



Research Article



Analisis Instrumen tes Pengambilan Keputusan Taksonomi Presseisen dan Abilitas Peserta Didik Menggunakan Model Rasch

(*Analysis of the Presseisen Taxonomy Decision-Making Test Instrument and Learners' Abilities Using the Rasch Model*)

Asih Triyanti, Anisyah Yuniarti*, Afandi

Program Studi Pendidikan Biologi, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Tanjungpura
Jl. Profesor Dokter H. Hadari Nawawi, Kec. Pontianak Tenggara, Kota Pontianak, Kalimantan Barat 78115

*Corresponding author: anisyah.yuniarti@fkip.untan.ac.id

Informasi Artikel	ABSTRACT
Submit: 29 – 08 – 2024 Diterima: 12 – 11 – 2024 Dipublikasikan: 05 – 03 – 2025	<p><i>This study aims to determine the quality of the Presseisen taxonomy decision-making test instrument and students' abilities using Rasch analysis. This research method is descriptive with a quantitative approach. The results showed item validity (item fit) and 12 valid questions. The question's difficulty level (item measure) shows that three questions are very difficult, three are complex, four are easy, and two are straightforward. Reliability shows that the Cronbach alpha value is excellent (0.89), the person reliability value is good (0.85), the item reliability value is excellent (0.98)—the person separation value is 2.41, and the item separation value is 6.85. The individual ability level shows a high category of 124 people, a medium of 49 people, and a low of 190 people. The level of individual suitability shows the fit criteria totalling 237 people, misfit 12 people, and outliers 114 people. Based on the analysis results, it can be concluded that the questions are valid and suitable for use.</i></p> <p>Keywords: Test Instrument, Decision Making, Presseisen Taxonomy, Rasch Model</p>
Penerbit	ABSTRAK
Program Studi Pendidikan Biologi FKIP Universitas Jambi, Jambi- Indonesia	<p>Penelitian ini bertujuan untuk mengetahui kualitas instrumen tes pengambilan keputusan taksonomi Presseisen dan abilitas peserta didik menggunakan analisis rasch. Metode penelitian ini deskriptif dengan pedekatan kuantitatif. Hasil penelitian menunjukkan Validitas item (<i>item fit</i>) menunjukkan 12 soal valid. Tingkat kesulitan soal (<i>item measure</i>) menunjukkan 3 soal sangat sulit, 3 soal sulit, 4 soal mudah, dan 2 soal sangat mudah. Reliabilitas menunjukkan nilai <i>alpha cronbach</i> bagus sekali (0.89), nilai <i>person reliability</i> bagus (0.85), dan nilai <i>item reliability</i> istimewa (0.98). nilai <i>person separation</i> 2.41 dan nilai <i>item separation</i> 6.85. Tingkat abilitas individu menunjukkan kategori tinggi berjumlah 124 orang, sedang 49 orang, dan rendah 190 orang. Tingkat kesesuaian individu menunjukkan kriteria fit berjumlah 237 orang, misfit 12 orang, dan outlier 114 orang. Berdasarkan hasil analisis, dapat disimpulkan bahwa soal dinyatakan valid dan layak untuk digunakan.</p> <p>Kata kunci: Instrumen Tes, Pengambilan Keputusan, Taksonomi Presseisen, Model Rasch</p>



This Biodik: Jurnal Ilmiah Pendidikan Biologi is licensed under a [CC BY-NC-SA \(Creative Commons Attribution-ShareAlike 4.0 International License\)](https://creativecommons.org/licenses/by-nc-sa/4.0/)

PENDAHULUAN

Instrumen tes adalah suatu alat ukur yang dapat digunakan untuk mengukur tingkat kemampuan berpikir peserta didik (Lestari & Setyarsih, 2020). Kualitas instrumen tes sangat berpengaruh terhadap akurasi pengukuran, validitas, reliabilitas, objektivitas, dan hasil belajar peserta didik. Menurut Agustini, dkk (2018) mencatat bahwa perancangan instrumen tes /penilaian yang baik dan sesuai dengan tingkat kemampuan berpikir dapat meningkatkan kemampuan berpikir peserta didik. Sejalan dengan pendapat Sutami (2021) bahwa untuk mencapai tujuan pembelajaran IPA secara efektif, diperlukan penggunaan instrumen tes yang dapat melatih keterampilan berpikir. Oleh karena itu, penting untuk selalu memastikan bahwa instrumen tes yang digunakan memenuhi standar kualitas yang tinggi. Sejalan dengan pendapat Magdalena dkk. (2021) menyatakan bahwa penggunaan instrumen tes yang digunakan oleh guru harus memenuhi standar kualitas ujian, yaitu validitas, reliabilitas, objektivitas, praktisitas, dan hemat biaya.

Salah satu cara untuk mengukur kemampuan pengambilan keputusan peserta didik yaitu dengan cara membuat instrumen tes yang mengarah pada kemampuan pengambilan keputusan. Instrumen tes pengambilan keputusan dibuat untuk mengetahui dan mengukur kemampuan peserta didik dalam berbagai aspek, baik akademik maupun non-akademik. Pembuatan instrumen tes untuk mengukur kemampuan pengambilan keputusan peserta didik masih jarang digunakan. Pendidik hanya membuat instrumen tes untuk mengukur pemahaman mereka saja. Padahal instrumen tes pengambilan keputusan sangat penting dalam proses pembelajaran karena dapat membantu mengevaluasi kemampuan siswa dalam berpikir kritis, memecahkan masalah, dan mengambil tanggung jawab atas keputusan mereka. Sejalan dengan Fitri (2023) menyatakan bahwa pengambilan keputusan yang efektif membutuhkan kemampuan berpikir kritis, kreatif dan dapat memecahkan suatu permasalahan.

