

PENGEMBANGAN INSTRUMEN TES KETERAMPILAN PROSES SAINS (KPS) PADA MATERI LAJU REAKSI: ANALISIS DENGAN MODEL RASCH

Nur Islamiah¹, Ramlawati², Sudding³

Universitas Negeri Makassar

e-mail: ramlawati@unm.ac.id

Diterima: 10/07/2026; Direvisi: 16/07/2026; Diterbitkan: 30/08/2026

ABSTRAK

Keterampilan proses sains (KPS) merupakan aspek penting dalam pembelajaran kimia pada Kurikulum Merdeka, namun penilaian pembelajaran kimia di sekolah masih lebih menitikberatkan pada aspek pengetahuan, sementara instrumen tes yang secara sistematis mengukur KPS masih terbatas, khususnya pada materi laju reaksi yang sarat aktivitas ilmiah. Kesenjangan ini diperbesar oleh belum terintegrasinya validasi ahli dengan analisis empiris berbasis Model Rasch pada pengembangan instrumen KPS materi tersebut, sehingga pengembangan instrumen yang valid, praktis, dan reliabel menjadi penting untuk dilakukan. Penelitian ini merupakan penelitian pengembangan yang bertujuan menghasilkan instrumen tes keterampilan proses sains (KPS) pada materi laju reaksi yang valid, praktis, dan reliabel. Pengembangan instrumen menggunakan model Plomp yang meliputi fase investigasi awal, fase desain, fase realisasi atau konstruksi, fase tes, evaluasi dan revisi, serta fase implementasi. Instrumen yang dikembangkan berupa butir soal pilihan ganda dengan lima alternatif jawaban yang meliputi indikator keterampilan proses sains. Uji coba dilakukan pada peserta didik kelas XI MK 1 SMA Negeri 10 Gowa. Data penelitian diperoleh melalui lembar validasi ahli, angket respons guru dan peserta didik, serta hasil tes peserta didik. Analisis data dilakukan secara deskriptif dan menggunakan Model Rasch. Hasil penelitian menunjukkan bahwa instrumen tes keterampilan proses sains yang dikembangkan dinyatakan valid berdasarkan penilaian ahli, reliabel berdasarkan analisis konsistensi teoretis maupun empiris, serta praktis berdasarkan respons guru yang sangat positif terhadap kemudahan penggunaannya. Dengan demikian, instrumen tes keterampilan proses sains pada materi laju reaksi yang telah dikembangkan telah memenuhi aspek kevalidan, reliabilitas, dan kepraktisan yang layak digunakan sebagai alat ukur keterampilan proses sains peserta didik.

Kata kunci: pengembangan plomp, instrumen tes, keterampilan proses sains, laju reaksi, model rasch.

ABSTRACT

Science process skills (SPS) are a key component of chemistry learning under the *Kurikulum Merdeka*, yet classroom assessment in chemistry still emphasizes knowledge over skills, and systematically designed SPS test instruments remain limited particularly for the reaction rate topic, which is closely tied to scientific inquiry activities. This gap is widened by the limited integration of expert validation with empirical Rasch Model analysis in developing SPS instruments for this topic, making the development of a valid, practical, and reliable instrument an important undertaking. This study is a development research project aimed at producing a valid, practical, and reliable Science Process Skills (SPS) test instrument covering the topic of reaction rates. The instrument was developed using the Plomp model, which comprises the preliminary investigation, design, realization/construction, test/evaluation/revision, and implementation phases. The developed instrument consists of multiple-choice questions with five answer options covering science process skill indicators. Field testing was conducted with

students from class XI MK 1 at SMA Negeri 10 Gowa. Data were collected via expert validation sheets, teacher and student response questionnaires, and student test results. Data analysis was performed using descriptive methods and the Rasch Model. The results indicate that the developed SPS test instrument was confirmed valid based on expert judgment, reliable based on both theoretical and empirical consistency analyses, and practical based on highly positive teacher responses regarding its ease of use. Consequently, the developed SPS test instrument for the reaction rate topic satisfies the requirements for validity, reliability, and practicality, making it suitable for measuring students' science process skills.

Keywords: Plomp development, test instrument, science process skills, reaction rate, rasch model.

PENDAHULUAN

Keterampilan proses sains (*science process skills*) merupakan kemampuan dasar yang memungkinkan peserta didik terlibat secara aktif dalam kegiatan ilmiah, mencakup kegiatan mengamati, mengklasifikasi, mengukur, mengomunikasikan, menyimpulkan, memprediksi, mengidentifikasi variabel, menyajikan data, mendeskripsikan hubungan antarvariabel, menyusun hipotesis, mendefinisikan variabel secara operasional, serta merancang percobaan (Rezba et al., 2007). Dalam konteks Kurikulum Merdeka, pembelajaran kimia menuntut peserta didik tidak hanya menguasai konsep, tetapi juga mampu mengaitkan fenomena nyata dengan penjelasan ilmiah, sebuah tuntutan yang muncul akibat kesenjangan antara pengalaman empiris yang dapat diamati secara langsung dan konsep teoritis yang bersifat abstrak (Taber, 2020). Melalui keterampilan proses sains, peserta didik dapat meningkatkan keterlibatan dalam proses pembelajaran, mengembangkan teknik penyelidikan dan strategi berpikir ilmiah, serta membangun pengetahuan melalui pengalaman ilmiah secara langsung (Ramlawati et al., 2025). Penguasaan keterampilan ini dinilai lebih esensial daripada sekadar menghafal fakta dan konsep sains, karena turut berkontribusi terhadap pengembangan sikap ilmiah, nilai, dan karakter sebagai dampak iringan dari proses pembelajaran.

Namun, praktik pembelajaran kimia di sekolah menunjukkan kesenjangan antara tuntutan kurikulum tersebut dan pelaksanaan penilaian di lapangan. Penilaian hasil belajar masih lebih menitikberatkan pada aspek pengetahuan dibandingkan aspek keterampilan (Ramlawati et al., 2025), dan keterampilan proses sains kerap dipersepsikan hanya dapat diukur melalui kegiatan praktikum, sehingga pengukurannya belum dilakukan secara komprehensif dan berkelanjutan. Di lapangan, kondisi ini tampak dari terbatasnya instrumen tes tertulis yang secara khusus dirancang untuk mengukur keterampilan proses sains pada materi kimia tertentu; guru pada umumnya masih mengandalkan soal-soal yang berorientasi pada penguasaan konsep, sementara instrumen yang secara sistematis memetakan indikator keterampilan proses sains apalagi yang telah diuji kualitasnya secara empiris jarang tersedia sebagai alat ukur siap pakai. Akibatnya, penguasaan keterampilan proses sains peserta didik belum memperoleh perhatian yang memadai dalam sistem penilaian pembelajaran kimia.

