



Analisis Kualitas Butir Soal Mata Pelajaran Ekonomi Berdasarkan Item Response Theory (Kasus Pada Sman 11, Sman 18, Dan Sman 23 Bandung)

Suchia Auliya Saputri, Dadang Dahlan, Leni Permana.*

Universitas Pendidikan Indonesia

*Correspondence: E-mail: suchia.aulia31@upi.edu

ABSTRACT

Penelitian ini membahas tentang kualitas butir soal berdasarkan Item Response Theory pada penilaian akhir semester Ganjil Kelas XI IPS di SMA Negeri 11, SMAN 18 dan SMAN 23 Bandung. Tujuan penelitian ini adalah untuk mengetahui kualitas butir soal mata pelajaran ekonomi ganjil pas. PAS merupakan salah satu tes yang digunakan untuk mengukur keberhasilan proses belajar mengajar dalam satu semester. Butir-butir tes yang diujikan merupakan butir-butir soal yang dibuat oleh guru dan harus berkualitas baik sehingga perlu dilakukan analisis mutu butir soal dengan menggunakan teori respon butir. Metode penelitian yang digunakan adalah analisis deskriptif melalui pendekatan kuantitatif. Penelitian melibatkan 427 respon siswa, dan 85 pertanyaan dari 3 sekolah. Analisis kualitas item didasarkan pada teori respon item yang mengukur 2 parameter logistik (tingkat kesulitan dan diferensiasi) dengan bantuan Software Quest. Hasil dari penelitian ini adalah kualitas butir pada mata pelajaran ekonomi kelas XI IPS SMA Negeri 11, SMAN 18, dan SMAN 23 Bandung berdasarkan Item Response Theory secara keseluruhan hanya terdapat 33 butir butir yang mutunya baik, 47 butir butir mutunya kurang baik, dan 5 item dengan kualitas buruk. Berdasarkan validitas dan reliabilitas, terdapat 80 item yang valid dan 5 item yang tidak valid, serta memiliki nilai reliabilitas item lebih dari 0,85 yang berarti item tes tersebut dapat digunakan untuk menentukan keputusan individu siswa.

ARTICLE INFO

Article History:

Submitted/Received 22 July 2022

First Revised 17 August 2022

Accepted 14 Sept 2022

First Available online 15 Sept 2022

Publication Date 01 Oct 2022

Keyword:

Analisis,
Ekonomi,
Item Response Theory.

1. INTRODUCTION

Pemerintah mengusahakan penyelenggaraan pendidikan yang sebaik-baiknya untuk dapat mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif dapat mengembangkan potensi dirinya, Proses pembelajaran tersebut meliputi tahap perencanaan, pelaksanaan hingga evaluasi pembelajaran. Hasil dari kegiatan evaluasi ini menjadi suatu gambaran proses dan hasil belajar yang telah dilakukan. Pembelajaran dapat dikatakan berhasil apabila siswa mampu menguasai kompetensi yang diharapkan, diantaranya pada aspek pengetahuan, keterampilan dan sikap. Salah satu cara untuk dapat melihat seberapa jauh ketercapaian tersebut adalah dengan dilakukannya penilaian hasil belajar dan penilaian proses belajar mengajar. Dalam melakukan penilaian tersebut, dibutuhkan suatu alat yang memudahkan untuk mengukurnya. Alat yang biasa digunakan untuk mengukur ketercapaian kompetensi pada aspek kognitif adalah tes, salah satunya melalui Penilaian Akhir Semester (PAS). Pada umumnya butir soal tes tersebut dibuat oleh guru mata pelajaran itu sendiri. Namun, tidak sedikit dari hasil penelitian yang telah dilakukan mengenai kualitas butir tes buatan guru, menunjukkan hasil bahwa soal tes buatan guru berkualitas buruk sedangkan tes yang diberikan kepada peserta didik haruslah memiliki kualitas yang baik. Agar tes buatan guru memiliki kualitas yang baik dan guna meningkatkan mutu tes yang telah diujikan, maka perlu dilakukan analisis kualitas butir soal atau analisis item (item analysis) sebagai bentuk evaluasi yang dapat dilakukan oleh guru dalam rangka mengendalikan mutu Pendidikan.

Terdapat dua pendekatan dalam bidang pengukuran pendidikan yang sering digunakan untuk menganalisis kualitas tes, yakni teori klasik (Classical Test Theory) dan teori responsi butir (Item Response Theory). Akan tetapi, terdapat banyak kelemahan pada teori klasik, yakni hasil dari pengukuran akan tergantung pada karakteristik tes yang dipakai, parameter item bergantung pada kemampuan peserta tes dan error measurement atau kesalahan pengukuran hanya dapat mengetahui untuk kelompok bukan individu (Mardapi, 2012).

Terdapat tiga model yang paling populer digunakan dalam melakukan analisis item berdasarkan teori responsi butir (Item Response Theory) yaitu model satu, dua dan tiga parameter (Hambleton, Swaminathan, dan H.J., 1991). Pada model satu parameter (1PL) memuat parameter tingkat kesukaran (b_i), model dua parameter (2PL) memuat parameter tingkat kesukaran (b_i), dan daya pembeda (a_i), dan pada model tiga parameter (3PL) memuat parameter tingkat kesukaran (b_i), daya pembeda (a_i), dan tebakan semu (c_i). Parameter tersebut digunakan untuk mengukur kemampuan siswa. Dalam melakukan analisis tersebut, item yang akan dianalisis harus memenuhi 3 asumsi diantaranya: 1) unidimensi, yang berarti setiap butir tes hanya mengukur satu kemampuan, 2) independensi lokal, yang berarti tidak ada hubungan antara respon peserta tes dengan butir soal yang berbeda, dan 3) invariansi parameter, yang berarti karakteristik butir soal tidak tergantung pada sebaran parameter peserta tes dan begitupun sebaliknya (Retnawati H. , 2014).

2. METHODS

Metode yang digunakan dalam penelitian ini adalah analisis deskriptif melalui pendekatan kuantitatif. Analisis deskriptif merupakan suatu metode analisis yang memiliki tujuan untuk mendeskripsikan atau menjelaskan suatu hal dengan apa adanya (Irawan, 2004, hlm. 1). Penelitian ini bersifat *ex post facto* karena objek yang dijadikan sumber data oleh peneliti merupakan dokumen yang kegiatannya telah selesai dilaksanakan.

Objek dan Subjek Penelitian

Objek penelitian dalam penelitian ini adalah kualitas butir. Sedangkan yang menjadi subjek dalam penelitian ini adalah set soal yang terdiri dari silabus, kisi-kisi, soal, kunci jawaban, respon siswa, dan nilai.

Populasi dan Sampel Penelitian

Populasi dalam penelitian ini adalah seluruh set soal Penilaian Akhir Semester Ganjil mata Pelajaran Ekonomi SMA Negeri Bandung, yaitu sebanyak 27 set soal. Setiap set soal terdiri dari 30 butir soal, sehingga didapatkan populasi berjumlah 810 butir soal.

Sedangkan Sampel dalam penelitian ini ditentukan berdasarkan pertimbangan peneliti tersendiri (teknik purposive sampling). Peneliti menentukan sampel set soal berdasarkan mutu sekolah dan zona/wilayah yang membagi persebaran SMA Negeri Kota Bandung. Dari seluruh SMA Negeri Kota Bandung berjumlah 27 sekolah terdapat 3 SMA Negeri yang terpilih menjadi sampel, yaitu SMA Negeri 11, SMA Negeri 18 dan SMA Negeri 23, dengan jumlah sampel respon siswa sebagai berikut:

Tabel 1. Sampel Respon Siswa

Nama Sekolah	Pembagian Kelas XI IPS	Jumlah Siswa	Jumlah Siswa XI IPS
SMAN 11 Bandung	XI IPS 1	36	144
	XI IPS 2	36	
	XI IPS 3	36	
	XI IPS 4	36	
SMAN 18 Bandung	XI IPS 1	35	139
	XI IPS 2	34	
	XI IPS 3	34	
	XI IPS 4	36	
SMAN 23 Bandung	XI IPS 1	36	144
	XI IPS 2	36	
	XI IPS 3	36	
	XI IPS 4	36	
Jumlah			427

Pada penelitian ini menggunakan sampel respon siswa SMAN 11 Bandung sebanyak 144 respon, SMAN 18 Bandung sebanyak 139 respon, dan SMAN 23 Bandung sebanyak 144 respon, sehingga didapatkan jumlah sampel respon siswa sebanyak 427 respon.

Teknik Analisis Data

Validitas

Analisis validitas dilakukan untuk melihat apakah tes sebagai alat penilaian dapat digunakan untuk meramalkan suatu ciri atau perilaku tertentu. Dalam pengujian validitas menggunakan program Quest, dapat dilihat pada rekapitulasi melalui hasil output program Quest dengan format file *NamaFile.it.out* dengan membandingkan melalui kriteria sebagai berikut:

Tabel 2. Kriteria Nilai INFIT MNSQ

Nilai INFIT MNSQ	Keterangan
>1.33	Tidak cocok dengan model
0.77 – 1.33	Cocok dengan model
<0.77	Tidak cocok dengan model

Reliabilitas

Analisis reliabilitas dilakukan untuk menunjukkan bahwa suatu instrumen cukup dapat dipercaya untuk dapat digunakan sebagai alat pengumpul data karena instrumen tersebut sudah baik (Arikunto, 2013). Dalam pengujian reliabilitas menggunakan program Quest, dapat dilihat pada rekapitulasi melalui hasil output program Quest dengan format file *NamaFile.sh.out* akan mendapatkan hasil estimasi reliabilitas item dan estimasi reliabilitas

subjek. Bentuk koefisien secara numerik sebesar $-1.00 \leq \rho \leq +1.00$. Semakin mendekati 1 maka reliabilitasnya semakin tinggi dan bila koefisien reliabilitas $+1.00$ menandakan reliabilitasnya sempurna. Tes yang digunakan untuk membuat keputusan pada siswa secara perseorangan, harus memiliki koefisien reliabilitas item minimal sebesar 0.85 (Mehrens & Lehmann dalam Retnawati, 2017).

Kualitas Butir Soal

Uji Prasyarat

Uji Unidimensi

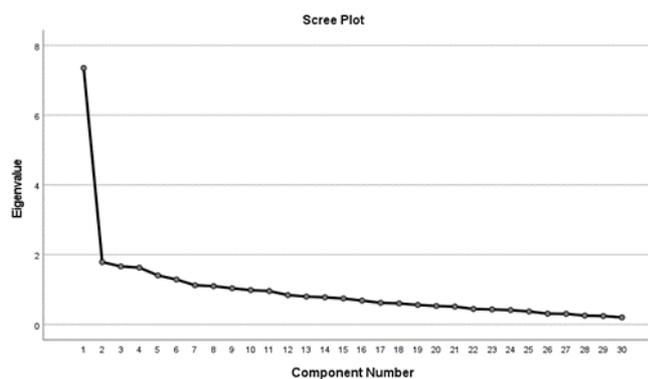
Uji unidimensi dilakukan untuk mengetahui bahwa tiap butir soal tes hanya mengukur satu kemampuan peserta tes. Uji ini terpenuhi jika tes terbukti hanya mengukur satu dimensi yang dominan atau kemampuan yang sama (Hambleton & Swaminathan, 1985). Unidimensi dapat diketahui melalui nilai Eigen matrik korelasi antar butir pada setiap tes yang kemudian dibuat grafik nilai Eigen terhadap jumlah komponen. Bila terdapat satu komponen paling dominan dibandingkan faktor lain, maka dapat dikatakan terdapat satu faktor dominan yang mendasari para peserta dalam memberikan respon pada butir item (Ramalis, 2016).

Dilakukan analisis untuk data dari SMA Negeri 11 Bandung. Berdasarkan hasil analisis dengan kecukupan sampel menunjukkan nilai khi-kuadrat pada uji Barlett sebesar 1330.114 dengan derajat kebebasan 435 dan nilai signifikansi <0.001 . Hasil ini menunjukkan bahwa ukuran sampel data set soal SMA Negeri 11 Bandung sebanyak 144 yang digunakan pada penelitian ini telah cukup dikarenakan syarat terpenuhi bila nilai KMO lebih dari 0.5.

Tabel 3. Hasil Uji KMO dan Barlett SMA Negeri 11 Bandung

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.806
Bartlett's Test of Sphericity	Approx. Chi-Square	1330.114
	Df	435
	Sig.	<.001

Berdasarkan hasil analisis faktor, diperjelas dengan scree plot yang menunjukkan bahwa terdapat 1 faktor yang paling dominan, sehingga hal ini menunjukkan data memenuhi syarat unidimensi.