Pengambilan keputusan adalah kemampuan berpikir seseorang untuk mengevaluasi berbagai pilihan yang ada. Menurut Preisseisen (1985) mengatakan bahwa pengambilan keputusan adalah kemampuan seseorang untuk memikirkan dan memilih pilihan terbaik dari banyak pilihan yang tersedia dengan mempertimbangkan manfaat dan risiko. Peserta didik yang memiliki kemampuan pengambilan keputusan yang baik dapat berdampak positif bagi dirinya sendiri maupun lingkungannya seperti, dapat menyelesaikan masalah dengan lebih efektif, memiliki hubungan sosial yang sehat, berani mengambil tanggung jawab, dapat memperkuat keterampilan berpikir kritis dan pemecahan masalah. Sedangkan peserta didik yang memiliki kemampuan pengambilan keputusan yang buruk lebih rentan terhadap kegagalan dan masalah di berbagai aspek kehidupan. Kemampuan pengambilan keputusan peserta didik masih tergolong rendah, ini ditunjukkan oleh beberapa siswa di salah satu SMA Sidoarjo mengenai kesulitan dalam pengambilan keputusan (Rahmasari, Noviekayati, & Pratitis, 2023).

Kemampuan pengambilan keputusan penting dilakukan untuk mengembangkan kemampuan mereka dalam memutuskan kehidupan dengan lebih baik dan bertanggung jawab. Sejalan dengan Maulana & Rochintaniawati (2021) menyatakan bahwa Keterampilan pengambilan keputusan sangat penting dalam pembelajaran IPA, karena jika siswa mengalami atau menghadapi masalah dalam hidup mereka, mereka dapat menggunakan berbagai solusi untuk membuat keputusan yang tepat. Kemampuan pengambilan keputusan seharusnya sudah dilatih sejak dini, agar siswa tidak salah arah dalam menentukan kehidupannya kedepan. Sejalan dengan Dianty, Supeno, dan Astutik (2020),

keputusan yang dibuat oleh siswa dapat berpengaruh pada kesejahteraan, dan kesuksesan mereka di masa depan. Proses pengambilan keputusan peserta didik memiliki beberapa aspek pengukuran yang mengacu pada taksonomi Presseisen.

Taksonomi Presseisen adalah salah satu taksonomi yang membahas tingkatan berpikir mulai dari dasar sampai berpikir kompleks. Salah satu keterampilan berpikir dalam taksonomi Presseisen adalah berpikir kompleks aspek pengambilan keputusan. Pengambilan keputusan adalah proses berpikir yang sesuai dengan perkembangan abad 21. Hal ini sejalan dengan Sanjaya, dkk (2020) bahwa Kemampuan pengambilan keputusan dalam proses pembelajaran sesuai dengan pergeseran paradigma belajar abad 21. Kemampuan pengambilan keputusan seharusnya sudah dilatih sejak dini, agar siswa tidak salah arah dalam menentukan kehidupannya kedepan. Menurut Dianty, Supeno, dan Astutik (2020), keputusan yang dibuat oleh siswa dapat berpengaruh pada kesejahteraan, hubungan, dan kesuksesan mereka di masa yang akan datang. Pengambilan keputusan menurut taksonomi Presseisen mencakup 6 indikator yaitu menentukan tujuan, mengidentifikasi hambatan untuk mencapai tujuan, mengidentifikasi alternatif, menganalisis alternatif, memberik peringkat alternatif, dan membuat alternatif terbaik (Presseisen 1985).

Model Rasch adalah salah satu model dalam teori respons butir (IRT) yang dapat digunakan untuk menganalisis hasil tes. Model teori respons butir dirancang untuk mengukur keterampilan atau kemampuan individu dengan mempertimbangkan tingkat kesulitan item tes dan kemampuan peserta didik (Halim, dkk 2014). Model Rasch memproses skor mentah untuk memberikan informasi yang lebih akurat dan jelas tentang kemampuan peserta didik dan kualitas tes. Analisis model rasch pada penelitian ini dilihat dari kualitas tes yaitu analisis tingkat kesesuaian butir soal (*item fit*), reliabilitas (*alpha Cronbach*, *person reliability*, dan *item reliability*), tingkat kesukaran butir soal (*item measure*) dan daya pembeda (*separation*). Selain itu, dilihat dari kemampuan peserta didik yaitu hasil analisis tingkat abilitas individu (*person measure*) dan tingkat kesesuaian individu (*person fit*). Ada beberapa kelebihan dari analisis model rasch yaitu, dapat memberikan ukuran linier dengan interval yang sama, kemampuan untuk mengatasi data yang hilang, kemampuan untuk menghasilkan estimasi yang lebih tepat, kemampuan untuk menemukan ketidaktepatan model, dan kemampuan untuk menyediakan instrumen pengukuran yang independen dari parameter yang diteliti (Sumintono & Widhiarso, 2015).

Berdasarkan permasalahan yang telah dijelaskan di atas, peneliti tertarik untuk melakukan analisis instrumen tes pengambilan keputusan menggunakan analisis model rasch. Penelitian ini bertujuan untuk mengetahui kualitas instrumen tes pengambilan keputusan taksonomi Presseisen dan abilitas peserta didik menggunakan analisis rasch. Analisis instrumen tes ini diharapkan dapat memberikan acuan kepada guru untuk menganalisis kualitas instrumen dan melihat kemampuan peserta didik, sehingga guru dapat menyusun instrumen tes yang lebih baik. Penggunaan model rasch juga dapat menjadi alternatif bagi guru untuk melakukan analisis instrumen tes atau melihat kemampuan peserta didik.

METODE PENELITIAN

Metode yang diterapkan pada penelitian ini menggunakan pendekatan deskriptif kuantitatif. Jenis penelitian yang digunakan R&D (*Research and Development*) dengan menggunakan model pengembangan 4D yang dimodifikasi menjadi 3D (*Define, Design, Develop*). Modifikasi model 4-D menjadi 3-D dilakukan karena adanya keterbatasan waktu, sumber dana, dan perangkat komputasi yang kurang memadai. Pengembangan instrumen tes mengacu pada Mardapi (2008), yang terdiri dari 6 tahapan: menyusun spesifikasi tes, menulis tes, menelaah tes, melakukan uji coba, menganalisis bagian

tes, dan memperbaiki tes. Populasi pada penelitian ini adalah seluruh peserta didik kelas X di SMA Negeri Pontianak tahun ajaran 2023/2024 sebanyak 3.946 orang. Kemudian dilakukan perhitungan jumlah peserta didik menggunakan rumus slovin Menurut Riyanto & Hatmawan (2020) sebagai berikut:

$$n = \frac{N}{N(e)^2 + 1}$$

Keterangan:

n= Jumlah sampel

N= Total populasi

e = Tingkat kesalahan dalam pengambilan sampel

Penelitian ini menggunakan tingkat kepercayaan sebesar 95% dan kesalahan sebesar 5%, diperoleh sampel 363 peserta didik. Pengambilan sampel dilakukan menggunakan teknik *cluster sampling* (area sampling) dan sekolah diambil secara acak (undi). Setiap Kecamatan yang di Pontianak mewakili satu SMA Negeri di Pontianak dan diperoleh 5 sekolah yaitu SMAN 3, SMAN 9, SMAN 5, SMAN 11, dan SMAN 8 Pontianak. Berikut hasil sampel dari *cluster sampling* dapat dilihat pada tabel 1.

