

ANALISIS BUTIR SOAL ULANGAN AKHIR SEMESTER MATA PELAJARAN EKONOMI SMA MENGGUNAKAN *RASCH MODEL*

Fitri Alfarisa¹, Dian Normalitasari Purnama²

Universitas Pendidikan Indonesia¹
Bandung, Indonesia
Universitas Negeri Yogyakarta²
Yogyakarta, Indonesia

email: alfarisa@upi.edu¹, dianpurnama@uny.ac.id²

Abstrak

Penelitian ini bertujuan untuk menganalisis dan mendeskripsikan karakteristik butir soal Ulangan Akhir Semester (UAS) mata pelajaran ekonomi siswa kelas XI SMA di Kota Yogyakarta menggunakan *Rasch model*. Penelitian ini merupakan penelitian deskriptif kuantitatif. Populasi penelitian ini adalah siswa kelas XI SMA di Kota Yogyakarta. Teknik pengambilan sampel menggunakan *stratified proportional random sampling* dan diperoleh jumlah sampel 306 siswa. Data yang digunakan adalah respon tes UAS mata pelajaran ekonomi yang berjumlah 40 butir pilihan ganda. Validitas instrumen tes dibuktikan dengan validitas isi menggunakan *expert judgment* dan dilanjutkan dengan perhitungan indeks *Aiken V*, sementara reliabilitas instrumen tes diestimasi menggunakan KR-20, dan diketahui besarnya koefisien reliabilitas instrument adalah 0,77. Hasil penelitian ini dianalisis menggunakan program *Winstep* yang berbasis *Rasch model*. Berdasarkan hasil analisis diketahui dari 40 butir terdapat 39 butir fit dengan model *Rasch*. Nilai tingkat kesukaran butir berada pada rentang -2,51 sampai dengan 1,77. Fungsi informasi maksimum tes adalah 18,3379 pada kemampuan (θ) siswa sekitar 0,2 logit dan SEM sebesar 0,2335.

Kata kunci: Ulangan Akhir Semester, Mata Pelajaran Ekonomi, *Rasch model*

Abstract

This research was aimed to analyze and describe the items characteristic of economics final semester examination of senior high school in Yogyakarta City using Rasch Model. This research was a quantitative-descriptive research. The populations of this research were the eleventh graders of senior high school in Yogyakarta City. The sampling technique used stratified proportional random sampling and obtained 306 students as samples. The data of this research was the students' responses of economics final semester examination with 40 multiple choice items. The instrument validity test was proven by using the content validity using expert judgment and detailed calculating of Aiken V index, while the reliability of instrument test was estimated using KR-20, and coefficient of the instrument reliability was found 0.77. The result was analyzed by Winstep based Rasch model. Based on the analysis, it found that from 40 items, there were 39 items fit onto Rasch model. The difficulty scores of the items was between -2,51 to 1,77. The maximum information from the test was 18,3379, the students' ability (θ) was 0,2 logit, and standard error measurement test was 0,21.

Keywords : final semester examination, economics subject, rasch model.

PENDAHULUAN

Kualitas pendidikan tidak terlepas dari sistem penilaian dan kualitas pembelajaran yang mendukungnya. Sistem pembelajaran yang baik akan menghasilkan kualitas belajar yang baik (Mardapi, 2016). Lebih lanjut Mardapi menyampaikan bahwa kualitas pembelajaran ini dapat dilihat dari hasil penilaian yang dilakukan oleh guru. Faktanya guru bukan hanya sebagai mengajar tetapi juga diupayakan dapat profesional dalam melakukan penilaian. Hal ini terdapat UU No. 14 Tahun 2005 tentang Guru dan Dosen, yang tercantum bahwa peran guru atau pendidik tidak hanya merancang dan melaksanakan pembelajaran saja, melainkan juga melakukan asesmen atau penilaian.

Penilaian hasil belajar siswa oleh pendidik hendaknya dilakukan secara berkesinambungan. Penilaian tidak hanya dimaksudkan untuk mengetahui tercapai tidaknya tujuan-tujuan yang telah ditetapkan, tetapi juga untuk mengetahui apakah tujuan-tujuan tersebut penting bagi siswa dan bagaimana siswa mencapainya. Hasil kajian menunjukkan bahwa sebanyak 87 persen guru masih kesulitan dalam memahami cara melakukan penilaian (Rusilowati, 2013). Tidak tersosialisasikannya langkah-langkah dalam melakukan penilaian menjadi salah satu kendalanya. Hal ini mengindikasikan bahwa kemampuan guru untuk melakukan penilaian secara baik masih perlu ditingkatkan.

