

**Submitted:** 2024-01-04**Published:** 2024-05-01

ANALISIS BUTIR SOAL MATEMATIKA PADA PENILAIAN AKHIR SEMESTER GANJIL MENGGUNAKAN RASCH MODEL KELAS X SMA NEGERI 2 BUNGURAN TIMUR

Rahma Gustiannur Simbolon^{a)}, Desi Rahmatina^{b)}, Febrian^{c)}

a,b,c) Program Studi Pendidikan Matematika, Universitas Maritim Raja Ali Haji, Tanjungpinang, Indonesia

Corresponding Author: desirahmatina@umrah.ac.id^b
190384202027@student.umrah.ac.id, febrian@umrah.ac.id

Article Info

Keywords: Analysis of Question Items, Item-Person, Rasch Model.

Abstract

The aim of this research is to analyze mathematical problems by measuring reliability, level of difficulty (item and person measure) and level of suitability (item and person fit) using the Rasch model. Respondents in this study were 77 students with 40 multiple choice questions. This research is classified as quantitative descriptive research. The sampling technique is purposive sampling and the data collection technique is documentation. The results of reliability analysis using Winsteps 3.73, obtained reliability seen from Alpha Cronbach KR 20 of 0.82 with person reliability of 0.80 and item reliability of 0.90. The results of the analysis of the difficulty level of the measuring questions showed that 12 questions were very difficult, 3 questions were difficult, 19 questions were easy, and 6 questions were very easy. Meanwhile, the number of people obtained was 36 students with high abilities, 33 students with medium abilities, and 8 students with low abilities. The results of the suitability level analysis showed that 1 item was not suitable and 39 items were found to be suitable. Meanwhile, in person fit, 8 students experienced different response patterns (misfit) and 69 students had ideal response patterns. The response pattern was different because students failed to answer very easy questions correctly, but were able to answer very difficult questions correctly. Based on this analysis, the use of the Rasch model is highly recommended for use by teachers so that it can help teachers find out the quality of assessment tools and students' abilities.

Kata Kunci:

Analisis Butir Soal; *Item-Person*; *Rasch Model*.

Tujuan penelitian ini adalah untuk menganalisis butir soal matematika dengan mengukur reliabilitas, tingkat kesulitan (*item dan person measure*) dan tingkat kesesuaian (*item dan person fit*) menggunakan Rasch model. Responden pada penelitian ini berjumlah 77 peserta didik dengan 40 soal berupa pilihan ganda. Penelitian ini tergolong penelitian deskriptif kuantitatif. Teknik pengambilan sampel yaitu purposive sampling dan teknik pengumpulan data berupa dokumentasi. Hasil analisis reliabilitas menggunakan bantuan *software* Winsteps 3.73, diperoleh reliabilitas dilihat dari Alpha Cronbach KR 20 sebesar 0,82 dengan person reliability 0,80 dan item reliability 0,90. Hasil analisis tingkat kesulitan, item measure diperoleh 12 soal sangat sulit, 3 soal sulit, 19 soal mudah dan 6 soal sangat mudah. Sedangkan person measure diperoleh 36 peserta didik abilitas tinggi, 33 peserta didik abilitas sedang dan 8 peserta didik abilitas rendah. Hasil analisis tingkat kesesuaian, item fit diperoleh 1 soal tidak fit dan 39 soal fit. Sedangkan person fit diperoleh 8 peserta didik mengalami pola respon berbeda (*misfit*) dan 69 peserta didik dengan pola respon ideal. Pola respon berbeda dikarenakan peserta didik tidak berhasil menjawab soal sangat mudah dengan benar, namun mampu menjawab soal sangat sulit dengan benar. Berdasarkan analisis tersebut, penggunaan Rasch model sangat dianjurkan untuk digunakan oleh guru agar dapat membantu guru dalam mengetahui kualitas alat penilaian dan kemampuan yang dimiliki oleh peserta didik.

PENDAHULUAN

Dunia pendidikan dituntut untuk terus meningkatkan kualitas agar dapat melahirkan generasi yang dapat bersaing dan dapat menghadapi perkembangan zaman. Upaya untuk mewujudkan hal tersebut, salah satunya ialah adanya pelaksanaan pendidikan yang berkualitas. Menurut Mirunnisa (2019), keterlaksanaan pendidikan yang baik merupakan hal penting dikarenakan akhir dari pendidikan bertujuan

untuk dapat generasi yang mempunyai kualitas. Ini berarti dengan terlaksananya pendidikan berkualitas ialah suatu keharusan dalam mempersiapkan generasi muda yang mampu bersaing pada tantangan dunia pendidikan.

Dunia Pendidikan selalu terhubung ke sekolah. Sekolah merupakan suatu Lembaga Pendidikan yang menjadi wadah untuk pembentukan karakter, sosialisasi, dan persiapan anak-anak menuju kehidupan

dewasa (Irwan, 2024). Sekolah memiliki tujuan utama untuk menyediakan pengetahuan, keterampilan, dan nilai-nilai moral kepada peserta didik melalui proses pembelajaran yang terstruktur. Menurut Susanto et al., (2015), ada tiga unsur yang mendukung dalam proses pembelajaran yaitu tujuan dari pengajaran, kegiatan belajar mengajar, dan nilai evaluasi. Umumnya, tugas seorang guru ialah merencanakan kegiatan belajar mengajar, melaksanakan proses pembelajaran dan menilai hasil dari proses pembelajaran. Menurut Sopian (2016), tugas guru meliputi mendidik, mengajar, dan melatih. Mendidik yaitu meneruskan dan mengembangkan nilai-nilai hidup. Mengajar yaitu meneruskan dan mengembangkan ilmu pengetahuan dan teknologi, sedangkan melatih yaitu mengembangkan keterampilan-keterampilan pada peserta didik. Selanjutnya dilakukan kegiatan pengukuran untuk melihat seberapa paham peserta didik akan materi yang telah dipelajari. Kegiatan pengukuran hasil belajar untuk mengetahui tingkat penguasaan materi disebut dengan evaluasi.