Tabel 1. Teknik Pengambilan Sampel secara Cluster Sampling

Cluster	Sekolah	Banyak Sampel
Kec. Pontianak Barat	SMA Negeri 11 Pontianak	73
Kec. Pontianak Timur	SMA Negeri 9 Pontianak	72
Kec. Pontianak Kota	SMA Negeri 8 Pontianak	73
Kec. Pontianak Utara	SMA Negeri 5 Pontianak	72
Kec. Pontianak Selatan	SMA Negeri 3 Pontianak	73
Kec. Pontianak Tenggara	-	-
Total Seluruh Sampel		363

Instrumen tes yang disusun berbentuk soal uraian berjumlah 12 butir soal yang mengacu pada indikator pengambilan keputusan menurut taksonomi Presseisen. Instrumen tes yang telah disusun divalidasi oleh 7 orang ahli yaitu 2 orang dosen Pendidikan Biologi FKIP UNTAN, dan 5 orang guru biologi SMA. Hasil analisis validitas isi menggunakan rumus Aiken V dengan skala 4 yaitu 0,94. Kriteria minimum Aiken's V skala 4 dengan 7 rater yaitu 0.76. Maka dapat dilihat dari hasil tersebut nilai V Aiken(hitung) > Aiken(tabel). Analisis reliabilitas inter rater menggunakan aplikasi SPSS versi 29 melalui *intarclass correlation coefficient* (ICC) ditunjukkan pada nilai *average measure* yaitu sebesar 0,806 termasuk ke dalam kategori reliabilitas bagus. Pengumpulan data dilakukan dengan menguji cobakan instrumen tes kepada sampel yang telah ditentukan. Hasil jawaban peserta didik adalah data kuantitatif berupa data mentah yang dianalisis menggunakan model rasch dengan bantuan software winstep versi 3.73.

Analisis model rasch pada penelitian yaitu analisis tingkat kesesuaian butir soal (*item fit*), tingkat kesukaran butir soal (*item measure*), reliabilitas (*alpha Cronbach*, *person reliability*, dan *item reliability*), daya pembeda (*separation*), hasil tingkat kesesuaian individu (*person fit*) dan tingkat abilitas individu (*person measure*). Analisis model Rasch dapat menunjukkan adanya individu yang memiliki pola respon tidak sesuai dan butir soal yang tidak valid (outliers atau misfits) (Aprilia dkk, 2020). Pemodelan Rasch memiliki beberapa keuntungan. Ini memungkinkan untuk menemukan jawaban yang salah, menemukan penilaian yang salah, dan membuat prediksi terhadap data yang hilang berdasarkan pola respon yang sistematis (Tyas et al., 2020). Tujuan penelitian yang menggunakan Model Rasch adalah untuk mengetahui kualitas soal dan kemampuan siswa.

HASIL PENELITIAN DAN PEMBAHASAN

Instrumen tes yang dikembangkan sebanyak 12 butir soal uraian yang mengacu pada indikator pengambilan keputusan menurut taksonomi Preseisen. Setiap indkiator pengambilan kputusan yang terdiri dari 6 indikator dibuat masing-masing dua butir soal. Kemudian soal diuji cobakan kepada 363 peserta didik dengan waktu pengerjaan 90 menit. Hasil jawaban peserta didik dilakukan peskoran sesuai dengan rubrik pesnkoran soal tes, kemudian, diperoleh data mentah yang akan dianalisis menggunakan model rasch melalui software winstep versi 3.73. Hasil analisis rasch dapat dilihat sebagai berikut.

1. Analisis Tingkat Kesesuaian Butir Soal (*Item Fit*)

Analisis tingkat kesesuaian butir soal bertujuan untuk melihat butir soal yang dapat diterima dan dibuang. Jika item soal yang dinyatakan memenuhi persyaratan berikut, maka soal dapat digunakan untuk pengukuran. Berikut kriteria menurut Sumintono & Widhiarso, (2015:72).

- Nilai Outfit MNSQ diterima adalah $0,5 < MNSQ < 1,5$;
- Nilai Outfit ZSTD diterima adalah $-2,0 < ZTSD < +2,0$;
- Nilai Corrsasi Rata-rata Pt adalah $0,4 < Corrsasi Rata-rata Pt < 0,85$.

Jika butir soal tidak memenuhi ketiga kriteria di atas, maka soal dinyatakan tidak valid atau tidak berfungsi normal untuk melakukan pengukuran. Sejalan dengan pendapat Purniasari, dkk (2021) menyatakan bahwa Soal dianggap valid jika memenuhi minimal salah satu dari ketiga kriteria, dan dibuang jika tidak memenuhi kriteria. Hasil analisis *item fit* dapat dilihat pada gambar 1.

