

**KARAKTERISTIK BUTIR SOAL TRY OUT UJIAN NASIONAL SMA
MATA PELAJARAN MATEMATIKA Se-KOTA MATARAM
DENGAN MENGGUNAKAN PROGRAM ITEMAN**

Habib Ratu Perwira Negara¹, Kiki Riska Ayu Kurniawati²

STMIK Bumigora Mataram¹, UIN Mataram²

e-mail: habib.ratu27@gmail.com¹, kikirak.2706@gmail.com²

Abstract

The purpose of this research was to know how characteristic items of Try Out National Exam Year 2010/2011 at Senior High School on Mathematics Subject in Mataram City that analyzed by using Iteman Program. The characteristic in question was whether the item can be used or not based on the difficulty level of the item, the difference in power, and the functioning of the answer option. The research method used was descriptive. Techniques of collecting data in the form of documents. Source of data in the form of answer sheet of computer package 61 of Try Out National Exam Year 2010/2011 at Senior High School on Mathematics Subject in Mataram City. Aspects analyzed by using this Iteman Program include the level of difficulty item, different power, and the function of whether or not the answer choice. Based on the results of Item of Analysis Try Out National Examination of School Year 2010/2011 Senior High School on Mathematics Subjects in Mataram City by using Iteman Program, it can be concluded that for the level of difficulty Package 61 was classified. Different power for package 61 has a good category. For the reliability of having a high category and a matter of Try Out Package 61 there are 23 problems received with the improvement of 4, 5, 8, 9, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 28, 30, 31, 32, 34, 35, 36, and 40. Improvements need to be made on the less functioning effluent. There are 15 accepted questions: 1, 2, 3, 7, 10, 11, 12, 13, 15, 27, 29, 33, 37, 38, and 39. And there are 2 rejected questions 6 and 25.

Keywords: analysis of items, Iteman Program.

Abstrak

Tujuan penelitian ini adalah untuk mengetahui bagaimana karakteristik butir soal Try Out Ujian Nasional (UN) Tahun Ajaran 2010/2011 di SMA pada mata pelajaran matematika Se-Kota Mataram yang dianalisis dengan menggunakan Program Iteman. Karakteristik yang dimaksud adalah apakah butir soal dapat digunakan atau tidak berdasarkan tingkat kesukaran butir soal, daya beda, dan berfungsi tidaknya pilihan jawaban. Metode penelitian yang digunakan adalah deskriptif kuantitatif. Teknik pengumpulan data berupa dokumen. Sumber data berupa lembar jawaban komputer (LJK) Soal Try Out UN Tahun Ajaran 2010/2011 paket 61 pada mata pelajaran matematika Se-Kota Mataram. Aspek yang dianalisis dengan menggunakan Program Iteman ini diantaranya tingkat kesukaran butir soal, daya beda, dan berfungsi tidaknya pilihan jawaban. Berdasarkan hasil analisis butir soal Try Out UN Tahun Ajaran 2010/2011 di SMA pada mata pelajaran matematika Se-Kota Mataram dengan menggunakan Program Iteman dapat disimpulkan bahwa pada paket 61 untuk tingkat kesukarannya tergolong sedang. Daya beda untuk paket 61 memiliki kategori yang baik. Untuk reliabilitasnya, soal ini memiliki kategori yang tinggi. Pada soal Try Out paket 61 terdapat 23 soal yang diterima dengan perbaikan yaitu nomor 4, 5, 8, 9, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 28, 30, 31, 32, 34, 35, 36, dan 40. Perbaikan perlu dilakukan pada pengecoh yang kurang berfungsi. Terdapat 15 soal yang diterima yaitu pada nomor 1, 2, 3, 7, 10, 11, 12, 13, 15, 27, 29, 33, 37, 38, dan 39. Serta terdapat 2 soal yang ditolak, yaitu soal nomor 6 dan 25.

Kata kunci : analisis soal, Program ITEMAN.