Evaluasi merupakan alat ukur yang dipakai untuk melihat seberapa paham peserta didik dalam menyerap materi yang telah diberikan (Hamimi et al., 2020). Sehingga pelaksanaan evaluasi yang dilaksanakan dapat mengetahui terkait tingkat capaian belajar peserta didik. Pada tingkat capaian belajar peserta didik yang

menjadi instrumen pendukung ialah tes. Tes ialah cara untuk mendapatkan nilai yang dilaksanakan pada waktu, tempat, dan situasi yang memenuhi syarat tertentu yang dirancang untuk peserta didik. Untuk peserta didik yang telah mengikuti rangkaian proses pendidikan selama waktu yang telah ditentukan baru bisa diberikan sebuah tes evaluasi pengukuran (Alam, 2017).

Salah satu pengukuran evaluasi di sekolah ialah kegiatan Penilaian Akhir Semester (PAS) dengan teknik penilaian berupa tes tertulis. Hasil penilaian akhir semester menjadi gambaran sejauh mana pemahaman kompetensi yang sudah dipelajari peserta didik melalui proses pembelajaran dalam satu semester. Kegiatan pengukuran nilai tersebut menggunakan instrumen yang dikenal dengan butir soal.

Menurut Alam (2017) mendefinisikan butir soal adalah kumpulan pertanyaan-pertanyaan seputar materi pembelajaran yang akan diujikan. Untuk mendapatkan kualitas soal baik dan tepat, guru perlu melakukan analisis butir soal. Analisis soal ialah suatu tahapan yang dilakukan secara sistematis, sehingga nantinya akan berguna untuk mendapatkan informasi khusus mengenai butir dari soal-soal yang diujikan (Anggraini, 2021). Sebagaimana dikatakan Suaedi (2016) bahwa dengan melakukan analisis terhadap butir-butir soal akan mendapatkan secara rinci tentang kelemahan dan juga kekuatan masing-

masing dari butir soal. Secara jelas, analisis butir soal diperlukan untuk melihat bagaimana tingkat kebermutuan soal tersebut yang telah digunakan. Guna dari mutu soal ialah dapat memberikan secara jelas dan tepat terkait penguasaan materi peserta didik untuk dapat membedakan peserta didik yang benar-benar paham akan materi dan peserta didik yang belum memahami materi.

Namun, fakta dilapangan didapati bahwa masih ada guru yang tidak melakukan analisis butir soal matematika penilaian akhir semester (PAS) sebagaimana mestinya. Guru tidak melakukan analisis butir soal sebelum diujikan yaitu kriteria butir soal ujian yang bermutu perlu adanya analisis secara validitas, reliabilitas, daya pembeda dan tingkat kesukaran dari soal evaluasi tes. Berdasarkan hasil wawancara, guru hanya mengembangkan dan menuliskan sesuai kisi-kisi soal dengan mengadaptasi soal-soal sebelumnya dan tanpa melakukan validasi. Guru juga beranggapan bahwa untuk menganalisis soal membutuhkan banyak waktu. Sejalan dengan pendapat Rahayu & Djazari (2016) bahwa guru beranggapan untuk dapat menganalisis butir soal yang akan digunakan dalam tes memerlukan banyak waktu dan tenaga. Sehingga tidak dilakukannya analisis butir soal yang berakibat soal yang digunakan untuk pengujian belum akurat dan tepat dalam mengukur pemahaman dari peserta didik.

Penelitian mengenai analisis butir soal sudah pernah dilakukan oleh beberapa peneliti sebelumnya. Diantaranya yaitu Amelia et al., (2021) yang meneliti tentang analisis butir soal menggunakan *software* Anates dari penelitiannya didapati pada tingkat kesukaran cenderung berkategori sukar, daya pembeda cenderung dalam kategori sangat buruk, dan untuk keefektifan pengecoh cenderung ke dalam kategori baik. Kemudian Astuti (2020) mengenai analisis butir soal didapati kualitas soal ujian semester yang telah diujikan tidak berkualitas. Ini dibuktikan dari hasil analisis menggunakan *software* SPSS dengan mengukur validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan penggunaan distraktor pada 30 soal, hanya dijumpai 8 soal yang termasuk baik.

Berdasarkan pemaparan-pemaparan tersebut, penelitian yang sering digunakan dalam mengevaluasi instrumen menggunakan analisis klasik ialah SPSS dan Anates. Namun, dengan perkembangan teknologi saat ini, penilaian klasik digantikan dengan penilaian modern yaitu salah satunya adalah analisis Rasch model (Anekawati et al., 2023). Rasch model muncul dari analisis yang dilakukan oleh Dr. Georg Rasch, seorang ahli matematika dari Denmark (Sumintono & Widhiarso, 2015). Rasch model adalah satu model teori respon butir (ITR) yang paling populer. Rasch model merupakan suatu alat analisis yang sangat berguna untuk menguji validitas, reliabilitas

instrument, serta person dan item secara sekaligus (Purnomo, 2016).

Sejalan dengan Febrian & Fera (2019) penggunaan Rasch model dalam dunia asesmen masih belum banyak atau jarang dipergunakan dalam menganalisis data secara spesifik mengenai kualitas suatu pembelajaran sebagai tolak ukur dalam proses pembelajaran. Selain itu, Sumintono & Widhiarso (2015) menyatakan bahwa dari berbagai pengembangan dalam menganalisis penilaian pendidikan, terkhusus instrumen tes, informasi akan lebih tepat jika menggunakan pemodelan Rasch. Sehingga dengan menggunakan Rasch model, informasi mengenai butir soal pengujian kemungkinan besar dapat diketahui secara akurat dan tepat.