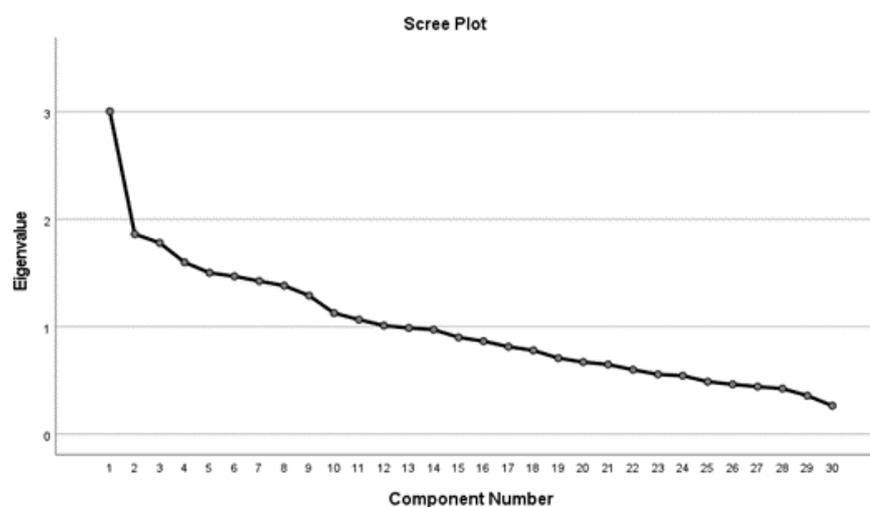
Gambar 1. Scree Plot SMA Negeri 11 Bandung

Dilakukan analisis untuk data dari SMA Negeri 18 Bandung. Berdasarkan hasil analisis dengan kecukupan sampel menunjukkan nilai khi-kuadrat pada uji Barlett sebesar 590.045 dengan derajat kebebasan 435 dan nilai signifikansi <0.001 . Hasil ini menunjukkan bahwa ukuran sampel data set soal SMA Negeri 11 Bandung sebanyak 139 yang digunakan pada penelitian ini telah cukup dikarenakan syarat terpenuhi bila nilai KMO lebih dari 0.5.

Tabel 4. Hasil Uji KMO dan Barlett SMA Negeri 18 Bandung

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.542
Bartlett's Test of Sphericity	Approx. Chi-Square	569.045
	Df	435
	Sig.	<.001

Berdasarkan hasil analisis faktor, diperjelas dengan scree plot yang menunjukkan bahwa terdapat 1 faktor yang paling dominan, serta 2 faktor lainnya yang cukup berpengaruh terhadap komponen varians, sehingga hal ini menunjukkan data memenuhi syarat unidimensi dikarenakan dominan mengukur 1 kemampuan, dapat dilihat melalui grafik *scree plot* sebagai berikut:

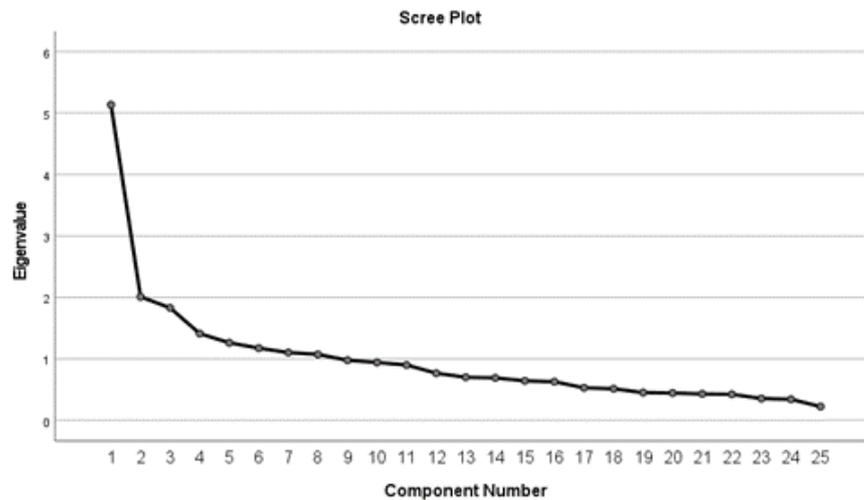
**Gambar 2.** Scree Plot SMA Negeri 18 Bandung

Dilakukan analisis terakhir untuk data dari SMA Negeri 23 Bandung. Berdasarkan hasil analisis dengan kecukupan sampel menunjukkan nilai khi-kuadrat pada uji Barlett sebesar 855.621 dengan derajat kebebasan 300 dan nilai signifikansi <0.001. Hasil ini menunjukkan bahwa ukuran sampel data set soal SMA Negeri 23 Bandung sebanyak 144 yang digunakan pada penelitian ini telah cukup dikarenakan syarat terpenuhi bila nilai KMO lebih dari 0.5.

Tabel 5. Hasil Uji KMO dan Barlett SMA Negeri 23 Bandung

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.763
Bartlett's Test of Sphericity	Approx. Chi-Square	855.621
	Df	300
	Sig.	<.001

Berdasarkan hasil analisis faktor, diperjelas dengan scree plot yang menunjukkan bahwa terdapat 1 faktor yang paling dominan, serta 1 faktor lainnya yang cukup berpengaruh terhadap komponen varians. Hal ini menunjukkan data memenuhi syarat unidimensi dikarenakan dominan mengukur 1 kemampuan, dapat dilihat melalui grafik *scree plot* sebagai berikut:



Gambar 3. Scree Plot SMA Negeri 23 Bandung

Uji Independensi Lokal

Uji Independensi Lokal dilakukan untuk mengetahui apakah benar salahnya peserta tes menjawab sebuah soal tidak dipengaruhi oleh benar salahnya peserta tes lain dalam menjawab benar soal dan tidak dipengaruhi oleh menjawab benar soal yang lain. Dalam teori responsi butir, terpenuhinya asumsi undiminesi melalui analisis faktor tersebut, secara tidak langsung juga sudah membuktikan terpenuhinya asumsi independensi lokal (Retnawati, 2014). Bila tidak memenuhi unidimensi, maka secara statistika. Pada penelitian ini telah memenuhi uji unidimensi pada data set soal 3 sekolah, sehingga syarat uji independensi lokal otomatis terbukti tanpa harus melakukan uji secara statistika.

Uji Invariansi Parameter

Uji Invariansi Parameter dilakukan untuk mengetahui parameter suatu butir tidak tergantung pada peserta tes dan sebaliknya parameter siswa tidak memiliki ketergantungan dengan parameter butir (Duskri & Suryanto, 2014). Invariansi parameter diketahui melalui pembagian peserta tes ke dalam dua kelompok, kemudian secara terpisah diestimasi melalui bentuk scree plot antar keduanya. Apabila sebaran parameter mendekati garis diagonal ($y=x$) maka dapat dikatakan bahwa parameter butir telah memenuhi asumsi invariansi parameter (Retnawati, 2014). Pembuktian pada penelitian ini menggunakan scree plot sebaran butir yang disesuaikan dengan kecocokan parameter butir terhadap model logistik.

Analisis Item Response Theory Parameter Tingkat Kesukaran

Analisis Item Response Theory 2PL digunakan untuk mengukur tingkat kesukaran butir dan daya diskriminasi, melalui persamaan:

$$P_i(\theta) = \frac{e^{Da(\theta-b_i)}}{1 + e^{Da(\theta-b_i)}}$$

Keterangan :

$P_i(\theta)$ = Probabilitas peserta dengan kemampuan θ yang dapat menjawab butir ke i dengan benar

θ = Tingkat kemampuan peserta tes, estimasi variabel laten

e = Bilangan natural yang nilainya mendekati 2,718

D = Faktor skala yang nilainya 1,7

a_i = Daya pembeda butir ke- i

b_i = Tingkat kesukaran butir ke- i

i = 1,2,3 ... n

(Ramalis, 2016)

Dalam pengujian parameter tingkat kesukaran menggunakan program Quest, dapat dilihat pada rekapitulasi melalui hasil output program Quest dengan format file NamaFile.sh.out. dan lebih detail pada hasil output dengan format NamaFile.it.out. (terlampir). Pada parameter tingkat kesukaran, butir terletak pada skala $-\infty \leq b \leq +\infty$. Butir dinyatakan baik bila memiliki tingkat kesukaran yang berkisar antara $-2 \leq b \leq +2$. Butir yang berskala dekat atau di bawah -2 menunjukkan butir soal termasuk dalam kategori mudah, dan bila dekat atau di atas 2 maka butir termasuk dalam kategori sukar (Friyatmi, 2018).

Analisis Item Response Theory Parameter Daya Pembeda

Dalam pengujian parameter daya pembeda/daya diskriminasi menggunakan program Quest, dapat dilihat pada rekapitulasi melalui hasil output program Quest dengan format file NamaFile.tn.out. Pada analisis Item Response Theory, butir terletak pada skala $-\infty \leq a \leq +\infty$. Butir dinyatakan baik bila memiliki daya pembeda yang berkisar antara 0-2 (Friyatmi, 2018).

Klasifikasi Kualitas Butir Soal

Menurut teori responsi butir, kualitas butir soal dapat ditentukan melalui kriteria sebagai berikut:

Tabel 6. Klasifikasi Kualitas Butir Soal

Kriteria	Kecocokan dengan model	Tingkat Kesukaran	Daya Pembeda
Baik	Cocok	$-2 \leq b \leq 2$ / Baik	$0 \leq a \leq 2$ / Baik
Kurang Baik	Cocok	$b > 2$ atau $b < -2$ / Mudah atau Sukar	$a > 2$ atau $a < -2$ / Kurang Baik
Tidak Baik	Tidak Cocok	$b > 2$ atau $b < -2$ $-2 \leq b \leq 2$	$a > 2$ atau $a < -2$ $-2 \leq a \leq 2$

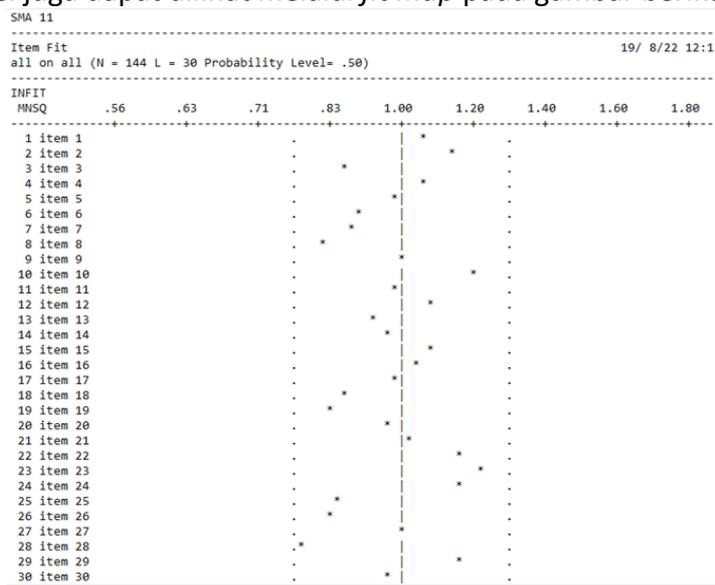
3. RESULTS AND DISCUSSION

Analisis Data

Validitas

SMA Negeri 11 Bandung

Melalui hasil output data set soal dari SMA Negeri 11 Bandung yang menghitung jumlah respons siswa sebanyak 144 dengan jumlah item sebanyak 30, dan tingkat probabilitas 0.50, mendapatkan informasi rentang nilai INFIT MNSQ berada antara 0.78 - 1.23, hal tersebut memberikan informasi bahwa seluruh item fit atau cocok dengan model. Selain itu, untuk melihat kecocokan model juga dapat dilihat melalui *fit map* pada gambar berikut:

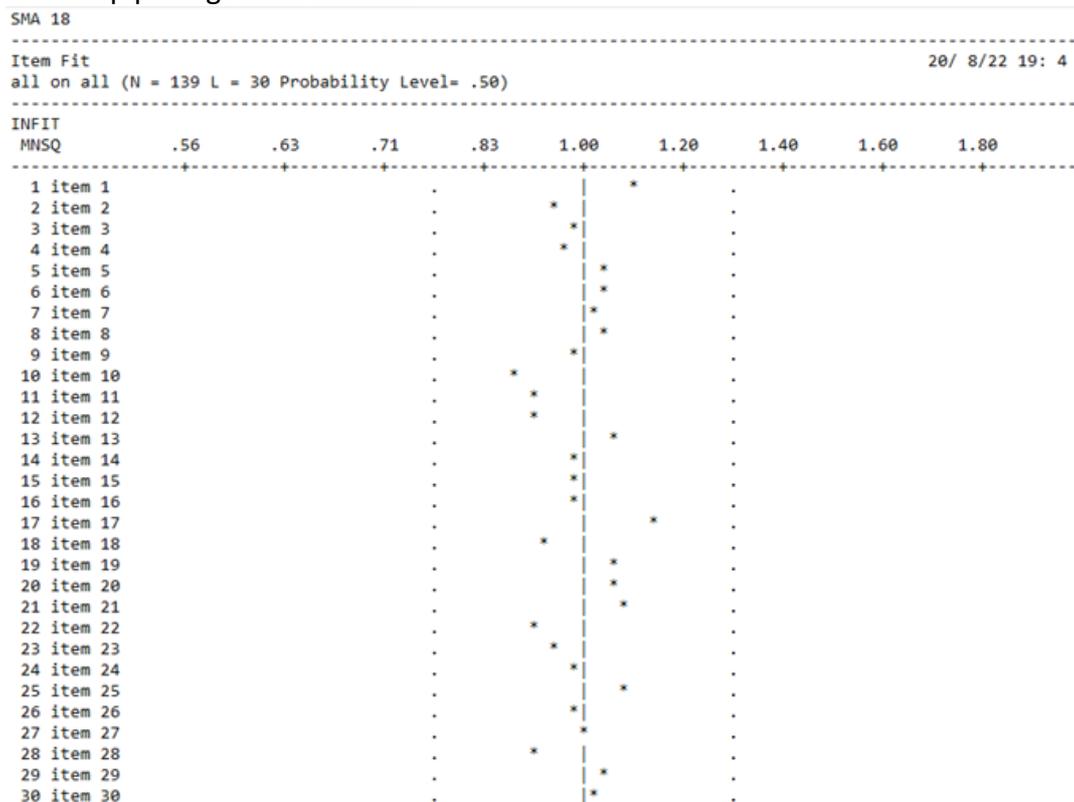


Gambar 4. Fit Map Soal SMA Negeri 11 Bandung

Bila dilihat melalui fit map atau persebaran fit pada gambar diatas dapat diketahui seluruh item berada pada rentang nilai INFIT MNSQ 0.77 – 1.30. Pada sebelah kiri dan kanan terdapat titik-titik yang menunjukkan nilai 0.77 di sebelah kiri, dan 1.33 di sebelah kanan. Bila dilihat secara sederhananya, dapat diketahui dengan melihat persebaran item (tanda bintang), seluruhnya berada di dalam tanda titik-titik tersebut (sebelah kiri dan kanan), tidak ada yang diluar titik sebelah kiri maupun kanan, sehingga dapat langsung disimpulkan bahwa keseluruhan item tersebut fit/cocok dengan model. Dalam analisis teori responsi butir bila mendapatkan hasil seluruh item fit dengan model, maka hal tersebut juga berarti bahwa seluruh item dinyatakan valid.