I. PENDAHULUAN

Pendidikan adalah suatu usaha atau kegiatan yang dijalankan dengan sengaja, teratur dan berencana dengan maksud mengubah atau mengembangkan perilaku yang diinginkan. Sekolah sebagai lembaga formal merupakan sarana dalam rangka pencapaian tujuan pendidikan tersebut. Belajar dalam pendidikan formal yaitu menunjukkan adanya perubahan yang sifatnya positif, sehingga pada akhirnya akan didapat keterampilan, kecakapan dan pengetahuan baru. Hasil dari proses belajar tersebut tercermin dalam prestasi belajarnya, dan untuk mengetahui apakah prestasi yang diperoleh baik atau tidak perlu adanya suatu penilaian. Begitu juga dengan yang terjadi pada seorang siswa yang mengikuti suatu pendidikan selalu diadakan penilaian dari hasil belajar seorang siswa untuk mengetahui sejauh mana telah mencapai sasaran belajar.

Banyak cara yang dilakukan pemerintah dalam mengendalikan mutu pendidikan di Indonesia. Salah satunya adalah dengan mengadakan evaluasi secara menyeluruh yang dikenal dengan Ujian Nasional (UN). Sesuai dengan undang-undang kependidikan mengenai evaluasi, akreditasi dan sertifikasi pasal 57 ayat satu dan dua yaitu:

- (1) Evaluasi dilakukan dalam rangka pengendalian mutu pendidikan secara nasional sebagai bentuk akuntabilitas penyelenggara pendidikan kepada pihak-pihak yang berkepentingan.
- (2) Evaluasi dilakukan terhadap siswa, lembaga, program pendidikan pada jalur formal dan nonformal untuk semua jenjang, satuan, dan jenis pendidikan [1].

Untuk itu banyak cara yang dilakukan oleh setiap satuan pendidikan dalam hal ini pemerintah setempat untuk mempersiapkan siswa agar siap menghadapi UN. Ini sesuai dengan pasal 59 ayat (1) yang menyatakan bahwa pemerintah daerah melakukan evaluasi terhadap pengelola, satuan, jalur, jenjang, dan jenis pendidikan (2) masyarakat dan/atau organisasi profesi dapat membentuk lembaga yang mandiri untuk melakukan evaluasi. Salah satu usaha pemerintah adalah mengadakan *Try Out* UN. Di setiap satuan pendidikan kegiatan

ini merupakan kegiatan rutin yang dilakukan menjelang UN. *Try Out* yang dilaksanakan sebulan penuh ini dilakukan tidak lain bertujuan untuk membiasakan siswa menghadapi UN, selain itu menambah pengetahuan tentang soal-soal UN, *Try-Out* diharapkan dapat membuat siswa semakin akrab dengan hal-hal seputar UN, Hal ini diperlukan karena pihak sekolah tentu tidak mau ada siswa yang tidak lulus hanya karena hal-hal administratif. Jadi, tujuan utama dari diadakannya *Try Out* UN adalah memberikan gambaran atau prediksi tentang soal-soal UN agar siswa nantinya dapat menjawab soal-soal UN tersebut.

Soal *Try Out* UN merupakan sebagai salah satu alat evaluasi yang digunakan untuk melihat kesiapan dari siswa sebelum menghadapi UN. Untuk itu agar evaluasi yang dilakukan sesuai dengan yang diharapkan, maka diperlukan adanya analisis soal *Try Out* UN dengan tujuan untuk mengetahui butir-butir soal yang sudah dapat menjalankan fungsinya sebagai alat pengukur yang baik atau tidak, sehingga akan mendapatkan berbagai informasi yang berharga yang dijadikan sebagai tolok ukur untuk melakukan perbaikan kembali terhadap butir-butir soal yang telah digunakan pada *Try Out* UN sebelumnya, sehingga tujuan utama diadakannya *Try Out* UN tercapai seperti yang dikemukakan sebelumnya. Hal ini sejalan dengan penelitian yang dilakukan oleh Wika Sevi Oktanin dan Sukirno (2015) yang memperoleh hasil bahwa soal ujian akhir semester genap pada mata pelajaran ekonomi akuntansi kelas XI IPS SMAN 1 Kalasan Tahun Ajaran 2013/2014 merupakan soal yang belum berkualitas baik [2]. Penelitian yang serupa yang dilakukan oleh Muslikah Purwanti (2014) yang memperoleh hasil bahwa soal ujian akhir mata pelajaran akuntansi keuangan menggunakan *Microsoft Office Excel* 2010 di SMK Negeri 1 Yogyakarta tahun ajaran 2013/2014, memiliki kualitas yang baik sebesar 23,33%, kualitas yang kurang baik sebesar 16,67% dan kualitas yang tidak baik sebesar 60% [3].