Sejumlah penelitian terdahulu telah berupaya menjawab persoalan ini, namun bila dicermati bersama-sama, penelitian tersebut memperlihatkan pola yang konsisten sekaligus celah yang belum tertutup. Yanti dan Fuadiyah (2025) serta Riani dan Ramalis (2020) menunjukkan bahwa instrumen keterampilan proses sains yang disusun berdasarkan indikator yang sistematis dapat dinyatakan valid dan layak digunakan, tetapi kelayakan tersebut umumnya ditetapkan berdasarkan penilaian ahli tanpa disertai analisis butir secara empiris. Di sisi lain, Tauhidah dan Rofi'ah (2023) memperlihatkan bahwa analisis Model Rasch mampu mengungkap reliabilitas instrumen, kesesuaian butir (*item fit*), dan tingkat kesukaran yang selaras dengan kemampuan peserta didik informasi yang tidak dapat diperoleh dari pendekatan

validasi ahli semata. Dengan kata lain, penelitian-penelitian tersebut secara bersama-sama menunjukkan bahwa validitas isi dan analisis empiris berbasis *Item Response Theory* saling melengkapi, namun keduanya jarang diterapkan secara terintegrasi dalam satu pengembangan instrumen, dan seluruhnya dikembangkan pada materi yang berbeda dari laju reaksi.

Kesenjangan ini menjadi lebih nyata pada materi laju reaksi. Materi ini memiliki karakteristik yang sangat erat dengan aktivitas ilmiah, seperti mengamati perubahan selama reaksi, mengidentifikasi variabel yang memengaruhi laju reaksi, merumuskan hipotesis, menganalisis data hasil percobaan, serta menarik kesimpulan berdasarkan bukti empiris, sehingga semestinya menjadi konteks yang ideal untuk mengukur keterampilan proses sains secara utuh. Namun, pengembangan instrumen tes keterampilan proses sains pada materi laju reaksi masih relatif terbatas, dan sejauh penelusuran penulis, belum ditemukan instrumen serupa yang divalidasi tidak hanya melalui penilaian ahli, tetapi juga dianalisis secara empiris menggunakan Model Rasch untuk memastikan kualitas setiap butir soal. Hasanah dan Aini (2025) menegaskan bahwa instrumen keterampilan proses sains perlu dikembangkan melalui prosedur yang sistematis dan pengujian empiris yang memadai agar menghasilkan pengukuran yang objektif dan akurat sebuah syarat yang pada materi laju reaksi belum sepenuhnya terpenuhi. Kesenjangan ini bersifat sekaligus empiris (belum tersedia data uji coba pada materi laju reaksi) dan metodologis (belum diterapkannya analisis Model Rasch secara terintegrasi dengan validasi ahli pada instrumen KPS materi ini).

Model Rasch dipilih untuk menutup kesenjangan metodologis tersebut karena mampu menempatkan kemampuan peserta didik dan tingkat kesukaran butir soal pada satu skala logit, menganalisis kesesuaian butir (*item fit*), serta menghasilkan parameter pengukuran yang lebih objektif dibandingkan pendekatan tes klasik. Model ini juga mampu memisahkan karakteristik butir soal dari kemampuan peserta didik sehingga hasil pengukuran menjadi lebih stabil dan dapat dipertanggungjawabkan secara empiris (Sumintono & Widhiarso, 2015), sekaligus memungkinkan identifikasi butir yang terlalu mudah dan terlalu sulit.

Berdasarkan kesenjangan tersebut, kebaruan (*novelty*) penelitian ini terletak pada pengembangan instrumen tes keterampilan proses sains yang secara khusus dirancang untuk materi laju reaksi, dengan proses validasi yang mengintegrasikan penilaian ahli dan analisis empiris berbasis Model Rasch secara menyeluruh mencakup kesesuaian butir, tingkat kesukaran, dan reliabilitas sehingga menghasilkan instrumen yang teruji baik dari sisi konten maupun dari sisi psikometrik, suatu kombinasi yang belum ditemukan pada penelitian-penelitian keterampilan proses sains sebelumnya. Oleh karena itu, penelitian ini bertujuan untuk mengembangkan instrumen tes keterampilan proses sains pada materi laju reaksi yang valid, praktis, dan reliabel, serta menganalisis kualitas butir soal instrumen tersebut menggunakan Model Rasch guna menghasilkan alat ukur keterampilan proses sains yang objektif dan akurat dalam pembelajaran kimia.

METODE PENELITIAN

Penelitian ini merupakan penelitian dan pengembangan (*Research and Development/R&D*) yang bertujuan mengembangkan instrumen tes keterampilan proses sains (KPS) pada materi laju reaksi serta menguji kualitasnya berdasarkan aspek validitas, reliabilitas, dan kepraktisan. Model pengembangan yang digunakan mengadaptasi model Plomp yang terdiri atas lima fase, yaitu investigasi awal (*preliminary investigation*), desain (*design*), realisasi/konstruksi (*realization/construction*), tes, evaluasi, dan revisi (*test, evaluation, and revision*), serta implementasi (*implementation*). Analisis kualitas instrumen

dilakukan menggunakan Model Rasch sehingga mampu memberikan informasi mengenai karakteristik butir soal dan kemampuan peserta didik secara objektif.

Penelitian dilaksanakan di SMA Negeri 10 Gowa dengan melibatkan 30 peserta didik kelas XI KM 1 sebagai subjek uji coba instrumen. Pada fase investigasi awal dilakukan observasi, wawancara dengan guru kimia, serta analisis capaian pembelajaran, materi laju reaksi, dan indikator keterampilan proses sains yang menjadi dasar pengembangan instrumen. Hasil analisis digunakan untuk menyusun spesifikasi instrumen yang mengintegrasikan enam indikator KPS dasar dan enam indikator KPS terintegrasi sesuai capaian pembelajaran Kurikulum Merdeka.

Fase desain menghasilkan rancangan instrumen berupa kisi-kisi, penyusunan stimulus, penulisan butir soal, kunci jawaban, dan pedoman penskoran. Instrumen dikembangkan dalam bentuk tes pilihan ganda sebanyak 25 butir dengan stimulus berupa tabel, grafik, gambar, data hasil percobaan, dan fenomena kontekstual yang dirancang untuk mengukur keterampilan proses sains pada materi laju reaksi. Selanjutnya, pada fase realisasi/konstruksi disusun prototipe awal instrumen beserta lembar validasi serta angket respon guru dan peserta didik.

