



Analisis Kualitas Instrumen Tes Untuk Mengukur Pemahaman Konsep Siswa Pada Materi Laju Reaksi Dengan Pemodelan Rasch

Hariyati Y. Latif^{1*}, Lukman Abdul Rauf Laliyo¹, Erga Kurniawati¹

¹Universitas Negeri Gorontalo, Gorontalo 96554, Indonesia

*Corresponding author: hariyatilatif164@gmail.com

DOI: <https://doi.org/10.34312/je.v21i1.36133>

Abstrak

Penelitian ini bertujuan untuk menganalisis kualitas instrumen tes *Two-Tier Multiple Choice* yang digunakan untuk mengukur pemahaman konsep siswa pada materi laju reaksi dengan pemodelan Rasch. Instrumen terdiri dari 15 butir soal yang telah divalidasi oleh ahli dan diujicobakan kepada 100 siswa kelas XI SMA Negeri 1 Telaga. Analisis data dilakukan menggunakan perangkat lunak *Winsteps* untuk menilai validitas butir, unidimensionalitas, dan reliabilitas instrumen. Hasil penelitian menunjukkan bahwa 12 dari 15 butir soal memenuhi kriteria item fit (*OUTFIT MNSQ*, *ZSTD*, dan *PT-Measure Corr*), sedangkan tiga butir lainnya berada dalam batas toleransi sehingga masih dapat digunakan tanpa revisi besar. Uji unidimensionalitas menghasilkan nilai *Raw Variance Explained by Measures* sebesar 35,0%, yang memenuhi persyaratan minimal bahwa instrumen mengukur satu dimensi konstruk. Reliabilitas instrumen termasuk tinggi dengan item reliability sebesar 0,94, person reliability sebesar 0,73, dan *Cronbach's Alpha* sebesar 0,76. Secara keseluruhan, instrumen ini memiliki kualitas yang baik sehingga layak digunakan untuk mengevaluasi pemahaman konsep siswa pada materi laju reaksi.

Kata kunci: Pemahaman konsep; Rasch Model; Validitas butir; Reliabilitas

Abstract

This study aimed to analyze the quality of a two-tier multiple-choice test instrument designed to measure students' conceptual understanding of reaction rate concepts using the Rasch model. The instrument consisted of 15 items that were validated by experts and administered to 100 eleventh-grade students in public high schools 1 Telaga. Data were analyzed using Winsteps to evaluate item validity, unidimensionality, and instrument reliability. The results showed that 12 out of 15 items met all item-fit criteria (OUTFIT MNSQ, ZSTD, and PT-Measure Corr), while the remaining items failed to meet only one or two criterion but were still considered acceptable without major revisions. The unidimensionality test indicated that the Raw Variance Explained by Measures was 35,0%, which met the minimum threshold, confirming that the instrument measures a single underlying construct. The instrument demonstrated high reliability, with an item reliability of 0.94, person reliability of 0.73, and a Cronbach's Alpha value of 0.76. These findings indicate that the instrument has good measurement quality and is suitable for assessing students' conceptual understanding of reaction rate material.

Keywords: Conceptual understanding; Rasch Model; Item validity; Reliability

The format cites this article in APA style:

Latif, H. Y., Laliyo, L. A. R., & Kurniawati, E. (2026). Analisis Kualitas Instrumen Tes Untuk Mengukur Pemahaman Konsep Siswa Pada Materi Laju Reaksi Dengan Pemodelan Rasch. *Jurnal Entropi*, 21(1), 1-5. <https://doi.org/10.34312/je.v21i1.36133>