Penilaian yang baik dapat dilakukan dengan cara mengumpulkan bukti akurat terkait pencapaian hasil belajar siswa dan menjadikan proses penilaian kelas dan hasilnya bermanfaat bagi siswa, yaitu mampu meningkatkan motivasi dan prestasi belajarnya (Stiggins & Chappuis, 2012: 3). Tes merupakan salah satu cara untuk menaksir besarnya tingkat kemampuan manusia secara tidak langsung, yaitu melalui *respons* seseorang terhadap sejumlah stimulus atau pertanyaan (Mardapi, 2008). Tes yang berkualitas baik memiliki karakteristik butir soal dan perangkat tes yang baik pula.

Hasil tes merupakan informasi tentang karakteristik individu atau sebuah

kelompok (Rasyid & Mansur, 2008). Oleh karena itu agar diperoleh informasi yang akurat dibutuhkan tes yang handal. Penilaian melalui teknik tes yang dibuat guru khususnya guru mata pelajaran ekonomi saat ini belum memadai. Hal ini diketahui berdasarkan wawancara dengan beberapa guru mata pelajaran ekonomi di SMA Kota Yogyakarta. Guru mengungkapkan bahwa tes buatan guru belum melalui pengujian empiris sehingga hasil yang diperoleh belum mampu merepresentasikan kemampuan siswa yang sebenarnya.

Analisis butir soal dapat dilakukan secara klasik dan modern. Hambleton & Swaminathan (1985: 1-3) juga mengungkapkan beberapa kelemahan teori tes klasik, diantaranya: (1) tingkat kesukaran serta daya beda butir soal tergantung pada kelompok peserta yang mengerjakan, (2) penggunaan metode dan teknik untuk analisis tes adalah dengan membandingkan kemampuan siswa pada pembagian kelompok atas, tengah, dan bawah, (3) konsep reliabilitas skor didefinisikan dari istilah tes paralel, (4) tidak ada dasar teori untuk menentukan bagaimana peserta tes memperoleh tes yang sesuai dengan kemampuan peserta tersebut, dan (5) *Standar Error of Measurement (SEM)* yang berlaku pada seluruh peserta tes.

Para ahli pengukuran berusaha mencari alternatif sebagai upaya mengatasi kelemahan-kelemahan yang ada pada teori klasik. Model alternatif yang mempunyai ciri-ciri itu adalah model pengukuran yang disebut teori respon butir (*Item Response Theory/ IRT*). Teori tes modern atau teori respon butir dikembangkan oleh para ahli pengukuran bidang psikologi dan pendidikan sebagai upaya meminimalkan kekurangan-kekurangan yang ada dalam teori tes klasik. Hal ini sejalan dengan Meyer & Shi Zhu (2013) yang menyatakan bahwa statistik IRT merupakan cara mengestimasi parameter dalam sebuah model.

Terdapat 4 model dalam IRT berdasarkan jumlah parameternya yakni *one parameter logistic model* (1PL), *two parameter logistic model* (2PL), *three*

parameter logistic model (3PL), four parameter logistic model (4PL). Pada penelitian ini hanya difokuskan pada model 1PL atau Rasch model. Rasch adalah orang pertama yang mengembangkan model logistic satu parameter (Wright & Stone, 1979: x). Ide sederhana dan jitu dari pengamatan hasil ujian siswa SD di tahun 1950-an membawa Georg Rasch pada suatu temuan baru, yaitu peluang bagi siswa untuk menjawab satu butir soal dengan benar sama dengan ketika kemampuan siswa dibandingkan dengan tingkat kesukaran soal (Sumintono & Widhiarso, 2015).

Pada model Rasch, orang diberi karakteristik tingkat kemampuan laten dan butir diberi karakteristik tingkat kesukaran. Probabilitas menjawab benar suatu fungsi adalah fungsi dari perbandingan antara tingkat kemampuan dan kesukaran butir. Ciri penting model Rasch adalah tidak mengandung parameter diskriminasi dan parameter terkaan. Pada model ini, diasumsikan bahwa kesukaran butir merupakan satu-satunya karakteristik butir yang mempengaruhi kinerja tes. Selain itu, pada model rasch, persoalan pemilihan butir soal untuk mengkonstruksi sebuah tes merupakan suatu upaya merencanakan tes yang berkualitas sesuai dengan kebutuhan dan tujuan pengujian (Sumintono & Widhiarso, 2015).