Fase tes, evaluasi, dan revisi diawali dengan validasi isi oleh dua dosen ahli Pendidikan Kimia yang menilai aspek kelayakan isi, konstruksi, dan kebahasaan instrumen. Hasil validasi dianalisis menggunakan rerata skor validitas yakni apabila validator memberikan kategori valid atau sangat valid sedangkan reliabilitas antarvalidator dihitung menggunakan *Percentage of Agreement* dengan rumus sebagai berikut:

$$R = \frac{A}{A + D} \times 100\%$$

Keterangan:

A = besarnya frekuensi kecocokan antara data dua pengamat.

D = besarnya frekuensi ketidakcocokan antara data dua pengamat.

R = koefisien (derajat) reliabilitas instrumen.

Instrumen penelitian dikatakan reliabel jika nilai reliabilitasnya $r_{xx} \geq 0,75$ atau $r_{xx} \geq 75\%$. Kemudian setelah itu masukan dari validator digunakan sebagai dasar revisi hingga diperoleh instrumen yang layak untuk diujicobakan. Pada fase implementasi, instrumen yang telah direvisi diujicobakan kepada 30 peserta didik kelas XI KM 1 SMA Negeri 10 Gowa. Data hasil uji coba dianalisis menggunakan Model Rasch untuk mengevaluasi unidimensionalitas, kesesuaian butir (*item fit*), tingkat kesukaran butir, serta reliabilitas item dan person. Kepraktisan instrumen diperoleh melalui angket respon guru dan peserta didik yang dianalisis secara deskriptif menggunakan persentase.

$$PRS = \frac{\sum A}{\sum B} \times 100\%$$

Keterangan:

PRS = Persentase respon pendidik/peserta didik

$\sum A$ = Jumlah skor perolehan respon pendidik/peserta didik

$\sum B$ = Jumlah maksimal angket respon

Skor persentase yang diperoleh selanjutnya diinterpretasikan berdasarkan rentang kategori berikut: 81–100% = sangat baik/sangat valid/sangat praktis; 61–80% = baik/valid/praktis; 41–60% = cukup; 21–40% = kurang; dan 0–20% = sangat kurang.

Instrumen penelitian meliputi lembar validasi ahli, instrumen tes KPS, lembar jawaban peserta didik, serta angket respon guru dan peserta didik. Instrumen dinyatakan valid apabila memenuhi kriteria validitas isi dan kesesuaian butir berdasarkan Model Rasch, reliabel apabila memenuhi kriteria *Percentage of Agreement* serta memiliki nilai *item reliability* dan *person*

reliability yang baik, serta praktis apabila memperoleh persentase respons guru pada kategori tinggi atau sangat tinggi.

HASIL DAN PEMBAHASAN

Hasil

Hasil Pengembangan Instrumen Tes Keterampilan Proses Sains

Pengembangan instrumen tes keterampilan proses sains (KPS) pada materi laju reaksi dilaksanakan menggunakan model Plomp yang meliputi lima fase, yaitu investigasi awal, desain, realisasi/konstruksi, tes, evaluasi dan revisi, serta implementasi. Pada fase investigasi awal dilakukan observasi dan wawancara dengan guru kimia untuk mengidentifikasi kebutuhan pengembangan instrumen. Hasil analisis menunjukkan bahwa penilaian yang digunakan masih berfokus pada penguasaan konsep dan belum secara khusus mengukur keterampilan proses sains peserta didik. Selain itu, instrumen yang tersedia masih terbatas sehingga diperlukan pengembangan instrumen yang mampu mengukur KPS secara lebih objektif melalui stimulus berupa data, tabel, grafik, gambar, ilustrasi percobaan, dan permasalahan kontekstual. Analisis capaian pembelajaran dan materi laju reaksi menghasilkan 12 indikator KPS yang terdiri atas enam indikator KPS dasar dan enam indikator KPS terintegrasi sebagai dasar penyusunan instrumen. Kemudian pada fase desain disusun spesifikasi instrumen yang meliputi kisi-kisi, stimulus, butir soal, kunci jawaban, dan pedoman penskoran. Instrumen dikembangkan dalam bentuk tes pilihan ganda sebanyak 25 butir yang dirancang untuk mengukur seluruh indikator KPS melalui berbagai stimulus kontekstual sesuai materi laju reaksi.

Fase realisasi/konstruksi menghasilkan prototipe I berupa instrumen tes KPS beserta lembar validasi dan angket respons guru. Selanjutnya, pada fase tes, evaluasi, dan revisi, instrumen divalidasi oleh dua dosen ahli Pendidikan Kimia berdasarkan aspek kelayakan isi, konstruksi, dan kebahasaan. Hasil validasi beserta masukan validator digunakan sebagai dasar penyempurnaan redaksi soal, kesesuaian indikator, penggunaan bahasa, dan pedoman penskoran sehingga dihasilkan prototipe II yang siap diimplementasikan. Terakhir fase implementasi, prototipe II diujicobakan kepada 30 peserta didik kelas XI KM 1 SMA Negeri 10 Gowa. Data hasil uji coba dianalisis menggunakan Model Rasch untuk mengevaluasi unidimensionalitas, kesesuaian butir (*item fit*), tingkat kesukaran butir, serta analisis reliabilitas yang meliputi *Cronbach Alpha (KR-20)*, *Person Reliability*, dan *Item Reliability*. Kepraktisan instrumen dievaluasi melalui angket respons guru untuk mengetahui kemudahan penggunaan instrumen dalam proses penilaian. Hasil analisis menunjukkan bahwa instrumen yang dikembangkan memenuhi kriteria valid, reliabel, dan praktis sehingga layak digunakan untuk mengukur keterampilan proses sains peserta didik pada materi laju reaksi.

Kualitas Instrumen Tes Keterampilan Proses Sains

Analisis Validasi Ahli

Validitas instrumen meliputi validitas logis (teoretis) dan validitas empiris. Validitas logis dilakukan oleh dua validator ahli terhadap kisi-kisi dan butir soal berdasarkan aspek isi, konstruksi, dan kebahasaan. Hasil validasi menunjukkan bahwa kisi-kisi instrumen memperoleh skor rata-rata 3,90 dengan reliabilitas antarvalidator sebesar 90,00%, sedangkan lembar soal memperoleh skor rata-rata 3,90 dengan reliabilitas sebesar 92,86%. Kedua hasil tersebut berada pada kategori sangat valid sehingga instrumen layak untuk diujicobakan setelah dilakukan penyempurnaan sesuai saran validator, terutama pada aspek redaksi, penyajian stimulus, dan penggunaan bahasa.