Berdasarkan penjelasan diatas, peneliti bermaksud untuk melakukan penelitian yang berjudul “Analisis Butir Soal Matematika pada Penilaian Akhir Semester Ganjil menggunakan Rasch Model Kelas X SMA Negeri 2 Bunguran Timur”. Penelitian ini diharapkan dapat memberikan bantuan dan masukan bagi guru dalam mengetahui kondisi dari soal yang telah atau akan digunakan dalam pengujian sehingga menjadi umpan balik untuk dapat memperbaiki sistem evaluasi pada kurun waktu mendatang.

METODE

Penelitian ini tergolong penelitian deskriptif kuantitatif. Menurut (Sulistyawati et

al., 2022) mendefinisikan penelitian deskriptif kuantitatif adalah penelitian yang mengkaji dan menjelaskan suatu fenomena dengan data atau angka secara apa adanya tanpa bermaksud untuk menguji suatu hipotesis tertentu.

Penelitian ini dilaksanakan pada bulan November tahun 2023. Populasi penelitian ialah seluruh peserta didik kelas X SMA Negeri 2 Bunguran Timur dengan sampel penelitian berjumlah 3 kelas yaitu X IIS 1, X IIS 2 dan X IIS 3 sebanyak 77 peserta didik. Pada penelitian ini teknik pengambilan sampel menggunakan teknik purposive sampling dan teknik pengumpulan data yang digunakan ialah dokumentasi. Menurut Arikunto (2014) dokumentasi adalah teknik mencari data berupa catatan, buku, transkrip, surat kabar serta foto-foto kegiatan.

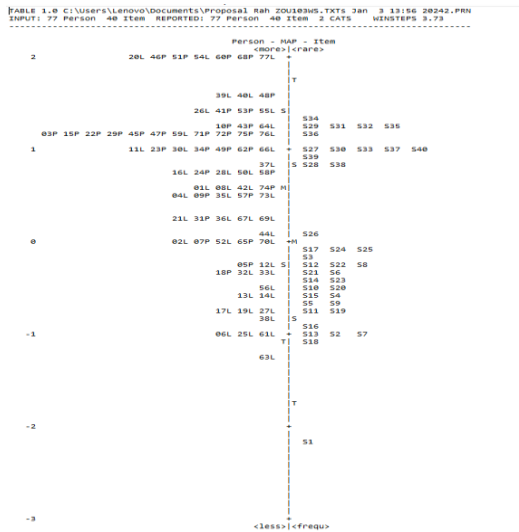
Dokumentasi digunakan untuk mendapatkan informasi mengenai data-data yang dibutuhkan selama peneliti. Adapun data yang digunakan berupa daftar nama peserta didik kelas X tahun pelajaran 2022/2023, soal Penilaian Akhir Semester Ganjil, kunci jawaban soal, dan jawaban peserta didik.

Teknik analisis data menggunakan analisis deskriptif. Analisis deskriptif digunakan untuk menggambarkan data yang telah dikumpulkan. Setelah data terkumpul, selanjutnya data di analisis menggunakan Rasch model dengan bantuan *software* Winsteps untuk menguji reliabilitas, tingkat kesulitan (item dan person measure), dan

tingkat kesesuaian (item dan person fit). Pengujian tersebut kemudian dijabarkan dan dijelaskan menggunakan pemeriksaan dari Suminthono & Widhiarso (2015).

HASIL DAN PEMBAHASAN

Berdasarkan hasil analisis, pada Peta Wright ditemukan adanya sebaran kemampuan peserta didik dan tingkat kesulitan butir soal. Pada sebelah kiri Peta Wright diperoleh gambaran abilitas peserta didik yang memiliki tujuh peserta didik dengan abilitas paling tinggi (+1,98 *logit*) dan peserta didik dengan abilitas paling rendah (-1,27 *logit*). Pada sebelah kanan Peta Wright menunjukkan sebaran nilai *logit* tiap butir soal. Butir S34 ialah butir soal dengan tingkat kesulitan tertinggi (+1,34 *logit*) dan butir S01 ialah butir soal dengan tingkat kesulitan terendah (-2,15 *logit*).



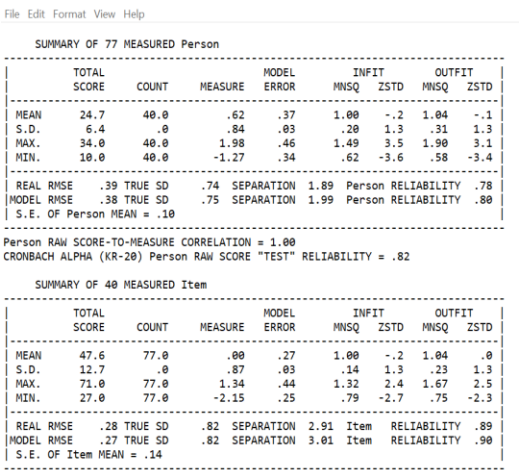
Gambar 1. Peta Wright

Dari hasil analisis tersebut jika digabungkan *logit person* dan *logit item*, terdapat temuan bahwa nilai *logit* peserta didik dari tingkat abilitasnya yang lebih tinggi dari nilai *logit* butir soal diartikan bahwa probabilitas untuk bisa mengerjakan soal yang nilai *logit*nya lebih rendah dengan benar adalah lebih dari 50%. Peserta didik dengan nilai *logit* lebih besar dari nilai *logit* butir soal tidak akan kesulitan menjawab soal dengan benar dikarenakan tingkat kesulitan butir soal ada dibawah abilitasnya (Abdullah et al., 2022).