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S. E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PT-MEASURE CORR	ASASURE EXP.	EXACT OBS%	MATCH EXP%	Item
8	1026	363	.04	.08	1.85	9.31	1.79	9.31	A .65	.65	32.0	56.0	E8
1	1099	358	-.58	.09	1.11	1.1	1.24	3.0	B .4	.64	53.4	58.2	E1
9	1024	358	-.03	.08	1.09	1.1	1.07	.9	C .7	.65	55.9	56.2	E9
12	898	352	.74	.08	1.09	1.2	1.07	1.0	D .6	.65	50.0	54.9	E12
2	1117	358	-.71	.09	.95	-.6	.97	-.3	E .5	.64	60.3	59.0	E2
11	940	360	.57	.08	.94	-.9	.92	-1.1	F .6	.65	60.6	55.1	E11
10	927	357	.61	.08	.93	-1.0	.91	-1.3	f .6	.65	58.8	55.0	E10
5	1000	360	.17	.08	.87	-1.0	.86	-2.1	g .7	.65	60.6	55.5	E5
7	1103	363	-.51	.09	.84	-2.1	.85	-2.2	d .6	.64	61.4	58.0	E7
3	1180	363	-1.10	.09	.84	-2.2	.79	-2.1	c .6	.62	68.6	61.1	E3
4	1035	358	-.12	.08	.74	-4.0	.73	-4.2	b .7	.65	64.0	56.2	E4
6	860	347	.94	.08	.66	-5.1	.65	-5.1	a .6	.64	70.6	55.4	E6
MEAN	1017.4	358.1	.00	.08	.99	-.4	.99	-.4			58.0	56.7	
S.D.	93.2	4.5	.61	.08	.29	3.6	.29	3.6			9.6	1.9	

Gambar 1. Output Item Fit

Berdasarkan Gambar 1, butir soal pengambilan keputusan yang dianalisis diterima semua atau valid. Jika diubah dalam bentuk persen diperoleh sebanyak 100% soal diterima atau *fit*. Namun pada salah satu soal yaitu soal nomor delapan memiliki nilai *outfit* MNSQ dan *outfit* ZSTD tidak memenuhi kriteria *item fit* tapi pada nilai *point measure correlation* memenuhi kriteria *item fit*, sehingga soal nomor delapan tetap dipertahankan. Suatu butir soal dinyatakan valid jika memenuhi satu dari tiga syarat kriteria *item fit* (Maulana, dkk 2023). Instrumen tes atau butir soal dikatakan valid ketika soal yang digunakan tidak mengandung bias (Kurniawan & Andriyani, 2018).

Deteksi butir soal bias atau *differential item functioning* (DIF) adalah suatu item yang ada pada tes memiliki keberfungsian yang berbeda. Butir soal dianggap bias jika nilai probabilitas (PROB) untuk butir soal tersebut kurang dari 0,05 (5%). (Sumintono & Widhiarso, 2015). Hasil analisis deteksi soal bias dapat dilihat pada gambar 2.

▲TABLE 30.4 C:\Users\USER\OneDrive\Desktop\DATA H ZOU001WS.TXT Aug 10 20:21 2024
INPUT: 363 Person 12 Item REPORTED: 363 Person 12 Item 4 CATS WINSTEPS 3.73

DIF class specification is: DIF=\$S1W1

Person CLASSES	SUMMARY DIF		BETWEEN-CLASS			Item	
	CHI-SQUARE	D.F.	PROB.	MEAN-SQUARE	t=ZSTD	Number	Name
2	.0000	1	1.0000	.0038	-1.3183	1	E1
2	.0000	1	1.0000	.0011	-1.4328	2	E2
2	.0403	1	.8408	.0299	-.9912	3	E3
2	.0000	1	1.0000	.0100	-1.1925	4	E4
2	.4439	1	.5053	.2206	-.3681	5	E5
2	.2187	1	.6400	.1085	-.6383	6	E6
2	.0820	1	.7747	.0414	-.9161	7	E7
2	.2581	1	.6114	.1282	-.5803	8	E8
2	2.2336	1	.1350	1.1163	.5507	9	E9
2	3.0874	1	.0789	1.5419	-.8008	10	E10
2	.5099	1	.4752	.2542	-.3061	11	E11
2	.0000	1	1.0000	.0034	-1.3309	12	E12

Gambar 2. Output Deteksi soal bias

Berdasarkan Gambar 2 butir soal tidak ada mengandung bias, nilai probabilitas pada 12 soal uraian diatas rata-rata nilai probabilitas yang telah ditetapkan yaitu 5% (0,05). Instrumen tes dikatakan valid jika butir soal yang dikembangkan tidak mengandung bias (Sumintono & Widhiarso, 2015). Sependapat dengan Kurniawan & Andriyani (2018) bahwa Instrumen tes atau butir soal dikatakan valid ketika soal yang digunakan bebas dari bias.

2. Analisis Tingkat Kesukaran Butir Soal (Item Measure)

Analisis butir soal ditinjau dari tingkat kesukaran bertujuan untuk mengetahui tingkat kesukaran soal yang akan diberikan dalam ujian. Hal ini menunjukkan semakin banyak siswa yang menjawab soal dengan benar, maka semakin mudah soal tersebut. Pengelompokan butir soal dilakukan menggunakan nilai standar deviasi (SD) dan nilai rata-rata *logit*. Nilai standar deviasi (SD) yaitu 0.61, dan nilai rata-rata *logit* yaitu 0.00. Berdasarkan nilai tersebut kategori sangat sulit > 0.61, kategori sulit >0.00 hingga <0.61, kategori mudah >-0.61 hingga <0.00, dan kategori sangat mudah <-0.61. Hasil *analisis item measure* dapat dilihat pada gambar 3.

13-465WS - Notepad

File Edit Format View Help

TABLE 13.1 C:\Users\USER\OneDrive\Desktop\DATA H ZOU465WS.TXT Jul 29 3:01 2024
INPUT: 363 Person 12 Item REPORTED: 363 Person 12 Item 4 CATS WINSTEPS 3.73

Person: REAL SEP.: 2.41 REL.: .85 ... Item: REAL SEP.: 6.85 REL.: .98

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	OUTFIT ZSTD	PT-MEASURE MNSQ	EXACT ZSTD	MATCH CORR.	EXP. EXP%	Item		
6	860	347	.94	.08	.66	-5.4	.65	-5.4	.69	.64	70.6	55.4	E6
12	898	352	.74	.08	1.09	1.2	1.07	1.0	.68	.65	50.0	54.9	E12
10	927	357	.61	.08	.93	-1.0	.91	-1.3	.67	.65	58.8	55.0	E10
11	940	360	.57	.08	.94	-.9	.92	-1.1	.66	.65	60.6	55.1	E11
5	1000	360	.17	.08	.87	-1.9	.86	-2.1	.70	.65	60.6	55.5	E5
8	1026	363	.04	.08	1.85	9.6	1.79	9.1	.61	.65	32.0	56.0	E8
9	1024	358	-.03	.08	1.09	1.3	1.07	.9	.71	.65	55.9	56.2	E9
4	1035	358	-.12	.08	.74	-4.0	.73	-4.2	.73	.65	64.0	56.2	E4
7	1103	363	-.51	.09	.84	-2.3	.85	-2.2	.67	.64	61.4	58.0	E7
1	1099	358	-.58	.09	1.11	1.5	1.24	3.0	.41	.64	53.4	58.2	E1
2	1117	358	-.71	.09	.95	-.6	.97	-.3	.57	.64	60.3	59.0	E2
3	1180	363	-1.10	.09	.84	-2.2	.79	-2.7	.66	.62	68.6	61.1	E3
MEAN	1017.4	358.1	.00	.08	.99	-.4	.99	-.4			58.0	56.7	
S.D.	93.2	4.5	.61	.00	.29	3.6	.29	3.6			9.6	1.9	