SMA Negeri 18 Bandung

Melalui hasil output data set soal dari SMA Negeri 18 Bandung yang menghitung jumlah respons siswa sebanyak 139 dengan jumlah item sebanyak 30, dan tingkat probabilitas 0.50, mendapatkan informasi rentang nilai INFIT MNSQ berada antara 0.87 - 1.13, hal tersebut memberikan informasi bahwa seluruh item fit atau cocok dengan model. Bila dilihat melalui fit map atau persebaran fit pada gambar diatas dapat diketahui seluruh item berada pada rentang nilai INFIT MNSQ 0.77 – 1.30. Pada sebelah kiri dan kanan terdapat titik-titik yang menunjukkan nilai 0.77 di sebelah kiri, dan 1.33 di sebelah kanan. Bila dilihat secara sederhananya, dapat diketahui dengan melihat persebaran item (tanda bintang), seluruhnya berada di dalam tanda titik-titik tersebut, tidak ada yang diluar titik sebelah kiri maupun kanan, sehingga dapat langsung disimpulkan bahwa keseluruhan item tersebut fit/cocok dengan model. Dalam teori responsi butir bila mendapatkan hasil seluruh item fit dengan model, maka hal tersebut juga berarti bahwa seluruh item dinyatakan valid, dapat dilihat melalui fit map pada gambar berikut:



Gambar 5. Fit Map Soal SMA Negeri 18 Bandung

Bila dilihat melalui fit map atau persebaran fit pada gambar diatas dapat diketahui sebanyak 20 item berada pada rentang nilai INFIT MNSQ 0.77 – 1.30 yang berarti fit dengan model dan sebanyak 5 item berada diluar rentang nilai tersebut yang berarti item tidak fit dengan model. Bila dilihat secara sederhananya, dapat diketahui dengan melihat persebaran item (tanda

bintang), terdapat 20 bintang yang berada di dalam titik-titik sebelah kiri dan kanan, dan item yang berada diluar titik sebelah kiri sebanyak 3 item yaitu item nomor 1,5 dan 10, dan diluar titik sebelah kanan sebanyak 2 item yaitu nomor 3 dan 11. Hal tersebut menunjukkan bahwa terdapat 5 item yang tidak fit dengan model. Maka disimpulkan bahwa sebanyak 20 item dinyatakan valid dan 5 item dinyatakan tidak valid.

Melalui hasil pengujian validitas dari 3 sekolah, dapat dilihat melalui tabel berikut:

Tabel 7. Rangkuman Validitas 3 Set Soal

Nama Sekolah	Jumlah Soal	Valid	Tidak Valid	Validitas (%)
SMAN 11	30	30	0	100%
SMAN 18	30	30	0	100%
SMAN 23	25	20	5	80%

Berdasarkan tabel tersebut dapat dilihat bahwa set soal SMA Negeri 11 dan SMA Negeri 18 Bandung dinyatakan 100% valid, dan untuk set soal pada SMA Negeri 23 Bandung validitasnya sebesar 80%.

Reliabilitas

SMA Negeri 11 Bandung

Berdasarkan data set soal dari SMA Negeri 11 Bandung mendapatkan nilai estimasi item (*item reliability*) sebesar 0,95 dan nilai reliabilitas berdasarkan subjek (*cases reliability*) sebesar 0,84. Melalui nilai tersebut didapatkan informasi bahwa kualitas item dalam aspek reliabilitasnya sangat tinggi karena dalam reliabilitas jika nilai semakin mendekati angka 1 maka reliabilitasnya semakin tinggi dan konsistensi jawaban dari subjek tinggi. Kemudian untuk set soal SMA Negeri 11 Bandung dapat digunakan untuk menentukan keputusan pada siswa secara perorangan karena nilai reliabilitas item nya lebih dari 0.85. Selain itu, mendapatkan informasi tambahan bahwa tidak ada item yang dijawab benar oleh semua siswa, namun terdapat 11 siswa yang dapat menjawab seluruh soal dengan benar. Hal tersebut dapat dilihat melalui tabel berikut:

Tabel 8. Reliabilitas SMAN 11 Bandung

Reliabilitas	Nilai	Informasi Tambahan
<i>Reliability of Item Estimate</i>	0.95	<i>0 items with perfect scores</i>
<i>Reliability of Case Estimate</i>	0.84	<i>11 cases with perfect scores</i>

SMA Negeri 18 Bandung

Berdasarkan data set soal dari SMA Negeri 18 Bandung mendapatkan nilai estimasi item (*item reliability*) sebesar 0,95 dan nilai reliabilitas berdasarkan subjek (*cases reliability*) sebesar 0,52. Melalui nilai tersebut didapatkan informasi bahwa kualitas item dalam aspek reliabilitasnya sangat tinggi karena dalam reliabilitas jika nilai semakin mendekati angka 1 maka reliabilitasnya semakin tinggi dan konsistensi jawaban dari subjek lemah.

Set soal SMA Negeri 18 Bandung dapat digunakan untuk menentukan keputusan pada siswa secara perorangan karena nilai reliabilitas item nya lebih dari 0.85. Selain itu, mendapatkan informasi tambahan bahwa tidak ada item yang dijawab benar oleh semua siswa, dan tidak ada siswa yang mampu menjawab semua item dengan benar. Hal tersebut dapat dilihat melalui tabel berikut:

Tabel 9. Reliabilitas SMAN 18 Bandung

Reliabilitas	Nilai	Informasi Tambahan
<i>Reliability of Item Estimate</i>	0.95	<i>0 items with perfect scores</i>
<i>Reliability of Case Estimate</i>	0.52	<i>0 cases with perfect scores</i>

SMA Negeri 23 Bandung

Berdasarkan data set soal dari SMA Negeri 23 Bandung mendapatkan nilai estimasi item (*item reliability*) sebesar 0,97 dan nilai reliabilitas berdasarkan subjek (*cases reliability*) sebesar 0,78. Melalui nilai tersebut didapatkan informasi bahwa kualitas item dalam aspek reliabilitasnya sangat tinggi dan hampir sempurna karena dalam reliabilitas angka 1 mengartikan sempurna dan konsistensi jawaban dari subjek dinyatakan cukup. Kemudian untuk set soal SMA Negeri 23 Bandung dapat digunakan untuk menentukan keputusan pada siswa secara perorangan karena nilai reliabilitas item nya lebih dari 0.85. Selain itu, mendapatkan informasi tambahan bahwa tidak ada item yang dijawab benar oleh semua siswa, dan tidak ada siswa yang mampu menjawab semua item dengan benar. Hal tersebut dapat dilihat melalui tabel berikut:

Tabel 10. Reliabilitas SMA 23 Bandung

Reliabilitas	Nilai	Informasi Tambahan
<i>Reliability of Item Estimate</i>	0.97	0 items with perfect scores
<i>Reliability of Case Estimate</i>	0.78	0 cases with perfect scores

Berdasarkan pengujian 3 set soal tersebut dapat ditarik kesimpulan bahwa ketiga set soal dari SMA Negeri Bandung memiliki nilai reliabilitas item yang tinggi dan set soal dari SMA Negeri 23 Bandung hampir sempurna, hal tersebut dikarenakan semakin mendekati angka 1 maka reliabilitasnya semakin tinggi, dan angka 1 pada reliabilitas dinyatakan sempurna. Ketiga set soal tersebut dapat digunakan untuk menentukan keputusan pada siswa secara perorangan karena nilai reliabilitas item nya lebih dari 0.85. Kemudian untuk reliabilitas subjek dapat diketahui bahwa SMA Negeri 11 Bandung mendapatkan nilai yang paling tinggi dan nilai reliabilitas subjek terendah pada SMA Negeri 18 Bandung, dan ketiganya memiliki nilai antara -1.00 sampai 1.00 yang berarti reliabilitas kemampuan siswa dalam menjawab soal tersebut adalah sedang.

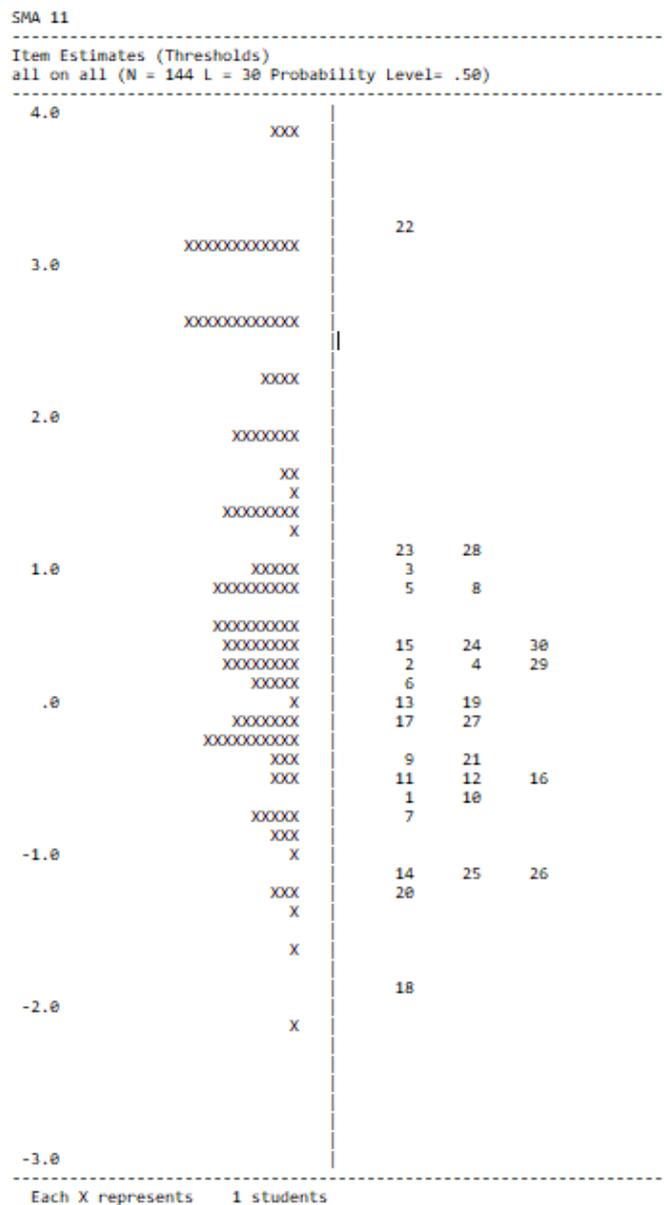
Tabel 11. Rangkuman Reliabilitas 3 Set Soal

Nama Sekolah Set Soal	<i>Reliability of Items Estimate</i>	<i>Reliability of Cases Estimate</i>
SMA Negeri 11	0.95	0.84
SMA Negeri 18	0.95	0.52
SMA Negeri 23	0.97	0.78
Interpretasi	Sempurna	Sedang

Parameter Tingkat Kesukaran

SMA Negeri 11 Bandung

Pada soal SMA Negeri 11, didapatkan informasi bahwa soal paling sukar adalah item nomor 22 dengan kategori sangat sukar karena nilainya melebihi 2.0 dan item yang paling mudah ialah item nomor 18 yang termasuk ke dalam kategori mudah.



