

ANALISIS KUALITAS INSTRUMEN ASESMEN FORMATIF PADA MATA PELAJARAN EKONOMI MENGGUNAKAN RASCH MODEL

Yanti Puspita¹, Astri Srigustini², Edi Fitriana Afriza³

^{1,2,3}Pendidikan Ekonomi FKIP Universitas Siliwangi

¹222165065@student.unsil.ac.id, ²astrisrigustini@unsil.ac.id

³edifitrianaafriza@unsil.ac.id

ABSTRACK

Formative assessment is an important part of the learning process, so the quality of the instruments used needs to be evaluated. This study aims to analyze the quality of formative assessment instruments in Economics using the Rasch Model. The study employed a quantitative approach with a descriptive method involving 86 students from classes X-6 and X-9 at Tasikmalaya State High School 2, who were selected through purposive sampling. The research instrument consisted of 40 multiple-choice items that had been validated by four experts using Aiken's V index. Data analysis was conducted using the Rasch Model. The results showed that the instrument had high content validity (Aiken's V for content = 0.947; for construct = 0.966; language = 0.966), generally met the assumptions of the Rasch Model, had item reliability of 0.79 in Class X-6 and 0.72 in Class X-9, and the majority of items met the item fit criteria with varying levels of difficulty and student response patterns consistent with the characteristics of the Rasch Model. Thus, the formative assessment instrument analyzed is of adequate quality and suitable for use as a learning evaluation tool in the Economics subject.

Keywords: formative assessment, instrument quality, Rasch Model

ABSTRAK

Asesmen formatif merupakan bagian penting dalam proses pembelajaran sehingga kualitas instrumen yang digunakan perlu dievaluasi. Penelitian ini bertujuan untuk menganalisis kualitas instrumen asesmen formatif pada mata pelajaran Ekonomi menggunakan Rasch Model. Penelitian menggunakan pendekatan kuantitatif dengan metode deskriptif terhadap 86 peserta didik kelas X-6 dan X-9 SMA Negeri 2 Tasikmalaya yang dipilih secara purposive. Instrumen penelitian berupa 40 butir soal pilihan ganda yang telah divalidasi oleh empat orang ahli menggunakan indeks Aiken's V. Analisis data dilakukan menggunakan Rasch Model. Hasil penelitian menunjukkan bahwa instrumen memiliki validitas isi yang tinggi (Aiken's V aspek materi = 0,947; konstruksi = 0,966; bahasa = 0,966), secara umum memenuhi asumsi Rasch Model, memiliki item reliability sebesar 0,79 pada kelas X-6 dan 0,72 pada kelas X-9, serta sebagian besar butir memenuhi kriteria item fit dengan tingkat kesulitan yang bervariasi dan pola respons peserta didik yang sesuai dengan karakteristik Rasch Model. Dengan demikian, instrumen asesmen formatif yang dianalisis memiliki kualitas yang memadai dan layak digunakan sebagai alat evaluasi pembelajaran pada mata pelajaran Ekonomi.

Kata kunci: asesmen formatif, kualitas instrumen, Rasch Model

A. Pendahuluan

Asesmen formatif merupakan bagian integral dalam proses pembelajaran yang berfungsi untuk memantau perkembangan belajar peserta didik sekaligus memberikan umpan balik bagi peserta didik dan pendidik dalam memperbaiki proses pembelajaran. Dalam implementasi Kurikulum Merdeka, asesmen formatif berperan penting dalam mendukung pembelajaran yang berpusat pada peserta didik dan berorientasi pada pencapaian tujuan pembelajaran (Kemendikbudristek, 2022). Oleh karena itu, efektivitas asesmen formatif tidak hanya ditentukan oleh pelaksanaannya, tetapi juga oleh kualitas instrumen yang digunakan. Instrumen yang berkualitas akan menghasilkan informasi yang akurat, objektif, dan dapat dipercaya sehingga mampu menjadi dasar dalam pengambilan keputusan pembelajaran.

Berdasarkan hasil observasi awal di SMA Negeri 2 Tasikmalaya, asesmen formatif pada mata pelajaran Ekonomi telah diterapkan sebagai bagian dari proses pembelajaran. Namun, instrumen asesmen yang digunakan belum dianalisis secara empiris sehingga kualitas instrumen

dan karakteristik butir soal belum diketahui secara komprehensif. Analisis instrumen masih didominasi oleh pendekatan *Classical Test Theory* (CTT), yang memiliki keterbatasan karena karakteristik butir soal dan kemampuan peserta didik dipengaruhi oleh sampel yang digunakan. Kondisi tersebut berpotensi menghasilkan informasi hasil asesmen yang kurang akurat sehingga diperlukan pendekatan analisis yang mampu mengevaluasi kualitas instrumen secara lebih objektif.