▲TABLE 13.3 C:\Users\USER\OneDrive\Desktop\DATA H ZOU465WS.TXT Jul 29 3:01 2024
INPUT: 363 Person 12 Item REPORTED: 363 Person 12 Item 4 CATS WINSTEPS 3.73

Gambar 3. Output item measure

Berdasarkan Gambar 3 analisis tingkat kesukaran soal yang dilakukan terhadap 12 soal (E adalah soal) uraian tersebut dapat diketahui bahwa 3 butir soal (E6,E10,E12) termasuk kategori sangat sulit dengan persentase (25%), 3 butir soal (E5,E8,E11) termasuk kategori sulit dengan persentase (25%), 4 butir soal (E1,E4,E7,E9) termasuk kategori mudah dengan persentase (33,3%), dan 2 butir soal (E2,E3) termasuk kategori sangat mudah mempunyai persentase (16,7%). Untuk hasil yang lebih jelas dapat dilihat pada tabel 2.

Tabel 2. Analisis *item measure*

Butir Soal	Jumlah	Kategori	Persentase(%)
6, 10, & 12	3	Sangat sulit	25%
5, 8, & 11	3	Sulit	25%
1, 4, 7, & 9	4	Mudah	33,3%
2 & 3	2	Sangat mudah	16,7%

Berdasarkan dari hasil analisis di atas bahwa butir soal uraian dilihat dari segi indikator pengambilan keputusan memiliki sebaran yang tidak konsisten. Hal ini dapat dipengaruhi dari kemampuan peserta didik dalam mengerjakan tes, sehingga memiliki sebaran yang tidak merata. Namun jika dilihat dari perbandingan butir soal dengan kategori sangat sulit, sulit, mudah dan sangat mudah sudah seimbang yaitu butir soal memuat semua kategori tingkat kesulitan soal, yang berarti soal tidak terlalu mudah dan tidak terlalu sulit. Menurut Suwanto (2021) menyatakan bahwa butir soal yang proporsional terdiri dari 25% soal mudah, 50% soal sedang, dan 25% soal sulit. Hal ini sejalan dengan Sukmawati & Kamarudin (2024) yang menyatakan soal yang dianggap baik adalah yang memiliki tingkat kesulitan yang seimbang (tidak terlalu mudah atau terlalu sulit).

3. Reliabilitas (Ringkasan *Summary Statistic*)

Uji reliabilitas instrumen penelitian adalah untuk menentukan apakah alat tes yang digunakan untuk mengumpulkan data penelitian dapat diandalkan. Istilah "reliabilitas" digunakan untuk menggambarkan seberapa konsisten hasil pengukuran jika dilakukan dua kali atau lebih dengan instrumen dan alat ukur yang sama (Sugiono, Noerdjanah, & Wahyu, 2020). Sejalan dengan Sumintono & Widhiarso (2015) menyatakan bahwa Instrumen dikatakan reliabel apabila hasilnya secara konsisten sama atau stabil. Reliabilitas dalam penelitian ini dinilai menggunakan analisis model rasch dengan program *winsteps*. Nilai *alpha cronbach*, *person reliability*, dan *item reliability* disajikan dalam tabel Summary Statistic. Hasil analisis reliabilitas dapat dilihat pada gambar 4.

File Edit Format View Help
TABLE 3.1 C:\Users\USER\OneDrive\Desktop\DATA HA ZOU465WS.TXTR Jul 29 3:01 2024
INPUT: 363 Person 12 Item REPORTED: 363 Person 12 Item 4 CATS WINSTEPS 3.73

SUMMARY OF 363 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	33.6	11.8	.79	.47	.99	-.1	.99	-.1
S.D.	6.7	.4	1.34	-.06	.49	1.3	.48	1.3
MAX.	47.0	12.0	4.76	1.04	3.35	3.9	3.29	3.9
MIN.	15.0	9.0	-3.42	.43	.14	-3.5	.14	-3.5
REAL RMSE	.51	TRUE SD	1.24	SEPARATION	2.41	Person RELIABILITY	.85	
MODEL RMSE	.48	TRUE SD	1.26	SEPARATION	2.04	Person RELIABILITY	.87	
S.E. OF Person MEAN = .07								
Person RAW SCORE-TO-MEASURE CORRELATION = .98								
CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .89								

SUMMARY OF 12 MEASURED Item

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	1017.4	358.1	.00	.08	.99	-.4	.99	-.4
S.D.	93.2	4.5	.61	.00	.29	3.6	.29	3.6
MAX.	1180.0	363.0	.94	.09	1.85	9.6	1.79	9.1
MIN.	860.0	347.0	-1.10	.08	.66	-5.4	.65	-5.4
REAL RMSE	.09	TRUE SD	.60	SEPARATION	6.85	Item RELIABILITY	.98	
MODEL RMSE	.08	TRUE SD	.61	SEPARATION	7.17	Item RELIABILITY	.90	
S.E. OF Item MEAN = .18								

Gambar 4. Analisis reliabilitas (ringkasan *summary statistic*)

Berdasarkan gambar 4 nilai *alpha Cronbach* sebesar 0.89 menunjukkan interaksi yang bagus. Selanjutnya untuk nilai *person reliability* sebesar 0.85 menunjukkan konsistensi jawaban peserta didik bagus dan *item reliability* sebesar 0.98 artinya kualitas butir-butir soal istimewa. Menurut Amalia & Dianingati (2022) menyatakan bahwa Nilai *cronbach's alpha* harus lebih atau sama dengan 0,6 untuk uji reliabilitas yang baik. Menurut Sumintono & Widhiarso (2015). Reliabilitas adalah interaksi antara individu dan soal secara keseluruhan, dapat diukur dengan menggunakan nilai *alpha Cronbach*. Nilai reliabilitas individu menunjukkan konsistensi jawaban siswa, dan nilai reliabilitas item menunjukkan kualitas soal. Hasil analisis reliabilitas menunjukkan bahwa instrumen yang dikembangkan reliabel. Untuk hasil yang lebih jelas dapat dilihat pada tabel 3.