Hasil-hasil penelitian tersebut menunjukkan bahwa perlunya dilakukan analisis butir soal untuk melihat apakah butir

soal sudah tepat atau belum dalam mengukur kemampuan siswa.

1.1 Teori Tes

Tes merupakan sejumlah pertanyaan yang memiliki jawaban yang benar atau salah. Tes diartikan juga sebagai sejumlah pertanyaan yang membutuhkan jawaban, atau sejumlah pernyataan yang harus diberikan tanggapan. Dengan tujuan mengukur tingkat kemampuan seseorang atau mengungkapkan aspek tertentu dari orang yang dikenai tes [4]. Nana Sudjana berpendapat bahwa tes sebagai alat penilaian adalah pertanyaan-pertanyaan yang diberikan kepada siswa untuk mendapatkan jawaban dari siswa dalam bentuk lisan (tes lisan), dalam bentuk tulisan (tes tulisan), atau dalam bentuk perbuatan (tes tindakan) [5]. Jadi, dari pengertian tersebut dapat disimpulkan bahwa tes adalah suatu alat untuk mengukur ketercapaian siswa apakah mereka sudah atau belum menguasai materi yang telah diajarkan selama proses belajar yang telah dilaluinya.

1.2 Analisis soal

Butir soal merupakan perangkat utama dalam sistem penilaian terhadap siswa di sekolah. Untuk itu sangat penting dilakukan analisis soal, untuk menentukan mana soal-soal yang cacat atau tidak berfungsi penggunaannya. Analisis soal mencakup analisis perangkat dan analisis butir soal. Dalam perangkat soal akan dianalisis mutunya dengan menggunakan dua spesifikasi yaitu validitas dan reliabilitas. Sedangkan analisis butir soal, difokuskan pada penilaian mutu butir soal yang berdasarkan pada spesifikasi butir soal yaitu parameter kualitatif dan karakteristik butir soal yaitu parameter kuantitatif.

1.3 Validitas

Pengertian validitas sangat erat berkaitan dengan tujuan pengukuran, validitas berasal dari kata *validity* yang mempunyai arti sejauh mana **ketepatan** dan **kecermatan** suatu instrumen pengukur tes dalam melakukan fungsi ukurannya [6]. Secara umum pengertian validitas adalah suatu konsep yang berkaitan dengan sejauh mana tes telah mengukur apa

yang seharusnya diukur [7]. Suatu tes mempunyai validitas yang tinggi apabila tes tersebut menjalankan fungsi ukuran, atau memberikan hasil ukur yang tepat dan akurat sesuai dengan maksud dari tes tersebut.

Untuk mengukur validitas konstruk dapat dilihat pada tabel spesifikasi kisi-kisi/*blueprint* yang dibuat guru dan dibandingkan dengan kurikulum yang berlaku di sekolah itu. Sedangkan untuk menghitung validitas empiris dengan *korelasi product moment* dan *korelasi point biserial*. *Korelasi product moment* digunakan untuk menghitung dua nilai hasil tes, sedangkan *korelasi point biserial* digunakan untuk menghitung validitas tiap butir item suatu tes.

Rumus *korelasi point biserial* [8]:

$$r_{pbi} = \frac{M_p - M_t}{SD_t} \sqrt{\frac{p}{q}}$$

Keterangan :

r_{pbi} : Koefisien *korelasi point biserial*

M_p : Skor rata-rata hitung yang dimiliki *testee* yang menjawab benar untuk butir soal yang dicari validitasnya

M_t : Skor rata-rata dari skor total

SD_t : Standar deviasi skor total

p : Proporsi *testee* yang menjawab benar

q : Proporsi *testee* yang menjawab salah ($q = 1 - p$)

Dengan interpretasi, jika $r_{pbi} > r$ tabel *product moment* pada taraf signifikansi 5%, maka item soal tersebut valid, tapi jika $r_{pbi} < r$ tabel *product moment* pada taraf signifikansi 5%, maka item soal tersebut invalid.