Analisis Pengujian Empirik Menggunakan Model Rasch

Pengujian empirik terhadap instrumen tes Keterampilan Proses Sains (KPS) pada materi laju reaksi dilakukan menggunakan Model Rasch dengan bantuan perangkat lunak Ministep. Analisis bertujuan mengevaluasi kualitas instrumen berdasarkan unidimensionalitas, kesesuaian butir (*item fit*), tingkat kesukaran butir (*item difficulty*), dan reliabilitas instrumen.

1. Unidimensionalitas

Analisis unidimensionalitas dilakukan untuk mengetahui sejauh mana instrumen tes KPS mengukur satu konstruk utama sesuai dengan tujuan pengukuran.

Tabel 1. Hasil Output Uji Unidimensionalitas

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = Item information units			
	Eigenvalue	Observed	Expected
Total raw variance in observations =	36.7934	100.0%	100.0%
Raw variance explained by measures =	11.7934	32.1%	31.7%
Raw variance explained by persons =	5.7340	15.6%	15.4%
Raw Variance explained by items =	6.0594	16.5%	16.3%
Raw unexplained variance (total) =	25.0000	67.9%	100.0%
Unexplned variance in 1st contrast =	2.9000	7.9%	11.6%
Unexplned variance in 2nd contrast =	2.6617	7.2%	10.6%
Unexplned variance in 3rd contrast =	2.4924	6.8%	10.0%
Unexplned variance in 4th contrast =	2.1924	6.0%	8.8%
Unexplned variance in 5th contrast =	1.9754	5.4%	7.9%

Tabel 2. Hasil Analisis Unidimensionalitas

Indikator	Hasil	Kriteria	Keterangan
Raw Variance Explained by Measures	32,1%	> 20%	Memenuhi
Unexplained Variance in Contrast	7,9%	< 15%	Memenuhi

(Sumber: Data Peneliti)

Hasil analisis menunjukkan bahwa *Raw Variance Explained by Measures* sebesar 32,1% dan *Unexplained Variance in Contrast* sebesar 7,9%. Kedua nilai tersebut memenuhi kriteria unidimensionalitas, sehingga instrumen mampu mengukur satu konstruk utama sesuai tujuan pengukuran dan layak dianalisis lebih lanjut menggunakan Model Rasch.

2. Kesesuaian Butir (*Item Fit*)

Analisis kesesuaian butir (*item fit*) dilakukan untuk mengetahui tingkat kesesuaian setiap butir soal dengan Model Rasch. Hasil outputnya dapat dilihat pada Tabel 3.

Tabel 3. Hasil Output Kesesuaian Butir (*Item Fit*)

Item STATISTICS: MISFIT ORDER														
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	EXACT EXP.	MATCH OBS	EXP%	Item	
3	26	30	-1.64	.57	.88	-.24	3.56	2.41	A .26	.32	90.0	86.7	S3	
10	19	30	-.05	.42	1.31	1.64	1.60	1.84	B .16	.44	63.3	71.8	S10	
12	25	30	-1.34	.53	1.21	.75	1.32	.68	C .17	.35	80.0	83.9	S12	
15	18	30	.13	.42	1.30	1.63	1.23	.91	D .24	.45	53.3	71.1	S15	
5	25	30	-1.34	.53	1.17	.63	1.22	.54	E .20	.35	86.7	83.9	S5	
24	19	30	-.05	.42	.97	-.13	1.22	.82	F .43	.44	70.0	71.8	S24	
8	20	30	-.23	.43	.88	-.57	1.18	.64	G .47	.44	83.3	73.0	S8	
19	19	30	-.05	.42	1.08	.49	1.14	.56	H .37	.44	76.7	71.8	S19	
13	20	30	-.23	.43	1.13	.72	1.12	.46	I .33	.44	70.0	73.0	S13	
1	20	30	-.23	.43	1.10	.56	1.06	.28	J .37	.44	70.0	73.0	S1	
11	19	30	-.05	.42	1.09	.55	1.04	.23	K .38	.44	76.7	71.8	S11	
18	10	30	1.51	.43	1.07	.39	1.02	.18	L .39	.44	70.0	73.4	S18	
4	24	30	-1.07	.50	1.04	.25	.92	.02	M .35	.38	80.0	81.1	S4	
7	11	30	1.33	.43	1.04	.28	.97	-.01	L .43	.45	73.3	72.3	S7	
25	20	30	-.23	.43	.99	.01	.85	-.40	k .47	.44	63.3	73.0	S25	
16	18	30	.13	.42	.97	-.14	.91	-.27	j .48	.45	80.0	71.1	S16	
17	8	30	1.90	.46	.96	-.12	.91	-.09	i .45	.42	80.0	76.8	S17	
20	18	30	.13	.42	.95	-.22	.83	-.60	h .51	.45	66.7	71.1	S20	
22	9	30	1.70	.44	.90	-.44	.74	-.69	g .53	.43	80.0	74.7	S22	
21	10	30	1.51	.43	.88	-.58	.73	-.82	f .56	.44	70.0	73.4	S21	
9	20	30	-.23	.43	.85	-.78	.74	-.76	e .56	.44	83.3	73.0	S9	
14	19	30	-.05	.42	.83	-.96	.79	-.70	d .58	.44	83.3	71.8	S14	
6	26	30	-1.64	.57	.73	-.70	.48	-.64	c .54	.32	90.0	86.7	S6	
2	27	30	-2.01	.64	.69	-.65	.32	-.77	b .55	.29	90.0	89.9	S2	
23	7	30	2.12	.47	.57	-2.03	.40	-1.57	a .74	.40	86.7	78.9	S23	
MEAN	18.3	30.0	.00	.46	.98	.01	1.05	.09			76.7	76.0		
P.SD	5.8	.0	1.13	.06	.18	.79	.59	.86			9.3	5.7		

Tabel 4. Hasil Analisis Item Fit

No.	Status Butir	Jumlah Butir	Persentase (%)
1	Fit	23	92%
2	Misfit	2	8%

(Sumber: Data Peneliti)

Hasil analisis menunjukkan bahwa 23 dari 25 butir (92%) memenuhi kriteria *fit*, sedangkan dua butir (S3 dan S10) tergolong *misfit*. Temuan ini menunjukkan bahwa sebagian besar butir telah sesuai dengan Model Rasch sehingga mampu mengukur konstruk keterampilan proses sains secara konsisten. Butir yang tergolong *misfit* menjadi bahan evaluasi untuk perbaikan instrumen.

3. Tingkat Kesukaran Butir

Analisis tingkat kesukaran butir dilakukan untuk mengetahui sebaran tingkat kesukaran setiap butir soal berdasarkan nilai logit pada Model Rasch.