Salah satu pendekatan yang banyak digunakan dalam evaluasi kualitas instrumen adalah *Item Response Theory* (IRT), khususnya Rasch Model. Dibandingkan dengan pendekatan CTT, Rasch Model mampu mengestimasi kemampuan peserta didik dan tingkat kesulitan butir soal secara terpisah sehingga menghasilkan pengukuran yang lebih objektif. Melalui berbagai prosedur analisis, Rasch Model memberikan informasi mengenai kualitas instrumen, karakteristik butir soal, kemampuan peserta didik, serta kesesuaian data dengan model sehingga evaluasi instrumen dapat dilakukan secara lebih komprehensif.

Berbagai penelitian menunjukkan bahwa Rasch Model efektif digunakan dalam analisis kualitas instrumen pendidikan. Azizah & Wahyuningsih (2020) menunjukkan bahwa Rasch Model mampu mengevaluasi validitas, reliabilitas, dan tingkat kesulitan instrumen secara objektif. Alfarisa & Purnama (2019) melaporkan bahwa Rasch Model memberikan informasi kualitas butir soal ekonomi yang lebih rinci dibandingkan analisis klasik. Munali & Alifah (2024) membuktikan bahwa Rasch Model efektif dalam mengidentifikasi butir soal yang memenuhi maupun yang belum memenuhi kriteria kelayakan sehingga dapat menjadi dasar penyempurnaan instrumen. Novriyanti & Arthur (2024) juga menunjukkan bahwa Rasch Model mampu memberikan informasi mengenai reliabilitas, tingkat kesulitan, dan kesesuaian butir soal secara objektif, sedangkan (Saputri et al., 2022) menyimpulkan bahwa pendekatan *Item Response Theory* mampu memberikan informasi kualitas butir secara komprehensif setelah memenuhi asumsi model.

Meskipun demikian, penelitian-penelitian tersebut umumnya

dilakukan pada asesmen sumatif, pengembangan instrumen, atau mata pelajaran selain Ekonomi. Penelitian yang secara khusus menganalisis kualitas instrumen asesmen formatif pada mata pelajaran Ekonomi menggunakan Rasch Model masih terbatas. Selain itu, penelitian terdahulu umumnya berfokus pada analisis kualitas butir soal atau indikator tertentu, sedangkan evaluasi instrumen secara komprehensif melalui pengujian validitas isi, uji asumsi Rasch Model, *summary statistics*, *item measure*, *Wright Map*, *item fit*, dan *scalogram* masih jarang dilakukan dalam satu kajian. Oleh karena itu, penelitian ini memberikan kontribusi dengan mengevaluasi kualitas instrumen asesmen formatif pada mata pelajaran Ekonomi secara komprehensif menggunakan Rasch Model sehingga menghasilkan informasi empiris yang dapat dimanfaatkan sebagai dasar penyempurnaan instrumen asesmen dan peningkatan kualitas pembelajaran.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk menganalisis kualitas instrumen asesmen formatif pada mata pelajaran Ekonomi berdasarkan Rasch Model.

Analisis dilakukan melalui pengujian validitas isi, uji asumsi Rasch Model, *summary statistics*, tingkat kesulitan butir (*item measure*), pemetaan person-item (*Wright Map*), kesesuaian butir (*item fit*), dan *scalogram* untuk memperoleh informasi yang komprehensif mengenai validitas, reliabilitas, karakteristik butir soal, serta kualitas instrumen asesmen formatif secara keseluruhan.

B. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskriptif yang bertujuan menggambarkan kualitas instrumen berdasarkan hasil analisis Rasch Model (Sugiyono, 2022). Penelitian bertujuan untuk menganalisis kualitas instrumen asesmen formatif pada mata pelajaran Ekonomi berdasarkan Rasch Model. Penelitian dilaksanakan di SMA Negeri 2 Tasikmalaya pada semester genap Tahun Ajaran 2025/2026.