1.4 Reliabel

Reliabel merupakan kata sifat dan reabilitas merupakan kata benda. Kata “reliabel” diartikan dapat dipercaya. Secara umum sebuah tes dikatakan dapat dipercaya jika memberikan hasil yang tetap apabila dites berkali-kali. Secara umum dapat diartikan sebagai sejauh mana suatu alat ukur dapat diyakini memberikan informasi yang konsisten dan tidak mendua tentang karakteristik peserta tes yang diujikan. Jika dikaitkan dengan tes hasil belajar, Anas Sudjiono berpendapat, “Sebuah tes hasil belajar dapat dinyatakan reliabel (reliable) apabila hasil-hasil pengukuran yang

dilakukan dengan menggunakan tes tersebut secara berulang kali terhadap subyek yang sama, senantiasa menunjukkan hasil yang tetap sama atau sifatnya ajeg dan stabil” [8]. Untuk menghitung besarnya reliabilitas digunakan beberapa metode, yaitu; *Test-retest methods* (Stabilitas), *Paralel* (Ekuivalen), *Solit-Half Methods* (Belah Dua), dan *Internal consistency*. Metode yang sering digunakan yakni *internal consistency*. Metode ini didasarkan pada homogenitas atau korelasi antar skor jawaban pada setiap butir soal. Jika korelasi rerata antar butir soal tinggi, maka reliabilitasnya juga tinggi. Jika korelasi mendekati nol, maka internal konsistensinya nol pula dan reliabilitasnya rendah [7]. Program Iteman dalam menganalisis soal juga menggunakan metode *internal consistency*. Persamaan yang digunakan untuk mencari reliabilitas dengan *internal consistency* yaitu Kuder-Richardson (KR-20).

$$r_{11} = \left(\frac{n}{n-1}\right) \left(\frac{S^2 - \sum pq}{S^2}\right)$$

Keterangan :

r_{11} : koefisien reliabilitas tes

n : banyaknya butir item

S^2 : standar deviasi tes

p : proporsi *testee* yang menjawab item benar

q : proporsi *testee* yang menjawab item salah

Dengan interpretasi, jika koefisien reliabilitas diantara 0,7 sampai 0,8 maka soal tersebut memiliki reliabilitas cukup tinggi.

1.5 Analisis Butir Soal

1) Analisis Kualitatif

Analisis kualitatif sering juga disebut sebagai validitas logis yaitu berupa penelaah yang dimaksudkan untuk menganalisis soal ditinjau dari segi teknis, isi, dan editorial [7]. Analisis ini dilakukan sebelum soal diujikan, untuk melihat apakah soal akan berfungsi atau tidak. Analisis teknis dimaksudkan sebagai penelaahan soal berdasarkan prinsip-prinsip pengukuran dan format penulisan soal. Analisis secara isi dimaksudkan sebagai penelaahan khusus yang berkaitan dengan kelayakan pengetahuan yang ditanyakan. Analisis secara editorial dimaksudkan sebagai penelaahan yang

khususnya berkaitan dengan keseluruhan format dan kejelasan editorial dari soal. Ada dua macam teknik analisis kualitatif, yaitu teknik moderator dan teknik panel.

2) Analisis Kuantitatif

Analisis soal secara kuantitatif menekankan pada analisis karakteristik internal tes melalui data yang diperoleh secara empiris. Karakteristik butir soal meliputi parameter soal kesukaran, daya beda, berfungsi tidaknya pilihan jawaban. Ketiga karakteristik butir ini secara bersama-sama akan menentukan mutu butir soal. Jika salah satu dari ketiga karakteristik ini tidak memenuhi persyaratan, maka mutu butir soal akan turun [9]. Analisis soal secara kuantitatif ini dilakukan setelah soal diujikan. Ada 2 (dua) pendekatan dalam analisis kuantitatif, yaitu pendekatan klasik dan pendekatan modern. Pada prakteknya pendekatan klasik masih dipilih untuk menganalisis butir soal dikarenakan tidak menggunakan parameter yang banyak.

Pendekatan secara klasikal/tradisional adalah proses penelaah butir soal melalui informasi dari jawaban siswa guna meningkatkan mutu butir soal yang bersangkutan. aspek yang diperhatikan dalam analisis butir soal secara klasik adalah setiap butir ditelaah dari tingkat kesukaran butir, daya beda dan efektivitas pengecoh.

1.6 Tingkat kesukaran butir

Tingkat kesukaran (TK) butir soal adalah proporsi siswa yang menjawab benar untuk butir soal tertentu. Tingkat kesukaran biasanya dibedakan menjadi tiga kategori yang ketiganya disajikan pada Tabel 1 berikut.