Table 3. Analisis reliabilitas

Summary Statistic	Reliability	Kategori
Alpha Cronbach	0.89	Bagus sekali
Person reliability	0.85	Bagus
Item reliability	0.98	Istimewa

Pada ringkasan *summary statistic* terdapat nilai daya pembeda (*separation*) yang berfungsi untuk untuk membedakan siswa dengan kemampuan menjawab soal yang baik dari siswa dengan kemampuan menjawab soal yang buruk. Untuk mengetahui pengelompokan daya pembeda secara lebih teliti digunakan persamaan strata (H). Berdasarkan gambar 4 diketahui bahwa nilai *person separation* sebesar 2.41 sehingga nilai H sebesar 3.54 sedangkan *item separation* nilainya sebesar 6.85 sehingga H sebesar 9.45. Data tersebut menunjukkan bahwa kelompok *person* dapat dibedakan menjadi 3 kelompok, sedangkan kelompok butir soal dapat dibedakan menjadi 9 kelompok berdasarkan nilai *separation*. Akan tetapi pada pemodelan Rasch hanya dapat mengidentifikasi jumlah pengelompokan *person* dan *item* namun tidak dapat diketahui pengelompokan yang diperoleh. Untuk hasil yang lebih jelas dapat dilihat pada tabel 4.

Tabel 4. Hasil analisis daya pembeda

Aspek yang diukur	Hasil analisis	Persamaan strata (H)	Kelompok
Person Separation	2.41	3.54	3
Item Separation	6.85	9.45	9

4. Tingkat Abilitas Individu (*Person Measure*)

Tingkat abilitas individu (*person measure*) adalah kemampuan individu dalam menyelesaikan soal. Tingkat abilitas individu digunakan untuk memberikan informasi mengenai peserta didik yang banyak menjawab dengan benar dan peserta didik yang paling sedikit menjawab benar. Analisis abilitas siswa menunjukkan pola jawaban siswa yang tidak sesuai atau berbeda dengan artian adanya siswa yang tidak hati-hati atau main tebakkan serta dapat mengidentifikasi siswa bekerja sama (menyontek) (Sumintono & Widhiarso, 2015). Hasil analisis tingkat abilitas, didapat nilai *standar deviasi* (SD) yaitu 1,34 dan nilai rata-rata logit (MEAN) yaitu 0,79 kemampuan siswa dikelompokkan menjadi tiga kelompok yaitu, tinggi, sedang, dan rendah. Hasil analisis tingkat abilitas individu dapat dilihat pada tabel 5.

Tabel 5. Analisis Tingkat Abilitas Individu

Kategori	Jumlah	Pesentase(%)
Tinggi	124	34,2%
Sedang	49	13,5%
Rendah	190	52,3%

Berdasarkan tabel 2 jumlah peserta didik yang memiliki kemampuan menjawab soal dengan benar kategori tinggi sebanyak 124 dengan persentase 34,2%, kategori sedang sebanyak 49 dengan persentase 13,5%, dan kategori rendah sebanyak 190 dengan persentase 52,3%. Untuk hasil yang lebih jelas dapat dilihat pada grafik 1.

Grafik 1. Hasil Tingkat Abilitas Individu



5. Tingkat Kesesuaian Individu (*Person Fit*)

Tingkat kesesuaian individu (*person fit*) pada pemodelan rasch dapat menentukan adanya pola respon yang berbeda. Pola respon yang berbeda adalah ketidaksesuaian jawaban yang diberikan berdasarkan tingkat abilitas siswa. Menurut Sumintono & Widhiarso (2015) kriteria yang diterapkan untuk menilai kesesuaian jawaban pada butir soal yang dianggap tidak sesuai (*outliers* atau *misfit*) adalah:

- a. *Outfit mean square* (MNSQ) diterima jika nilai: $0,5 < MNSQ < 1,5$
- b. *Outfit Z-standard* (ZSTD) diterima jika nilai: $-2,0 < ZSTD < +2,0$
- c. *Point measure correlation* (Pt Mean Corr) diterima jika nilai: $0,4 < Pt \text{ Measure Corr} < 0,85$

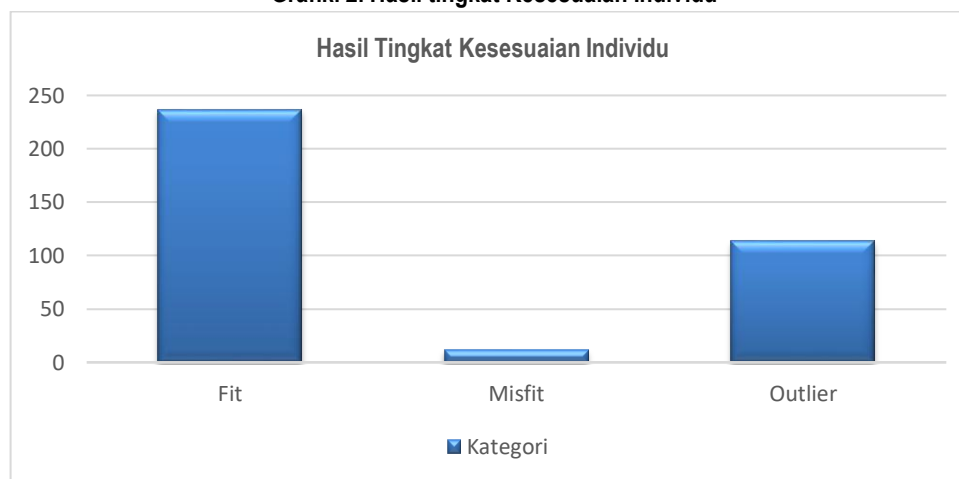
Kriteria fit atau valid jika memenuhi tiga atau dua kriteria yang ada, Namun, jika memenuhi satu kriteria, butir soal tersebut dapat dianggap cocok atau valid dan masih bisa dipertahankan untuk digunakan dalam penelitian. (Muntazhimah, Putri & Khusna, 2020). Hasil analisis tingkat kesesuaian individu dapat dilihat pada tabel 6.