METODE

Penelitian ini merupakan penelitian deskriptif dengan pendekatan kuantitatif. Subyek penelitian adalah siswa SMA di Kota Yogyakarta berdasarkan respon UAS siswa kelas XI pada mata pelajaran ekonomi. Penentuan sampel menggunakan teknik *proportional random sampling*. Langkah penentuan sekolah yakni dengan mengelompokkan sekolah menjadi tiga kategori berdasarkan rerata nilai UN yakni kategori tinggi, kategori sedang dan kategori rendah. Adapun pengelompokan kategori yaitu kategori tinggi (rerata UN $\geq 5,87$), kategori sedang (rerata UN 4,11 s.d 5,86), dan kategori rendah (rerata UN $< 4,1$). Jumlah sampel yang diambil berdasarkan tabel *Isaac & Michael* dengan taraf kepercayaan 95%

adalah 306 siswa. Hal ini sudah memenuhi syarat minimal sampel untuk analisis dengan Rasch model yakni 250 responden (Linacre, 1994).

Variabel yang diteliti adalah kemampuan siswa yang dilihat dalam hasil ulangan akhir semester pada mata pelajaran ekonomi. Teknik tes digunakan sebagai instrumen pengumpulan data yang dalam penelitian ini. Instrumen tes menggunakan model dikotomis berbentuk pilihan ganda berjumlah 40 butir dengan lima pilihan jawaban. Hal ini digunakan karena bentuk pilihan ganda lebih objektif dan handal dalam melihat respon siswa, tanpa dipengaruhi subyektivitas penilai. Pengujian validitas dalam penelitian ini menggunakan uji validitas isi. Validitas isi pada instrumen tes ini sangat ditekankan, sehingga dilakukan dengan dua cara yakni telaah kualitatif kemudian hasilnya diolah dengan rumus Aiken (V). Telaah kualitatif yang digunakan berdasarkan *judgment* dari ahli dengan kriteria telaah dilihat berdasarkan aspek materi, konstruksi, dan bahasa (Kartowagiran, 2012). Telaah kualitatif dilakukan oleh lima ahli, yang terdiri dari tiga orang praktisi guru ekonomi SMA yang telah memiliki pengalaman dan kompeten dibidangnya, satu dosen ekonomi, serta satu dosen ahli pengukuran dan penilaian.

Pembuktian validitas instrumen tes diketahui semua butir sudah valid. Hal ini diketahui dari rerata nilai V sebesar 0,90. Senada dengan itu, Aiken (1985: 134) mengungkapkan nilai V untuk setiap butir dikatakan baik adalah 0,87 apabila menggunakan lima rater dengan empat kategori pilihan yang digunakan. Estimasi reliabilitas instrumen tes menggunakan Kuder Richardson-20 diketahui memiliki koefisien 0,77. Hasil ini menunjukkan reliabilitas instrumen tes sudah baik dengan memiliki reliabilitas lebih dari 0,70 (Linn, 1989: 106). Analisis data instrumen tes dilakukan dengan pendekatan teori respon butir model Rasch menggunakan program *WINSTEPS*.

HASIL DAN PEMBAHASAN

Hasil penelitian ini memuat karakteristik tingkat kesukaran butir tes mata pelajaran ekonomi berdasarkan teori

respon butir dengan pendekatan Rasch Model. Berdasarkan pendekatan teori respon butir yang digunakan, perlu dilakukan uji asumsi teori respon butir. Apabila uji asumsi telah terpenuhi maka analisis teori respon butir dapat dilanjutkan. Adapun uji asumsi teori respon butir itu meliputi tiga hal yakni uji asumsi unidimensi, independensi lokal, dan invariansi parameter (Hambleton, Swaminathan, & Rogers; 1991).

Uji Asumsi Unidimensi

Uji asumsi unidimensi dilakukan dengan analisis faktor (Dai-TrangLe, 2013). Hasil analisis ini yang pertama dilihat adalah nilai *Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO-MSA) and Bartlett's Test* yang merupakan angka kecukupan sampel dalam analisis faktor. Hasil uji *KMO-MSA and Bartlett's Test* instrumen tes UAS mata pelajaran

ekonomi kelas XI diketahui sebesar 0,656 dengan signifikansi *Bartlett's Test* sebesar 0,000. Nilai *KMO-MSA* > 0,5 dan signifikansi *Bartlett's Test of Sphericity* < 0,05 menunjukkan bahwa tes tersebut memenuhi syarat untuk diolah menggunakan analisis faktor. Hal ini senada dengan Field (2009) yang menyatakan bahwa Nilai *KMO-MSA* dianggap mencukupi jika lebih dari 0,5.