Populasi penelitian adalah seluruh peserta didik kelas X SMA Negeri 2 Tasikmalaya yang mengikuti asesmen formatif mata pelajaran Ekonomi. Sampel penelitian dipilih menggunakan teknik *purposive sampling*, yaitu kelas X-6 dan X-9

yang masing-masing berjumlah 43 peserta didik sehingga total sampel sebanyak 86 peserta didik. Pemilihan kedua kelas dilakukan secara *purposive* dari keseluruhan kelas yang mengikuti uji coba instrumen untuk menyajikan analisis kualitas instrumen secara lebih mendalam sesuai dengan ruang lingkup artikel ilmiah.

Instrumen penelitian berupa asesmen formatif mata pelajaran Ekonomi berbentuk pilihan ganda sebanyak 40 butir soal. Sebelum diujicobakan kepada peserta didik, instrumen divalidasi oleh empat orang ahli berdasarkan aspek materi, konstruksi, dan bahasa. Validitas isi dianalisis menggunakan indeks Aiken's V. Selanjutnya, data respons peserta didik dianalisis menggunakan perangkat lunak Winsteps versi 5.6.0.0 berdasarkan prosedur Rasch Model yang dijelaskan oleh Bond & Fox (2015) serta Sumintono & Widhiarso (2015).

Tahapan analisis meliputi pengujian validitas isi, uji asumsi Rasch Model, *summary statistics*, analisis tingkat kesulitan butir (*item measure*), pemetaan person-item (*Wright Map*), analisis kesesuaian butir (*item fit*), dan *scalogram*. Hasil analisis digunakan untuk memperoleh

informasi mengenai validitas, reliabilitas, karakteristik butir soal, serta kualitas instrumen asesmen formatif secara komprehensif.

C. Hasil dan Pembahasan

1. Validitas Isi Instrumen

Validitas isi instrumen dianalisis menggunakan indeks Aiken's V berdasarkan penilaian empat orang ahli terhadap aspek materi, konstruksi, dan bahasa. Hasil validitas isi disajikan pada Tabel 1.

Tabel 1. Hasil Validitas Isi Instrumen

Aspek Penilaian	Rata-rata Aiken's V	Kategori
Materi	0,947	Tinggi
Konstruksi	0,966	Tinggi
Bahasa	0,966	Tinggi

Berdasarkan Tabel 1, seluruh aspek penilaian memperoleh nilai Aiken's V di atas 0,80 sehingga termasuk kategori validitas tinggi. Hasil tersebut menunjukkan bahwa butir instrumen telah sesuai dengan indikator pembelajaran, memiliki konstruksi yang tepat, serta menggunakan bahasa yang jelas sehingga layak digunakan pada tahap uji coba. Nilai Aiken's V di atas 0,80 menunjukkan validitas isi yang tinggi berdasarkan penilaian para ahli (Utami et al., 2024). Meskipun seluruh

aspek telah memenuhi kategori validitas tinggi, beberapa butir soal tetap direvisi berdasarkan masukan para validator untuk menyempurnakan redaksi soal dan meningkatkan kualitas instrumen sebelum digunakan pada tahap uji coba.

2. Uji Asumsi Rasch Model

a. Unidimensionalitas

Uji unidimensionalitas dilakukan untuk memastikan bahwa instrumen asesmen formatif mengukur satu konstruk utama yang sama. Analisis dilakukan menggunakan output *Standardized Residual Loadings for Item* pada perangkat lunak Winsteps. Butir yang memiliki nilai *loading residual* dengan nilai absolut lebih dari 0,50 diidentifikasi sebagai butir yang berpotensi berkaitan dengan dimensi residual di luar konstruk utama. Hasil analisis disajikan pada Tabel 2.

Tabel 2. Hasil Uji Unidimensionalitas

Kelas	Butir dengan Loading Residual > 0,50 	Jumlah Butir
X-6	I31 (0,60); I22 (-0,57); I2 (-0,55); I16 (-0,53); I15 (-0,52)	5

X-9	I25 (0,74); I7 (0,70); I39 (-0,56); I17 (- 0,55)	4
-----	--	---

Berdasarkan Tabel 2, pada kelas X-6 terdapat lima butir yang memiliki nilai *loading residual* lebih dari 0,50, sedangkan pada kelas X-9 ditemukan empat butir. Jumlah tersebut relatif kecil dibandingkan keseluruhan 40 butir yang dianalisis sehingga secara umum instrumen masih menunjukkan kecenderungan mengukur satu konstruk utama. Namun demikian, butir-butir tersebut perlu dievaluasi lebih lanjut karena mengindikasikan adanya variasi yang tidak sepenuhnya dijelaskan oleh konstruk yang diukur.