Tabel 6. Analisis Tingkat Kesesuaian Individu

Kriteria	Jumlah	Persentase(%)
Fit	237	65,3%
Misfit	12	3,3%
Outlier	114	31,4%

Berdasarkan tabel 3 terdapat tiga kriteria pada *person fit* yaitu *fit*, *misfit*, dan *outlier*. Pada kriteria *fit* terdapat 237 siswa dengan presentase 65,3%, kriteria *misfit* terdapat 12 siswa dengan presentase 3,3%, dan kriteria *outlier* terdapat 114 siswa dengan presentase 31,4%. Peserta didik yang memenuhi satu atau tiga kriteria nilai tersebut dapat dikatakan fit, namun peserta didik yang tidak memenuhi tiga kriteria nilai tersebut dapat dikatakan misfit. Hasil analisis tingkat kesesuaian individu juga dapat di lihat pada grafik 2.

Grafiki 2. Hasil tingkat Kesesuaian Individu



Peserta didik yang memiliki pola respon yang tidak fit adalah peserta didik yang tidak memenuhi ketiga kriteria pola respon ideal, dan peserta didik yang hanya memenuhi satu kriteria masih dianggap pola responnya ideal (Aprilia dkk, 2020). Sedangkan peserta didik yang memiliki nilai diatas rata-rata atau skor ekstrem dapat dikatakan *outlier* (pencilan). Menurut Budiarti, Tarno, & Warsito, (2013) mengatakan bahwa *outlier* adalah sebuah data pengamatan yang tidak konsisten atau sangat tidak biasa pada seriesnya. Data outlier (pencilan) dapat muncul ketika terjadi kesalahan dalam memasukkan data, analisis, atau kesalahan dalam pengukuran (Nurdin dkk, 2014).

SIMPULAN

Berdasarkan hasil penelitian dapat disimpulkan bahwa instrumen tes pengambilan keputusan taksonomi Presseisen yang telah dikembangkan memenuhi kriteria instrumen yang layak. Hal ini dapat dilihat dari hasil analisis yang telah dilakukan sebagai berikut. Analisis validitas item menunjukkan 12 soal yang dikembangkan diterima. Analisis tingkat kesukaran menunjukkan sebanyak 3 soal sangat sulit, 3 soal sulit, 4 soal mudah, dan 2 soal sangat mudah. Analisis reliabilitas juga menggunakan model Rasch yang menunjukkan nilai *alpha Cronbach* sebesar 0.89 artinya interaksi antara person dan butir atau item soal secara keseluruhan bagus sekali. Selanjutnya untuk nilai person reliability sebesar 0.85 artinya konsistensi jawaban dari peserta didik bagus dan item reliability sebesar 0.98 artinya kualitas butir-butir soal istimewa. Analisis daya pembeda menunjukkan nilai *person separation* sebesar 2.41 sehingga nilai H sebesar 3.54 sedangkan *item separation* nilainya sebesar 6.85 sehingga H sebesar 9.45. Tingkat abilitas individu menunjukkan peserta didik kategori tinggi sebanyak 124 dengan persentase 34,2%, kategori sedang sebanyak 49 dengan persentase 13,5%, dan kategori rendah sebanyak 190 dengan persentase 52,3%. Tingkat kesesuaian individu menunjukkan peserta didik yang termasuk dalam kriteria fit atau valid sebanyak 237 dengan persentase 65,3%, kriteria misfit atau tidak valid sebanyak 12 dengan persentase 3,3 %, dan kriteria outlier sebanyak 114 dengan persentase 31,4%. Pengembangan instrumen tes ini hanya mencakup tiga tahapan 4D, yaitu *define*, *design*, dan *develop*. Oleh karena itu, diperlukan penelitian lebih lanjut terkait pengembangan instrumen tes 4D pada tahap *disseminate*.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih kepada semua pihak yang telah membantu dan mendukung penelitian ini, termasuk dosen pembimbing yang selalu memberikan arahan dan motivasi, serta guru biologi dan siswa kelas X SMA Pontianak yang telah berkontribusi dalam uji coba soal yang dikembangkan. Terima kasih juga disampaikan kepada kedua orang tua, saudara, pacar, dan teman-teman yang selalu memberikan dukungan dan semangat sehingga penelitian ini dapat berjalan dengan baik.

RUJUKAN

- Agustin, N., Sudarmin, S., Sumarti, S. S., & Addiani, A. K. (2018). Desain instrumen tes bermuatan etnosains untuk mengukur kemampuan berpikir kritis siswa SMA. *Jurnal Inovasi Pendidikan Kimia*, 12(2). doi: <https://doi.org/10.15294/jipk.v12i2.15475>
- Ardhani, Y. (2020). Kualitas Butir Soal Penilaian Akhir Tahun Mata Pelajaran Teknologi Dasar Otomotif Kelas X Teknik Kendaraan Ringan Otomotif Di Smk Muhammadiyah Gamping. *Jurnal Pendidikan Vokasi Otomotif*, 3(1), 85-94. doi: <https://doi.org/10.21831/jpvo.v3i1.34917>
- Aprilia, N., Susilaningsih, E., Priatmoko, S., & Kasmui, K. (2020). Desain Instrumen Tes Pemahaman Konsep Berbasis Hot Dengan Analisis Model Rasch. *Chemistry in Education*, 9(2), 1-8. Diperoleh dari: <http://journal.unnes.ac.id/sju/index.php/chemined>
- Budiarti, L., Tarno, T., & Warsito, B. (2013). Analisis Intervensi Dan Deteksi Outlier Pada Data Wisatawan Domestik (Studi Kasus Di Daerah Istimewa YOGYAKARTA). *Jurnal Gaussian*, 2(1), 39-48. doi: <https://doi.org/10.14710/j.gauss.2.1.39-48>