Gambar 7. Tingkat Kesukaran Soal SMAN 11 Bandung

Berdasarkan gambar tersebut, dapat dilihat bahwa terdapat 3 tanda X yang letaknya lebih tinggi bila dibandingkan dengan nomor item yang paling sukar yaitu nomor 22, hal tersebut menandakan bahwa terdapat 3 subjek/siswa yang memiliki kemampuan lebih tinggi dibandingkan kemampuan yang harus dimiliki untuk menjawab item nomor 22, yang berarti 3 subjek tersebut dapat mengerjakan item yang paling sukar dengan benar. Sebaliknya dapat dilihat bahwa terdapat 1 tanda X yang letaknya paling rendah dan berada di bawah bila dibandingkan dengan nomor item yang paling mudah yaitu item nomor 18, hal tersebut menandakan bahwa terdapat 1 subjek/siswa tidak memiliki kemampuan untuk menjawab item dengan benar bahkan pada item yang paling mudah. Begitupun dengan penafsiran antara butir soal yang lain dengan kemampuan subjek.

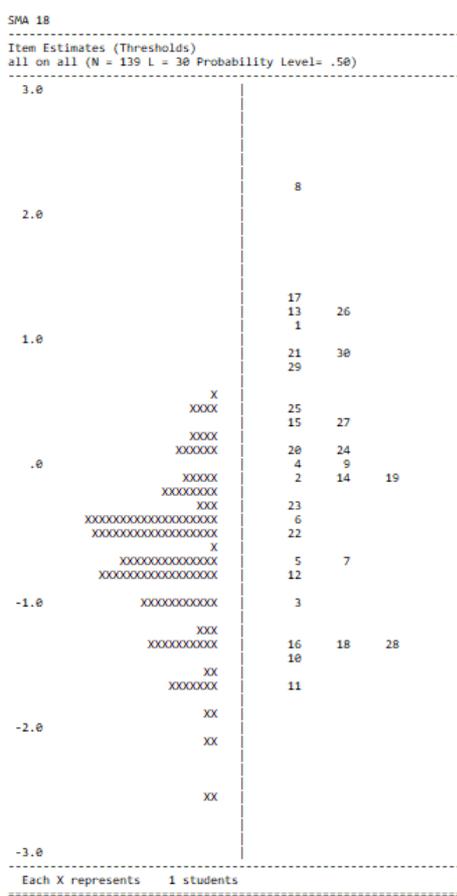
Melalui nilai THRSH diketahui bahwa terdapat 1 item dalam kategori sukar yaitu pada nomor item 22 atau sebesar 3.33% dari keseluruhan item. Terdapat 14 item yang memiliki tingkat kesukaran yang baik atau sebesar 46.67% yaitu pada nomor item 2, 4, 3, 5, 6, 8, 13, 15, 19, 23, 24, 28, 29, dan 30. Kemudian sebesar 50% atau setengah item dari tes tersebut memiliki tingkat kesukaran yang mudah, yaitu pada nomor item 1, 7, 9, 10, 11, 12, 14, 16, 17, 18, 20, 21, 25, 26, dan 27.

Tabel 12. Klasifikasi Parameter Tingkat Kesukaran Soal SMA Negeri 11

Klasifikasi	Kategori	Jumlah Item	Nomor Item	Persentase
> 2.0	Sukar	1	22	3.33%
0 – 2.0	Baik	14	2, 3, 4, 5, 6, 8, 13, 15, 19, 23, 24, 28, 29, 30	46.67%
< 0	Mudah	15	1, 7, 9, 10, 11, 12, 14, 16, 17, 18, 20, 21, 25, 26, 27	50%

SMA Negeri 18 Bandung

Hasil pengujian soal SMA Negeri 18 Bandung, didapatkan informasi bahwa soal yang paling sukar adalah item nomor 8 dengan kategori sangat sukar karena nilainya melebihi 2.0 dan item yang paling mudah ialah item nomor 11 yang termasuk ke dalam kategori mudah.

**Gambar 8.** Tingkat Kesukaran Soal SMA Negeri 18 Bandung

Berdasarkan pengujian soal SMA Negeri 23 Bandung, dapat dilihat bahwa terdapat 1 tanda X yang letaknya paling tinggi dibandingkan dengan tanda X yang lain. Namun, letak X tertinggi tersebut berada lebih rendah dibandingkan dengan nomor item yang letaknya lebih tinggi, terutama pada item yang memiliki tingkat kesukaran tertinggi yaitu item nomor 8. Tidak ada tanda X (subjek) yang letaknya lebih tinggi dibandingkan item nomor 8, maka peluang item tersebut untuk dijawab dengan benar oleh subjek hampir tidak ada. Begitupun dapat dilihat terdapat 2 tanda X yang letaknya paling rendah, berada lebih rendah pula bila dibandingkan dengan nomor item yang paling mudah yaitu nomor 11. Kedua hal tersebut menandakan bahwa secara keseluruhan subjek lebih rendah dibandingkan dengan itemnya, dengan kata lain secara keseluruhan kemampuan subjek lebih rendah dibandingkan tingkat kesukaran item. Maka dapat disimpulkan bahwa diperlukan subjek yang memiliki

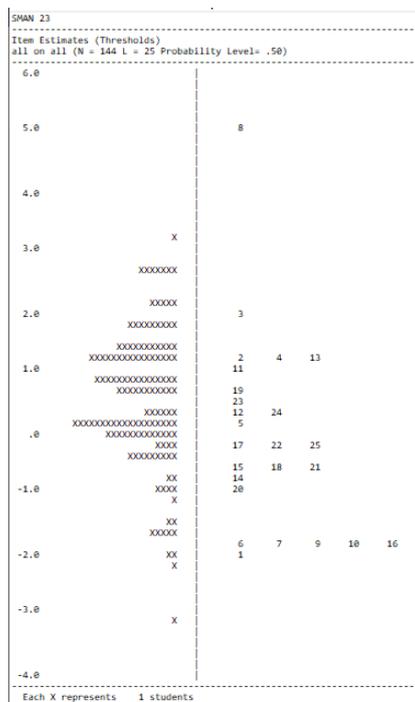
kemampuan lebih tinggi untuk dapat mengerjakan item tes tersebut ataupun item tersebut lebih baik digunakan kepada subjek yang memiliki kemampuan lebih tinggi. Melalui nilai THRSH diketahui bahwa terdapat 1 item (3,33%) yaitu pada nomor item 8 dinyatakan sukar, terdapat 14 item item (46.67%) dinyatakan baik yaitu pada nomor item 1,4, 9, 13, 15, 17, 20, 21, 24, 25, 26, 27, 29, dan 30. Kemudian sebesar 50% atau setengah item dari tes tersebut memiliki tingkat kesukaran yang mudah, yaitu pada nomor item 2, 3, 5, 6, 7, 10, 11, 12, 14, 16, 18, 19, 22, 23, dan 28.

Tabel 13. Klasifikasi Parameter Tingkat Kesukaran Soal SMA Negeri 18

Klasifikasi	Kategori	Jumlah Item	Nomor Item	Persentase
> 2.0	Sukar	1	8	3.33%
0 – 2.0	Baik	14	1,4, 9, 13, 15, 17, 20, 21, 24, 25, 26, 27, 29, 30	46.67%
< 0	Mudah	15	2, 3, 5, 6, 7, 10, 11, 12, 14, 16, 18, 19, 22, 23, 28	50%

SMA Negeri 23 Bandung

Hasil pengujian data set soal SMA Negeri 23 Bandung dapat dilihat melalui grafik berikut ini:



Gambar 9. Tingkat Kesukaran Soal SMA Negeri 23 Bandung

Berdasarkan grafik didapatkan informasi bahwa soal yang paling sukar adalah item nomor 8 dengan kategori sangat sukar karena nilainya melebihi 2.0 dan item yang paling mudah ialah item nomor 1 yang termasuk ke dalam kategori mudah. Sama halnya seperti pengujian soal SMA Negeri 18, pada grafik SMA Negeri 23 pun tidak ada tanda X yang letaknya lebih tinggi dibandingkan item tersulit (nomor 8), dan terdapat 1 tanda X yang letaknya lebih rendah bila dibandingkan dengan nomor item yang termudah (nomor 1). Secara keseluruhan diketahui kemampuan subjek lebih rendah dibandingkan tingkat kesukaran item. Maka diperlukan subjek yang memiliki kemampuan lebih tinggi untuk dapat mengerjakan item tes ataupun item item tersebut lebih baik dikerjakan oleh subjek yang memiliki kemampuan lebih tinggi.

Melalui nilai THRSH diketahui bahwa terdapat 1 item atau sebesar 4% yang memiliki tingkat kesukaran sukar yaitu pada item nomor 8. Sebesar 40% atau sebanyak 10 item memiliki tingkat kesukaran yang baik yaitu pada nomor 2, 3, 4, 5, 11, 12, 13, 19, 23, dan 24. Kemudian sebesar 56% atau lebih dari setengah item dari tes tersebut memiliki tingkat kesukaran yang mudah, yaitu pada nomor 1, 6, 7, 9, 10, 14, 15, 18, 16, 17, 20, 21, 22, dan 25.

Tabel 14. Klasifikasi Parameter Tingkat Kesukaran Soal SMA Negeri 23

Klasifikasi	Kategori	Jumlah Item	Nomor Item	Persentase
> 2.0	Sukar	1	8	4%
0 – 2.0	Baik	10	2, 3, 4, 5, 11, 12, 13, 19, 23, 24	40%
< 0	Mudah	14	1, 6, 7, 9, 10, 14, 15, 18, 16, 17, 20, 21, 22, 25	56%

Berdasarkan hasil pengujian 3 set soal tersebut, diketahui bahwa sekolah yang memiliki soal sangat sukar adalah SMA Negeri 23 Bandung sebanyak 1 soal dengan nilai THRSH (tingkat kesukaran) tertinggi yaitu sebesar 5.18 dan SMA Negeri 18 memiliki nilai THRSH (tingkat kesukaran) terendah untuk soal yang paling sukar sebesar 2.20. Hasil dari pengujian ketiga sekolah tersebut menunjukkan bahwa lebih banyak butir soal yang memiliki tingkat kesukaran yang mudah, yaitu sebanyak 51.76% dari keseluruhan butir gabungan sebanyak 85 soal. Rangkuman tersebut dapat dilihat melalui tabel berikut:

Tabel 15. Rangkuman Tingkat Kesukaran 3 Set Soal

Nama Sekolah Set Soal	Jumlah Soal	Tingkat Kesukaran		
		Sukar	Baik	Mudah
SMA Negeri 11	30	1	14	15
SMA Negeri 18	30	1	14	15
SMA Negeri 23	25	1	10	14
Persentase Keseluruhan		3.52%	44.7%	51.76%

Parameter Daya Pembeda

SMA Negeri 11 Bandung

Melalui pengujian dari 30 soal SMA Negeri 11 Bandung mendapatkan hasil bahwa besaran nilai p_t biserial bergerak pada angka 0.29-0.68, menandakan bahwa seluruh butir dinyatakan baik. Secara lebih rinci nilai p_t -biserial per item dapat dilihat melalui tabel berikut:

Tabel 16. Daya Pembeda Soal SMA Negeri 11 Bandung

Item	Nilai p_t -biserial	Keterangan	Item	Nilai p_t -biserial	Keterangan
1	0.40	Baik	16	0.43	Baik
2	0.44	Baik	17	0.52	Baik
3	0.65	Baik	18	0.42	Baik
4	0.47	Baik	19	0.62	Baik
5	0.57	Baik	20	0.40	Baik
6	0.59	Baik	21	0.45	Baik
7	0.54	Baik	22	0.33	Baik
8	0.66	Baik	23	0.42	Baik
9	0.47	Baik	24	0.42	Baik
10	0.29	Baik	25	0.50	Baik
11	0.48	Baik	26	0.51	Baik

12	0.39	Baik	27	0.49	Baik
13	0.55	Baik	28	0.68	Baik
14	0.43	Baik	29	0.40	Baik
15	0.47	Baik	30	0.54	Baik

Bila ditinjau lebih dalam lagi, hasil nilai *pt-biserial* atau dalam artian daya pembeda, terdapat hal yang mempengaruhi nilai tersebut yaitu efektivitas distraktor pada pilihan jawaban selain kunci jawaban. Pada set soal SMAN 11 Bandung tidak ada *pt-biserial* kunci jawaban yang memiliki nilai dibawah 0.20, hal tersebut berarti pilihan jawaban tidak perlu diperbaiki. Kemudian bila melihat *pt-biserial* seluruh pilihan jawaban selain kunci jawaban, mendapatkan informasi bahwa dari 30 item terdapat 4 item yang distraktornya belum berfungsi dengan baik, yaitu pada nomor item 1, 14, 18, dan 22, dikarenakan ada pilihan jawaban pengecoh yang tidak dipilih oleh subjek dan terdapat 26 item yang memiliki keberfungsian distraktor baik karena keseluruhan pilihan pengecoh dipilih oleh subjek. Dapat disimpulkan bahwa item yang memiliki distraktor yang berfungsi dengan baik sebesar 86.67% dan sebesar 13.33% distraktornya belum berfungsi dengan baik.