Adapun hasil analisis dijelaskan sebagai berikut.

1. Reliabilitas

Berdasarkan hasil analisis reliabilitas, diperoleh *item reliability* 0,89 dan 0,90 dengan kategori bagus, *person reliability* 0,78 dan 0,80 dengan kategori cukup serta *Alpha Cronbach KR 20* sebesar 0,82



Gambar 2. Output summary statistic

Dari Gambar 2 menunjukkan nilai item, person reliability dan nilai *Alpha Cronbach*. Berdasarkan nilai yang diperoleh, diketahui bahwa kualitas butir soal yang digunakan memiliki reliabilitas bagus dan tingkat konsistensi jawaban peserta didik cukup. Kemudian, nilai *Alpha Cronbach* yang menunjukkan interaksi antara peserta didik dan soal secara keseluruhan bernilai bagus sekali. Hal ini berarti, instrument soal matematika pada penilaian akhir semester ganjil memiliki koefisien reliabilitas yang bagus sekali.

Sejalan dengan Azizah & Wahyuningsih (2020), *Alpha Cronbach* adalah hasil uji tingkat reliabilitas secara keseluruhan. Sehingga dengan memanfaatkan nilai *Alpha Cronbach* akan memudahkan guru untuk mengetahui tingkat reliabilitas secara sekaligus dari peserta didik maupun butir soal yang diujikan.

2. Tingkat Kesulitan Item Measure

Berdasarkan hasil analisis *item measure*, diperoleh butir soal yang diurutkan dari butir soal sangat sulit sampai sangat mudah berdasarkan nilai *logit* tiap butir soal. Pemeriksaan butir soal ini menggunakan nilai rata-rata dan nilai Standar Deviasi (SD) yang dikelompokkan menjadi empat kelompok tergantung pada tingkat kesulitan butir soal. Nilai Standar Deviasi pada penelitian ini adalah 0,87.

Pengelompokan kriteria *item measure* dapat dilihat pada Tabel 1.

Tabel 1. Kriteria *item measure*

Nilai Logit	Kategori
Lebih besar dari +0,87	Sangat Sulit
0,0 logit +0,87	Sulit
0,0 logit -0,87	Mudah
Lebih kecil dari - 0,87	Sangat Mudah

Sumber: Sumintono & Widhiarso (2015)

Adapun dari hasil keseluruhan item measure, 40 butir soal yang diujikan diperoleh 12 soal dengan kategori sangat sulit yaitu ketentuan nilai *logit* lebih besar dari +0,87, 3 soal dengan kategori sulit yaitu ketentuan 0,0 *logit* sampai dengan +0,87, 19 soal dengan kategori mudah yaitu ketentuan 0,0 *logit* sampai -0,87 dan 6 soal kategori sangat mudah dengan ketentuan nilai *logit* lebih kecil dari -0,87.

Item STATISTICS: MEASURE ORDER														
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MODEL MEASURE	S.E.	INFSQ	ZSTD	INFSQ	ZSTD	CORR.	EXP.	IPF-MEASURE	EXACT	WATCH	Item
34	27	77	1.34	-.25	-.90	-1.01	-.85	-1.01	-.46	-.34	67.5	68.5	534	
29	29	77	1.21	-.25	-.83	-1.91	-.76	-1.91	-.55	-.35	74.0	67.1	529	
31	29	77	1.21	-.25	-.81	-2.21	-.75	-2.01	-.57	-.35	74.0	67.1	531	
32	29	77	1.21	-.25	-.88	-1.31	-.82	-1.41	-.49	-.35	74.0	67.1	532	
35	29	77	1.21	-.25	-.92	-.91	-.92	-.61	-.44	-.35	74.0	67.1	535	
36	30	77	1.15	-.25	-.91	-1.01	-.90	-.81	-.45	-.35	70.1	66.4	536	
27	32	77	1.02	-.25	-1.02	-.31	-.97	-.21	-.35	-.36	61.0	65.8	527	
30	32	77	1.02	-.25	-1.01	-1.11	-.81	-.11	-.35	-.36	61.0	65.8	530	
40	32	77	1.02	-.25	-.83	-2.11	-.77	-2.11	-.56	-.36	74.0	65.8	540	
33	33	77	-.96	-.25	-.84	-2.01	-.86	-1.31	-.54	-.36	79.2	65.7	533	
37	33	77	-.96	-.25	-.82	-2.31	-.76	-2.21	-.58	-.36	74.0	65.7	537	
39	34	77	-.90	-.25	-.79	-2.71	-.76	-2.31	-.60	-.36	77.9	65.7	539	
28	35	77	-.84	-.25	-.91	-1.01	-.90	-.91	-.46	-.37	72.7	65.6	528	
38	35	77	-.84	-.25	-.83	-2.21	-.78	-2.21	-.57	-.37	75.3	65.6	538	
26	47	77	-.11	-.25	-1.03	-1.11	-.82	-.21	-.35	-.37	67.5	68.8	526	
17	50	77	-.09	-.26	-1.13	1.21	-.26	1.81	-.20	-.37	72.7	70.6	517	
24	50	77	-.09	-.26	-1.02	-1.21	-.08	-.61	-.33	-.37	72.7	70.6	524	
25	50	77	-.09	-.26	-.96	-.31	1.01	-.11	-.40	-.37	72.7	70.6	525	
3	51	77	-.15	-.26	-.99	-.81	-.98	-.11	-.38	-.37	68.8	71.2	513	
8	53	77	-.29	-.26	-.93	-.51	-.89	-.61	-.44	-.36	74.0	72.5	518	
12	53	77	-.29	-.26	-1.14	1.21	1.28	1.71	-.18	-.36	66.2	72.5	512	
22	53	77	-.29	-.26	-.96	-.31	-.99	-.81	-.39	-.36	79.2	72.5	522	
6	54	77	-.36	-.27	-1.32	2.41	1.46	2.51	-.01	-.36	61.0	73.2	516	
21	54	77	-.36	-.27	1.00	-1.71	-.87	-.81	-.55	-.36	81.8	73.2	521	
14	55	77	-.43	-.27	-1.20	1.51	1.37	2.01	-.10	-.36	67.5	74.0	514	
23	55	77	-.43	-.27	-.89	-.81	-.91	-.51	-.46	-.36	77.9	74.0	523	
10	56	77	-.51	-.27	-1.23	1.71	1.50	2.41	-.05	-.35	70.1	74.9	510	
20	56	77	-.51	-.27	-.96	-1.21	1.11	-.61	-.36	-.35	77.9	74.9	519	
4	57	77	-.58	-.28	-1.04	-1.31	1.08	-.41	-.30	-.35	74.0	75.7	514	
15	57	77	-.58	-.28	-1.18	1.31	1.37	1.08	-.11	-.35	71.4	75.7	515	
5	58	77	-.65	-.28	-1.13	1.01	1.15	-.81	-.20	-.34	72.7	76.5	519	
9	58	77	-.66	-.28	-1.05	-.41	1.08	-.51	-.28	-.34	77.9	76.5	519	
11	59	77	-.74	-.29	-1.12	-.81	1.26	1.21	-.18	-.34	72.7	77.5	511	
19	59	77	-.74	-.29	-.98	-.71	-.83	-.81	-.46	-.34	77.9	77.5	519	
16	61	77	-.91	-.30	-1.23	1.31	1.42	1.61	-.05	-.33	79.2	79.6	516	
2	62	77	-1.00	-.30	-1.06	-.41	1.09	-.41	-.24	-.32	81.8	88.6	512	
7	62	77	-1.00	-.30	-.95	-.21	1.02	-.21	-.35	-.32	81.8	88.6	517	
13	62	77	-1.00	-.30	-1.10	-.61	1.22	-.91	-.20	-.32	79.2	88.6	513	
18	63	77	-1.10	-.31	-1.26	1.41	1.07	2.11	-.04	-.31	83.1	81.9	518	
1	71	77	-2.15	-.44	-1.00	-.81	1.00	-.21	-.24	-.31	92.2	92.2	515	
MEAN	47.6	77.0	-.00	-.27	-1.00	-2.11	1.04	-.01			74.1	72.4		
S.D.	12.7	0	-.87	-.03	1.14	1.31	1.23	1.31			6.2	5.91		

