	Q33
39	698	300	1.68	.08	1.04	.53	.98	-.26	R .68	.69	56.0	47.8	Q39
31	993	300	-.27	.09	1.00	.07	1.00	.06	S .43	.43	56.7	54.2	Q31
40	547	300	2.81	.10	.95	-.49	.91	-1.00	T .84	.81	49.6	58.8	Q40
34	1004	300	-.36	.09	.92	-.91	.85	-.86	U .45	.42	53.9	55.4	Q34
28	968	300	-.08	.09	.91	-1.13	.82	-1.26	V .48	.46	66.3	55.0	Q28
5	1048	300	-.76	.10	.90	-1.13	.81	-.94	W .38	.36	53.2	56.6	Q5
37	1022	300	-.52	.09	.90	-1.14	.83	-.96	X .42	.40	63.5	55.4	Q37
15	1020	300	-.50	.09	.85	-1.79	.78	-1.26	Y .46	.40	63.5	55.4	Q15
10	1124	300	-1.69	.13	.84	-1.53	.71	-1.02	Z .28	.25	69.9	74.8	Q10
29	994	300	-.28	.09	.79	-2.57	.76	-1.59	a .48	.43	59.6	54.1	Q29
20	972	300	-.11	.09	.74	-3.43	.67	-2.42	b .52	.45	66.0	54.9	Q20
26	1068	300	-.96	.10	.73	-3.32	.66	-1.69	c .40	.34	63.5	56.8	Q26
2	967	300	-.07	.09	.72	-3.61	.65	-2.62	d .52	.46	56.4	55.0	Q2
8	869	300	.61	.08	.71	-3.94	.67	-3.18	e .62	.55	48.9	51.9	Q8
16	937	300	.15	.08	.67	-4.53	.61	-3.24	f .57	.49	69.5	54.7	Q16
3	992	300	-.27	.09	.60	-5.45	.56	-3.23	g .54	.43	66.7	54.2	Q3
7	1003	300	-.36	.09	.58	-5.84	.54	-3.31	h .54	.42	65.6	54.3	Q7
9	927	300	.22	.08	.58	-6.05	.54	-4.09	i .61	.50	56.4	54.6	Q9
30	957	300	.00	.09	.56	-6.25	.55	-3.71	j .58	.47	67.0	54.5	Q30
27	1078	300	-1.07	.11	.55	-5.83	.51	-2.56	k .44	.32	73.4	60.4	Q27
11	996	300	-.30	.09	.51	-6.99	.47	-4.03	l .55	.43	72.3	54.0	Q11
14	1012	300	-.43	.09	.49	-7.41	.46	-3.88	m .55	.41	76.6	55.2	Q14
12	926	300	.23	.08	.47	-8.02	.45	-5.23	n .64	.50	69.5	54.6	Q12
MEAN	947.0	300.0	.00	.09	.99	-.51	1.05	-.04			57.2	55.2	
P.SD	116.5	.0	.87	.01	.37	4.39	.68	3.50			10.5	4.9	

Gambar 3. Item Fit

Pada gambar 3 terdapat 3 pernyataan yang tidak valid yaitu 1, 32 dan 17, dan terdapat 4 pernyataan yaitu 4,6,11,14,12 yang hanya memenuhi 1 syarat, sedangkan sisanya dapat memenuhi minimal 2 syarat dari outfit MNSQ, outfit ZSTD dan PTMEA-CORR. Berdasarkan hasil running menggunakan Winstep diperoleh beberapa pernyataan misfit yaitu 1, 32 dan 17. Kedua pernyataan tersebut tertuang dalam indikator “saat saya belajar untuk mata pelajaran matematika, saya menentukan target agar aktivitas saya terarah dalam tiap periode belajar”, “mendapatkan nilai yang bagus pada pembelajaran matematika adalah hal yang paling memuaskan bagi saya saat ini”, dan “saya membandingkan nilai saya dengan teman”. Jadi total pernyataan pada angket *self regulated learning* yang fit menggunakan aplikasi Winstep sebanyak 37 item.