Tabel 5. Hasil Output Kesukaran Butir Soal

Item STATISTICS: MEASURE ORDER													
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	JMLE MEASURE	MODEL S.E.	INFIT MNSQ	OUTFIT ZSTD	PTMEASUR-AL CORR.	EXACT EXP.	EXACT MATCH OBSN	EXPN	Item		
23	7	30	2.12	.47	.57	-2.03	.40	-.157	.74	.40	86.7	78.9	S23
17	8	30	1.90	.46	.96	-.12	.91	-.09	.45	.42	80.0	76.8	S17
22	9	30	1.70	.44	.90	-.44	.74	-.69	.53	.43	80.0	74.7	S22
18	10	30	1.51	.43	1.07	.39	1.02	.18	.39	.44	70.0	73.4	S18
21	10	30	1.51	.43	.88	-.58	.73	-.82	.56	.44	70.0	73.4	S21
7	11	30	1.33	.43	1.04	.28	.97	-.01	.43	.45	73.3	72.3	S7
15	18	30	.13	.42	1.30	1.63	1.23	.91	.24	.45	53.3	71.1	S15
16	18	30	.13	.42	.97	-.14	.91	-.27	.48	.45	80.0	71.1	S16
20	18	30	.13	.42	.95	-.22	.83	-.60	.51	.45	66.7	71.1	S20
10	19	30	-.05	.42	1.31	1.64	1.60	1.84	.16	.44	63.3	71.8	S10
11	19	30	-.05	.42	1.09	.55	1.04	.23	.38	.44	76.7	71.8	S11
14	19	30	-.05	.42	.83	-.96	.79	-.70	.58	.44	83.3	71.8	S14
19	19	30	-.05	.42	1.08	.49	1.14	.56	.37	.44	76.7	71.8	S19
24	19	30	-.05	.42	.97	-.13	1.22	.82	.43	.44	70.0	71.8	S24
1	20	30	-.23	.43	1.10	.56	1.06	.28	.37	.44	70.0	73.0	S1
8	20	30	-.23	.43	.88	-.57	1.18	.64	.47	.44	83.3	73.0	S8
9	20	30	-.23	.43	.85	-.78	.74	-.76	.56	.44	83.3	73.0	S9
13	20	30	-.23	.43	1.13	.72	1.12	.46	.33	.44	70.0	73.0	S13
25	20	30	-.23	.43	.99	.01	.85	-.40	.47	.44	63.3	73.0	S25
4	24	30	-1.07	.50	1.04	.25	.92	.02	.35	.38	80.0	81.1	S4
5	25	30	-1.34	.53	1.17	.63	1.22	.54	.20	.35	86.7	83.9	S5
12	25	30	-1.34	.53	1.21	.75	1.32	.68	.17	.35	80.0	83.9	S12
3	26	30	-1.64	.57	.88	-.24	3.56	2.41	.26	.32	90.0	86.7	S3
6	26	30	-1.64	.57	.73	-.70	.48	-.64	.54	.32	90.0	86.7	S6
2	27	30	-2.01	.64	.69	-.65	.32	-.77	.55	.29	90.0	89.9	S2
MEAN	18.3	30.0	.00	.46	.98	.01	1.05	.09			76.7	76.0	
P.SD	5.8	.0	1.13	.06	.18	.79	.59	.86			9.3	5.7	

Tabel 6. Distribusi Tingkat Kesukaran Butir

No.	Kategori	Jumlah Butir	Persentase
1	Sangat Sukar	1	4%
2	Sukar	5	20%
3	Sedang	13	52%
4	Mudah	5	20%
5	Sangat Mudah	1	4%

(Sumber: Data Peneliti)

Hasil analisis menunjukkan bahwa sebagian besar butir berada pada kategori sedang (52%), sedangkan masing-masing 20% berada pada kategori sukar dan mudah. Sementara itu, hanya satu butir (4%) yang termasuk kategori sangat sukar dan satu butir (4%) berkategori sangat mudah. Distribusi tersebut menunjukkan bahwa instrumen memiliki variasi tingkat kesukaran yang baik sehingga dapat mengukur kemampuan peserta didik pada berbagai tingkat kemampuan.

4. Reliabilitas Instrumen

Analisis reliabilitas dilakukan untuk mengetahui tingkat konsistensi instrumen dalam mengukur keterampilan proses sains peserta didik.

Tabel 7. Summary Statistics

SUMMARY OF 30 MEASURED Person								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	15.2	25.0	.64	.51	1.01	-.02	1.05	-.06
SEM	.9	.0	.21	.01	.04	.18	.12	.17
P.SD	4.7	.0	1.13	.08	.20	.97	.62	.90
S.SD	4.8	.0	1.15	.08	.21	.99	.63	.92
MAX.	23.0	25.0	2.94	.77	1.44	1.73	4.11	1.99
MIN.	5.0	25.0	-1.72	.45	.67	-1.83	.57	-1.69

REAL RMSE	.54	TRUE SD	.99	SEPARATION	1.84	Person	RELIABILITY	.77
MODEL RMSE	.51	TRUE SD	1.00	SEPARATION	1.95	Person	RELIABILITY	.79
S.E. OF Person MEAN = .21								

Person RAW SCORE-TO-MEASURE CORRELATION = .99								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .81 SEM = 2.07								
STANDARDIZED (50 ITEM) RELIABILITY = .88								

SUMMARY OF 25 MEASURED Item								
	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT	
					MNSQ	ZSTD	MNSQ	ZSTD
MEAN	18.3	30.0	.00	.46	.98	.01	1.05	.09
SEM	1.2	.0	.23	.01	.04	.16	.12	.18
P.SD	5.8	.0	1.13	.06	.18	.79	.59	.86
S.SD	5.9	.0	1.15	.06	.18	.80	.60	.88
MAX.	27.0	30.0	2.12	.64	1.31	1.64	3.56	2.41
MIN.	7.0	30.0	-2.01	.42	.57	-2.03	.32	-1.57

REAL RMSE	.48	TRUE SD	1.02	SEPARATION	2.13	Item	RELIABILITY	.82
MODEL RMSE	.47	TRUE SD	1.03	SEPARATION	2.21	Item	RELIABILITY	.83
S.E. OF Item MEAN = .23								

Item RAW SCORE-TO-MEASURE CORRELATION = -1.00								
Global statistics: please see Table 44.								
UMEAN=.0000 USCALE=1.0000								

Tabel 8. Hasil Analisis Reliabilitas

Aspek	Nilai	Kategori
Cronbach Alpha (KR-20)	0,81	Bagus Sekali
Person Reliability	0,79	Cukup
Item Reliability	0,83	Bagus

(Sumber: Data Peneliti)

Nilai Cronbach Alpha sebesar 0,81 menunjukkan konsistensi internal instrumen yang sangat baik. Nilai *Person Reliability* sebesar 0,79 menunjukkan kemampuan instrumen yang cukup dalam membedakan kemampuan peserta didik, sedangkan *Item Reliability* sebesar 0,83 menunjukkan konsistensi tingkat kesukaran butir yang baik. Secara keseluruhan, hasil tersebut menunjukkan bahwa instrumen memiliki reliabilitas yang baik dan layak digunakan untuk mengukur keterampilan proses sains pada materi laju reaksi.