Hambleton & Swaminathan (1985: 16) mengungkapkan uji unidimensi terpenuhi jika tes terbukti hanya mengukur satu dimensi yang dominan yakni kemampuan yang sama. Ketentuan unidimensi ini dapat terlihat dalam kolom nilai *eigen* yang diperoleh untuk masing-masing instrumen tes. Disajikan nilai *eigen* untuk instrumen tes mata pelajaran ekonomi kelas XI pada Tabel 1.

Tabel 1. Nilai Eigen Tes Ekonomi Kelas XI

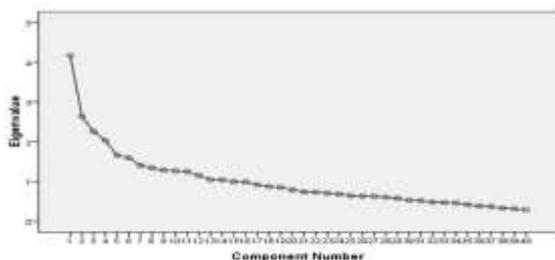
Faktor	Nilai Eigen	Varian (%)	Kumulatif (%)	Faktor	Nilai Eigen	Varian (%)	Kumulatif (%)
1	4.171	10.427	10.427	21	0.738	1.845	75.670
2	2.636	6.590	17.017	22	0.73	1.824	77.494
3	2.256	5.640	22.657	23	0.704	1.759	79.253
4	2.029	5.072	27.729	24	0.682	1.705	80.959
5	1.658	4.145	31.874	25	0.644	1.609	82.567
6	1.595	3.988	35.862	26	0.632	1.580	84.148
7	1.402	3.505	39.367	27	0.622	1.555	85.702
8	1.340	3.350	42.718	28	0.603	1.507	87.209
9	1.282	3.205	45.922	29	0.572	1.431	88.640
10	1.268	3.170	49.093	30	0.528	1.321	89.961
11	1.251	3.128	52.221	31	0.515	1.288	91.250
12	1.152	2.881	55.102	32	0.479	1.198	92.448
13	1.044	2.611	57.712	33	0.464	1.161	93.609
14	1.037	2.592	60.304	34	0.46	1.149	94.758
15	0.995	2.487	62.791	35	0.416	1.041	95.798
16	0.986	2.464	65.255	36	0.381	0.953	96.752
17	0.918	2.294	67.550	37	0.366	0.914	97.666
18	0.87	2.175	69.724	38	0.333	0.832	98.498
19	0.852	2.130	71.854	39	0.31	0.775	99.273
20	0.788	1.970	73.825	40	0.291	0.727	100.000

Faktor yang memiliki nilai eigen > 1 merupakan faktor yang dapat dipakai sebagai indikator suatu sifat atau *trait* (Wagiran, 2014: 302). Hasil analisis faktor pada Tabel 3 menunjukkan terdapat 14 komponen yang mempunyai nilai eigen > 1, sehingga dikatakan bahwa dari 40 butir

soal instrumen tes ekonomi kelas XI dapat membentuk 14 faktor. Hasil analisis juga menunjukkan bahwa faktor 1 merupakan faktor yang dominan karena memiliki nilai eigen sebesar 4,171 dimana nilai eigen itu paling besar atau dominan dari yang lain, sehingga dapat dikatakan bahwa

instrumen tes ekonomi kelas XI bersifat unidimensi.

Kriteria unidimensi dapat juga terlihat pada *scree plot* yang terbentuk. Furr (2008: 74) mengungkapkan bahwa suatu tes dikatakan unidimensi apabila komponen 1 dan 2 dalam *scree plot* mempunyai jarak yang cukup jauh. *Scree plot* pada Gambar 1 menggambarkan bahwa faktor 1 jaraknya jauh dengan faktor 2, sedangkan jarak faktor 2 ke faktor 3 dan seterusnya sangat dekat. Berdasarkan *scree plot* yang terdapat pada Gambar 1, menguatkan kesimpulan bahwa instrumen tes ekonomi kelas XI bersifat unidimensi.