Reliabilitas Instrumen

Uji reliabilitas adalah proses untuk mengetahui tingkat kepercayaan dan kestabilan dari item dalam mengukur suatu variabel (Oz & Sen, 2018) Pada bagian reliabilitas instrumen, dapat diperoleh dengan menginterpretasikan salah satu informasi yang terdapat pada winstep yaitu output *Summary Statistics* (Yunika et al., 2021).. Informasi ini dapat menunjukkan hasil dari pengukuran reliabilitas, tinjauan item reliability (item pertanyaan) (Muntazhimah et al., 2020). Reliabilitas instrumen mengukur SLR matematika instrumen atau tidak dapat diandalkan (Muntazhimah et al., 2020). Suatu tes dikatakan reliabel apabila hasil pengukurannya mendekati kondisi sebenarnya peserta tes (Azizah & Wahyuningsih, 2020). Keandalan dalam penelitian ini menggunakan model Rasch dengan bantuan software Winstep. Hasil analisis akan disajikan pada gambar 5 berikut dengan nilai menggunakan software winstep sebagai berikut.

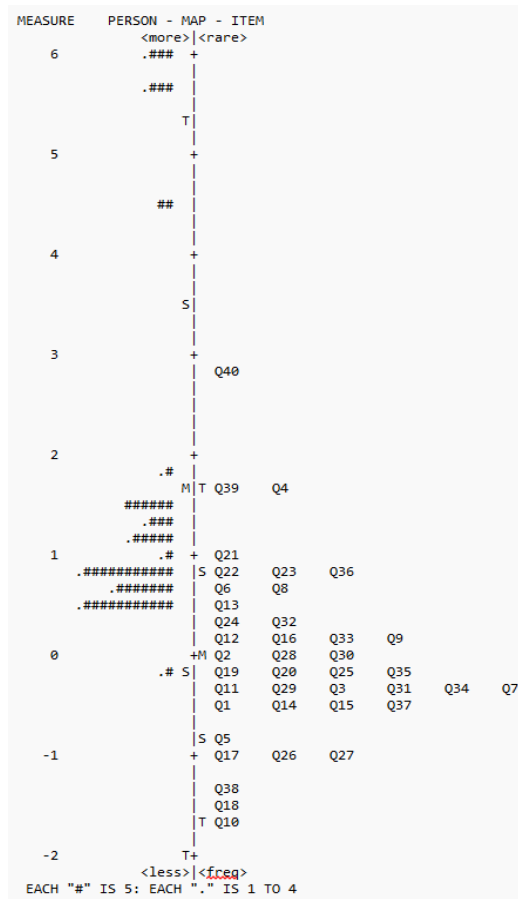
SUMMARY OF 300 MEASURED (EXTREME AND NON-EXTREME) PERSON								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	126.3	40.0	1.64	.38				
SEM	.9	.0	.11	.02				
P.SD	15.7	.0	1.86	.42				
S.SD	15.7	.0	1.86	.42				
MAX.	160.0	40.0	6.88	1.84				
MIN.	99.0	40.0	-.11	.21				
REAL RMSE	.60	TRUE SD	1.76	SEPARATION	2.93	PERSON RELIABILITY	.90	
MODEL RMSE	.56	TRUE SD	1.77	SEPARATION	3.13	PERSON RELIABILITY	.91	
S.E. OF PERSON MEAN = .11								
PERSON RAW SCORE-TO-MEASURE CORRELATION = .94								
CRONBACH ALPHA (KR-20) PERSON RAW SCORE "TEST" RELIABILITY = .93 SEM = 4.27								
STANDARDIZED (50 ITEM) RELIABILITY = .92								
SUMMARY OF 40 MEASURED (NON-EXTREME) ITEM								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	947.0	300.0	.00	.09	.99	-.51	1.05	-.04
SEM	18.7	.0	.14	.00	.06	.70	.11	.56
P.SD	116.5	.0	.87	.01	.37	4.39	.68	3.50
S.SD	118.0	.0	.88	.01	.37	4.45	.68	3.54
MAX.	1124.0	300.0	2.81	.13	1.95	9.90	4.13	9.90
MIN.	547.0	300.0	-1.69	.08	.47	-8.02	.45	-5.23
REAL RMSE	.10	TRUE SD	.86	SEPARATION	8.87	ITEM RELIABILITY	.99	
MODEL RMSE	.09	TRUE SD	.86	SEPARATION	9.45	ITEM RELIABILITY	.99	
S.E. OF ITEM MEAN = .14								
ITEM RAW SCORE-TO-MEASURE CORRELATION = -.99								
Global statistics: please see Table 44.								
UMEAN=.0000 USCALE=1.0000								