Tabel 17. Efektivitas Distraktor Soal SMA Negeri 11 Bandung

Kategori	Jumlah Item	Persentase
Berfungsi dengan baik	26	86.67%
Belum berfungsi dengan baik	4	13.33%

SMA Negeri 18 Bandung

Melalui pengujian dari 30 soal SMA Negeri 18 Bandung mendapatkan hasil bahwa besaran nilai *pt biserial* bergerak pada angka -0.20-0.47. Dapat disimpulkan bahwa hampir keseluruhan item yaitu sebanyak 27 item memiliki daya pembeda yang baik dengan nilai *pt-biserial* tertinggi adalah 0.47 pada item nomor 10 dan 3 item memiliki daya pembeda yang tidak baik dikarenakan nilai *pt-biserial* mendapatkan hasil yang negatif yaitu item nomor 1 dengan nilai -0.03, item nomor 8 dengan nilai -0.08, dan item nomor 17 dengan nilai -0.20.

Tabel 18. Daya Pembeda Soal SMA Negeri 18 Bandung

Item	Nilai <i>pt-biserial</i>	Keterangan	Item	Nilai <i>pt-biserial</i>	Keterangan
1	-0.03	Tidak Baik	16	0.33	Baik
2	0.38	Baik	17	-0.20	Tidak Baik
3	0.33	Baik	18	0.41	Baik
4	0.32	Baik	19	0.18	Baik
5	0.22	Baik	20	0.15	Baik
6	0.20	Baik	21	0.04	Baik
7	0.27	Baik	22	0.43	Baik
8	-0.08	Tidak Baik	23	0.36	Baik
9	0.30	Baik	24	0.28	Baik
10	0.47	Baik	25	0.09	Baik
11	0.38	Baik	26	0.26	Baik
12	0.44	Baik	27	0.23	Baik
13	0.04	Baik	28	0.42	Baik
14	0.28	Baik	29	0.11	Baik
15	0.27	Baik	30	0.17	Baik

Berdasarkan hasil analisis, parameter daya pembeda pada soal SMA Negeri 18 Bandung didominasi dengan butir soal dengan kategori baik dengan jumlah item sebanyak 27, dan terdapat 3 item yang termasuk ke dalam kategori tidak baik.

Tabel 19. Klasifikasi Parameter Daya Pembeda Soal SMA Negeri 18 Bandung

Klasifikasi	Kategori	Jumlah Item	Persentase
0 – 2.00	Baik	27	90%
<0	Tidak Baik	3	10%
>2.00	Tidak Baik	-	0%

Bila ditinjau berdasarkan efektivitas distractor, pada set soal SMA Negeri 18 Bandung terdapat 8 item dengan pt-biserial kunci jawaban yang memiliki nilai dibawah 0.20 yaitu pada item nomor 13, 19, 20, 21, 25, 29, dan 30. Hal tersebut berarti nomor item tersebut masih perlu diperbaiki. Terdapat 4 item yang distraktornya belum berfungsi dengan baik, yaitu nomor 5, 6, 14, dan 15, dikarenakan ada pilihan jawaban pengecoh yang tidak dipilih oleh subjek dan terdapat 26 item yang memiliki keberfungsian distraktor baik.

Tabel 20. Efektivitas Distraktor Soal SMA Negeri 18 Bandung

Kategori	Jumlah Item	Persentase
Berfungsi dengan baik	26	86.67%
Belum berfungsi dengan baik	4	13.33%

SMA Negeri 23 Bandung

Melalui pengujian dari 25 soal SMA Negeri 23 Bandung mendapatkan hasil bahwa besaran nilai pt biserial bergerak pada angka -0.09-0.68, sehingga dapat diketahui bahwa hampir keseluruhan item yaitu sebanyak 24 item memiliki daya pembeda yang baik dengan nilai pt-biserial tertinggi adalah 0.68 pada item nomor 5 dan 1 item memiliki daya pembeda yang tidak baik dikarenakan nilai pt-biserial mendapatkan hasil yang negatif yaitu item nomor 3 dengan nilai -0.09, dapat dilihat melalui tabel berikut:

Tabel 21. Daya Pembeda Soal SMA Negeri 23 Bandung

Item	Nilai pt-biserial	Keterangan	Item	Nilai pt-biserial	Keterangan
1	0.61	Baik	14	0.44	Baik
2	0.32	Baik	15	0.43	Baik
3	-0.09	Tidak Baik	16	0.49	Baik
4	0.28	Baik	17	0.47	Baik
5	0.68	Baik	18	0.36	Baik
6	0.53	Baik	19	0.38	Baik
7	0.32	Baik	20	0.47	Baik
8	0.02	Baik	21	0.60	Baik
9	0.46	Baik	22	0.51	Baik
10	0.52	Baik	23	0.35	Baik
11	0.11	Baik	24	0.48	Baik
12	0.36	Baik	25	0.47	Baik
13	0.53	Baik			

Bila diklasifikasikan secara keseluruhan dapat diketahui bahwa terdapat 24 item dalam kategori baik dan 1 item tidak baik, dapat dilihat melalui tabel berikut:

Tabel 22. Klasifikasi Parameter Daya Pembeda Soal SMA Negeri 23 Bandung

Klasifikasi	Kategori	Jumlah Item	Persentase
0 – 2.00	Baik	24	96%
<0	Tidak Baik	1	4%
>2.00	Tidak Baik	-	0%

Pada set soal SMA Negeri 23 Bandung terdapat 2 item dengan ρ -biserial kunci jawaban yang memiliki nilai dibawah 0.20 yaitu pada item nomor 8 dan 11. Hal itu berarti nomor item tersebut masih perlu diperbaiki. Kemudian terdapat 6 item yang distraktornya belum berfungsi dengan baik, yaitu pada nomor item 2, 7, 8, 12, 14, dan 20, dikarenakan ada pilihan jawaban pengecoh yang tidak dipilih oleh subjek dan terdapat 19 item yang memiliki keberfungsian distraktor baik karena keseluruhan pilihan pengecoh dipilih oleh subjek. Dapat disimpulkan bahwa item yang memiliki distraktor yang berfungsi dengan baik sebesar 76% dan sebesar 24% distraktornya belum berfungsi dengan baik.

Tabel 23. Efektivitas Distraktor Soal SMA Negeri 23 Bandung

Kategori	Jumlah Item	Persentase
Berfungsi dengan baik	19	76%
Belum berfungsi dengan baik	6	24%

Bila hasil pengujian 3 set soal tersebut dibandingkan, diketahui bahwa seluruh butir soal SMA Negeri 11 Bandung memiliki daya pembeda yang baik sebesar 100%, dan yang terendah pada SMA Negeri 18 Bandung sebesar 90%, begitupun sebaliknya yang memiliki jumlah item dengan daya pembeda tidak baik terbanyak adalah SMA Negeri 18 Bandung yaitu sebesar 10%, dan terendah pada SMA Negeri 11 Bandung sebesar 0%. Bila dilihat dari keberfungsian distraktor, SMA Negeri 11 dan 18 memiliki nilai yang sama sebesar 86.67% dengan distraktor yang berfungsi dengan baik, dan terendah SMA Negeri 23 sebesar 76% distraktor berfungsi dengan baik, begitupun sebaliknya SMA Negeri 23 Bandung memiliki distraktor yang tidak berfungsi dengan baik terbanyak yaitu sebesar 24% dan SMA Negeri 11 dan 18 memiliki nilai 13.33% distraktor yang tidak berfungsi dengan baik.

Tabel 24. Rangkuman Daya Pembeda dan Efektivitas Distraktor 3 Set Soal

Klasifikasi	Kategori	SMAN 11	SMAN 18	SMAN 23
Daya Pembeda	Baik	100%	90%	96%
	Tidak baik	0%	10%	4%
Efektivitas Distraktor	Baik	86.67%	86.67%	76%
	Tidak Baik	13.33%	13.33%	24%

Kualitas Butir Soal

SMA Negeri 11 Bandung

Pada analisis berdasarkan *Item Response Theory* pengujian 2 Parameter Logistik (2PL), kualitas butir soal dapat ditentukan melalui kecocokan model, tingkat kesukaran dan daya pembeda. Pada soal SMA Negeri 11 bila ditinjau berdasarkan tingkat kesukaran dan daya pembeda pada masing masing item maka dapat diketahui terdapat 14 butir memiliki kualitas yang baik dan dapat disimpan di bank soal, dan 16 butir lainnya memiliki kualitas yang kurang baik sehingga perlu direvisi, dapat dilihat melalui tabel berikut :

Tabel 25. Kualitas Butir Soal SMA Negeri 11 Bandung

Item	Kecocokan Model	Tingkat Kesukaran	Kriteria	Daya Pembeda	Kriteria	Keputusan
1	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
2	Cocok	Baik	Baik	Baik	Baik	Disimpan
3	Cocok	Baik	Baik	Baik	Baik	Disimpan
4	Cocok	Baik	Baik	Baik	Baik	Disimpan
5	Cocok	Baik	Baik	Baik	Baik	Disimpan

6	Cocok	Baik	Baik	Baik	Baik	Disimpan
7	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
8	Cocok	Baik	Baik	Baik	Baik	Disimpan
9	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
10	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
11	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
12	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
13	Cocok	Baik	Baik	Baik	Baik	Disimpan
14	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
15	Cocok	Baik	Baik	Baik	Baik	Disimpan
16	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
17	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
18	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
19	Cocok	Baik	Baik	Baik	Baik	Disimpan
20	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
21	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
22	Cocok	Sukar	Kurang Baik	Baik	Baik	Direvisi
23	Cocok	Baik	Baik	Baik	Baik	Disimpan
24	Cocok	Baik	Baik	Baik	Baik	Disimpan
25	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
26	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
27	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
28	Cocok	Baik	Baik	Baik	Baik	Disimpan
29	Cocok	Baik	Baik	Baik	Baik	Disimpan
30	Cocok	Baik	Baik	Baik	Baik	Disimpan

SMA Negeri 18 Bandung

Pada analisis berdasarkan *Item Response Theory* pengujian 2 Parameter Logistik (2PL), kualitas butir soal dapat ditentukan melalui kecocokan model, tingkat kesukaran dan daya pembeda. Pada soal SMA Negeri bila ditinjau berdasarkan tingkat kesukaran dan daya pembeda pada masing masing item maka dapat diketahui terdapat 12 butir memiliki kualitas yang baik dan dapat disimpan di bank soal, dan 18 butir lainnya memiliki kualitas yang kurang baik sehingga perlu direvisi, dapat dilihat melalui tabel berikut :

Tabel 26. Kualitas Butir Soal SMA Negeri 18 Bandung

Item	Kecocokan Model	Tingkat Kesukaran	Kriteria	Daya Pembeda	Kriteria	Keputusan
1	Cocok	Baik	Baik	Tidak Baik	Kurang Baik	Direvisi
2	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
3	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
4	Cocok	Baik	Baik	Baik	Baik	Disimpan
5	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
6	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
7	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
8	Cocok	Sukar	Kurang Baik	Tidak Baik	Kurang Baik	Direvisi
9	Cocok	Baik	Baik	Baik	Baik	Disimpan
10	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
11	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
12	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
13	Cocok	Baik	Baik	Baik	Baik	Disimpan
14	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi

15	Cocok	Baik	Baik	Baik	Baik	Disimpan
16	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
17	Cocok	Baik	Baik	Tidak Baik	Kurang Baik	Direvisi
18	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
19	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
20	Cocok	Baik	Baik	Baik	Baik	Disimpan
21	Cocok	Baik	Baik	Baik	Baik	Disimpan
22	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
23	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
24	Cocok	Baik	Baik	Baik	Baik	Disimpan
25	Cocok	Baik	Baik	Baik	Baik	Disimpan
26	Cocok	Baik	Baik	Baik	Baik	Disimpan
27	Cocok	Baik	Baik	Baik	Baik	Disimpan
28	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
29	Cocok	Baik	Baik	Baik	Baik	Disimpan
30	Cocok	Baik	Baik	Baik	Baik	Disimpan

SMA Negeri 23 Bandung

Pada analisis berdasarkan *Item Response Theory* pengujian 2 Parameter Logistik (2PL), kualitas butir soal dapat ditentukan melalui kecocokan model, tingkat kesukaran dan daya pembeda. Pada soal SMA Negeri 23 bila ditinjau berdasarkan tingkat kesukaran dan daya pembeda pada masing masing item maka dapat diketahui terdapat 7 butir memiliki kualitas yang baik dan dapat disimpan di bank soal, 13 butir memiliki kualitas yang kurang baik sehingga perlu direvisi, dan 5 butir memiliki kualitas yang tidak baik sehingga perlu dibuang dan diganti, dapat dilihat melalui tabel berikut :

Tabel 27. Kualitas Butir Soal SMA Negeri 23 Bandung

Item	Kecocokan Model	Tingkat Kesukaran	Kriteria	Daya Pembeda	Kriteria	Keputusan
1	Tidak Cocok	-	Tidak Baik	-	Tidak Baik	Dibuang
2	Cocok	Baik	Baik	Baik	Baik	Disimpan
3	Tidak Cocok	-	Tidak Baik	-	Tidak Baik	Dibuang
4	Cocok	Baik	Baik	Baik	Baik	Disimpan
5	Tidak Cocok	-	Tidak Baik	-	Tidak Baik	Dibuang
6	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
7	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
8	Cocok	Sukar	Kurang Baik	Baik	Baik	Direvisi
9	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
10	Tidak Cocok	-	Tidak Baik	-	Tidak Baik	Dibuang
11	Tidak Cocok	-	Tidak Baik	-	Tidak Baik	Dibuang
12	Cocok	Baik	Baik	Baik	Baik	Disimpan
13	Cocok	Baik	Baik	Baik	Baik	Disimpan
14	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
15	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
16	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
17	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
18	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
19	Cocok	Baik	Baik	Baik	Baik	Disimpan
20	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
21	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi

22	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi
23	Cocok	Baik	Baik	Baik	Baik	Disimpan
24	Cocok	Baik	Baik	Baik	Baik	Disimpan
25	Cocok	Mudah	Kurang Baik	Baik	Baik	Direvisi

Bila membandingkan hasil dari ketiga analisis, soal SMAN 11 memiliki jumlah terbanyak yang memiliki kualitas baik yang dapat disimpan pada bank soal dan hanya pada soal SMAN 23 terdapat butir yang memiliki kualitas yang tidak baik dan harus dibuang.