Gambar 3. Output *item measure Person Measure*

Berdasarkan hasil analisis *person measure* pada Gambar 3, diperoleh tingkat abilitas peserta didik diurutkan berdasarkan peserta didik dengan kemampuan menyelesaikan soal tinggi hingga rendah. Kriteria *person measure* menggunakan nilai Standar Deviasi dan titik tolak penentuan abilitas yaitu nilai *mean logit person*. Adapun nilai Standar Deviasi pada analisis *person measure* ini adalah 0,84 dan nilai *mean logit person* adalah 0,62.

Pengelompokkan kriteria *person measure* sebagai berikut.

Tabel 2. Kriteria *person measure*

Nilai <i>Logit</i>	Kategori
Lebih besar dari 0,84	Tinggi
Lebih kecil dari 0,84	Sedang
Lebih kecil dari -0,62	Rendah

Sumber: Sumintono & Widhiarso (2015)

Adapun dari 77 peserta didik yang mengikuti ujian diperoleh 36 peserta didik dengan abilitas menyelesaikan soal tinggi (nilai *logit* lebih besar dari nilai Standar Deviasi), 33 peserta didik dengan abilitas menyelesaikan soal sedang (nilai *logit* lebih kecil dari nilai Standar Deviasi), dan 8 peserta didik dengan abilitas menyelesaikan soal rendah (lebih kecil dari – nilai rata-rata *logit* yaitu 0,62). Data tersebut dapat dilihat pada Tabel 2.

Sejalan dengan Abdullah et al., (2022), analisis *item measure* pada *Rasch model* memberikan gambaran nilai *logit* soal secara terperinci. Hasil *output* dari tabel

menampilkan informasi mengenai butir soal yang terurut. Nilai *logit* pada tabel dimulai dari yang memiliki nilai *logit measure* tertinggi ke nilai *logit* terendah serta menunjukkan butir soal tersukar sampai pada butir soal termudah secara berurutan.