Analisis Kepraktisan

Kepraktisan instrumen dievaluasi melalui angket respons guru setelah implementasi instrumen. Hasil analisis menunjukkan bahwa guru 1 memberikan persentase respons sebesar 96,87%, sedangkan guru 2 sebesar 93,75%, yang keduanya termasuk dalam kategori sangat tinggi. Begitupun respon peserta didik berada pada kategori baik hingga sangat baik. Hasil tersebut menunjukkan bahwa instrumen mudah dipahami, mudah diadministrasikan, memiliki pedoman penskoran yang jelas, serta membantu guru dalam mengukur keterampilan proses sains peserta didik. Dengan demikian, instrumen yang dikembangkan memenuhi kriteria praktis untuk digunakan dalam pembelajaran.

Pembahasan

Validitas

Hasil penelitian menunjukkan bahwa instrumen tes keterampilan proses sains (KPS) yang dikembangkan telah memenuhi kriteria valid secara teoretis maupun empiris. Validitas teoretis ditunjukkan oleh skor rata-rata validasi sebesar 3,90 dengan kategori sangat valid, yang menunjukkan bahwa isi, konstruksi, dan kebahasaan instrumen telah sesuai dengan indikator capaian pembelajaran dan indikator KPS. Hasil ini mengindikasikan bahwa butir-butir soal telah merepresentasikan keterampilan proses sains sebagaimana dikemukakan oleh Rezba dkk. (2007), baik pada aspek keterampilan proses sains dasar maupun keterampilan proses sains terintegrasi.

Temuan tersebut diperkuat oleh hasil analisis Model Rasch yang menunjukkan nilai *Raw Variance Explained by Measures* sebesar 32,1% dan *Unexplained Variance in 1st Contrast* sebesar 7,9%, sehingga memenuhi asumsi unidimensionalitas. Artinya, instrumen mampu mengukur satu konstruk utama, yaitu keterampilan proses sains pada materi laju reaksi. Selain itu, sebanyak 23 dari 25 butir soal (92%) memenuhi kriteria *item fit*, yang menunjukkan bahwa sebagian besar butir telah berfungsi sesuai dengan model pengukuran. Meskipun terdapat dua butir yang belum memenuhi kriteria, proporsi tersebut tidak memengaruhi kualitas instrumen secara keseluruhan dan justru menjadi dasar untuk penyempurnaan lebih lanjut.

Analisis *Point Measure Correlation* juga menunjukkan bahwa seluruh butir memiliki korelasi positif sehingga mampu membedakan peserta didik sesuai tingkat kemampuannya. Di samping itu, distribusi tingkat kesukaran yang terdiri atas butir sangat mudah, mudah, sedang, sukar, dan sangat sukar menunjukkan bahwa instrumen mampu mengukur kemampuan peserta didik pada berbagai tingkat kemampuan. Dengan demikian, hasil validitas logis dan empiris saling mendukung bahwa instrumen memiliki validitas isi dan validitas konstruk yang baik sehingga layak digunakan sebagai alat ukur keterampilan proses sains.

Reliabilitas

Selain valid, instrumen yang dikembangkan juga menunjukkan tingkat reliabilitas yang baik. Reliabilitas teoretis berdasarkan *Percentage of Agreement* memperoleh nilai 90,00% pada kisi-kisi dan 92,86% pada lembar soal. Tingginya tingkat kesepakatan antarvalidator menunjukkan bahwa instrumen memiliki konsistensi yang baik dari aspek isi, konstruksi, dan kebahasaan sebelum diujicobakan.

Hasil tersebut diperkuat oleh analisis reliabilitas empiris menggunakan Model Rasch. Nilai *Cronbach's Alpha* sebesar 0,81, *person reliability* sebesar 0,79, dan *item reliability* sebesar 0,83 menunjukkan bahwa instrumen memiliki konsistensi internal yang baik. Selain itu, nilai *person separation* sebesar 1,95 dan *item separation* sebesar 2,21 mengindikasikan bahwa instrumen mampu membedakan kemampuan peserta didik maupun tingkat kesukaran butir

secara memadai. Temuan ini sejalan dengan Sumintono dan Widhiarso (2015) yang menyatakan bahwa reliabilitas yang baik dalam Model Rasch menunjukkan kestabilan hasil pengukuran serta kualitas butir soal yang konsisten. Oleh karena itu, instrumen yang dikembangkan layak digunakan sebagai alat ukur keterampilan proses sains.

Kepraktisan

Kepraktisan instrumen ditunjukkan melalui respons guru dan peserta didik setelah implementasi. Respons guru memperoleh persentase 96,87% dan 93,75%, yang keduanya berada pada kategori sangat tinggi. Hasil ini menunjukkan bahwa instrumen mudah dipahami, mudah diadministrasikan, memiliki petunjuk yang jelas, serta membantu guru dalam mengevaluasi keterampilan proses sains peserta didik. Respons peserta didik juga menunjukkan bahwa instrumen diterima dengan baik, dengan persentase respons berada pada rentang 55%–100% dan mayoritas berada pada kategori tinggi hingga sangat tinggi. Hal ini menunjukkan bahwa bahasa, penyajian stimulus, serta penggunaan data, tabel, dan grafik dalam soal dapat dipahami dengan baik serta mendorong peserta didik untuk berpikir, menganalisis, dan menarik kesimpulan sesuai karakteristik keterampilan proses sains. Meskipun beberapa peserta didik masih mengalami kesulitan pada aspek waktu pengerjaan, kondisi tersebut menunjukkan bahwa instrumen mampu menantang kemampuan berpikir ilmiah tanpa mengurangi kemudahan penggunaannya.