Tabel 1. Kategori Tingkat Kesukaran Soal

Nilai P	Kategori
$p < 0.3$	Sukar
$0.3 \leq p \leq 0.7$	Sedang
$p > 0.7$	Mudah

Semakin besar kategori tingkat kesukaran yang diperoleh, berarti semakin mudah soal tersebut. Secara umum, menurut teori klasik tingkat kesukaran dapat dinyatakan melalui beberapa cara diantaranya (1) proporsi menjawab benar, (2) skala kesukaran linier, (3) indeks Davis,

dan (4) skala bivariat [7]. Tingkat kesukaran butir soal tidaklah menunjukkan bahwa butir soal tertentu itu baik atau tidak baik. Tingkat kesukaran butir soal hanya menunjukkan bahwa butir soal itu sukar atau mudah untuk kelompok peserta tes tertentu. Butir soal yang terlalu sukar atau terlalu mudah tidak banyak memberi informasi tentang butir soal atau peserta tes.

1.7 Daya beda butir

Salah satu tujuan analisis kuantitatif soal adalah untuk menentukan dapat tidaknya suatu soal membedakan kelompok dalam aspek yang diukur sesuai dengan perbedaan yang ada dalam kelompok itu [7]. Semakin tinggi daya pembeda suatu butir soal, semakin baik butir soal tersebut membedakan peserta tes yang pandai dan yang kurang pandai atau membedakan antara peserta tes yang berkemampuan tinggi dengan peserta tes yang berkemampuan rendah. Indeks daya pembeda soal-soal yang ditetapkan dari selisih proporsi yang menjawab dari masing-masing kelompok. Indeks ini menunjukkan kesesuaian antara fungsi soal dengan fungsi tes secara keseluruhan. Dengan demikian validitas soal ini sama dengan daya pembeda soal, yaitu daya dalam membedakan antara peserta tes yang berkemampuan tinggi dengan peserta tes yang berkemampuan rendah.

Indeks yang menunjukkan besarnya daya beda berkisar antara -1 sampai dengan angka +1. Tanda negatif menunjukkan bahwa peserta tes yang kemampuannya rendah dapat menjawab benar, sedangkan peserta tes yang kemampuannya tinggi menjawab salah, atau memberikan gambaran yang terbalik mengenai kualitas peserta tes. Dari sini dapat diambil kesimpulan bahwa soal yang baik jika soal itu hanya dijawab benar oleh siswa yang pandai saja. Adapun interpretasi daya beda disajikan pada Tabel 2 berikut.

Tabel 2.

Interpretasi Indeks Daya Beda Butir Soal

Koefisien	Keputusan soal
> 0,30	Diterima
0,10 s.d. 0,29	Direvisi
	Ditolak

< 0,10	
--------	--

Berbeda dari tingkat kesukaran, daya beda butir soal secara langsung menentukan kualitas butir soal, dalam arti kualitas konstruksi butir soal. Oleh karena itu, jika akan merevisi butir soal, patokan utama yang akan digunakan ialah daya beda butir soal. Perhitungan daya bedanya berlaku untuk setiap *option* dalam suatu butir soal. Hanya untuk *option* yang bukan kunci jawaban (*option* pengecoh) daya bedanya harus negatif.

1.8 Option Pengecoh

Penyebaran pilihan jawaban dijadikan dasar dalam penelaahan soal. Hal ini dimaksudkan untuk mengetahui berfungsi tidaknya jawaban yang tersedia. Dalam menganalisis pola jawaban soal cukup ditinjau dari jumlah pemilih *option* yang benar, jika distraktor (pengecoh) jawaban tersebut dipilih paling sedikit 5% dari *testee*, maka distraktor tersebut baik.

1.9 Kriteria Pemilihan soal

Setelah perhitungan statistik berupa tingkat kesukaran, daya beda, dan distribusi jawaban selesai, selanjutnya kita lakukan pemilihan soal. Sampai saat ini belum ada ketentuan yang dapat dijadikan patokan yang pasti dalam penentuan validitas dan tingkat kesukaran bagi soal yang baik [7]. Menurut Nitko, kriteria pemilihan soal bergantung pada tujuan penggunaan soal. Jika tujuan tes adalah untuk mengukur satu aspek kemampuan, maka tingkat kesukaran sebaiknya berkisar antara 0,26 sampai 0,84. Apabila yang akan diukur adalah sekumpulan aspek kemampuan, sebaiknya tingkat kesukaran berkisar antara 0,30 sampai 0,70 dan validitas yang disarankan adalah lebih besar dari 0,30. Adapun kriteria pemilihan soal disajikan pada Tabel 3 berikut.