Tabel 28 . Kualitas Butir Soal SMA Negeri 11, 18, dan 23 Bandung

Nama Sekolah	Kualitas	Keputusan	Jumlah	No Item
SMAN 11	Baik	Disimpan	14	2, 3, 4, 5, 6, 8, 13, 15, 19, 23, 24, 28, 29, 30
	Kurang Baik	Direvisi	16	1, 7, 9, 10, 11, 12, 14, 16, 17, 18, 20, 21, 22, 25, 26, 27
	Tidak Baik	Dibuang	0	-
SMAN 18	Baik	Disimpan	12	4, 9, 13, 15, 20, 21, 24, 25, 26, 27, 29, 30
	Kurang Baik	Direvisi	18	1, 2, 3, 5, 6, 7, 8, 10, 11, 12, 14, 16, 17, 18, 19, 22, 23, 28
	Tidak Baik	Dibuang	0	-
SMAN 23	Baik	Disimpan	7	2, 4, 12, 13, 19, 23, 24
	Kurang Baik	Direvisi	13	6, 7, 8, 9, 14, 15, 16, 17, 18, 20, 21, 22, 25
	Tidak Baik	Dibuang	5	1, 3, 5, 10, 11

PEMBAHASAN

Pada penelitian ini menganalisis mengenai kualitas butir soal Penilaian Akhir Semester Ganjil mata pelajaran ekonomi tahun ajaran 2021/2022 berdasarkan Item Response Theory. Seperti yang dikemukakan oleh Nitko dan Anthony dalam Mahendra (2019), kegiatan analisis butir soal merupakan kegiatan yang wajib dilakukan guru dalam usaha mengetahui tingkat hasil belajar serta berguna untuk informasi peningkatan mutu soal yang telah disusun. Selain itu tes hasil belajar juga haruslah memenuhi syarat sebagai alat ukur yang baik.

Analisis kualitas butir soal yang telah dilakukan berdasarkan Item Response Theory terhadap 3 set soal dengan jumlah item sebanyak 85 butir dan testi sebanyak 427 siswa, diketahui bahwa terdapat 80 butir soal dinyatakan valid dan 5 butir dinyatakan tidak valid, hal tersebut diketahui melalui kecocokan model. Kemudian untuk estimasi nilai reliabilitas item sebesar 0.96 yang menandakan tes tersebut memiliki reliabilitas yang tinggi dan dapat digunakan untuk menentukan keputusan pada siswa secara perorangan karena nilai reliabilitas item nya lebih dari 0.85. Terdapat reliabilitas subjek juga sebesar 0.71 yang menandakan konsistensi dari subjek yang menjawab soal masih perlu ditingkatkan lagi dan didapatkan informasi juga bahwa dari keseluruhan subjek yang berjumlah 427 siswa, terdapat 11 testi/subjek yang memiliki skor sempurna, dengan kata lain 11 subjek tersebut memiliki kemampuan untuk menjawab seluruh soal dengan benar. Berdasarkan parameter tingkat kesukaran, terdapat 3 butir soal dalam kategori sukar, 38 butir soal dengan kategori baik dan 44 butir soal dengan kategori soal mudah. Testi yang mampu menjawab butir soal paling sukar yaitu berjumlah 28, dan testi yang mampu menjawab butir soal paling mudah berjumlah 350 dari total keseluruhan 427 testi. Berdasarkan hasil tingkat kesukaran, peneliti juga menganalisis mengenai materi yang diujikan, bila ditinjau dari segi isi soal berdasarkan Kompetensi Dasar serta indikator, pada SMA Negeri 11 Bandung item yang paling sukar adalah item nomor 22 yang memuat kompetensi dasar menganalisis permasalahan

ketenagakerjaan dalam pembangunan ekonomi, serta menyajikan hasil analisis masalah ketenagakerjaan dalam pembangunan ekonomi dan cara mengatasinya. Item tersebut mengukur indikator menganalisis akibat pengangguran terhadap pembangunan ekonomi. Untuk soal yang paling mudah dari SMA Negeri 11 Bandung yaitu nomor 18 yang memuat Kompetensi Dasar yang sama dengan mengukur indikator menganalisis cara meningkatkan kualitas tenaga kerja. Kemudian pada SMA Negeri 18 Bandung item yang paling sukar adalah item yang mengukur kompetensi dasar yaitu mendeskripsikan konsep dan metode penghitungan pendapatan nasional dalam konteks pertumbuhan dan pembangunan serta menyajikan hasil penghitungan pendapatan nasional dalam konteks pertumbuhan dan pembangunan ekonomi dengan indikator yang diukur ialah mengenai membandingkan pendapatan per kapita setiap negara dengan tepat, dan untuk item yang paling mudah yaitu item nomor 11 yang memuat kompetensi dasar menganalisis permasalahan ketenagakerjaan dalam pembangunan ekonomi dan menyajikan hasil analisis masalah ketenagakerjaan dalam pembangunan ekonomi dan cara mengatasinya dengan indikator yang diukur adalah menjelaskan penyebab perbedaan tingkat upah di setiap daerah dengan tepat. Sedangkan pada SMA Negeri 23 Bandung, item yang paling sukar ialah item nomor 8 yang memuat kompetensi dasar menganalisis inflasi, kebijakan moneter dan kebijakan fiskal serta menyajikan hasil analisis inflasi, kebijakan moneter, dan kebijakan fiskal yang mengukur indikator mengidentifikasi penyebab terjadinya inflasi, dan untuk item yang paling mudah adalah item nomor 1 yang memuat kompetensi dasar yang sama dengan indikator yang diukur adalah mengidentifikasi jenis-jenis inflasi. Sehingga dapat disimpulkan bahwa bila ditinjau dari segi materi yang diujikan, dari ketiga sekolah memiliki kompetensi dasar yang berbeda dengan indikator yang berbeda juga pada item yang paling sukar dan pada item paling mudah SMA Negeri 11 dan SMA Negeri 18 mengukur materi yang sama. Untuk soal paling sukar pada SMA Negeri 11 mengenai ketenagakerjaan yaitu akibat pengangguran, SMA Negeri 18 mengenai penghitungan pendapatan nasional yaitu pendapatan per kapita dan SMA Negeri 23 mengenai penyebab inflasi. Untuk soal paling mudah pada SMA Negeri 11 mengenai cara meningkatkan kualitas tenaga kerja dan SMA Negeri 18 mengenai penyebab perbedaan tingkat upah, dan SMA Negeri 23 mengenai jenis inflasi. Kemudian berdasarkan parameter daya pembeda sebanyak 81 butir memiliki daya pembeda yang baik dan 4 butir memiliki daya pembeda yang tidak baik. Diketahui pula efektivitas distraktor yang baik terdapat pada 71 butir dan efektivitas distraktor yang kurang baik terdapat pada 14 butir.

Bila dikaitkan dengan penelitian terdahulu yang dilakukan oleh Dadan Rosana dan Sukardiyono (2015) yang berjudul Analisis Butir dan Identifikasi Ketidkewajaran Skor Ujian Akhir Sekolah untuk Standarisasi Penilaian, memiliki persamaan yaitu melakukan analisis kualitas butir soal berdasarkan Item Response Theory menggunakan program software Quest dan meneliti siswa SMA Negeri. Hasil analisis program Quest juga memiliki kesamaan yaitu meneliti bagian output yang terdiri dari kecocokan dengan model, tingkat kesukaran butir, reliabilitas soal. Hasil reliabilitas yang hampir sama, pada penelitian Dadan dan Sukardiyono mendapat nilai 0.97, dan pada penelitian ini mendapat nilai 0.96. Berdasarkan hasil temuan analisis Dadan Rosana dan Sukardiyono, perbedaan dengan hasil penelitian ini sebagai berikut. Pertama, parameter logistik yang digunakan adalah 1PL sedangkan pada penelitian ini menggunakan 2PL. Kedua, pada penelitian ini tidak menguji mengenai ketidakwajaran skor ujian melainkan penelitian ini menganalisis validitas, daya pembeda serta efektivitas distraktor. Ketiga, kategori yang digunakan untuk menentukan tingkat kesukaran terbagi menjadi 5 yaitu sangat mudah, mudah, sedang, sukar dan sangat sukar, sedangkan pada penelitian ini hanya membagi kategori menjadi 3, yaitu sukar, baik dan mudah.

Selain menganalisis kualitas butir soal berdasarkan Item Response Theory yang meninjau validitas, reliabilitas, dan 2 parameter logistik (tingkat kesukaran dan daya pembeda), pada penelitian ini juga melakukan analisis mengenai ketepatan butir dalam mengukur indikator yang ingin dicapai serta gambaran butir soal dalam mengukur tingkat keterampilan berpikir. Analisis tersebut menggunakan expert judgement dari 5 validator, serta menggunakan perhitungan formula Aiken untuk menentukan ketepatan butir dalam mengukur indikator. Pada soal SMA Negeri 11 Bandung terdapat 30 butir soal yang memuat 4 Kompetensi Dasar dan mengukur 30 indikator berbeda. Butir soal dinyatakan tepat dalam mengukur indikator bila nilai perhitungan berdasarkan formula Aiken mendapatkan hasil minimal 0.8, berdasarkan ketentuan tersebut maka didapatkan hasil sebanyak 9 butir tepat dalam mengukur indikator dan 21 butir belum tepat dalam mengukur indikator.

Butir yang tepat dalam mengukur indikator adalah butir nomor 6, 14, 16, 18, 21, 23, 25, 26 dan 27. Butir nomor 6 dan nomor 21 mendapatkan nilai paling tinggi dibandingkan butir soal yang lain yaitu sebesar 1. Butir nomor 6 memuat kompetensi dasar mendeskripsikan konsep dan metode penghitungan pendapatan nasional dan mengukur indikator menghitung pendapatan per kapita. Pada soal tersebut ditampilkan terlebih dahulu tabel jumlah penduduk dari beberapa negara serta besaran GNP pada masing masing negara, kemudian siswa diminta untuk menentukan negara mana yang memiliki pendapatan per kapita terendah. Butir ini tepat dalam mengukur indikator dikarenakan kesesuaian antara indikator yang ingin dicapai ialah siswa mampu menghitung dan pada butir soal pun diberikan informasi yang dapat digunakan siswa untuk menghitung yang kemudian siswa diminta untuk menentukan berdasarkan hasil perhitungan yang telah dilakukan. Kemudian untuk butir soal nomor 21 memuat kompetensi dasar menganalisis permasalahan ketenagakerjaan dalam pembangunan ekonomi dengan indikator yang diukur ialah menghitung tingkat pengangguran. Pada butir nomor 21 juga memiliki kesamaan dengan butir nomor 6, dimana indikator yang diukur ialah menghitung, dan pada butir soal hanya diberikan informasi yang akan digunakan siswa untuk melakukan perhitungan sendiri dan kemudian menentukan pernyataan yang benar berdasarkan atas perhitungan yang telah dilakukan.

Butir yang belum tepat dalam mengukur indikator adalah butir nomor 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 15, 17, 19, 20, 22, 24, 28, 29, dan 30. Butir nomor 3 mendapatkan nilai paling rendah yaitu sebesar 0.25. Butir ini memuat kompetensi dasar mendeskripsikan konsep dan metode penghitungan pendapatan nasional dengan mengukur indikator mengidentifikasi manfaat pendapatan nasional. Pada butir ini diberikan sebuah pernyataan mengenai TKI yang bekerja pada perusahaan negara lain, kemudian diberikan pilihan beberapa informasi mengenai akan masuk kemana penghasilan TKI tersebut, kemudian siswa diminta untuk mengidentifikasi pernyataan mana yang tepat sesuai dengan pernyataan soal. Pada butir ini, soal dinyatakan belum tepat dalam mengukur indikator dikarenakan pada indikator yang ingin dicapai ialah mengenai identifikasi manfaat, tetapi pada butir soal yang muncul ialah menanyakan penghasilan tersebut akan masuk ke dalam penghitungan mana pendapatan negara mana. Dapat disimpulkan bahwa soal SMA Negeri 11 dengan jumlah soal 30 butir, hanya terdapat 9 butir yang tepat dalam mengukur indikator dan terdapat 21 butir tidak tepat. Sehingga masih perlu dilakukan perbaikan dari segi isi soal agar soal benar benar dapat mengukur indikator.