Temuan ini sejalan dengan pendapat Sumintono & Widhiarso (2015) yang menyatakan bahwa analisis *standardized residual loadings* digunakan untuk mengidentifikasi kemungkinan adanya dimensi lain di luar konstruk utama. Keberadaan beberapa butir dengan *loading residual* yang tinggi tidak secara langsung menunjukkan bahwa instrumen bersifat multidimensional, tetapi menjadi indikator perlunya peninjauan terhadap substansi maupun redaksi butir agar tetap konsisten dengan

konstruk yang diukur. Dengan demikian, hasil uji unidimensionalitas menunjukkan bahwa instrumen asesmen formatif secara umum telah memenuhi asumsi unidimensionalitas.

b. Independensi Lokal

Uji independensi lokal dilakukan menggunakan output *Largest Standardized Residual Correlation* pada Winsteps untuk mengidentifikasi pasangan butir yang memiliki korelasi residual lebih dari 0,50. Nilai korelasi residual di atas batas tersebut mengindikasikan adanya ketergantungan local (*local dependence*) antarpasangan butir sehingga perlu dilakukan peninjauan lebih lanjut. Hasil analisis disajikan pada Tabel 3.

Tabel 3. Hasil Uji Independensi Lokal

Kelas	Pasangan Butir	Standardized Residual Correlation
X-6	I16-I22	0,55
X-9	I22-I24 dan I27-I39	0,58 dan 0,54

Berdasarkan Tabel 3, hanya ditemukan satu pasangan butir pada kelas X-6 dan dua pasangan butir pada kelas X-9 yang memiliki *standardized residual correlation* di atas 0,50. Jumlah tersebut relatif kecil sehingga secara umum asumsi

independensi lokal masih terpenuhi. Sumintono & Widhiarso (2015) menjelaskan bahwa korelasi residual yang tinggi dapat menunjukkan adanya kesamaan isi, konteks, atau redaksi antarpasangan butir, sedangkan Bond & Fox (2015) menyatakan bahwa setiap butir seharusnya memberikan informasi secara independen setelah kemampuan peserta didik dikendalikan. Temuan ini konsisten dengan hasil penelitian Novriyanti dan Arthur (2024) yang menekankan pentingnya evaluasi independensi lokal untuk menghasilkan estimasi parameter yang lebih akurat.

3. Summary Statistics

Analisis *summary statistics* bertujuan untuk memberikan gambaran umum mengenai kualitas instrumen berdasarkan nilai *mean measure*, *item reliability*, dan *separation*. Hasil analisis pada kelas X-6 dan X-9 disajikan pada Tabel 4.

Tabel 4. Hasil Summary Statistics Instrumen

Indikator	Kelas X-6	Kelas X-9
<i>Mean measure</i> (logit)	0,00	0,00
<i>Item Reliability</i>	0,79	0,72
<i>Separation</i>	1,93	1,62

Berdasarkan Tabel 4, kedua kelas memiliki nilai mean measure sebesar 0,00 logit dengan item reliability masing-masing sebesar 0,79 dan 0,72 serta separation sebesar 1,93 dan 1,62. Nilai tersebut menunjukkan bahwa instrumen memiliki konsistensi yang cukup baik dalam mengestimasi tingkat kesulitan butir serta mampu membedakan tingkat kesulitan ke dalam lebih dari satu kelompok. Menurut Sumintono & Widhiarso (2015), nilai item reliability dan separation mencerminkan stabilitas estimasi karakteristik butir. Boone dan Staver (2020) menambahkan bahwa semakin tinggi nilai separation, semakin baik kemampuan instrumen dalam mengelompokkan tingkat kesulitan butir. Hasil ini juga sejalan dengan Nur et al., (2023) yang menunjukkan bahwa reliabilitas dan separation yang memadai mendukung kualitas pengukuran instrumen berbasis Rasch.

4. Tingkat Kesulitan Butir (*Item Measure*)

Analisis *item measure* dilakukan untuk mengetahui tingkat kesulitan setiap butir soal berdasarkan nilai logit yang dihasilkan oleh Rasch Model. Pengelompokan tingkat

kesulitan didasarkan pada nilai simpangan baku (*standard deviation*) masing-masing kelas sehingga butir diklasifikasikan ke dalam kategori sangat sulit, sulit, mudah, dan sangat mudah. Ringkasan hasil analisis disajikan pada Tabel 5.