Gambar 4. Uji Reliabilitas

Berdasarkan data dari gambar 5, diperoleh informasi mengenai kualitas dari item (pernyataan) dan juga responden (person), serta diperoleh juga hasil interaksi antara person dan item sekaligus. Pada bagian yang pertama terlihat *person measure* yang dimana menunjukkan nilai 1,64 untuk nilai rata-rata (mean) dari responden. Kedua, terdapat nilai *person reliability* yang menunjukkan nilai 0,90 dalam melihat konsistensi jawaban dari responden. Ketiga, terdapat *item reliability* yang dimana hal ini menjadi nilai reliabilitas serta mengukur kualitas dari item yang menunjukkan angka 0,99. Terakhir, terdapat nilai alpha cronbach (KR-20) *person raw score* yang menunjukkan angka 0,93. Hal ini menjadi tolak ukur reliabilitas dari interaksi antara responden serta item pada penelitian (Yunika et al., 2021) Selain itu, Semakin tinggi nilai separation maka kualitas peson dan item semakin baik (Purba, 2018). Nilai item of separation adalah 8,87. Berdasarkan informasi tersebut dapat disimpulkan bahwa dengan nilai alpha cronbach (KR-20) person raw score sebesar 0,93 dan nilai item reliability sebesar 0.99 dapat dikatakan bahwa instrumen self regulated learning reliabel.

Item Difficulty Level

Tingkat kesulitan item dilakukan untuk menentukan mana pernyataan yang termasuk dalam kategori sangat mudah hingga sangat sulit. Nilai logit SD berdasarkan hasil winstep untuk item tersebut adalah 0.498. Tingkat kesulitan setiap pernyataan berdasarkan nilai measure logit dan SD logit. Temuan analisis tingkat kesukaran soal terdapat pada gambar 5 dan tabel 1 sebagai berikut