PENDAHULUAN

Kimia merupakan salah satu cabang sains yang mempelajari materi, sifat, perubahan, dan energi yang menyertainya,

serta menjadi mata pelajaran peminatan di jenjang MIPA. Pembelajaran kimia memiliki karakteristik yang kompleks karena melibatkan konsep abstrak dan konkret yang saling

berkaitan, sehingga penguasaan konsep dasar menjadi landasan penting sebelum mempelajari konsep yang lebih tinggi (Maghfiroh, dkk 2016). Namun, berbagai penelitian menunjukkan bahwa banyak siswa masih menganggap kimia sebagai mata pelajaran yang sulit dan kurang diminati (Hatimah, dkk 2021). Salah satu topik yang menuntut pemahaman tinggi adalah materi laju reaksi, yang berkaitan dengan perubahan konsentrasi zat terhadap waktu dan memiliki keterkaitan erat dengan fenomena alam maupun aplikasi industri.

Pemahaman konsep laju reaksi menuntut kemampuan menghubungkan level representasi makroskopik, submikroskopik, dan simbolik. Kesulitan mengintegrasikan ketiga representasi tersebut sering menyebabkan miskonsepsi siswa. Safitri dkk, (2019) menemukan bahwa sekitar 37% siswa belum mampu mengkoneksikan ketiga level representasi ini dengan baik. Kompleksitas materi laju reaksi juga disebabkan oleh keterkaitannya dengan konsep-konsep lain seperti teori tumbukan, energi aktivasi, orde reaksi, dan katalis (Arifanti, 2019). Kondisi ini menyebabkan rendahnya pemahaman siswa, yang berdampak pada capaian hasil belajar (Suteno dkk., 2021). Oleh karena itu, diperlukan strategi evaluasi yang mampu mengungkap pemahaman konseptual siswa secara mendalam, tidak hanya sekadar menilai jawaban benar-salah.

Instrumen *Two-Tier Multiple Choice* menjadi salah satu alternatif evaluasi yang mampu menilai dua aspek sekaligus, yaitu pilihan jawaban dan alasan yang mendasarinya. Instrumen ini memberikan informasi yang lebih komprehensif tentang kedalaman pemahaman siswa. Namun, instrumen yang baik memerlukan analisis kualitas butir yang dilakukan menggunakan pendekatan pengukuran yang tepat. Pendekatan yang selama ini umum digunakan adalah Classical Test Theory (CTT), yang menilai butir melalui tingkat kesukaran dan daya pembeda. Namun, karakteristik butir

dalam CTT cenderung tidak stabil dan sangat bergantung pada karakteristik peserta tes (Hardianti dkk., 2023), sehingga dapat memengaruhi akurasi interpretasi hasil tes.

Sebagai solusi, model Rasch hadir sebagai pendekatan modern yang mampu memberikan analisis lebih objektif karena mempertimbangkan hubungan antara kemampuan peserta dan karakteristik item dalam satu skala interval (Azizah & Wahyuningsih, 2020). Model ini mampu memetakan person dan item secara bersamaan serta memberikan prediksi yang valid meskipun terdapat data yang hilang (Sumintono & Widhiarso, 2015). Penelitian terbaru juga menunjukkan bahwa model Rasch efektif menganalisis pemahaman konsep secara lebih akurat (Widia Linta Nurjanah dkk, 2024) Berdasarkan urgensi tersebut, penelitian ini bertujuan untuk menganalisis kualitas instrumen two-tier pada materi laju reaksi menggunakan pemodelan Rasch.

METODE PENELITIAN

Jenis Penelitian

Jenis penelitian ini adalah penelitian kuantitatif dengan desain evaluatif, yang bertujuan menganalisis kualitas instrumen two-tier melalui pemodelan Rasch. Penelitian ini berfokus pada analisis karakteristik butir, reliabilitas, dan unidimensionalitas instrumen berdasarkan respons 100 siswa.

Waktu dan Tempat Penelitian

Penelitian dilaksanakan pada semester genap tahun ajaran 2024/2025, bulan Mei 2025. Lokasi penelitian di SMA Negeri 1 Telaga, Kabupaten Gorontalo, Provinsi Gorontalo.