Tabel 5. Distribusi Tingkat Kesulitan Butir Soal

Kategori	Kelas X-6	Kelas X-9
Sangat Sulit	5	8
Sulit	15	12
Mudah	16	13
Sangat Mudah	4	7
Jumlah	40	40

Berdasarkan Tabel 5, distribusi tingkat kesulitan butir pada kedua kelas menunjukkan pola yang relatif seimbang. Pada kelas X-6 terdapat 5 butir berkategori sangat sulit, 15 butir sulit, 16 butir mudah, dan 4 butir sangat mudah. Sementara itu, pada kelas X-9 ditemukan 8 butir sangat sulit, 12 butir sulit, 13 butir mudah, dan 7 butir sangat mudah.

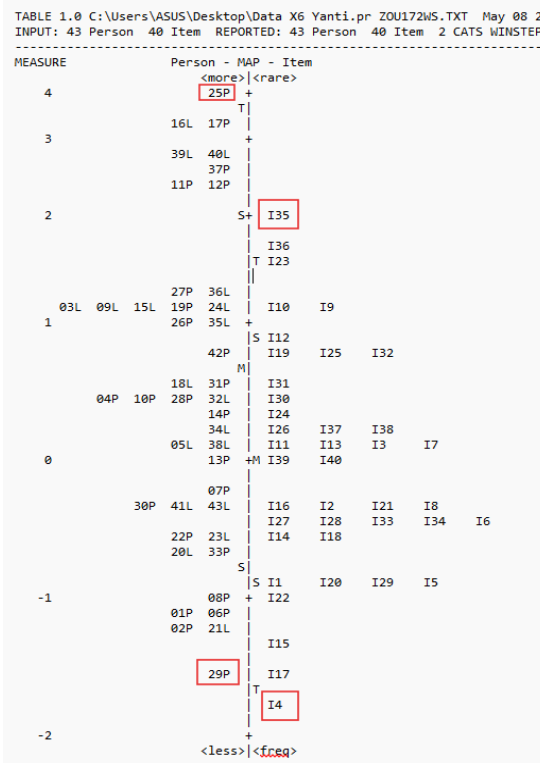
Butir tersulit pada kelas X-6 adalah I35 dengan nilai *measure* sebesar 2,00 logit, sedangkan butir termudah adalah I4 dengan nilai -1,73 logit. Pada kelas X-9, butir tersulit adalah I10 dengan nilai 2,08 logit, sedangkan butir termudah adalah I20 dan I34 dengan nilai -1,29 logit. Perbedaan nilai logit tersebut

menunjukkan bahwa setiap butir memiliki tingkat kesulitan yang bervariasi sehingga instrumen mampu mengukur kemampuan peserta didik pada berbagai tingkat kemampuan.

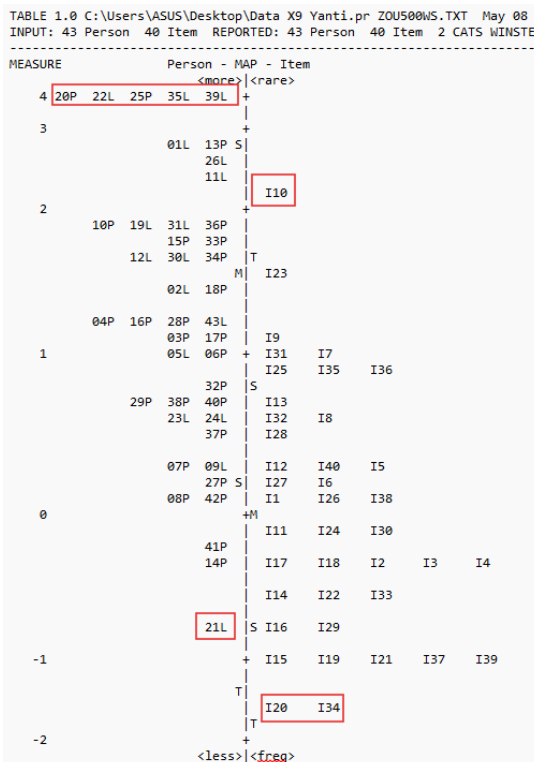
Boone & Staver (2020) juga menjelaskan bahwa distribusi tingkat kesulitan yang beragam meningkatkan kualitas pengukuran. Temuan ini sejalan dengan Saputri et al. (2022) yang menunjukkan bahwa analisis Item Response Theory mampu menggambarkan karakteristik tingkat kesulitan butir secara lebih komprehensif.