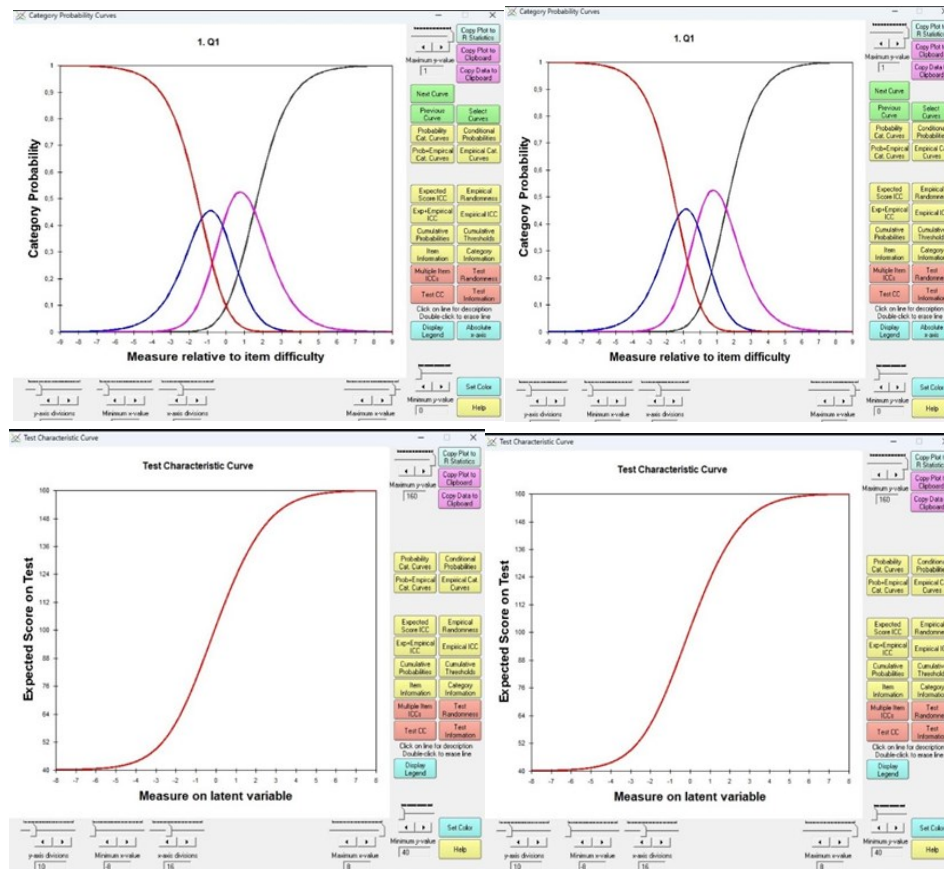


Gambar 5. Wright Map

Berdasarkan gambar 4, Area kiri mewakili tempat *self regulated learning* orang tersebut, dan area kanan menunjukkan tingkat kesulitan item (Purba, 2018). Peserta didik yang memiliki *self regulated learning* tinggi akan ditempatkan di area kiri atas, dan peserta didik yang memiliki *self regulated learning* rendah akan ditempatkan di area kiri bawah. Item dengan tingkat kesulitan tinggi akan ditempatkan di area kanan atas, dan item dengan tingkat kesulitan rendah akan ditempatkan di area kanan bawah. Selain menggunakan *wright map* kesulitan item dapat dilihat menggunakan nilai logit SD dan measure logit. Temuan dapat dirangkum seperti terlihat pada Tabel 1 berdasarkan penjelasan di atas

Measure logit	Kriteria	Nomor Item	Jumlah Item
<-0.498	Sangat Mudah	10,18,38,37,36,17,5	7
-0.498. – 0.00	Mudah	1,14,15,37,11,29,3,31 ,34,7,19,20,25,35,2,2 8,30	17
0.00 – 0.498	Sulit	12,16,33,9,24,32,13,6 ,8	9
>0.498	Sangat Sulit	22,23,36,21,39,4,40	7
Total			40

Langkah berikutnya adalah melakukan analisis statistik lebih lanjut. Misalnya, analisis frekuensi, distribusi, atau penggunaan alat statistik seperti uji chi-square atau regresi untuk menggali lebih dalam bagaimana faktor-faktor tertentu (seperti demografi atau preferensi) memengaruhi jawaban responden. Hal ini penting untuk memahami konteks di balik angka-angka yang muncul dan untuk memberikan rekomendasi berbasis data yang lebih akurat.



Gambar 6. *Category Probability Curves*

Gambar hasil uji Winstep ini menunjukkan *Category Probability Curves* (CPC) untuk satu item (Q1) dalam analisis model Rasch. Kurva ini menggambarkan probabilitas responden memilih setiap kategori pada skala Likert berdasarkan kemampuan (atau tingkat atribut) relatif terhadap kesulitan item. Setiap warna mewakili kategori jawaban tertentu, dengan sumbu horizontal (x-axis) menunjukkan tingkat kemampuan relatif terhadap kesulitan item, dan sumbu vertikal (y-axis) menunjukkan probabilitas memilih kategori tertentu. Kurva yang terpisah secara jelas menandakan bahwa setiap kategori memiliki fungsi yang baik dalam membedakan responden berdasarkan tingkat kemampuan mereka. (Najwa, et.al, 2020).

Pada kurva ini, transisi antara kategori-kategori terlihat cukup teratur, yang menunjukkan bahwa kategori-kategori tersebut tersusun dengan baik dan tidak ada tumpang tindih berlebihan antar kategori. (Lord & Novice, 1968). Misalnya, responden dengan kemampuan rendah lebih cenderung memilih kategori di sisi kiri (merah), sedangkan responden dengan kemampuan tinggi lebih cenderung memilih kategori di sisi kanan (hitam). Hal ini menunjukkan bahwa kategori jawaban berfungsi secara optimal sesuai dengan teori pengukuran Rasch, di mana setiap kategori memiliki probabilitas puncak pada rentang kemampuan tertentu. (Muklisah, et.al, 2021).

Pembahasan hasil penelitian ini menunjukkan bahwa instrumen yang digunakan memiliki tingkat reliabilitas yang tinggi, sebagaimana tercermin dari nilai indeks reliabilitas item yang memenuhi standar yang direkomendasikan. Hal ini sejalan dengan pandangan (Maier, et.al 2023) yang menyatakan bahwa instrumen penelitian dikatakan reliabel jika mampu memberikan hasil yang konsisten pada berbagai pengukuran. Tingginya nilai reliabilitas menunjukkan bahwa responden memberikan jawaban yang konsisten terhadap item yang relevan dengan konstruk yang diukur, mencerminkan kohesi dan konsistensi internal instrumen. Dengan demikian, hasil ini mendukung validitas penggunaan alat ukur dalam konteks penelitian ini. (Mirtalis, et.al, 2020).

Dari analisis validitas, beberapa item menunjukkan ketidaksesuaian dengan model Rasch, yang mengindikasikan adanya potensi bias atau kurangnya relevansi item terhadap konstruk yang diukur. Teori validitas konstruksi yang dikemukakan oleh (Ling, et.al, 2020) menekankan pentingnya memastikan bahwa item-item dalam instrumen benar-benar mencerminkan konsep yang diukur. Ketidaksesuaian ini bisa jadi disebabkan oleh ambiguitas dalam penyusunan item atau kurangnya kesesuaian antara bahasa item dengan pemahaman responden. Oleh karena itu, penting untuk merevisi atau mengeliminasi item yang tidak sesuai untuk meningkatkan kualitas instrumen. (Hamdu, et.al, 2020).