Target/Subjek Penelitian

Target dalam penelitian ini adalah siswa kelas XI SMA Negeri 1 Telaga. Subjek penelitian terdiri atas kelas XI-1, XI-2, XI-3.

Prosedur

Prosedur penelitian ini meliputi tiga tahap utama. Pertama, peneliti menyusun kisi-kisi instrumen dan butir soal *Two-Tier Multiple Choice* berdasarkan kurikulum dan materi laju

reaksi. Kedua, instrumen divalidasi oleh tiga validator ahli untuk menilai aspek isi, konstruk, dan bahasa. Setelah dinyatakan layak, instrumen diujicobakan pada 100 siswa untuk memperoleh data empiris mengenai respons siswa. Ketiga, data hasil uji coba dianalisis menggunakan pemodelan Rasch melalui Winsteps untuk menilai validitas butir, unidimensionalitas, dan reliabilitas instrumen.

Instrumen Penelitian

Instrumen yang digunakan dalam penelitian ini adalah tes *Two-Tier Multiple Choice* sebanyak 15 butir soal.

Teknik Analisis Data

Data hasil tes direkap dalam Microsoft Excel sebagai data mentah, kemudian dianalisis menggunakan Model Rasch melalui perangkat lunak Winsteps versi 3.73.

HASIL DAN PEMBAHASAN

Hasil disajikan secara deskriptif melalui pengujian validitas dan reliabilitas.

a) Validitas Instrumen

Validitas butir dianalisis menggunakan Item Statistic Misfit Order pada Model Rasch dengan memeriksa tiga indikator utama, yaitu OUTFIT MNSQ, ZSTD, dan PT-Measure Correlation. Suatu butir dianggap fit apabila nilai OUTFIT MNSQ berada dalam rentang 0,5–1,5, nilai ZSTD berada di kisaran –2,0 hingga +2,0, dan nilai PT-Measure Corr bernilai positif serta minimal 0,20 (Sumintono & Widhiarso, 2015). Hasil analisis Misfit Order ditampilkan pada Tabel 1.

Tabel 1. Hasil Analisis Misfit Order

Item	OUTFIT		PT-MEASURE CORR
	MNSQ	ZSTD	
Soal 1	1,13	0,7	0,49
Soal 2	0,72	-2,0	0,32
Soal 3	0,79	-0,8	0,60
Soal 4	0,76	-0,9	0,59
Soal 5	0,97	-0,1	0,55
Soal 6	1,52	1,6	0,22
Soal 7	0,68	-1,6	0,62
Soal 8	0,79	-0,8	0,60
Soal 9	1,41	2,4	0,28
Soal 10	0,99	0,0	0,49
Soal 11	0,86	-0,8	0,56
Soal 12	0,82	-1,1	0,50

Item	OUTFIT		PT-MEASURE CORR
	MNSQ	ZSTD	
Soal 13	0,95	-0,3	0,55
Soal 14	1,25	1,3	0,44
Soal 15	1,25	1,5	0,43

Berdasarkan Tabel 1, 12 dari 15 butir soal (nomor 1, 3, 4, 5, 7, 8, 10, 11, 12, 13, 14, dan 15) memenuhi seluruh kriteria fit. Tiga butir lainnya (nomor 2, 6, dan 9) menunjukkan ketidaksesuaian pada satu atau dua indikator saja.

Namun, dalam pendekatan analisis Rasch, keputusan revisi tidak hanya didasarkan pada satu parameter yang menyimpang, melainkan pada kecenderungan pola keseluruhan butir (Bond & Fox, 2015). Ketiga butir tersebut masih berada dalam rentang yang dapat diterima dan tidak menunjukkan misfit ekstrem.