Pada soal SMA Negeri 18 Bandung terdapat 30 butir soal yang memuat 3 Kompetensi Dasar yang mengukur 30 indikator berbeda. Didapatkan hasil sebanyak 9 butir tepat dalam mengukur indikator dan 21 butir belum tepat dalam mengukur indikator. Butir yang tepat dalam mengukur indikator adalah butir nomor 8, 16, 18, 21, 22, 25, 26, 29, dan 30. Butir nomor 9 mendapatkan nilai paling tinggi yaitu nomor sebesar 1. Butir ini memuat kompetensi

dasar yaitu mendeskripsikan konsep dan metode penghitungan pendapatan nasional dalam konteks pertumbuhan dan pembangunan ekonomi dengan indikator yang diukur ialah menyimpulkan hasil penghitungan pendapatan nasional menggunakan berbagai metode dengan tepat. Pada butir ini diberikan informasi mengenai komponen pendapatan yang menjadi sumber data bagi siswa dalam menghitung pendapatan nasional. Setelah melakukan penghitungan tersebut siswa diminta untuk menyimpulkan berdasarkan hasil penghitungan yang telah dikerjakan. Butir ini tepat dalam mengukur indikator dikarenakan indikator dan soal sesuai, meminta siswa untuk mampu menyimpulkan berdasarkan hasil dari penghitungan pendapatan nasional tersebut. Butir soal yang belum tepat dalam mengukur indikator diantaranya butir nomor 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 17, 19, 20, 21, 23, 24, 27, dan 28. Butir yang mendapatkan nilai terendah adalah butir nomor 1 dengan angka sebesar 0.4. Butir ini memuat kompetensi dasar mendeskripsikan konsep dan metode penghitungan pendapatan nasional dalam konteks pertumbuhan dan pembangunan ekonomi dengan indikator yang diukur ialah mengidentifikasi pengertian pendapatan nasional dengan tepat. Pada soal diberikan pertanyaan sebagai berikut : "Pendapatan nasional merupakan...". Butir soal tersebut tidak tepat dalam mengukur indikator dikarenakan tidak sesuai dengan indikatornya. Soal tersebut lebih tepat untuk mengukur indikator dengan menjelaskan dibandingkan mengidentifikasi dikarenakan belum menunjukkan kemungkinan siswa dalam melakukan identifikasi mengenai pengertian tersebut.

Kemudian, pada soal SMA Negeri 23 Bandung terdapat 25 butir soal yang memuat 4 Kompetensi Dasar yang mengukur 25 indikator berbeda. Butir soal dinyatakan tepat dalam mengukur indikator bila nilai perhitungan berdasarkan formula Aiken mendapatkan hasil minimal 0.8, berdasarkan ketentuan tersebut maka didapatkan hasil sebanyak 4 butir yang tepat dalam mengukur indikator dan 21 butir yang belum tepat dalam mengukur indikator. Butir yang tepat dalam mengukur indikator adalah butir nomor 3, 16, 20, dan 25. Keempat butir tersebut memperoleh nilai yang sama yaitu 0.95 dan sama sama memuat kompetensi dasar yang sama yaitu mendeskripsikan konsep dan metode penghitungan pendapatan nasional dalam konteks pertumbuhan dan pembangunan ekonomi. Pada butir nomor 3, 16 dan 20 pun sama sama mengukur indikator penghitungan pendapatan nasional, hanya berbeda pada pendekatan dan distribusi tetapi masih sama mengenai penghitungan pendapatan nasional. Dari ketiga butir tersebut juga sama sama memberikan informasi maupun keterangan yang dapat digunakan siswa dalam melakukan penghitungan pendapatan nasional kemudian siswa menjawab melalui hasil penghitungan yang telah dilakukan. Kemudian pada butir nomor 25, mengukur indikator mengenai mendeskripsikan konsep PDB, PNB, PNN, PI, DI, dan PDB. Pada butir soal diberikan suatu pengertian dari salah satu konsep tersebut, kemudian siswa diminta untuk menentukan konsep mana yang sesuai dengan soal. Butir yang belum tepat dalam mengukur indikator ialah butir nomor 1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17, 18, 19, 21, 22, 23, dan 24. Butir nomor 8 mendapatkan nilai yang paling rendah yaitu sebesar 0.45. Butir ini memuat kompetensi dasar menganalisis inflasi, kebijakan moneter dan kebijakan fiskal dengan mengukur indikator mengidentifikasi penyebab terjadinya inflasi. Pada soal tersebut diberikan satu pernyataan mengenai salah satu penyebab inflasi, kemudian siswa diminta untuk menentukan jenis inflasi berdasarkan penyebab yang telah disebutkan. Sehingga butir tersebut dinilai tidak tepat dalam mengukur indikator dikarenakan pada indikator diharapkan mampu mengidentifikasi penyebab inflasi, tetapi pada soal siswa diminta menentukan jenis inflasi.

Secara keseluruhan mendapatkan informasi bahwa mayoritas butir soal belum tepat dalam mengukur indikator berdasarkan hasil analisis dari keseluruhan butir soal yang berjumlah 85 hanya terdapat 22 butir yang tepat dalam mengukur indikator dan 63 butir soal

lainnya belum tepat dalam mengukur indikator. Nilai ketepatan tertinggi yaitu sebesar 1 pada butir soal yang membahas mengenai penghitungan pendapatan per kapita serta materi pendapatan nasional dalam mengidentifikasi pengertiannya, kemudian untuk nilai ketepatan terendah yaitu 0.25 yang memuat materi pendapatan nasional mengenai identifikasi manfaat pendapatan nasional. Hal tersebut mengartikan bahwa butir-butir soal tersebut harus direvisi/diperbaiki supaya dapat mengukur indikator yang ingin dicapai dengan tepat.

Bila ditinjau melalui kemampuan butir dalam mengukur kemampuan tingkat berpikir, terbagi menjadi 2, yaitu butir mampu mengukur kemampuan berpikir tingkat rendah yang mencakup C1 (mengingat), C2 (memahami), dan C3 (mengaplikasikan), serta butir mampu mengukur kemampuan berpikir tingkat tinggi yang mencakup C4 (menganalisis), C5 (mengevaluasi), dan C6 (membuat).

Pada soal SMA Negeri 11 Bandung terdiri dari 30 butir. Secara keseluruhan terdapat 7 butir soal yang disepakati oleh kelima validator yaitu pada butir nomor 1, 2, 9, 14, 15, 16, dan 26, butir-butir tersebut mampu mengukur kemampuan berpikir siswa dengan tingkat C1 yaitu mengingat, dimana C1 merupakan tingkatan berpikir paling rendah yang memuat proses mengenali dan mengingat kembali. Dimana pada butir nomor 1 dan 2 mengenai materi penghitungan pendapatan nasional, butir nomor 9 mengenai pertumbuhan dan pembangunan ekonomi, butir nomor 14, 15 dan 16 mengenai ketenagakerjaan dan butir nomor 26 mengenai indeks harga dan inflasi. Terdapat 4 kompetensi dasar yang diukur, pada kompetensi dasar pertama membahas materi mengenai penghitungan pendapatan nasional, disimpulkan secara keseluruhan terdapat 2 butir yang mengukur kemampuan C1 yaitu mengingat, sebanyak 3 butir mengukur kemampuan C3 yaitu mengaplikasikan, dan sebanyak 1 butir mengukur kemampuan C4 atau menganalisis. Pada kompetensi dasar kedua membahas materi mengenai pertumbuhan dan pembangunan ekonomi, disimpulkan secara keseluruhan terdapat 4 butir yang mengukur kemampuan berpikir tingkat C1 atau mengingat, sebanyak 1 butir yang mengukur kemampuan berpikir C3 atau mengaplikasikan dan soal lainnya mendapatkan perbedaan pendapat yang cukup besar dari para validator. Pada kompetensi dasar ketiga membahas mengenai materi ketenagakerjaan, sebanyak 4 butir mengukur kemampuan berpikir C1 atau mengingat, sebanyak 3 butir mengukur kemampuan berpikir C3 atau mengaplikasikan dan 1 butir mengukur kemampuan C4 atau menganalisis. Pada kompetensi terakhir membahas materi mengenai indeks harga dan inflasi, sebanyak 4 butir soal mengukur kemampuan berpikir C1 atau mengingat, sebanyak 1 butir mengukur kemampuan C3 atau mengaplikasikan, dan butir lainnya memiliki perbedaan pendapat yang cukup besar dari kelima validator. Didapatkan kesimpulan bahwa pada masing masing kompetensi dasar, butir soal paling banyak mengukur kemampuan berpikir tingkat terendah yaitu C1 atau mengingat.

Berdasarkan hasil tersebut, dapat diketahui bahwa butir soal yang mengukur kemampuan berpikir tingkat rendah yaitu pada level kognitif C1 (mengingat), C2 (memahami) dan C3 (mengaplikasikan) dengan jumlah persentase sebesar 86.67% sedangkan kemampuan berpikir tingkat tinggi, yaitu pada level kognitif C4 (menganalisis), C5 (mengevaluasi), dan C6 (mencipta) dengan jumlah persentase sebesar 6.67%. Sebesar 6.67% mendapatkan perbedaan pendapat cukup besar dari kelima validator sehingga tidak dapat ditentukan kepada salah satu level kognitif. Dapat disimpulkan bahwa soal dari SMA Negeri 11 Bandung masih didominasi dengan soal yang mengukur kemampuan berpikir tingkat rendah.

Pada soal SMA Negeri 18 Bandung terdiri dari 30 butir. Secara keseluruhan hanya terdapat 1 butir soal yang disepakati oleh kelima validator yaitu pada butir nomor 12, butir tersebut mampu mengukur kemampuan berpikir siswa dengan tingkat C2 yaitu memahami. Terdapat 4 kompetensi dasar yang diukur pada soal SMAN 18. Pada kompetensi dasar pertama yaitu

membahas materi mengenai penghitungan pendapatan nasional, disimpulkan secara keseluruhan terdapat 3 butir yang mengukur kemampuan C1 yaitu mengingat, sebanyak 1 butir yang mengukur kemampuan C2 yaitu memahami, sebanyak 3 butir mengukur kemampuan C3 yaitu mengaplikasikan, dan sebanyak 2 butir mengukur kemampuan C4 atau menganalisis. Pada kompetensi dasar kedua yaitu membahas materi mengenai ketenagakerjaan, disimpulkan secara keseluruhan terdapat 3 butir yang mengukur kemampuan berpikir tingkat C2 atau memahami, sebanyak 2 butir yang mengukur kemampuan berpikir C3 atau mengaplikasikan, sebanyak 1 butir mengukur kemampuan C4 atau menganalisis dan soal lainnya mendapatkan perbedaan pendapat yang cukup besar dari para validator. Kemudian pada kompetensi dasar terakhir yaitu membahas materi mengenai inflasi, kebijakan moneter dan kebijakan fiskal, sebanyak 2 butir soal mengukur kemampuan berpikir C1 atau mengingat, sebanyak 1 butir mengukur kemampuan berpikir C2 atau memahami, sebanyak 4 butir mengukur kemampuan C3 atau mengaplikasikan, sebanyak 1 soal mengukur kemampuan berpikir C4 atau menganalisis dan C5 atau mengevaluasi, dan butir lainnya memiliki perbedaan pendapat yang cukup besar dari kelima validator. Didapatkan kesimpulan bahwa pada materi mengenai penghitungan pendapatan nasional lebih banyak butir soal yang mengukur kemampuan C1 atau mengingat dan C3 atau mengaplikasikan. Pada materi ketenagakerjaan, lebih banyak butir soal yang mengukur kemampuan berpikir C2 atau memahami, dan pada materi mengenai inflasi, kebijakan moneter dan kebijakan fiskal lebih banyak butir soal yang mengukur kemampuan C3 atau mengaplikasikan. Bila dilihat secara keseluruhan berdasarkan kesepakatan validator, keseluruhan butir soal terbanyak yaitu mengukur kemampuan berpikir C3 atau mengaplikasikan. Berdasarkan hasil tersebut, dapat diketahui bahwa butir soal yang mengukur kemampuan berpikir tingkat rendah yaitu pada level kognitif C1 (mengingat), C2 (memahami) dan C3 (mengaplikasikan) dengan jumlah persentase sebesar 63.33% sedangkan kemampuan berpikir tingkat tinggi, yaitu pada level kognitif C4 (menganalisis), C5 (mengevaluasi), dan C6 (mencipta) dengan jumlah persentase sebesar 20%. Sebesar 16.67% mendapatkan perbedaan pendapat cukup besar dari kelima validator sehingga tidak dapat ditentukan kepada salah satu level kognitif. Sehingga dapat disimpulkan bahwa soal dari SMA Negeri 18 Bandung belum mengukur kemampuan berpikir tingkat tinggi.