3. Tingkat Kesesuaian

Item Fit

Item STATISTICS: MISFIT ORDER												
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MODEL MEASURE	S.E.	INFIT MNSQ	OUTFIT ZSTD	PT-MEASURE ZSTD	CORR. EXP.	EXACT MATCH OBS% EXP%	Item		
18	63	77	-1.10	.31	1.26	1.4	1.67	2.1	A-.04 .31	83.1 81.9	S18	
10	56	77	-.51	.27	1.23	1.7	1.50	2.4	B-.05 .35	70.1 74.9	S10	
6	54	77	-.36	.27	1.32	2.4	1.46	2.5	C-.01 .36	61.0 73.2	S6	
16	61	77	-.91	.30	1.23	1.3	1.42	1.6	D-.05 .33	79.2 79.6	S16	
14	55	77	-.43	.27	1.20	1.5	1.37	2.0	E-.10 .36	67.5 74.0	S14	
15	57	77	-.58	.28	1.18	1.3	1.37	1.8	F-.11 .35	71.4 75.7	S15	
12	53	77	-.29	.26	1.14	1.2	1.28	1.7	G-.18 .36	66.2 72.5	S12	
11	59	77	-.74	.29	1.12	.8	1.26	1.2	H-.18 .34	72.7 77.5	S11	
17	50	77	-.09	.26	1.13	1.2	1.26	1.8	I-.20 .37	72.7 70.6	S17	
13	62	77	-1.00	.30	1.10	.6	1.22	.9	J-.20 .32	79.2 80.6	S13	
5	58	77	-.66	.28	1.13	.9	1.15	.8	K-.20 .34	72.7 76.5	S5	
20	56	77	-.51	.27	.96	-.2	1.11	.6	L-.36 .35	77.9 74.9	S20	
2	62	77	-1.00	.30	1.06	.4	1.09	.4	M-.24 .32	81.8 80.6	S2	
24	50	77	-.09	.26	1.02	.2	1.08	.6	N-.33 .37	72.7 70.6	S24	
9	58	77	-.66	.28	1.05	.4	1.08	.5	O-.28 .34	77.9 76.5	S9	
4	57	77	-.58	.28	1.04	.3	1.08	.4	P-.30 .35	74.0 75.7	S4	
26	47	77	-.11	.25	1.03	.3	1.02	.2	Q-.35 .37	67.5 68.8	S26	
7	62	77	-1.00	.30	.95	-.2	1.02	.2	R-.35 .32	81.8 80.6	S7	
27	32	77	1.02	.25	1.02	.3	.97	-.2	S-.35 .36	61.0 65.8	S27	
25	50	77	-.09	.26	.96	-.3	1.01	.1	T-.40 .37	72.7 70.6	S25	
30	32	77	1.02	.25	1.01	.1	1.01	.1	U-.35 .36	61.0 65.8	S30	
1	71	77	-2.15	.44	1.00	.1	.82	-.2	S-.24 .23	92.2 92.2	S1	
3	51	77	-.15	.26	.99	.0	.98	-.1	r-.38 .37	68.8 71.2	S3	
22	53	77	-.29	.26	.96	-.3	.99	.0	q-.39 .36	79.2 72.5	S22	
8	53	77	-.29	.26	.93	-.5	.89	-.6	p-.44 .36	74.0 72.5	S8	
35	29	77	1.21	.25	.92	-.9	.92	-.6	o-.44 .35	74.0 67.1	S35	
28	35	77	.84	.25	.91	-1.0	.90	-.9	n-.46 .37	72.7 65.6	S28	
36	30	77	1.15	.25	.91	-1.0	.90	-.8	m-.45 .35	70.1 66.4	S36	
23	55	77	-.43	.27	.89	-.8	.91	-.5	l-.46 .36	77.9 74.0	S23	
34	27	77	1.34	.25	.90	-1.0	.85	-1.0	k-.46 .34	67.5 68.5	S34	
19	59	77	-.74	.29	.90	-.7	.83	-.8	j-.46 .34	77.9 77.5	S19	
32	29	77	1.21	.25	.88	-1.3	.82	-1.4	i-.49 .35	74.0 67.1	S32	
21	54	77	-.36	.27	.80	-1.7	.87	-.8	h-.55 .36	81.8 73.2	S21	
33	33	77	.96	.25	.84	-2.0	.86	-1.3	g-.54 .36	79.2 65.7	S33	
29	29	77	1.21	.25	.83	-1.9	.76	-1.9	f-.55 .35	74.0 67.1	S29	
40	32	77	1.02	.25	.83	-2.1	.77	-2.1	e-.56 .36	74.0 65.8	S40	
38	35	77	.84	.25	.83	-2.2	.78	-2.2	d-.57 .37	75.3 65.6	S38	
37	33	77	.96	.25	.82	-2.3	.76	-2.2	c-.58 .36	74.0 65.7	S37	
31	29	77	1.21	.25	.81	-2.2	.75	-2.0	b-.57 .35	74.0 67.1	S31	
39	34	77	.90	.25	.79	-2.7	.76	-2.3	a-.60 .36	77.9 65.7	S39	
MEAN	47.6	77.0	.00	.27	1.00	-.2	1.04	.0		74.1 72.4		
S.D.	12.7	.0	.87	.03	.14	1.3	.23	1.3		6.2 5.9		