Temuan kepraktisan ini memiliki relevansi penting dalam konteks pendidikan kimia. Instrumen yang dikembangkan dapat membantu guru memperoleh informasi yang lebih komprehensif mengenai penguasaan keterampilan proses sains peserta didik, sehingga penilaian tidak hanya berfokus pada aspek pengetahuan. Dengan tersedianya instrumen yang mudah diadministrasikan sekaligus teruji secara empiris, guru memiliki alat bantu konkret untuk mengintegrasikan penilaian keterampilan proses sains ke dalam praktik pembelajaran kimia sehari-hari, tanpa harus bergantung sepenuhnya pada kegiatan praktikum sebagai satu-satunya cara pengukuran. Secara keseluruhan, hasil penelitian menunjukkan bahwa instrumen yang dikembangkan telah memenuhi tiga karakteristik utama produk pengembangan, yaitu valid, reliabel, dan praktis. Dengan demikian, instrumen tes keterampilan proses sains pada materi laju reaksi layak digunakan sebagai alat penilaian dalam pembelajaran kimia dan berpotensi dimanfaatkan pada penelitian sejenis dengan karakteristik peserta didik yang serupa.

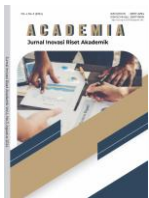
KESIMPULAN

Penelitian ini berhasil mengembangkan instrumen tes keterampilan proses sains (KPS) pada materi laju reaksi menggunakan model pengembangan Plomp melalui fase investigasi awal, desain, realisasi/konstruksi, tes, evaluasi dan revisi, serta implementasi. Produk yang dihasilkan berupa instrumen tes pilihan ganda sebanyak 25 butir yang disusun berdasarkan 12 indikator keterampilan proses sains, meliputi enam indikator KPS dasar dan enam indikator KPS terintegrasi. Hasil penelitian menunjukkan bahwa instrumen yang dikembangkan memenuhi tiga karakteristik utama produk pengembangan, yaitu valid, reliabel, dan praktis. Validitas teoretis yang tinggi serta hasil analisis Model Rasch yang memenuhi kriteria unidimensionalitas, item fit, dan distribusi tingkat kesukaran menunjukkan bahwa instrumen mampu mengukur konstruk keterampilan proses sains secara tepat dan memberikan informasi pengukuran yang akurat. Reliabilitas yang baik mengindikasikan bahwa instrumen menghasilkan pengukuran yang konsisten, sedangkan respons positif dari guru dan peserta didik menunjukkan bahwa instrumen mudah dipahami dan diterapkan dalam kegiatan penilaian di kelas.

Dengan demikian, instrumen yang dikembangkan tidak hanya layak digunakan sebagai alat ukur keterampilan proses sains pada materi laju reaksi, tetapi juga dapat menjadi alternatif instrumen penilaian yang mendukung implementasi pembelajaran kimia berbasis keterampilan proses sains. Ke depan, instrumen ini berpotensi dikembangkan pada materi kimia lainnya serta diuji pada jumlah responden dan karakteristik sekolah yang lebih beragam untuk memperoleh bukti empiris yang lebih luas mengenai kualitas dan generalisasi instrumen.

DAFTAR PUSTAKA

- Abdullah, M. (2022). Lembaga pendidikan sebagai suatu sistem sosial (Studi tentang peran lembaga pendidikan di Indonesia dalam perspektif Undang-Undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional). *Mamba'ul 'Ulum*, 18(1), 38–48. <https://doi.org/10.54090/mu.56>
- Allen, M. J., & Yen, W. M. (2002). *Introduction to measurement theory*. Waveland Press.
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Dordrecht, The Netherlands: Springer. <https://doi.org/10.1007/978-94-007-6857-4>
- Borg, W. R., & Gall, M. D. (1983). *Educational research: An introduction* (4th ed.). Longman.
- Chang, R., & Goldsby, K. (2016). *Chemistry* (12th ed.). McGraw-Hill Education.
- De Ayala, R. J. (2013). *The theory and practice of item response theory*. The Guilford Press.
- Harun, H., Prayitno, P., Sudaryanti, S., Rolina, N., & Manaf, A. (2023). Karakteristik butir instrumen asesmen dinamis capaian karakter anak usia dini: Analisis Rasch model. *Jurnal Penelitian Ilmu Pendidikan*, 16(2), 154–163. <https://doi.org/10.21831/jpipfip.v16i2.63358>
- Hasanah, H., & Aini, F. Q. (2025). Pengembangan instrumen tes keterampilan proses sains pada topik kesetimbangan kimia: Analisis dengan model Rasch. *JagoMIPA: Jurnal Pendidikan Matematika dan IPA*, 5, 68–82. <https://doi.org/10.53299/jagomipa.v5i1.1043>
- Juandi, D., Tamur, M., Adem, A. M. G., & Wijaya, T. T. (2023). Application of Rasch model in educational research: A systematic literature review. *European Journal of Educational Research*, 12(1), 515–530. <https://doi.org/10.12973/eu-jer.12.1.515>
- Lestari, N., Nisa, C., Mutazam, & Soeharto. (2023). Developing two tier test of critical thinking ability using Rasch model. *JIPF (Jurnal Ilmu Pendidikan Fisika)*, 8(2), 259–268. <https://doi.org/10.26737/jipf.v8i2.4029>
- Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16(2), 878. <https://www.rasch.org/rmt/rmt162f.htm>
- Magdalena, I., Syariah, E. N., Mahromiyati, M., & Nurkamilah, S. (2021). Analisis instrumen tes sebagai alat evaluasi pada mata pelajaran SBDP siswa kelas II SDN Duri Kosambi 06 Pagi. *Jurnal Pendidikan dan Ilmu Sosial*, 3(2), 276–287. <https://ejournal.stitpn.ac.id/index.php/nusantara/article/view/1343>
- Ndruru, E., & Sabani. (2021). Development of student learning outcomes test instruments of 11th grade high school on sound wave material. *Indonesian Physics Education Research*, 2(1), 1–6. <https://jurnal.uhn.ac.id/index.php/inspire/article/view/534>
- Nisfatulsanah, A., & Sugiharto, B. (2024). Comparing Classical Test Theory and Item Response Theory for respiratory system question instruments. *PIONIR: Jurnal Pendidikan*, 13(3). <https://doi.org/10.22373/pjp.v13i3.24043>