Pada soal SMA Negeri 23 Bandung terdiri dari 25 butir soal yang mengukur 4 kompetensi dasar serta 25 indikator yang berbeda. Secara keseluruhan, terdapat 12 butir soal, yaitu butir nomor 4, 5, 6, 7, 8, 10, 11, 12, 14, 17, 21, dan 25 yang disepakati oleh kelima validator bahwa butir tersebut mampu mengukur kemampuan berpikir siswa dengan tingkat C1 atau mengingat. Dari kedua belas butir tersebut, butir nomor 11, 12, dan 25 mengukur mengenai materi penghitungan pendapatan nasional, kemudian butir nomor 4, 5, 7, dan 14 mengukur mengenai materi ketenagakerjaan, sedangkan butir nomor 8, 17, dan 21 mengukur mengenai materi inflasi, kebijakan moneter dan kebijakan fiskal, dan butir nomor 6, dan 10 mengukur mengenai materi APBN dan APBD. Pada kompetensi dasar pertama yaitu membahas materi mengenai penghitungan pendapatan nasional, disimpulkan secara keseluruhan terdapat 4 butir yang mengukur kemampuan C1 yaitu mengingat dan sebanyak 2 butir mengukur kemampuan C3 yaitu mengaplikasikan. Pada kompetensi dasar kedua yaitu membahas materi mengenai ketenagakerjaan, disimpulkan secara keseluruhan terdapat 5 butir yang mengukur kemampuan berpikir C1 atau mengingat dan sebanyak 3 butir yang mengukur kemampuan berpikir C2 atau memahami. Pada kompetensi dasar ketiga yang membahas materi mengenai inflasi, kebijakan moneter dan kebijakan fiskal terdapat 7 butir yang mengukur kemampuan berpikir C1 atau mengingat dan sebanyak 2 butir mengukur kemampuan berpikir C2 atau memahami. Kemudian pada kompetensi dasar terakhir yaitu membahas materi mengenai

APBN dan APBD, sebanyak 3 butir soal mengukur kemampuan berpikir C1 atau mengingat, dan sebanyak 1 butir mengukur kemampuan berpikir C2 atau memahami. Didapatkan kesimpulan bahwa pada masing masing kompetensi dasar, butir soal paling banyak mengukur kemampuan berpikir tingkat terendah yaitu C1 atau mengingat. Berdasarkan hasil tersebut, dapat diketahui bahwa butir soal yang mengukur kemampuan berpikir tingkat rendah yaitu pada level kognitif C1 (mengingat), C2 (memahami) dan C3 (mengaplikasikan) dengan jumlah persentase sebesar 100%, sedangkan untuk butir soal yang mengukur keterampilan berpikir tingkat tinggi tidak ada satupun. Sehingga dapat disimpulkan bahwa soal dari SMA Negeri 23 Bandung belum mengukur kemampuan berpikir tingkat tinggi.

Melalui ketiga analisis mengenai kemampuan butir dalam mengukur tingkat keterampilan berpikir yang diukur pada soal Penilaian Akhir Semester Ganjil kelas XI IPS SMAN 11, SMAN 18 dan SMAN 23 Bandung, diketahui belum mampu untuk mengukur keterampilan berpikir tingkat tinggi, dikarenakan hampir keseluruhan butir soal masih mengukur kemampuan berpikir tingkat rendah, yaitu pada level kognitif C1 (mengingat), C2 (memahami) dan C3 (mengaplikasikan) dengan jumlah 70 butir soal, sedangkan kemampuan berpikir tingkat tinggi, yaitu pada level kognitif C4 (menganalisis), C5 (mengevaluasi), dan C6 (mencipta) hanya berjumlah 8 butir soal. Kemudian sebanyak 7 butir soal mendapatkan perbedaan pendapat cukup besar dari kelima validator sehingga tidak dapat ditentukan kepada salah satu level kognitif, namun perbedaan level kognitif tersebut masih termasuk ke dalam kategori kemampuan berpikir tingkat rendah, sehingga dapat disimpulkan bahwa sebanyak 77 butir soal mampu mengukur keterampilan berpikir tingkat rendah, dan hanya terdapat 8 butir soal yang mampu mengukur keterampilan berpikir tingkat tinggi. Berdasarkan seluruh analisis yang telah dilakukan, bila ditentukan kualitas butir soal berdasarkan parameter tingkat kesukaran, jumlah butir soal terbanyak adalah butir yang berkualitas kurang baik, tetapi bila ditentukan berdasarkan parameter daya pembeda, jumlah butir soal terbanyak termasuk ke dalam kategori butir soal yang baik. Apabila ditinjau dari ketepatan butir dalam mengukur indikator, butir soal terbanyak adalah butir yang belum tepat dalam mengukur indikator. Sehingga dapat disimpulkan bahwa butir perlu dilakukan perbaikan sehingga dapat meningkatkan kualitas butir menjadi baik serta meningkatkan kemampuan butir dalam mengukur indikator.

Pada penelitian ini, penemuan yang memerlukan perhatian khusus adalah informasi kemampuan subjek/testi (dalam penelitian ini ialah siswa yang mengerjakan butir soal tersebut) yang belum mendapatkan hasil tinggi, sehingga kemampuan siswa perlu ditingkatkan lagi untuk dapat mengerjakan dan menjawab butir soal dengan benar. Kemudian, hasil temuan kecocokan model terdapat satu sekolah yang butir soalnya tidak cocok dengan model sehingga sudah otomatis berkualitas tidak baik dan harus dibuang/diganti, terlebih lagi secara keseluruhan kualitas butir soal masih kurang baik serta analisis ketepatan butir dalam mengukur indikator yang masih menunjukkan butir tidak tepat dalam mengukur indikator harus diperbaiki atau direvisi agar benar benar memiliki kualitas butir soal yang baik dan mampu dengan tepat dalam mengukur indikator. Terdapat evaluasi keterbatasan pada penelitian ini yaitu tidak lengkapnya set soal yang didapatkan dari sekolah sehingga analisis yang dilakukan tidak maksimal, khususnya pada penentuan indikator yang digunakan pada masing masing butir soal serta pembagian soal sesuai dengan kompetensi dasar. Ketidakeengkapan set soal yang didapatkan tersebut disebabkan oleh data yang hilang pada guru yang bersangkutan dan tidak tersip pada bidang kurikulum sekolah, serta respons siswa yang sulit didapatkan dikarenakan penerapan ujian yang dilakukan secara online berbasis sistem yang hanya memberikan hasil skor benar, skor salah dan nilai siswa, tidak lengkap dengan respons siswa per butir.

4. CONCLUSION

Hasil penelitian mengenai kualitas butir soal mata pelajaran ekonomi di SMAN 11, SMAN 18 dan SMAN 23 Bandung berdasarkan Item Response Theory pada parameter tingkat kesukaran sebanyak 35 butir berkualitas baik, 16 butir berkualitas kurang baik dan 5 butir berkualitas tidak baik. Pada parameter daya pembeda sebanyak 76 butir berkualitas baik, 4 butir berkualitas kurang baik dan 5 butir berkualitas tidak baik dengan butir sebanyak 71 memiliki efektivitas distraktor baik dan 14 butir memiliki efektivitas distraktor tidak baik. Secara keseluruhan hanya terdapat 33 butir berkualitas baik, 47 butir berkualitas kurang baik, dan 5 butir berkualitas tidak baik. Berdasarkan validitas dan reliabilitasnya, terdapat 80 butir valid dan 5 butir tidak valid, dan memiliki nilai reliabilitas item lebih dari 0.85 yang berarti butir tes dapat digunakan untuk menentukan keputusan pada siswa secara perorangan.

5. REFERENCES

- Alfarisa, F., & Purnama, D. N. (2019). Analisis butir soal ulangan akhir semester mata pelajaran ekonomi SMA menggunakan Rasch model. *Jurnal Pendidikan Ekonomi Undiksha*, 11(2), 366-374.
- Amalia, A. N., & Widayati, A. (2012). Analisis butir soal tes kendali mutu kelas XII SMA mata pelajaran ekonomi akuntansi di kota Yogyakarta tahun 2012. *Jurnal Pendidikan Akuntansi Indonesia*, 10(1).
- Arlinwibowo, J., Hadi, S., & Firdaus, E. M. (2020). Validasi Persepsi Siswa Terhadap Pembelajaran Sains Dengan Teori Respon Butir Multidimensi. *Jurnal Ilmu Komputer Dan Matematika*, 1(1), 7-14.
- Danni, R., Wahyuni, A., & Tauratiya, T. (2021). Item response theory approach: Kalibrasi butir soal penilaian akhir semester mata pelajaran bahasa Arab. *Arabi: Journal of Arabic Studies*, 6(1), 93-104.
- Fajar, R., & Junaidi, J. (2022). Analisis Butir Soal Tipe Multiple Choices Questions (MCQ) pada Ujian Akhir Semester Ganjil Mata Pelajaran Sosiologi Kelas XII SMA. *Jurnal Sikola: Jurnal Kajian Pendidikan dan Pembelajaran*, 4(1), 23-34.
- Fauziana, A., & Wulansari, A. D. (2021). Analisis Kualitas Butir Soal Ulangan Harian di Sekolah Dasar dengan Model Rasch. *Jurnal Ibriez: Jurnal Kependidikan Dasar Islam Berbasis Sains*, 6(1), 10-19.
- Fatkhudin, A., Surarso, B., & Subagio, A. (2014). Item response theory model empat parameter logistik pada computerized adaptive test. *Jurnal Sistem Informasi Bisnis*, 2, 121-129.
- Friyatmi. (2018). Estimasi parameter tes dengan penskoran politomus menggunakan graded response model pada sampel kecil. *Jurnal Inovasi Pendidikan Ekonomi*, 22-31.
- Kurniawan, T. K. (2015). Analisis butir soal ulangan akhir semester gasal mata pelajaran IPS Sekolah Dasar. *Journal of Elementary Education*, 4(1), 1-6.
- Kusumayanti, A., & Jannah, N. (2022). Analisis Butir Soal Ujian Masuk Mandiri Uin Alauddin Makassar Dengan Teori Tes Modern. *MaPan: Jurnal matematika dan Pembelajaran*, 10(1), 159-174.
- Noptihalia, Y. H. (2010). Meneropong kualitas soal tes buatan guru biologi MTS Negeri se-Jakarta Selatan. *Jurnal Penelitian dan Evaluasi Pendidikan*. 175-198.
- Pratama, D. (2020). Analisis kualitas tes buatan guru melalui pendekatan item response theory (IRT) model rasch. *Tarbawy: Jurnal Pendidikan Islam*, 7(1), 61-70.

- Purnama, D. N., & Alfarisa, F. (2020). Karakteristik butir soal try out Teori Kejuruan Akuntansi SMK berdasarkan teori Tes Klasik dan teori Respons Butir. *Jurnal Pendidikan Akuntansi Indonesia*, 18(1), 36-46.
- Purniasari, L., Masykuri, M., & Ariani, S. R. D. (2021). Analisis Butir Soal Ujian Sekolah Mata Pelajaran Kimia SMA N 1 Kutowinangun Tahun Pelajaran 2019/2020 Menggunakan Model Iteman dan Rasch. *Jurnal Pendidikan Kimia*, 10(2), 205-214.
- Rahim, A., & Haryanto, H. (2021). Implementation of Item Response Theory (IRT) Rasch Model in Quality Analysis of Final Exam Tests in Mathematics. *Journal of Research and Educational Research Evaluation*, 10(2), 57-65.
- Rahmayani, A., & Alfarisa, F. (2022). Analisis Kualitas Butir Soal Ulangan Harian Matematika Di Kelas IV MI Al-Islamiyah Menggunakan Rasch Model. *Jurnal Perseda: Jurnal Pendidikan Guru Sekolah Dasar*, 5(3), 170-177.
- Sarea, M. S., & Ruslan, R. (2019). Karakteristik butir soal: *classical test theory vs item response theory?*. *Didaktika: Jurnal Kependidikan*, 13(1), 1-16.
- Siregar, M. N. N., & Aghni, R. I. (2021). Pengembangan Perangkat Pembelajaran Berbasis Problem Based Learning (PBL) untuk Meningkatkan Higher Order Thinking Skill (HOTS). *Jurnal Pendidikan Akuntansi (JPAK)*, 9(2), 292-301.
- Sudaryono, S. (2011). Implementasi Teori Responsi Butir (Item Response Theory) Pada Penilaian Hasil Belajar Akhir di Sekolah. *Jurnal Pendidikan dan Kebudayaan*, 17(6), 719-732.
- Wahyuni, W., & Fahmi, M. (2021). Sistem Analisis Butir Soal Pada Cbt Dengan Metode Teori Respon Butir. *Jurnal Ilmiah MATRIK*, 23(2), 201-208.