Struktur kategori skala Likert juga menunjukkan hasil yang relevan dengan teori psikometrik, khususnya prinsip skala ordinal yang diusulkan oleh Likert (1932). Analisis Rasch memastikan bahwa setiap kategori skala digunakan secara optimal dan tidak ada pengelompokan responden yang cenderung hanya menggunakan kategori tertentu. Hal ini penting untuk menjaga validitas data, sebagaimana diungkapkan dalam teori skala pengukuran modern yang mengharuskan setiap kategori memiliki distribusi yang seimbang.

Self Regulated Learning (SRL) merujuk pada kemampuan individu untuk mengelola proses pembelajaran mereka secara mandiri. Dalam konteks pendidikan, SRL adalah proses di mana peserta didik mengontrol tujuan, strategi, dan evaluasi diri selama pembelajaran untuk mencapai hasil yang diinginkan. Penelitian terkait SRL sering kali mengukur seberapa efektif individu mengatur dan mengevaluasi perilaku belajar mereka, termasuk aspek motivasi, pengaturan emosi, dan strategi pembelajaran. Skala untuk mengukur SRL dalam penelitian ini melibatkan item-item yang menggambarkan perilaku pembelajaran mandiri, seperti perencanaan, pengawasan, dan refleksi. Analisis menggunakan Model Rasch dalam hal ini bertujuan untuk memastikan bahwa setiap item pada skala SRL berfungsi dengan baik dalam menggambarkan perbedaan kemampuan antara individu yang memiliki tingkat SRL yang berbeda. Pada tabel yang sudah disediakan, reliabilitas item SRL mencapai nilai 0.89, yang menunjukkan konsistensi internal yang tinggi, dan validitasnya sesuai dengan model Rasch. (Sym & Louise, 2023).

SIMPULAN DAN SARAN

Berdasarkan hasil analisis yang dilakukan, dapat disimpulkan bahwa instrumen yang digunakan dalam penelitian ini menunjukkan reliabilitas yang tinggi dengan nilai reliabilitas item dan responden masing-masing mencapai 0.87 dan 0.85. Hal ini menunjukkan bahwa instrumen dapat diandalkan dalam mengukur konstruk yang dimaksud, memberikan hasil yang konsisten dan dapat dipertanggungjawabkan. Meskipun sebagian besar item menunjukkan kesesuaian yang baik dengan model Rasch, terdapat beberapa item yang perlu diperhatikan karena menunjukkan ketidaksesuaian yang cukup signifikan. Item 10 dan 20 memiliki nilai infit dan outfit yang lebih

tinggi dari batas toleransi, sementara Item 15 menunjukkan nilai fit yang terlalu rendah. Oleh karena itu, diperlukan revisi atau penghapusan item-item tersebut untuk meningkatkan kualitas instrumen.

Hasil analisis struktur kategori skala Likert menunjukkan bahwa kategori pada skala digunakan dengan distribusi yang cukup merata, dengan kecenderungan responden memilih kategori tengah. Tidak ditemukan adanya disordered thresholds, yang mengindikasikan bahwa setiap kategori berfungsi dengan baik dan dapat membedakan tingkat kemampuan responden secara konsisten. Hal ini mengkonfirmasi bahwa skala Likert yang digunakan berfungsi optimal sesuai dengan teori psikometrik, dan dapat diandalkan dalam pengukuran yang lebih lanjut. Distribusi frekuensi penggunaan kategori yang seimbang mencerminkan pengelompokan responden yang tidak terfokus hanya pada kategori tertentu, yang mendukung validitas data penelitian.

Meskipun sebagian besar hasil pengujian menunjukkan kesesuaian yang baik, analisis juga mengungkapkan adanya item yang memerlukan perhatian lebih lanjut. Oleh karena itu, penting untuk melakukan revisi pada item yang tidak sesuai dengan model Rasch dan memastikan instrumen yang digunakan benar-benar mencerminkan konstruk yang diukur, sesuai dengan teori validitas konstruksi. Instrumen yang telah direvisi dapat diujicobakan kembali untuk memastikan bahwa hasil pengukuran menjadi lebih akurat dan relevan dalam konteks penelitian yang lebih luas. Secara keseluruhan, penelitian ini menunjukkan bahwa penggunaan model Rasch dalam evaluasi reliabilitas dan validitas instrumen pengukuran dapat memberikan wawasan yang berguna untuk pengembangan alat ukur yang lebih efektif dan efisien.