5. Pemetaan Person-Item (Wright Map)

Pemetaan kemampuan peserta didik dan tingkat kesulitan butir dilakukan menggunakan Wright Map. Hasil pemetaan untuk kelas X-6 dan kelas X-9 disajikan pada Gambar 1 dan Gambar 2.



Gambar 1. Wright Map Kelas X-6



Gambar 2. Wright Map Kelas X-9

Berdasarkan Gambar 1, peserta didik dengan kemampuan

tertinggi pada kelas X-6 adalah 25P, sedangkan peserta didik dengan kemampuan terendah adalah 29P. Pada sisi butir, I35 merupakan butir dengan tingkat kesulitan tertinggi, sedangkan I4 merupakan butir dengan tingkat kesulitan terendah. Posisi butir I4 yang berada di bawah sebagian besar peserta didik menunjukkan bahwa butir tersebut relatif mudah dijawab oleh mayoritas responden.

Pada kelas X-9 (Gambar 2), peserta didik dengan kemampuan tertinggi adalah 20P, 22L, 25P, 35L, dan 39L, sedangkan peserta didik dengan kemampuan terendah adalah 21L. Dari sisi tingkat kesulitan butir, I10 merupakan butir yang paling sulit, sedangkan I20 dan I34 merupakan butir yang paling mudah. Sebagian besar peserta didik berada pada rentang kemampuan yang berdekatan dengan sebaran tingkat kesulitan butir, sehingga instrumen dinilai memiliki tingkat penargetan (*targeting*) yang cukup baik.

Wright Map memberikan gambaran hubungan antara kemampuan peserta didik dan tingkat kesulitan butir dalam satu skala logit. Menurut Bond & Fox (2015), pemetaan ini memudahkan peneliti

mengevaluasi apakah tingkat kesulitan butir telah sesuai dengan kemampuan responden. Pada penelitian ini, distribusi kemampuan peserta didik dan tingkat kesulitan butir menunjukkan bahwa sebagian besar butir berada pada rentang kemampuan responden, meskipun masih terdapat beberapa butir yang sangat mudah maupun sangat sulit.

6. Kesesuaian Butir (Item Fit)

Analisis *item fit* dilakukan untuk mengevaluasi kesesuaian setiap butir soal terhadap Rasch Model berdasarkan tiga indikator, yaitu *Outfit Mean Square* (MNSQ), *Outfit Z-Standard* (ZSTD), dan *Point Measure Correlation* (Pt Measure Corr.). Butir yang tidak memenuhi ketiga kriteria tersebut dikategorikan sebagai *misfit* dan direkomendasikan untuk ditinjau kembali. Ringkasan hasil analisis *item fit* disajikan pada Tabel 6.

Tabel 6. Hasil Analisis Item Fit

Kelas	Butir Misfit	Jumlah Butir Fit	Jumlah Butir Misfit
X-6	I1 dan I35	38	2
X-9	I5 dan I11	38	2

Berdasarkan Tabel 6, pada kelas X-6 terdapat dua butir yang tidak memenuhi seluruh kriteria *item fit*, yaitu butir I1 dan I35. Butir I1 memiliki

nilai *Outfit MNSQ* sebesar 3,80, *Outfit ZSTD* sebesar 3,57, dan *Pt Measure Corr.* sebesar 0,25, sedangkan butir I35 memiliki nilai *Outfit MNSQ* sebesar 1,93, *Outfit ZSTD* sebesar 2,02, dan *Pt Measure Corr.* sebesar 0,34. Kedua butir tersebut dikategorikan sebagai *misfit* karena menunjukkan pola respons yang kurang sesuai dengan ekspektasi Rasch Model.

Pada kelas X-9 juga ditemukan dua butir *misfit*, yaitu I5 dan I11. Butir I5 memiliki nilai *Outfit MNSQ* sebesar 1,85, *Outfit ZSTD* sebesar 2,26, dan *Pt Measure Corr.* sebesar 0,28, sedangkan butir I11 memiliki nilai *Outfit MNSQ* sebesar 2,07, *Outfit ZSTD* sebesar 2,16, dan *Pt Measure Corr.* sebesar 0,31. Kondisi tersebut menunjukkan bahwa respons peserta didik terhadap kedua butir tersebut kurang konsisten dengan model pengukuran yang digunakan.