- Nitko, A. J., & Brookhart, S. M. (2014). *Educational assessment of students* (6th ed.). Pearson Education.
- Nuridin. (2007). *Model pembelajaran matematika yang menumbuhkan kemampuan metakognitif untuk menguasai bahan ajar*. UNESA. <https://repository.unesa.ac.id/>
- Nurpaizah, N., Maison, M., & Risnita, R. (2023). *Pengembangan Instrumen Penilaian pada Materi Sistem Pernapasan dengan Menggunakan Model Keterpaduan Nested (Keterpaduan) Kelas XI Sekolah Menengah Atas (SMA)*. *Jurnal Ilmiah Profesi Pendidikan*, 8(4), 2681–2686. <https://doi.org/10.29303/jipp.v8i4.1823>
- Plomp, T., & Nieveen, N. (Eds.). (2013). *Educational design research: Part A – An introduction*. Netherlands Institute for Curriculum Development (SLO).
- Qomariyah, R. S., Ghina, R., Putri, A., Putri, D. R., Putri, D. S., P, R. T., Pascasarjana, P., & Marga, U. P. (2022). Analisis Tingkat Kesukaran Dan Daya Pembeda Pada Butir Soal Pilihan Ganda Mata Pelajaran Bahasa Indonesia Kelas V Semester 1 SDN Kedungdalem 2. *Jurnal Pendidikan, Sains, Dan Teknologi*, 1(2), 74–80. <https://doi.org/10.47233/jpst.v1i2.278>
- Rahayu, A. (2025). Metode penelitian dan pengembangan (R&D): Pengertian, jenis dan tahapan. *DIAJAR: Jurnal Pendidikan dan Pembelajaran*, 4(3), 459–470. <https://doi.org/10.54259/diajar.v4i3.5092>
- Ramli, M., Saridewi, N., Budhi, T. M., & Suhendar, A. (2022). *Kimia untuk SMA/MA kelas XI*. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi. <https://static.buku.kemdikbud.go.id/content/pdf/bukuteks/kurikulum21/Kimia-BS-KLS-XI.pdf>
- Ramlawati, Sari, N. I., Kusumawati, R., Yesin, M., Ilmi, N., & Arsyad, A. A. (2025). The effect of differentiated science inquiry learning model based on teaching at the right level on students' critical thinking and science process skills. *Jurnal Pendidikan IPA Indonesia*, 14, 1–16. <https://doi.org/10.15294/jpii.v14i1.19479>
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Danish Institute for Educational Research.
- Rezba, R. J., Sprague, C., McDonough, J. T., & Matkins, J. J. (2007). *Learning and assessing science process skills* (5th ed.). Kendall/Hunt Publishing Company.
- Riani, N., & Ramalis, T. R. (2020). Pengembangan instrumen tes keterampilan proses sains berbentuk pilihan ganda pada pembelajaran fisika. *WaPFI (Wahana Pendidikan Fisika)*, 5(2), 131–136. <https://ejournal.upi.edu/index.php/WapFi/article/view/28173>
- Riduwan. (2012). *Belajar mudah penelitian untuk guru, karyawan, dan peneliti pemula*. Bandung: Alfabeta.
- Rohmatika, R., Fatmawati, S., & Suhartono. (2025). Analisis butir soal instrumen tes keterampilan berpikir kreatif pada materi energi terbarukan menggunakan Rasch model. *Jurnal EDUCATIO (Jurnal Pendidikan Indonesia)*, 11(1), 92–10. <https://doi.org/10.29210/1202525556>
- Rustaman, N. (2005). *Modul 1: Bekerja ilmiah*. Jakarta: Direktorat Tenaga Kependidikan, Departemen Pendidikan Nasional. <https://repository.upi.edu/>
- Safitri, I., Lestarani, D., Imtikhanah, R. D. N. W., Akbarini, N. R., Sari, M. W., Fitrah, M., Ilyas, I., Sista, T. R., Setyono, I. D., Setyaningtyas, R. F., Ndayizeye, O., & Hapsan, A. (2024). *Teori pengukuran dan evaluasi*. Padang: CV Ruang Tentor. <https://books.google.com/books?id=KxT7EQAAQBAJ>
- Sari, E. D. K., & Mahmudi, I. (2024). *Analisis pemodelan Rasch pada asesmen pendidikan (Analisis dengan menggunakan aplikasi Winsteps)*. PT. Pena Persada Kerta Utama.



- Subali, B. (2010). Pengembangan tes pengukur keterampilan proses sains pola divergen mata pelajaran biologi SMA. Dalam *Prosiding Seminar Nasional Biologi, Lingkungan, dan Pembelajarannya dalam Pengembangan Pendidikan Karakter* (hlm. 581–593). Universitas Negeri Yogyakarta. <https://staffnew.uny.ac.id/>
- Subhaktiyasa, P. G. (2024). Evaluasi validitas dan reliabilitas instrumen penelitian kuantitatif: Sebuah studi pustaka. *Journal of Education Research*, 5(5), 5599–5606. <https://doi.org/10.37985/jer.v5i5.1703>
- Sudikan, S. Y., Indarti, T., & Faizin. (2023). *Metode penelitian dan pengembangan (Research and Development) dalam pendidikan dan pembelajaran*. Malang: Universitas Muhammadiyah Malang Press. <https://ummpress.umm.ac.id/>
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan Rasch pada assessment pendidikan*. Trim Komunikata Publishing House.
- Taber, K. S. (2020). Conceptual confusion in the chemistry curriculum. *Foundations of Chemistry*, 22, 309–334. <https://doi.org/10.1007/s10698-019-09346-y>
- Tauhidah, D., & Rofi'ah, N. L. (2023). Validation of undergraduate science process skills tests: Rasch model analysis. *Research and Development in Education (RaDEn)*, 3(1), 51–57. <https://doi.org/10.22219/raden.v3i1.25917>
- Tauhidah, D., & Rofi'ah, N. L. (2023). Validation of undergraduate science process skills tests: Rasch model analysis. *Research and Development in Education (RaDEn)*, 3(1), 51–57. <https://doi.org/10.22219/raden.v3i1.14366>
- Tyas, E. H., Hamdu, G., & Pranata, O. H. (2020). Analisis soal pilihan ganda dengan menggunakan pemodelan Rasch untuk mengukur kemampuan siswa dalam mengurutkan bilangan pecahan di sekolah dasar. *PEDADIDAKTIKA: Jurnal Ilmiah Mahasiswa Pendidikan Guru Sekolah Dasar*, 7(2), 1–12. <https://ejournal.upi.edu/index.php/pedadidaktika/article/view/25344>
- Yanti, S. R., & Fuadiyah, S. (2025). *Validitas Instrumen Tes Keterampilan Proses Sains pada Materi Sistem Pernapasan Kelas XI SMA/MA*. Jurnal Pengabdian Masyarakat dan Riset Pendidikan (JERKIN), 3(4), 4590–4597. <https://doi.org/10.31004/jerkin.v3i4.1039>
- Zamsiswaya, Z., Syawaluddin, S., & Syahrizul, S. (2024). Pengembangan model ADDIE (Analysis, Design, Development, Implementation, Evaluation). *Jurnal Pendidikan Tambusai*, 8(3), 46363–46369. <https://jptam.org/index.php/jptam/article/view/17094>