Dengan demikian, dapat disimpulkan bahwa seluruh butir tetap dapat digunakan, dan instrumen secara umum valid untuk mengukur pemahaman konsep siswa pada materi laju reaksi.

b) Unidimensionality

Uji unidimensionalitas digunakan untuk memastikan bahwa instrumen hanya mengukur satu konstruk, yaitu pemahaman konsep siswa pada materi laju reaksi. Instrumen dinyatakan unidimensional apabila nilai Raw Variance Explained by Measures dan proporsi Unexplained Variance memenuhi kriteria yang direkomendasikan dalam analisis komponen utama (PCA). Hasil uji unidimensionalitas instrumen ditampilkan pada Tabel 2.

Tabel 2. Item Unidimensionality

	Empirical	Modeled
Total raw variance in observations	23,1	100,0%
Raw Variance explained by measures	8,1	35,0%
Raw Variance explained by persons	3,0	13,2%
Raw Variance explained by items	5,0	21,7%

	Empirical		Modeled	
Raw unexplained variance (total)	15.0	65,0%	100.0%	65,2%
Unexplned variance in 1st contrast	3,2	14,0%	21,5%	
Unexplned variance in 2nd contrast	1.9	8,1%	12,5%	

Nilai *Raw Variance Explained by Measures* sebesar 35,0%, yang sedikit di atas nilai ekspektasi model (34,8%). Nilai ini telah memenuhi batas minimal >20% sehingga instrumen dinilai memenuhi kriteria unidimensionalitas. Kategori penilaian yang umum digunakan adalah: >20% (diterima), >40% (baik), dan >60% (sangat baik) (Fadhilah et al., 2024).

Nilai Unexplained Variance pada 1st contrast sebesar 14,0% dan pada 2nd contrast sebesar 8,1%, dengan rata-rata 11,05%. Persentase ini menunjukkan bahwa keberadaan dimensi kedua atau ketiga cukup rendah, sehingga item-item dalam instrumen menunjukkan independensi yang memadai dan tidak membentuk subdimensi yang mengganggu konstruk utama.

c) Reliabilitas Instrumen

Reliabilitas instrumen dianalisis menggunakan Rasch Model, yang ditinjau melalui Summary Statistics sebagaimana disajikan pada Tabel 3.

Tabel 3. Summary Statistic

	Responden(N=100)	Item (N=15)
Mean	0,28	0.00
Standar Defiasi	0.77	0.60
Reliability	0.73	0.94
Separation Indeks (Reliability)	1.66	3,95
Cronbach Alpha (KR-20) = 0.76		
Chi-Square : 2383.78 with 1384d.f. p=.0000		

Person reliability sebesar 0.73 berada pada kategori cukup, sedangkan item reliability sebesar 0.94 termasuk kategori bagus sekali

(Sumintono, 2015). Temuan ini menunjukkan bahwa konsistensi respons siswa berada pada tingkat memadai, dan kualitas butir soal dari sisi reliabilitas sangat baik.

Nilai Cronbach's Alpha = 0.76 mengindikasikan konsistensi internal instrumen yang kuat, sehingga interaksi antara responden dan item dinilai stabil.

Sementara itu, nilai Separation Index untuk responden (1.66) termasuk kategori cukup, sedangkan untuk item (3.95) termasuk kategori baik. Tingginya nilai separation index pada item menunjukkan bahwa instrumen mampu membedakan level kesulitan butir secara jelas, sehingga meningkatkan kualitas pengukuran secara keseluruhan (Sumintono & Widhiarso, 2015).

KESIMPULAN

Berdasarkan analisis kualitas instrumen dua tingkat (*Two-Tier Multiple Choice*) menggunakan Model Rasch, dapat disimpulkan bahwa instrumen yang dikembangkan memiliki karakteristik psikometrik yang baik untuk mengukur pemahaman konsep siswa pada materi laju reaksi. Dari 15 butir soal yang dianalisis, 12 butir memenuhi seluruh kriteria item fit, sedangkan tiga butir lainnya tetap berada dalam rentang kelayakan berdasarkan evaluasi holistik, sehingga seluruh butir dapat dipertahankan tanpa revisi.