Secara keseluruhan, sebanyak 38 dari 40 butir soal pada masing-masing kelas memenuhi kriteria *item fit*. Temuan ini menunjukkan bahwa sebagian besar butir telah sesuai dengan Rasch Model sehingga layak digunakan untuk mengukur kemampuan peserta didik. Hanya sebagian kecil butir yang perlu ditinjau kembali melalui revisi redaksi,

kesesuaian indikator, maupun tingkat kesulitan agar kualitas instrumen menjadi lebih optimal. Hasil ini sejalan dengan penelitian Azizah & Wahyuningsih (2020) yang menyatakan bahwa analisis Rasch mampu mengidentifikasi butir soal yang sesuai maupun yang memerlukan perbaikan sehingga kualitas instrumen dapat dievaluasi secara lebih objektif.

7. Analisis Scalogram

Analisis *scalogram* dilakukan untuk mengidentifikasi pola respons peserta didik terhadap butir soal berdasarkan tingkat kesulitan item. Analisis ini digunakan untuk mendeteksi konsistensi pola jawaban serta mengidentifikasi kemungkinan adanya respons yang tidak lazim, seperti ketidaktelitian atau *guessing*. Hasil *scalogram* pada kelas X-6 dan X-9 disajikan pada Gambar 3 dan Gambar 4.

[illegible]

Gambar 3. Hasil Scalogram Kelas X-6

[illegible]

Gambar 3. Hasil Scalogram Kelas X-6

Berdasarkan hasil analisis *scalogram*, sebagian besar peserta didik pada kedua kelas menunjukkan pola respons yang teratur, yaitu mampu menjawab butir yang lebih mudah sebelum menjawab butir yang lebih sulit. Pola tersebut menunjukkan bahwa secara umum respons peserta didik telah sesuai dengan karakteristik Rasch Model.

Pada kelas X-6 masih ditemukan beberapa pola respons yang tidak sepenuhnya konsisten. Peserta didik 25P menunjukkan ketidakkonsistenan dalam menjawab beberapa butir soal yang mengindikasikan kemungkinan kurang teliti. Selain itu, peserta didik 07P, 30P, dan beberapa peserta didik lainnya memperlihatkan pola respons yang mengindikasikan kemungkinan penggunaan strategi *guessing*. Beberapa peserta didik lainnya, seperti 39L, 40L, 11P, dan 12P, menunjukkan pola respons yang sangat tidak lazim sehingga mengindikasikan adanya respons yang tidak konsisten dengan pola kemampuan yang diharapkan. Pola tersebut memerlukan perhatian lebih lanjut, meskipun tidak dapat dijadikan bukti adanya perilaku tertentu hanya berdasarkan hasil *scalogram*.

Pada kelas X-9, pola respons peserta didik cenderung lebih konsisten. Peserta didik 22L menunjukkan pola respons yang paling sesuai dengan tingkat kemampuannya. Sementara itu, peserta didik 20P dan 25P memperlihatkan beberapa ketidakkonsistenan dalam menjawab butir soal yang mengindikasikan

kemungkinan kurang teliti. Selain itu, peserta didik 14P dan 21L menunjukkan pola respons yang mengindikasikan kemungkinan penggunaan strategi *guessing*.

Secara keseluruhan, hasil analisis *scalogram* menunjukkan bahwa pola respons peserta didik pada kedua kelas didominasi oleh pola yang sesuai dengan ekspektasi Rasch Model. Penyimpangan yang ditemukan hanya terjadi pada sebagian kecil peserta didik sehingga tidak memengaruhi kualitas instrumen secara keseluruhan. Hasil penelitian ini juga sejalan dengan Novriyanti & Arthur (2024) yang menyatakan bahwa analisis Rasch tidak hanya mengevaluasi kualitas butir, tetapi juga memberikan informasi mengenai pola respons peserta didik. Selain itu, Azizah & Wahyuningsih (2020) menegaskan bahwa informasi dari *scalogram* dapat dimanfaatkan sebagai dasar penyempurnaan instrumen dan pelaksanaan asesmen.