Gambar 1. *Scree Plot* Eigen value for Tes Ekonomi Kelas XI

Uji Asumsi Independensi Lokal

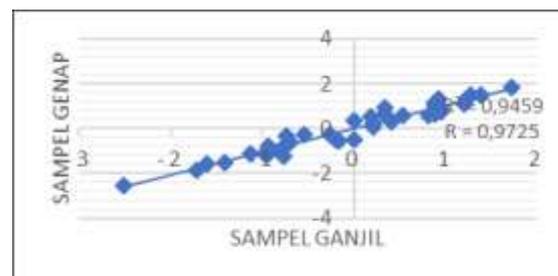
Uji asumsi independensi lokal dapat dilihat berdasarkan kemampuan peserta tes dan terhadap butir tes. Independensi lokal terhadap respons peserta tes memiliki arti bahwa betul salahnya peserta tes menjawab sebuah butir tidak terpengaruh oleh betul salahnya peserta tes yang lain dalam menjawab butir tersebut. Independensi lokal terhadap butir, memiliki arti bahwa betul salahnya seorang peserta tes menjawab sebuah butir tidak terpengaruh oleh betul salahnya peserta tes dalam menjawab butir yang lain.

Uji asumsi independensi lokal otomatis terbukti, setelah dibuktikan dengan unidimensionalitas dan respons peserta terhadap suatu tes (Lord, 1980: 19; Retnawati, 2014: 7; DeMars, 2010: 48). Uji asumsi independensi lokal juga dapat dihitung dengan menggunakan matriks kovarian berdasarkan kemampuan atau *ability* siswa dari masing-masing instrumen tes yang diujikan (Greiff et al., 2013: 8). Berdasarkan hasil perhitungan

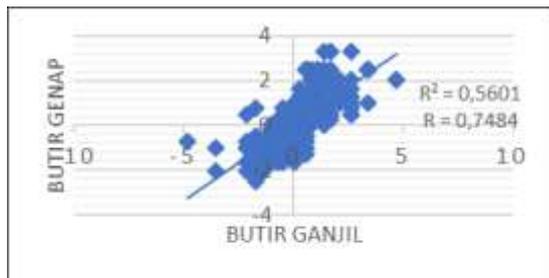
matriks kovarian berdasarkan kemampuan siswa dengan mengelompokkan kemampuan siswa menjadi 10 kelompok dari yang paling tinggi ke yang paling rendah. Diketahui bahwa nilai kovarian masing-masing kelompok kemampuan siswa kelas XI (nilai diluar garis diagonal) adalah mendekati nol, yang artinya asumsi independensi lokal terpenuhi. Hal ini selaras dengan Hambleton & Swaminathan (1984: 25) yang mengungkapkan bahwa apabila nilai kovarian dari kelompok kemampuan siswa mendekati nol, maka uji asumsi independensi lokal telah terpenuhi.

Uji Asumsi Invariansi Parameter

Uji asumsi invariansi parameter terdiri dari invariansi parameter butir dan invariansi kemampuan. Uji invariansi parameter butir dilakukan untuk mengetahui konsistensi karakteristik butir yang dijawab oleh kelompok siswa yang berbeda. Begitu pula uji invariansi kemampuan, dilakukan untuk mengestimasi kemampuan yang tidak berubah meskipun butir soalnya berubah-ubah. Hal ini senada dengan Hambleton, Swaminathan, & Rogers (1991: 18) yang menyatakan bahwa invariansi parameter kemampuan dapat diselidiki dengan mengajukan dua seperangkat tes atau lebih yang memiliki tingkat kesukaran yang berbeda pada sekelompok peserta tes. Hasil pengujian dapat dilakukan dengan menggunakan diagram pencar atau *scree plot* untuk parameter butir dan kemampuan.



Gambar 2. *Scree Plot* Invariansi Parameter Tingkat Kesukaran Butir Tes Kelas XI



Gambar 3. Scree Plot Invariansi Kemampuan Siswa Kelas XI

Gambar 2 dan Gambar 3 menggambarkan *scree plot* estimasi invariansi parameter tingkat kesukaran butir dan kemampuan siswa setelah mengerjakan soal ekonomi pada butir ganjil dan genap. Kedua gambar menunjukkan secara keseluruhan *plot* menyebar dan mendekati garis diagonal. Selain itu, hasil perhitungan korelasi antara tingkat kesukaran dari respon siswa dengan butir nomor genap dan ganjil sebesar 0,9725 (sangat tinggi) dan 0,7484 (tinggi) untuk hasil perhitungan korelasi antara kemampuan siswa dari respon siswa dengan nomor genap dan ganjil sebesar (Best & Kahn, 1998: 372). Berdasarkan *scree plot* dan perhitungan korelasi tersebut dapat disimpulkan bahwa parameter tingkat kesukaran butir tes dan kemampuan siswa kelas XI bersifat invariansi. Hal ini senada dengan Retnawati (2014) menyatakan asumsi invariansi terpenuhi apabila memiliki koefisien korelasi positif dan tinggi.