- Dianty, AP, Supeno, S., & Astutik, S. (2020). Kemampuan pengambilan keputusan siswa SMA dalam pembelajaran fisika berbasis inkuiri terbimbing. *Jurnal Pembelajaran Fisika*, 9 (1), 1-10. doi: <http://dx.doi.org/10.19184/jpf.v9i1.17935>
- Fatimah, L. U., & Alfath, K. (2019). Analisis kesukaran soal, daya pembeda, dan fungsi distraktor. *Jurnal Komunikasi Dan Pendidikan Islam*, 8(2), 37–64. Diperoleh dari: <https://journal.staimsyk.ac.id/index.php/almanar/article/view/115/104>
- Fitri, S. (2023). Pengaruh model pembelajaran decision making terhadap hasil belajar siswa pada mata pelajaran IPA-Biologi materi pokok system gerak pada tumbuhan: *Jurnal Pendidikan Biologi*, 4(2), 73-88. Diperoleh dari: <https://jurnal.uniraya.ac.id/index.php/Tunas/article/view/1175>
- Halim, R. A., Zaharim, A., Rashid, R. A., & Masodi, M. S. (2014). *Application of Logistic Regression Model in Rasch Measurement to Establish a Performance Index: A Case in Audits on Malaysian Institute of Higher Learning*. The 12th WSEAS Int Conf. on Applied Mathematics, Cairo, Egypt, 29-31 December. doi: <https://doi.org/10.5555/1376368.1376424>
- Kurniawan, U., & Andriyani, K. D. K. (2018). Analisis soal pilihan ganda dengan RASCH Model. *Jurnal Statistika Universitas Muhammadiyah Semarang*, 6(1). Doi: 10.26714/jsunimus.6.1.2018.%25p
- Lestari, D., & Setyarsih, W. (2020). Kelayakan instrumen penilaian formatif berbasis literasi sains peserta didik pada materi pemanasan global. *Inovasi Pendidikan Fisika*, 9(03), 561-570. doi: <https://doi.org/10.26740/ipf.v9n3.p561-570>
- Magdalena, I., Mahromiyati, M., & Nurkamilah, S. (2021). Analisis Instrumen Tes Sebagai Alat Evaluasi Pada Mata Pelajaran SBdP Siswa Kelas II SDN Duri Kosambi 06 Pagi. *NUSANTARA*, 3(2), 276-287. Diperoleh dari: <https://ejournal.stitpn.ac.id/index.php/nusantara/article/view/1264>
- Maulana, A. K., & Rochintaniawati, D. (2021). Analisis keterampilan pengambilan keputusan siswa kelas XI SMAN 1 Cihaurbeuti. *ISEJ: Indonesian Science Education Journal*, 2(2), 83-89. doi: <https://doi.org/10.62159/isej.v2i2.265>
- Maulana, S., Rusilowati, A., Nugroho, S. E., & Susilaningsih, E. (2023, June). Implementasi rasch model dalam pengembangan instrumen tes diagnostik. In *Prosiding Seminar Nasional Pascasarjana* (Vol. 6, No. 1, pp. 748-757). Diperoleh dari: <https://proceeding.unnes.ac.id/snpsasca/article/view/2214>
- Muntazhimah, M., 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. doi: <http://dx.doi.org/10.30998/jkpm.v6i1.8144>
- Nurdin, N. N., Raupong, R., & Islamiyati, A. (2014). Penggunaan regresi robust pada data yang mengandung pencilan dengan metode momen. *Jurnal Matematika, Statistika dan Komputasi*, 10(2), 114-123. doi: <https://doi.org/10.20956/jmsk.v10i2.3418>
- Presseisen, B.Z., 1985. *Glossary of Thinking Skill*, in A.L. Costa (ed). *Developing Minds: A Resource Book for Teaching Thinking*, Alexandria: ASCD. Diperoleh dari: <https://files.eric.ed.gov/fulltext/ED257858.pdf>
- Purniasari, L., Masykuri, M., & Ariani, S. R. D. (2021). Analisis Butir Soal Ujian Sekolah Mata Pelajaran Kimia SMA N 1 Kutowinangun Tahun Pelajaran 2019/2022 Menggunakan Model Iteman dan Rasch. *Jurnal Pendidikan Kimia*, 10(2), 205–214. doi: <https://doi.org/10.20961/jpkim.v10i2.48244>

- Rahmasari, Y., Noviekayati, I. G. A. A., & Pratitis, N. T. (2023). Self-Determination and Conformity with Student Career Decision Making: How Are They Related? *International Journal of Social and Management Studies*, 4(6), 27-32. doi: <https://doi.org/10.5555/ijosmas.v4i6.376>
- Sanjaya, M. S. M., Wahidin, W., & Maryuningsih, Y. (2019). Penerapan Pembelajaran Biologi Berbasis Entrepreneurship pada Materi Pencemaran Lingkungan untuk Meningkatkan Keterampilan Pengambilan Keputusan Siswa. *Jurnal Ilmu Alam Indonesia*, 2(1). Diperoleh dari: <https://www.syekhnurjati.ac.id/jurnal/index.php/jia/article/view/6272>
- Sugiono, S., Noerdjanah, N., & Wahyu, A. (2020). Uji validitas dan reliability alat ukur SG posture evaluation. *Jurnal Keterampilan Fisik*, 5(1), 55-61. doi: <https://doi.org/10.37341/jkf.v5i1.167>
- Sukmawati, D., & Kamarudin, K. (2024). Analisis Tingkat Kesukaran Butir Soal Penilaian Tengah Semester pada Pelajaran PPKn di Kelas V SD Negeri 1 Baadia. *Prosa: Jurnal Penelitian Pendidikan Guru Sekolah Dasar*, 2(3), 836-843. doi: <https://doi.org/10.35326/prosa.v8i4.5335>
- Sumintono, B dan Widhiarso, W. (2015). Aplikasi Pemodelan Rasch pada Assessment Pendidikan. Cimahi: Trim Komunikata Publishing House.
- Suwarto. (2021). The Characteristics of Indonesia Second- Semester Final Test for Eight-Grade Students. *Turkish Online Journal of Qualitative Inquiry*. 12(9). Diperoleh dari: <https://www.tojqi.net/index.php/journal/article/download/5499/3928/5981>
- 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 Pendidikan Guru Sekolah Dasar*, 7(2), 1-12. doi: <https://doi.org/10.17509/pedadidaktika.v7i2.24773>