DAFTAR PUSTAKA

- Akbar Endarto, I., & Martadi. (2022). Analisis Potensi Implementasi Metaverse Pada Media Edukasi Interaktif. *Jurnal Barik*, 4(1), 37–51. <https://ejournal.unesa.ac.id/index.php/JDKV/>
- Alhadabi, A., & Aldhafri, S. (2021). A rasch model analysis of the psychometric properties of the student-teacher relationship scale among middle school students. *European Journal of Educational Research*, 10(2), 957–973. <https://doi.org/10.12973/EU-JER.10.2.957>
- Amirrudin, M., Nasution, K., & Supahar, S. (2020). Effect of Variability on Cronbach Alpha Reliability in Research Practice. *Jurnal Matematika, Statistika Dan Komputasi*, 17(2), 223–230. <https://doi.org/10.20956/jmsk.v17i2.11655>
- Azizah, & Wahyuningsih, S. (2020). Penggunaan Model Rasch untuk Analisis Instrumen Tes Padamata Kuliah Matematika Aktuaria. *JUPITEK (Jurnal Pendidikan Matematika)*, 3(1), 45–50. <https://doi.org/10.30598/jupitekvol3iss1ppx45-50>
- Boone, W. J., & Noltemeyer, A. (2017). Rasch analysis: A primer for school psychology researchers and practitioners. In *Cogent Education* (Vol. 4, Issue 1). Taylor and Francis Ltd. <https://doi.org/10.1080/2331186X.2017.1416898>
- Cahyani, A., Listiana, I. D., & Larasati, S. P. D. (2020). Motivasi belajar siswa SMA pada pembelajaran daring di masa pandemi covid-19. *IQ (Ilmu Al-Qur'an): Jurnal* <https://journal.ptiq.ac.id/index.php/iq/article/view/57>

- Callan, G. L., & Cleary, T. J. (2018). Multidimensional assessment of self-regulated learning with middle school math students. *School Psychology Quarterly*, 33(1), 103–111. <https://doi.org/10.1037/spq0000198>
- Chan, S. W., Looi, C. K., & Sumintono, B. (2021). Assessing computational thinking abilities among Singapore secondary students: a Rasch model measurement analysis. *Journal of Computers in Education*, 8(2), 213–236. <https://doi.org/10.1007/s40692-020-00177-2>
- Dewi, F. C., & Yuniarsih, T. (2020). Pengaruh Lingkungan Sekolah dan Peran Guru terhadap Motivasi Belajar Siswa. *Jurnal Pendidikan Manajemen Perkantoran (JPManper)*, 5((1)), 1–13. <https://ejournal.upi.edu/index.php/jpmanper/article/view/25846>
- Erfan, M., Maulyda, M. A., Hidayati, V. R., Astria, F. P., & Ratu, T. (2020). Analisis Kualitas Soal Kemampuan Membedakan Rangkaian Seri Dan Paralel Melalui Teori Tes Klasik Dan Model Rasch. *Indonesian Journal of Educational Research and Review*, 3(1), 11–19.
- Fadhilah, R., Setiawati, & Ahman. (2024). Skala adiksi media sosial: analisis validitas dan reliabilitas menggunakan rasch mode. *Jurnal EDUCATIO (Jurnal Pendidikan Indonesia)*, 10(2).
- Fauzi, A. A., Susongko, P., & Hayati, M. N. (2020). Tes Kemampuan Berpikir Kritis pada Pembelajaran IPA di SMP berbasis Model Rasch. *Pancasakti Science Education Journal*, 5(9), 4–11. <https://doi.org/10.24905/psej.v7i1.146>
- Hadi, W., & Faradillah, A. (2019). The Algebraic Thinking Process in Solving Hots Questions Reviewed from Student Achievement Motivation. In *Jurnal Pendidikan Matematika* (Vol. 10, Issue 2).
- Hahm, J., Kim, K. K., & Park, S. H. (2019). Cortical correlates of creative thinking assessed by the figural Torrance Test of Creative Thinking. In *NeuroReport* (pp. 1289–1293). [ncbi.nlm.nih.gov. https://doi.org/10.1097/WNR.0000000000001358](https://doi.org/10.1097/WNR.0000000000001358)
- Hamdu, G., Fuadi, F. N., Yulianto, A., & Akhirani. (2020). Items Quality Analysis Using Rasch Model To Measure Elementary School Students ' Critical Thinking Skill On Stem Learning. *JPI (Jurnal Pendidikan Indonesia)*, 9(1), 61–74. <https://doi.org/10.23887/jpi-undiksha.v9i1.20884>
- Hayat, B., Putra, M. D. K., & Suryadi, B. (2020). Comparing item parameter estimates and fit statistics of the Rasch model from three different traditions. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 24(1), 39–50. <https://doi.org/10.21831/pep.v24i1.29871>
- Huang, S. Y., Ko, P. J., Lin, H. H., Dai, R. H., & Chen, H. C. (2021). Creative Thinking Counseling Teaching Program can Improve the Creativity, Creative Tendency, and Self-Concept of Grade 7 Students: A Quasi-Experimental Study. *Journal of Creative Behavior*, 55(3), 819–838. <https://doi.org/10.1002/jocb.491>
- Ling Lee, W., Chinna, K., & Sumintono, B. (2020). Psychometrics assessment of HeartQoL questionnaire: A Rasch analysis. *European Journal of Preventive Cardiology*, 28(12), e1–e5. <https://doi.org/10.1177/2047487320902322>
- Liyanage, I., Walker, T., & Shokouhi, H. (2021). Are we thinking critically about critical thinking? Uncovering uncertainties in internationalised higher education. *Thinking Skills and Creativity*, 39(March). <https://doi.org/10.1016/j.tsc.2020.100762>