Tabel 3. Kriteria Pemilihan Soal

Kriteria Butir Soal	Koefesien	Keputusan soal
Tingkat Kesukaran	0,30 s.d. 0,70	Diterima
	0,10 s.d. 0,29	Direvisi
	Atau 0,70 s.d. 0,90 < 0,10 dan > 0,90	Ditolak
Daya Beda	> 0,30	Diterima
	0,10 s.d. 0,29 < 0,10	Direvisi Ditolak
Proporsi Jawaban	> 0,005	Diterima

Kelebihan analisis butir soal secara klasik adalah murah, dapat dilaksanakan sehari-hari, dengan kecepatan tinggi dengan bantuan kalkulator *scientific* dan program komputer, banyak program komputer yang murah untuk menganalisisnya. Program yang digunakan untuk menganalisis soal yang sudah umum digunakan untuk menganalisis adalah *Microsoft Excel*, *SPSS*, dan *IteMan*. *Microsoft Excel* cukup baik digunakan dalam menganalisis soal akan tetapi *Excel* hanya sedikit lebih baik jika dibandingkan dengan kalkulator, sebab pengoperasiannya juga masih manual. *SPSS* untuk menganalisis data dengan perakitan soal *essay* dan *IteMan* untuk menganalisis soal pilihan ganda. Pada kesempatan ini peneliti menggunakan analisis *IteMan* untuk menganalisis soal *Try Out* UN.

1.10 Program ITEMAN

IteMan merupakan program komputer yang digunakan untuk menganalisis butir soal secara klasik. Program ini termasuk satu paket program dalam *MicroCATn* yang dikembangkan oleh *Assessment Systems Corporation* mulai tahun 1982 dan mengalami revisi pada tahun 1984, 1986, 1988, dan 1993 mulai dari versi 2.00 sampai dengan versi 3.50 [10].

IteMan berguna dalam melakukan analisis statistik untuk menentukan kualitas butir soal mencakup informasi mengenai tingkat kesukaran, daya pembeda soal dan statistik sebaran jawaban berdasarkan pendekatan teori tes klasik. Selain informasi yang dihasilkan di atas, juga dapat mengetahui tentang statistik

tes, kesalahan pengukuran, kurtosis, skewness, standar deviasi serta distribusi skor. Oleh karena itu, informasi yang terdapat dari hasil analisis *IteMan* akan sangat membantu dalam melihat kualitas sebuah tes secara kuantitatif. Pada akhirnya yang diharapkan dengan mengetahui kualitas tes tersebut, dapat dilakukan upaya-upaya perbaikan untuk meningkatkan kualitas pembelajaran serta secara tidak langsung tercapainya tujuan yang ingin dicapai oleh peneliti. Program ini dapat menskor dan menganalisis data soal pilihan ganda untuk 30.000 siswa dan 250 butir soal.

II. METODE PENELITIAN

2.1 Desain dan Pendekatan Penelitian

Desain penelitian adalah semua proses yang diperlukan dalam perencanaan dan pelaksanaan penelitian [11]. Dalam penelitian ini peneliti menggunakan desain deskriptif kuantitatif, karena penelitian ini merupakan analisis dokumen. Dalam penelitian ini peneliti menggunakan pendekatan deskriptif jenis *study* evaluasi. Jenis *study* evaluasi digunakan untuk pengumpulan data atau informasi untuk dibandingkan dengan kriteria, kemudian diambil kesimpulan [12]. Ini sesuai dengan tujuan peneliti, yaitu untuk mengetahui kualitas karakteristik butir soal *Try Out* UN di SMA pada mata pelajaran matematika Se-Kota Mataram.

2.2 Populasi dan Sampel

Populasi adalah wilayah generalisasi yang terdiri atas obyek/subyek yang mempunyai kualitas dan karakteristik tertentu yang ditetapkan oleh peneliti untuk dipelajari dan kemudian ditarik kesimpulannya [13]. Adapun yang menjadi populasi dalam penelitian ini adalah semua siswa SMA yang mengikuti *Try Out* UN Tahun Ajaran 2010/2011 Se-Kota Mataram pada Paket soal 61. Adapun jumlah populasi dalam penelitian ini disajikan pada Tabel 4.