Kecocokan Butir dengan Model

Analisis kecocokan butir pada model dilakukan berbantu program *Winstep* dengan pendekatan *Rasch Model*. Butir soal yang cocok berarti berperilaku secara konsisten dengan apa yang diharapkan oleh model. Butir soal yang dikatakan cocok adalah butir yang memiliki nilai *outfit mean square* (MNSQ) $0,5 < MNSQ < 1,5$ dan nilai *Point Measure Correlation* (*Pt Mean Corr*) tidak negatif (Sumintono & Widiharso, 2015: 72; Linacre, 2002).

Berdasarkan hasil analisis kecocokan model instrumen tes kelas XI dengan program *Winstep* diketahui dari 40 butir yang dianalisis terdapat 39 butir (97,5%) yang cocok dengan model (*fit model*) dan yang tidak cocok dengan model adalah 1 butir (2,5%) yakni butir nomor 4. Selanjutnya pada model *Rasch* dapat diketahui estimasi tingkat kesukaran butir pada instrumen UAS mata pelajaran ekonomi kelas XI.

Parameter Tingkat Kesukaran Butir Tes

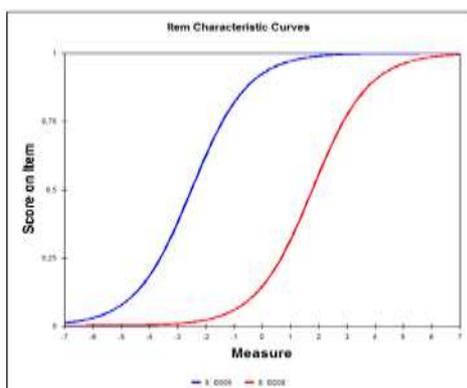
Karakteristik UAS ekonomi kelas XI dianalisis berdasarkan teori respon butir dengan model *Rasch* (1PL) berbantu program *Winstep*. Berdasarkan model *Rasch* (1PL) maka karakteristik butir yang dapat dilihat adalah tingkat kesukaran butir, sedangkan daya beda dianggap konstan. Kriteria tingkat kesukaran butir terbagi dalam lima kategori yakni kategori butir mudah, sedang, sukar, sangat mudah, dan sangat sukar. Penyebaran tingkat kesukaran butir instrumen tes ekonomi kelas XI terletak pada rentang -2,00 sampai dengan 2,00. Butir dikategorikan mudah apabila memiliki nilai mendekati -2,00 logit, butir dikategorikan sedang apabila $-1,00 \text{ logit} < b < +1,00 \text{ logit}$ dan butir dikategorikan sukar apabila nilai mendekati +2,00 logit (Hambleton & Swaminathan: 1985). Sedangkan butir dengan nilai $b > +2,00 \text{ logit}$ masuk kategori sangat sukar dan $b < -2,00 \text{ logit}$ masuk dalam kategori sangat mudah.

Berdasarkan hasil analisis 40 butir instrumen tes UAS ekonomi kelas XI, diketahui butir dengan tingkat kesukaran paling tinggi adalah butir no 8 dengan tingkat kesukaran 1,77 logit dan butir yang masuk dalam kategori sangat mudah adalah butir no 5 dengan nilai tingkat kesukaran -2,51 logit. Kategori tingkat kesukaran butir soal UAS mata pelajaran ekonomi kelas XI dan rangkumannya disajikan pada Tabel 2 berikut ini.

Tabel 2. Kategori Tingkat Kesukaran Butir Tes UAS Ekonomi Kelas XI

No	Tingkat Kesukaran (logit)	Kategori	No	Tingkat Kesukaran (logit)	Kategori
1	-0.99	Sedang	21	1.11	Sukar
2	-1.11	Mudah	22	0.95	Sedang
3	1.3	Sukar	23	-0.36	Sedang
4	0.88	Sedang	24	-0.53	Sedang
5	-2.51	Sangat mudah	25	1.37	Sukar
6	0.86	Sedang	26	-0.65	Sedang
7	-1.78	Mudah	27	0.71	Sedang
8	1.77	Sukar	28	1.14	Sukar
9	0.13	Sedang	29	-0.26	Sedang
10	-1.59	Mudah	30	-1.01	Mudah
11	-0.36	Sedang	31	0.76	Sedang
12	0.34	Sedang	32	-0.24	Sedang
13	-0.88	Sedang	33	1.02	Sukar
14	0.88	Sedang	34	-0.84	Sedang
15	-0.41	Sedang	35	-0.73	Sedang
16	-1.47	Mudah	36	-0.36	Sedang
17	-1.03	Mudah	37	1.43	Sukar
18	0.57	Sedang	38	0.2	Sedang
19	0.34	Sedang	39	0.49	Sedang
20	0.64	Sedang	40	0.22	Sedang