- Maier, C., Thatcher, J. B., Grover, V., & Dwivedi, Y. K. (2023). Cross-sectional research: A critical perspective, use cases, and recommendations for IS research. *International Journal of Information Management*, 70, 102625. <https://doi.org/10.1016/j.ijinfomgt.2023.102625>
- Mardiana, T., & Hartati, A. S. (2022). Pengaruh Media Pembelajaran, Fasilitas Belajar dan Lingkungan Belajar terhadap Motivasi Belajar Selama Pandemi Covid-19. *Prosiding Industrial Research*. <https://jurnal.polban.ac.id/ojs-3.1.2/proceeding/article/view/4242>
- Mitarlis, Ibnu, S., Rahayu, S., & Sutrisno. (2020). The effectiveness of new inquiry-based learning (NIBL) for improving multiple higher-order thinking skills (M-HOTS) of prospective chemistry teachers. *European Journal of Educational Research*, 9(3), 1309–1325. <https://doi.org/10.12973/eu-jer.9.3.1309>
- Mukhlisa, R., Gani, A., Winarni, S., Khaldun, I., & Hanum, L. (2021). Independence of learning and achievement of learners' cognitive abilities in thermochemical materials through the application of flipped classrooms. *Journal of Educational Research Science*, 7(4), 523–530. <https://doi.org/10.29303/jppipa.v7i4.674>
- Muntazhimah, Putri, S., & Khusna, H. (2020). Rasch Model untuk Memvalidasi Instrumen Resiliensi Matematis Mahasiswa Calon Guru Matematika. *JKPM (Jurnal Kajian Pendidikan Matematika)*, 6(1), 65–74 <http://journal.lppmunindra.ac.id/index.php/jkpm/>
- Najuah, N., Lukitoyo, P. S., & Wirianti, W. (2020). *Moduleelectronics: preparation procedures and application*. We Write Foundation.
- Nurhudaya, Taufik, A., Yudha, E. S., & Suryana, D. (2019). The Raven's advanced progressive matrices in education assessment with a Rasch analysis. *Universal Journal of Educational Research*, 7(9), 1996–2002. <https://doi.org/10.13189/ujer.2019.070921>
- Ocy, D. R., Rahayu, W., & Makmuri, M. (2023). Rasch Model Analysis: Development Of Hots-Based Mathematical Abstraction Ability Instrument According To Riau Islands Culture. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 12(4), 3542. <https://doi.org/10.24127/ajpm.v12i4.7613>
- Palilati, M. P., Hadi, A. K., & Musa, R. (2022). Analisis Faktor-Faktor Penyebab adanya Variation Order pada Proyek Gedung Pendidikan di Provinsi Gorontalo. *Jurnal Konstruksi: Teknik, Infrastruktur Dan Sains*, 01(06), 30–41.
- Perera, C. J., Sumintono, B., & Na, J. (2018). The Psychometric Validation Of The Principal Practices Questionnaire Based On Item Response Theory. In *International Online Journal of Educational Leadership* (Vol. 2, Issue 1).
- Pichotac, N., Bonettoabc, E., Pavana, J. B., Arciszewskia, T., Bonnardel, N., & Weisbergd, R. W. (2020). The Construct Validity of Creativity: Empirical Arguments in Favor of Novelty as the Basis for Creativity. *Review of Research in Education*, 46(1), 288–323. <https://doi.org/10.31234/osf.io/wtd8n>
- Prasetya, W. A., & Pratama, A. T. (2023). Item quality analysis using the Rasch model to measure critical thinking ability in the material of the human digestive system of Biology subject in high school. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 27(1), 76–91.
- Rahmawati, Y., Hadinugrahaningsih, T., Ridwan, A., Palimbunga, U. S., & Mardiah, A. (2021). Developing the critical thinking skills of vocational school students in electrochemistry