Uji unidimensionalitas menunjukkan bahwa instrumen mampu mengukur satu konstruk secara konsisten, dengan nilai Raw Variance Explained by Measures berada di atas standar minimal yang direkomendasikan. Reliabilitas instrumen juga berada pada kategori tinggi, tercermin dari nilai item reliability sebesar 0,94, person reliability sebesar 0,73, serta Cronbach's Alpha sebesar 0,76.

Dengan demikian, instrumen *Two-Tier Multiple Choice* yang dikembangkan dapat dinyatakan valid, reliabel, dan layak digunakan sebagai alat evaluasi untuk mengidentifikasi

tingkat pemahaman konseptual siswa terkait laju reaksi.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih kepada para dosen pembimbing atas arahan, bimbingan, dan masukan yang sangat berharga selama proses penyusunan dan pelaksanaan penelitian ini. Ucapan terima kasih juga disampaikan kepada kepala sekolah, guru, serta siswa SMAN 1 Telaga yang telah memberikan kesempatan, dukungan, dan partisipasi sehingga penelitian ini dapat terlaksana dengan baik.

DAFTAR PUSTAKA

- Arifanti Deviia Resma. (2019). *Analisis Pemahaman materi laju reaksi berdasarkan Knoeledge-Space Theory (KST) pada siswa kelas XI SMA Nasional Malang*. Universitas Negeri Malang.
- Azizah, & Wahyuningsih, S. (2020). Use of Rasch Model for Test Instrument Analysis in Actuarial Mathematics Course. *JUPITEK: Jurnal Pendidikan Matematika*, 3(1), 45–50.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch Model: Fundamental Measurement in the HumBond, T. G., & Fox, C. M. (2007). Applying the Rasch Model: Fundamental Measurement in the Human Sciences Second Edition University of Toledo.an Sciences Second Edition University of Toledo*.
- 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>
- Hatimah, H., Khery, Y., Studi, P., & Kimia, P. (2021). *Pemahaman Konsep dan Literasi Sains dalam Penerapan Media Pembelajaran Kimia Berbasis Android*. 8(1).
- Maghfiroh, N., Susilo, H., & Gofur, A. (2016). Pengaruh Project Based Learning Terhadap Kerampilan Proses Sains Siswa Kelas X SMA Negeri Sidoarjo. *Jurnal Pendidikan*, 1(8), 1588–1593.
- Safitri, N. C., Nursaadah, E., & Wijayanti, I. E. (2019). Analisis Multipel Representasi Kimia Siswa pada Konsep Laju Reaksi. *EduChemia (Jurnal Kimia Dan Pendidikan)*, 4(1), 1. <https://doi.org/10.30870/educhemia.v4i1.5023>
- Sumintono, B., & Widhiarso, W. (2013). *Aplikasi Model Rasch Untuk Penelitian Ilmu-Ilmu Sosial*.
- Sumintono, B., & Widhiarso, W. (2015). Aplikasi Pemodelan Rasch Pada Assessment Pendidikan. *Aplikasi Rasch Pemodelan Pada Assessment Pendidikan, September*, 1–24.
- Suteno, I. K., Laliyo*, L. A. R., Iyabu, H., & Abdullah, R. (2021). Mengevaluasi Level Pemahaman Konsep Hidrolisis Garam Peserta Didik Menggunakan Tes Diagnostik Pilihan Ganda Empat Tingkat. *Jurnal Pendidikan Sains Indonesia*, 9(3), 482–497. <https://doi.org/10.24815/jpsi.v9i3.20543>
- Widia Linta Nurjanah, Ika Mustika Sari, & Duden Saepuzaman. (2024). Analisis Pemahaman Konsep Peserta Didik Pada Topik Perambatan Kalor Menggunakan Rasch Model. *Didaktika: Jurnal Kependidikan*, 13(2), 1361–1374. <https://doi.org/10.58230/27454312.442>