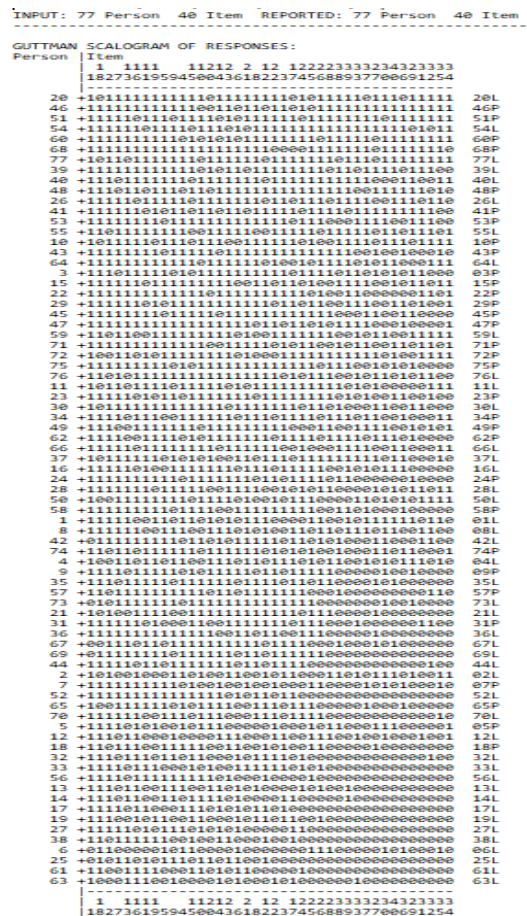
Gambar 4. Output item fit

Berdasarkan analisis *item fit* pada Gambar 4, batas atas *item* dinyatakan *fit* dengan model jika memenuhi salah satu atau keduanya dari tiga syarat (Azizah & Wahyuningsih, 2020). Adapun kriteria syaratnya yaitu Pertama, nilai *Outfit* MNSQ terletak pada 0,5 sampai dengan 1,5. Kedua, nilai *Outfit* ZSTD terletak pada -2,0 sampai dengan 2,0. Ketiga, nilai korelasi butir soal dengan skor total (*Pt Measure Corr*) terletak diantara 0,4 sampai dengan 0,85 (Sumintono & Widhiarso, 2015). Adapun dari 40 butir soal diperoleh 1 soal *misfit* (tidak *fit*) dikarenakan tidak memenuhi ketiga kriteria tingkat kesesuaian dan 39 soal *fit*. Sehingga butir soal yang tidak *fit* memerlukan perbaikan ataupun diganti.

Person Fit

Berdasarkan analisis *person fit*, kriteria pemeriksaan *person fit* sama dengan *item*. Hasil analisis tersebut, dari 77 peserta didik yang mengikuti ujian diperoleh 8 peserta didik terdeteksi memiliki pola respon berbeda. Pola respon berbeda disebabkan adanya ketidaksesuaian jawaban peserta didik berdasarkan tingkat abilitasnya. Setelah dilakukan uji lebih lanjut menggunakan scalogram dengan matriks Guttman, diperoleh bahwa peserta didik tersebut mengalami pola respon berbeda yaitu peserta didik tidak mampu menyelesaikan soal kategori sangat mudah. Namun, mampu menyelesaikan soal dengan kategori sangat sulit. Menurut R. N. Amelia (2021), jika didefinisikan dengan *Rasch*

model, peserta didik dengan kemampuan rendah tidak akan memiliki peluang untuk menyelesaikan soal yang lebih sulit dari kemampuannya. Sehingga, jawaban yang diberikan oleh peserta didik berkemungkinan besar bahwa peserta didik tersebut memberikan jawaban tidak cermat, jawaban tebakan yang kebetulan benar ataupun hasil contekan.



Gambar 5. Output scalogram

Hasil penelitian ini, sejalan dengan temuan Nudin & Hidayatullah (2023) yang mengemukakan bahwa butir soal yang sudah sesuai dengan *Rasch model* dapat digunakan sebagai instrumen tes, dan untuk soal tidak *fit* harus diperbaiki ataupun diganti. Sedangkan peserta didik yang terdeteksi mengalami pola respon berbeda akan diperiksa lebih lanjut untuk mendapatkan bimbingan dari guru.

Berdasarkan pembahasan diatas, pengukuran analisis butir soal dan peserta didik secara terperinci terjabarkan jika menggunakan pemodelan *Rasch*. Menurut Chan et al., dalam Nuryanti et al., (2018), analisis *Rasch model* dapat memberikan informasi secara keseluruhan dimulai dari kualitas instrument, kualitas respon dari peserta didik, serta interaksi antar responden dengan item. Sehingga, penggunaan *Rasch model* sebagai teknik analisis merupakan hal yang tepat. Apalagi analisis dengan *Rasch model* cukup mudah dilakukan tetapi dapat menghasilkan analisis yang akurat sehingga memudahkan bagi guru untuk mengetahui kemampuan dari peserta didik beserta butir soal yang diujikan. Sejalan dengan saran Marjiastuti & Wahyuni (2014) dalam penelitiannya, mengungkapkan bahwa penggunaan analisis dengan *Rasch model* disarankan untuk digunakan oleh guru agar dapat membantu guru dalam melakukan pengukuran dan penilaian hasil belajar peserta didik untuk mengetahui kemampuan peserta didik secara apa adanya. Oleh karena itu, penggunaan *Rasch model* dalam

menganalisis instrumen penilaian Pendidikan terkhusus ujian, akan menghasilkan informasi yang lebih tepat dan sesuai sasaran yang dibutuhkan oleh guru.

PENUTUP

Simpulan

Berdasarkan hasil penelitian, kesimpulan dari penelitian ini yaitu soal Matematika pada Penilaian Akhir Semester Ganjil menggunakan *Rasch model* didapati soal memiliki nilai koefisien reliabilitas yang bagus sekali. Dari tingkat kesulitan, *item measure* didapati sebanyak 12 soal tipe sangat sulit, 3 soal dengan tipe soal sulit, 19 soal mudah dan 6 soal tipe soal sangat mudah. Kemudian *person measure*, didapati sebanyak 36 orang memiliki tingkat kemampuan menyelesaikan soal tinggi, 33 peserta didik berkemampuan menyelesaikan soal sedang dan 8 orang peserta didik dengan kemampuan menyelesaikan soal rendah. Dari tingkat kesesuaian, *item fit* didapati 39 soal *fit* dan 1 butir soal kurang bagus yang memerlukan perbaikan ataupun diganti. Berdasarkan *person fit*, didapati 2 peserta didik dengan pola respon berbeda, 6 orang peserta didik diperiksa lebih lanjut untuk melihat adakah kesalahan pola respon, dan 69 peserta didik lainnya memiliki pola respon dengan model ideal. Berdasarkan skala Guttman, pola respon berbeda dikarenakan peserta didik tidak mampu mengerjakan soal sangat mudah, namun mampu mengerjakan soal dengan kategori sulit.

Saran

Saran yang dapat dilakukan bagi peneliti selanjutnya untuk dapat melanjutkan penelitian terkait analisis butir soal dengan menambah pengujian lainnya seperti daya pembeda dan tingkat keefektifan pengecoh, serta variable lain sebagai pendukung penelitian ini.