- through STEM - Project-based learning (STEM-PjBL). The 2nd Science and Mathematics International Conference (SMIC 2020), 2331. <https://doi.org/10.1063/5.0041915>
- Rusmana, N., Suryana, D., Kurniasih, H. S., & Almigo, N. (2020). The development of speaking Skill's instrument in elementary school with rasch model analysis. *Universal Journal of Educational Research*, 8(7), 2758–2765. <https://doi.org/10.13189/ujer.2020.080702>
- Setiawan, M. A., Susongko, P., & Hayati, M. N. (2020). Pendeteksian DIF Pada Perangkat Tes Objektif Penilaian Akhir Semester IPA dengan Menggunakan Permodelan Rasch. *Pancasakti Science Education Journal*, 5(2), 23–29.
- Suratman, A., Afyaman, D., & Rakhmasari, R. (2019). Pembelajaran berbasis TIK terhadap Hasil Belajar Matematika dan Motivasi Belajar Matematika Siswa. *Jurnal Analisa*, 5(1), 41–50. <https://doi.org/10.15575/ja.v5i1.4828>
- Sym, I. K., & Louise, I. S. Y. (2023). Implementation of electronic module based multiple representation on buffer solution materials to improve students' self-regulated learning and cognitive learning outcomes. *Jurnal Penelitian Pendidikan IPA*, 9(10), 7816-7825. <https://doi.org/10.29303/jppipa.v9i10.3656>
- Toering, T., Elferink-Gemser, M. T., Jonker, L., van Heuvelen, M. J. G., & Visscher, C. (2012). Measuring self-regulation in a learning context: Reliability and validity of the Self-Regulation of Learning Self-Report Scale (SRL-SRS). *International Journal of Sport and Exercise Psychology*, 10(1), 24–38. <https://doi.org/10.1080/1612197X.2012.645132>
- van der Lans, R. M., van de Grift, W. J. C. M., & van Veen, K. (2018). Developing an instrument for teacher feedback: Using the rasch model to explore teachers' development of effective teaching strategies and behaviors. *Journal of Experimental Education*, 86(2), 247–264. <https://doi.org/10.1080/00220973.2016.1268086>
- Warlinda, Y. A., Yerimadesi, Y., Hardeli, H., & Andromeda, A. (2022). Implementation of guided discovery learning model with sets approach assisted by e-modul chemistry on scientific literacy of students. *Jurnal Penelitian Pendidikan IPA*, 8(2), 507–514. <https://doi.org/10.29303/jppipa.v8i2.1264>
- Yastuti, D. F., & Suwatno, S. (2020). Upaya Meningkatkan Hasil Belajar Siswa Melalui Motivasi Guru. *Jurnal Pendidikan Manajemen Perkantoran*, 2(1), 252. <https://doi.org/10.17509/jpm.v2i1.14605>
- Yunika, V. R., Rohmah, A. N., Istiqomah, S. N., & Faradillah, A. (2021a). Validitas dan Reliabilitas Anxiety Questioner dalam Pembelajaran Matematika dengan menggunakan Rasch Models. *Seminar Nasional Pendidikan Matematika Universitas Pattimura*, 2, 161–169.
- Zainudin, M., Subali, B., & Jailani. (2019). Construct validity of mathematical creativity instrument: First-order and second-order confirmatory factor analysis. *International Journal of Instruction*, 12(3), 595–614. <https://doi.org/10.293>
- Zimmerman, B. J., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology*, 80(3), 284–290. <https://doi.org/10.1037/0022-0663.80.3.284>