Instrumen tes UAS ekonomi kelas XI yang telah diujikan terdapat 6 butir (15%) masuk dalam kategori sukar, 29 butir (72,5%) masuk kategori sedang, 4 butir (10%) yang masuk kategori mudah, dan 1 butir (2,5%) masuk kategori sangat mudah. Rerata tingkat kesukaran instrumen tes ekonomi kelas XI adalah 0,00 logit (kategori sedang) dengan standar deviasi 1,00 logit. Karakteristik butir soal UAS Ekonomi kelas XI juga tersaji dalam bentuk ICC yang terdapat dalam Gambar 4.



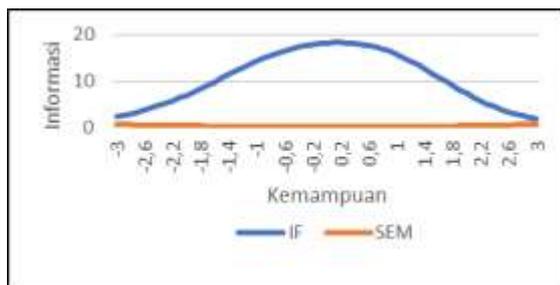
Gambar 4. ICC Instrumen UAS Ekonomi Kelas X

Gambar 4 tersaji ICC (*Item Characteristic Curve*) instrumen UAS Ekonomi kelas XI. ICC ini menggambarkan hubungan antara *ability* atau kemampuan siswa dengan peluang siswa menjawab benar. ICC berbentuk *S-shape* difungsikan untuk mendeskripsikan hubungan antara probabilitas menjawab benar suatu butir pada rentang kemampuan yang diukur (Baker, 2001:7). Pergerakan kurva yang semakin kekanan, menunjukkan kemampuan yang dimiliki siswa semakin tinggi sehingga peluang menjawab benar semakin tinggi, begitu pula sebaliknya. Pada Gambar 4 diketahui butir no 8 menggambarkan butir yang memiliki tingkat kesukaran paling tinggi dalam instrumen UAS Ekonomi kelas XI sedangkan butir no 5 menggambarkan butir yang memiliki tingkat kesukaran paling rendah.

Fungsi Informasi

Fungsi informasi butir merupakan cara untuk menyatakan kekuatan suatu butir pada suatu instrumen tes. Fungsi informasi butir ini berkaitan dengan sumbangan butir dalam mengungkapkan

latent trait yang diukur dengan tes tersebut. Pengukuran koefisien parameter butir dan kemampuan merupakan estimasi, sehingga kebenarannya bersifat probabilistik dan tidak bebas dari kesalahan pengukuran. *Standard Error of Measurement* (SEM) dan fungsi informasi ini mempunyai hubungan yang berbanding terbalik secara kuadrat, dimana semakin besar fungsi informasi, maka semakin kecil SEM atau sebaliknya. Adapun grafik hubungan fungsi informasi butir dan SEM disajikan pada Gambar 5 berikut.



Gambar 5. Hubungan Fungsi Informasi Butir dan SEM UAS Ekonomi Kelas XI

Gambar 5 menggambarkan bahwa nilai fungsi informasi (TIF=*Test Information Function*) maksimum dari suatu tes ekonomi kelas XI dengan 40 butir adalah 18,3379 pada kemampuan siswa (θ) sekitar 0,2 logit dan SEM sebesar 0,2335. Menurut Hambleton (dalam Wiberg, 2004) tes yang baik (reliabel) memiliki nilai TIF = 10. Hal ini menunjukkan bahwa siswa kelas XI memberikan informasi yang baik dengan kesalahan pengukuran terkecil, apabila dikerjakan oleh siswa yang memiliki kemampuan sekitar 0,2 logit.