D. Kesimpulan

Berdasarkan hasil analisis menggunakan Rasch Model, instrumen asesmen formatif pada mata pelajaran Ekonomi memiliki kualitas yang baik dan layak

digunakan sebagai alat evaluasi pembelajaran. Hasil validitas isi menunjukkan seluruh aspek memperoleh nilai Aiken's V pada kategori tinggi. Instrumen juga secara umum memenuhi asumsi Rasch Model, yang ditunjukkan oleh terpenuhinya asumsi unidimensionalitas dan independensi lokal meskipun masih ditemukan beberapa butir yang memerlukan peninjauan lebih lanjut. Analisis *summary statistics* menunjukkan nilai *item reliability* sebesar 0,79 pada kelas X-6 dan 0,72 pada kelas X-9 dengan kategori cukup. Analisis *item measure* dan *Wright Map* memperlihatkan bahwa instrumen memiliki variasi tingkat kesulitan butir yang mampu membedakan kemampuan peserta didik. Selain itu, sebanyak 38 dari 40 butir pada masing-masing kelas memenuhi kriteria *item fit*, sedangkan hasil *scalogram* menunjukkan bahwa sebagian besar peserta didik memiliki pola respons yang sesuai dengan karakteristik Rasch Model. Dengan demikian, Rasch Model mampu memberikan informasi yang komprehensif mengenai kualitas instrumen asesmen formatif sehingga instrumen yang dikembangkan layak

digunakan sebagai dasar evaluasi pembelajaran pada mata pelajaran Ekonomi, dengan beberapa butir yang direkomendasikan untuk direvisi agar kualitas pengukurannya semakin optimal.

DAFTAR PUSTAKA

- Alfarisa, F., & Purnama, D. N. (2019). *ANALISIS BUTIR SOAL ULANGAN AKHIR SEMESTER MATA PELAJARAN EKONOMI SMA MENGGUNAKAN RASCH MODEL*.
- Azizah, & Wahyuningsih, S. (2020). *Nomor 1 | Hal. 3, 50*. <https://doi.org/10.30598/jupitekvo13iss1ppx45-50>
- Bond, T. G. ., & Fox, C. M. . (2015). *Applying the Rasch model: fundamental measurement in the human sciences*. Routledge, Taylor and Francis Group.
- Boone, W. J., & Staver, J. R. (2020). *Advances in Rasch Analyses in the Human Sciences*.
- Kemendikbudristek. (2022). *Pembelajaran dan Asesmen*.
- Munali, & Alifah, S. (2024). *ANALISIS BUTIR SOAL STATISTIKA MODEL SUMMATIVE EVALUATION ASSESSMENT MENGGUNAKAN MODEL RASCH*. 5(3). <https://doi.org/10.46306/lb.v5i3>
- Novriyanti, E., & Arthur, R. (2024). Analisis Kualitas Butir Soal Ujian Tengah Semester Biologi Umum Menggunakan Model Rasch. *JagoMIPA: Jurnal Pendidikan Matematika Dan IPA*, 4(4), 718–733. <https://doi.org/10.53299/jagomipa.v4i4.927>
- Nur, A. S., Suryani, D. R., & Mayasari, D. (2023). ANALISIS KUALITAS BUTIR SOAL PADA KOMPETISI MATEMATIKA MENGGUNAKAN TEORI RESPON ITEM . *SIGMA: JURNAL PENDIDIKAN MATEMATIKA*, Volume 15 Nomor 1.
- Saputri, S. A., Dahlan, D., & Leni, P. (2022). *Analisis Kualitas Butir Soal Mata Pelajaran Ekonomi Berdasarkan Item Response Theory (Kasus Pada Sman 11, Sman 18, Dan Sman 23 Bandung)*. <https://ejournal.upi.edu/index.php/JPEI>
- Sugiyono. (2022). *METODE PENELITIAN KUANTITATIF, KUALITATIF DAN R & D*. ALFABETA.
- Sumintono, B., & Widhiarso, W. (2015). *Aplikasi pemodelan*. www.trimkomunikata.com
- Utami, L., Festiyed, Dian Purnama Ilahi, Arista Ratih, Elvi yenti, & Lazulva. (2024). ANALISIS INDEKS AIKEN UNTUK MENGETAHUI VALIDITAS ISI INSTRUMEN SCINETIFIC HABBITS OF MIND. *Journal of Research and Education Chemistry*, 6(1), 59. [https://doi.org/10.25299/jrec.2024.vol6\(1\).17430](https://doi.org/10.25299/jrec.2024.vol6(1).17430)