DAFTAR PUSTAKA

- Abdullah, N., Jahja, M., & Setiawan, D. G. E. (2022). Analisis kualitas butir soal pada mata pelajaran fisika di jurusan fisika fakultas mipa universitas negeri gorontalo tahun ajaran 2021/2022. *Jurnal Sains Dan Pendidikan Fisika*, 18(1), 44–52. <https://doi.org/10.35580/jspf.v18i1.32358>
- Alam, S. (2017). *Analisis soal matematika ujian akhir semester genap di kelas VIII SMPN 1 sinjai timur*. Universitas Islam Negeri (UIN) Alauddin Makassar.
- Amelia, N., Paridjo, & Sina, I. (2021). Analisis butir soal penilaian akhir tahun mata pelajaran matematika menggunakan software anates. *AKSIOMA: Jurnal Matematika Dan Pendidikan Matematika*, 12(3), 447–456. <https://doi.org/10.26877/aks.v12i3.8356>
- Amelia, R. N. (2021). Identifikasi item fit dan person fit dalam pengukuran hasil belajar kimia. *Jurnal Ilmiah WUNY*, 3(1), 13–26.
- Anekawati, A., Rofik, M., & Isyanto. (2023). Pelatihan analisis rasch untuk MGMP guru matematika kabupaten sumenep. *Reswara: Jurnal Pengabdian Kepada Masyarakat*, 4(1).
- Anggraini, W. (2021). *Analisis butir soal pada pembelajaran e-learning mata pelajaran ipa kelas iv (alat indera pada manusia) di 3 sekolah kabupaten seluma*. Institut Agama Islam Negeri Bengkulu.
- Arikunto, S. (2014). *Prosedur Penelitian suatu pendekatan praktek* (Edisi Revi). Rineka Cipta.
- Astuti. (2020). Analisis soal ujian akhir semester genap mata pelajaran matematika siswa SD negeri 005 binuang ajaran 2017/2018. *Jurnal Pendiidikan Tambusai*, 4(1), 867–880.
- Azizah, & Wahyuningsih, S. (2020). Penggunaan model rasch untuk analisis instrumen tes pada mata kuliah matematika aktuaria. *JUPITEK: Jurnal Pendidikan Matematika*, 3(1), 45–50. <https://doi.org/10.30598/jupitekvol3iss1pp45-50>
- Febrian, & Fera, M. (2019). Kualitas perangkat dan keterampilan mengajar mahasiswa pendidikan matematika pada mata kuliah micro teaching menggunakan analisis model rasch. *Jurnal Gantang*, 4(1), 87–95. <https://doi.org/10.31629/jg.v4i1.1065>
- Hamimi, L., Zamharirah, R., & Rusydy, R. (2020). Analisis butir soal ujian matematika kelas vii semester ganjil tahun pelajaran 2017/2018. *Mathema: Jurnal Pendidikan Matematika*, 2(1), 57. <https://doi.org/10.33365/jm.v2i1.459>
- Irwan, A. (2024). *Pengertian sekolah beserta unsur, fungsi, dan jenjang pendidikan*. Maxmanroe.Com.
- Marjiastuti, K., & Wahyuni, S. (2014).

- Analisis kemampuan peserta didik dengan model racsh. *Jurnal Indonesia*, 121–128.
- Mirunnisa. (2019). Peningkatan kemampuan penalaran matematis siswa sekolah menengah atas dengan model discovery learning. *Jurnal Numeracy*, 6(1), 76–85. <https://doi.org/https://doi.org/10.46244/numeracy.v6i1.434>
- Nudin, I., & Hidayatullah, R. S. (2023). Analisis butir soal penilaian tengah semester menggunakan model Rasch di SMK Negeri 5 Surabaya. *JPTM*, 12(2), 132–139.
- Nuryanti, S., Masykuri, M., & Susilowati, E. (2018). Analisis iteman dan model Rasch pada pengembangan instrumen kemampuan berpikir kritis peserta didik sekolah menengah kejuruan. *Jurnal Inovasi Pendidikan IPA*, 4(2), 224–233. <http://journal.uny.ac.id/index.php/jipi>
- Purnomo, S. (2016). *Pengembangan soal matematika model pisa konten space and shape untuk mengetahui level kemampuan berpikir tingkat tinggi berdasarkan analisis model rasch*. Universitas Jember.
- Rahayu, R., & Djazari, M. (2016). Analisis kualitas soal pra ujian nasional mata pelajaran ekonomi akuntansi. *Jurnal Pendidikan Akuntansi Indonesia*, 14(1), 85–94. <https://doi.org/10.21831/jpai.v14i1.11370>
- Sopian, A. (2016). Tugas, peran, dan fungsi guru dalam pendidikan. *Raudhah Proud To Be Professionals: Jurnal Tarbiyah Islamiyah*, 1(1), 88–97. <https://doi.org/10.48094/raudhah.v1i1.10>
- Suaedi, M. (2016). Analisis terhadap kualitas butir soal buatan guru ipa kelas viii mts negeri slawi tahun pelajaran 2015/2016. *Jurnal Kamil*, 3, 1–23.
- Sulistiyawati, W., Wahyudi, & Trinuryono, S. (2022). Analisis (deskriptif kuantitatif) motivasi belajar siswa dengan model blended learning di masa pandemi covid19. *Kadikma*, 13(1), 68–73. <https://doi.org/10.19184/kdma.v13i1.31327>
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi Pemodelan Rasch pada Assessment Pendidikan* (Cetakan I). Trim Komunika.
- Susanto, H., Rinaldi, A., & Novalia. (2015). Analisis validitas reabilitas tingkat kesukaran dan daya beda pada butir soal ujian akhir semester ganjil mata pelajaran matematika. *Aljabar: Jurnal Pendidikan Matematika*, 6(2), 203–217.