SIMPULAN DAN SARAN

Hasil analisis diperoleh UAS mata pelajaran ekonomi kelas XI mempunyai rerata tingkat kesukaran yang sedang. Hal ini didukung dengan nilai fungsi informasi maksimum yang diperoleh untuk kelas XI sebesar 18,3379 pada θ sekitar 0,2 logit dan SEM sebesar 0,2335.

Berdasarkan kesimpulan ada beberapa saran yang diajukan, diantaranya: (1) Bagi guru ekonomi SMA, hasil analisis butir soal UAS ini dapat dijadikan bank soal, (2) Bagi kepala sekolah, pentingnya penilaian yang baik

menjadikan perlu adanya pelatihan analisis butir soal bagi guru

DAFTAR PUSTAKA

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Journal of Educational and Psychological Measurement*. Vol (45). 134
- Baker, F.B. (2001). *The basics of item response theory (2ed.)*. Wisconsin: ERIC Clearinghouse on Assessment and Evaluation.
- Best, J.W & Kahn, J.V. (1998). *Research in education 8th edition*. London: Allyn & Bacon
- Dai, Trang Le. (2013). *Applying item response theory modeling in educational research*. Graduate Theses and Dissertation; tidak diterbitkan, Iowa State University
- DeMars, C. (2010). *Item response theory: understanding statistics measurement*. New York: Oxford University Press, Inc.
- Field, A. (2009). *Discovering statistics using SPSS (3rd ed.)*. London: Sage Publication Ltd.
- Furr, R. M. & Bacharach, V. R. (2008). *Psychometrics an introduction*. Thousand Oaks: Sage Publications, Inc
- Greiff, S. et al. (2013). Complex problem solving in educational contexts (Something beyond: concept, assessment, measurement invariance, and construct validity). *Journal of Educational Psychology*, 18, p. 1-15.
- Hambleton R.K. & Swaminathan H., (1985). *Items response theory: principles and application*. Boston: Kluwer-Nijhoff Publish.
- Hambleton, R.K, Swaminathan, H & Rogers, H.J. (1991). *Fundamentals of item response theory*. Newbury Park: SAGE Publications, Inc

- Kartowagiran, B. (Oktober 2012). *Penulisan butir soal*. Makalah disampaikan pada Pelatihan Penulisan dan Analisis Butir Soal bagi Sumber Daya PNS Dik-Rekinpeg, di Hotel Kawanua Aerotel
- Liin, R.L. (1989). *Educational measurement*. New York: Mc Millan Publish
- Linacre, J.M. (1994a). Sample size and item calibration stability. *Rasch Measurement Transaction*, 7 (4). 328.
- Linacre, J.M. (2002b). What do infit and outfit mean-square and standardized mean?. *Rasch Measurement Transaction*, 16(2). 878.
- Lord, F.M. (1980). *Application of item response theory to practice testing problem*. Hillsdale: Lawrence Elbaum Associates.
- Mardapi, D. (2008a). *Teknik penyusunan instrumen tes dan non tes*. Yogyakarta: Mitra Cendekia Offset
- Mardapi, D. (2016b). *Pengukuran, penilaian dan evaluasi pendidikan edisi 2*. Yogyakarta: Nuha Litera
- Meyer, J. P. & Shi Zhu. (2013). Fair and equitable measurement of student learning in moocs: an introduction to item response theory, scale linking, and score equating. *Journal Research & Practice in Assessment*, Vol. 8, p.26-39
- Rasyid, H. & Mansur. (2008). *Penilaian hasil belajar*. Bandung: CV Wacana Prima
- Republik Indonesia. (2005). *Undang-undang no. 14 Tahun 2005 tentang Guru dan Dosen*
- Retnawati, H. (2014). *Item response theory & aplikasi*. Yogyakarta: Nuha Medika
- Rusilowati. (2013). *Kurikulum 2013, 87 persen guru kesulitan cara penilaian*. (Online). Diambil pada tanggal 6 Januari 2016 (<http://unnes.ac.id>)
- Stiggins, R.J. & Chappuis, J. (2012). *An introduction to student involved assessment for learning (6th ed.)*. Boston: Pearson
- Sumintono, B. & Widhiarso, W. (2015). *Aplikasi permodelan RASCH pada assessment pendidikan*. Cimahi: Trim Komunikata
- Wagiran. (2014). *Metodologi penelitian pendidikan (teori dan implementasi)*. Yogyakarta: Depublish
- Wiberg, M. (2004). *Classical test theory vs item response theory*. Stockholm: Umea Universitet
- Wright, B.D., & Stone, M.H. (1979). *Best test design*. Chicago: MESA Press.