Tabel 4.

Jumlah Peserta Try Out UN SMA Se-Kota Mataram

No	Nama Sekolah	Jumlah
1	SMA Negeri 1 Mataram	139
2	SMA Negeri 2 Mataram	86
3	SMA Negeri 3 Mataram	71
4	SMA Negeri 4 Mataram	51
5	SMA Negeri 5 Mataram	80
6	SMA Negeri 6 Mataram	63
7	SMA Negeri 7 Mataram	75

8	SMA Negeri 8 Mataram	28
9	SMA Muhammadiyah	27
10	SMA Hangtuah	10
11	SMA Al-Marif	4
Jumlah		634

Penentuan sampel dalam penelitian ini menggunakan teknik *proportional random sampling*. Dari populasi yang disebutkan, yakni 11 SMA Se-Kota Mataram akan diambil 50% untuk setiap sekolah pada setiap sekolah. Berikut ini disajikan jumlah sampel yang digunakan pada Tabel 5.

Tabel 5.
Sampel Peserta Try Out UN SMA Se-Kota Mataram

No	Nama Sekolah	Jumlah
1	SMA Negeri 1 Mataram	69
2	SMA Negeri 2 Mataram	43
3	SMA Negeri 3 Mataram	35
4	SMA Negeri 4 Mataram	25
5	SMA Negeri 5 Mataram	40
6	SMA Negeri 6 Mataram	31
7	SMA Negeri 7 Mataram	37
8	SMA Negeri 8 Mataram	14
9	SMA Muhammadiyah	13
10	SMA Hangtuah	5
11	SMA Al-Marif	2
Jumlah		314

2.3 Teknik Analisis Data

Program ITEMAN merupakan salah satu *software* yang tersedia untuk melakukan analisis terhadap butir-butir soal. Setelah data terkumpul dari hasil proses pengumpulan data, maka selanjutnya dianalisis dan diinterpretasikan dengan teliti serta kecakapan, sehingga diperoleh suatu kesimpulan yang objektif. Adapun langkah-langkah yang ditempuh dalam analisis data ini dengan menggunakan Program ITEMAN terbagi menjadi tiga tahap, yaitu membuat file data,

menjalankan program ITEMAN, dan melakukan interpretasi hasil.

Selanjutnya, menggunakan kriteria yang telah ditentukan, yaitu taraf kesukaran, daya beda dan efektivitas pengecoh pada tiap butir soal untuk mendeskripsikan hasil analisis tersebut. Dalam melakukan interpretasi butir soal ada beberapa catatan yang perlu diperhatikan, diantaranya yaitu:

1. Ditinjau dari tujuan pelaksanaan tes, perlu diperhatikan bahwa soal yang terlalu mudah atau sulit mungkin memang kurang memberikan informasi yang berguna bagi peserta tes pada umumnya, diantaranya kemungkinan karena belum berfungsinya pengecoh dengan baik, namun demikian pada soal yang terlalu mudah atau sukar ini, apabila pengecohnya berfungsi dengan baik yakni *Prop Endorsing* memenuhi ($> 0,05$) serta daya beda pembeda negatif (kecuali kunci) maka soal tersebut masih memenuhi untuk diterima, terlebih jika memang soal tersebut akan digunakan untuk seleksi yang sifatnya sangat menentukan.
2. Soal yang ekstrem mudah atau soal yang ekstrem sukar tidak memberikan informasi yang berguna bagi sebagian besar peserta tes. Sekalipun soal ini terlalu sukar atau mudah apabila setiap pengecoh (distribusi jawaban) pada soal tersebut menunjukkan kondisi distribusi jawaban merata atau distribusi jawaban logis dan daya pembedanya negatif (kecuali kunci), maka soal tersebut masih memenuhi syarat untuk diterima dan sebagai satu alternatif untuk disimpan dalam bank soal.
3. Apabila soal ekstrem mudah atau ekstrem sukar serta daya pembeda dan statistik pengecohnya belum memenuhi kriteria, maka soal tersebut perlu direvisi dan diujicoba lagi.

III. HASIL DAN PEMBAHASAN

4.1 Hasil Penelitian

Hasil analisis dengan menggunakan program ITEMAN Versi 3.00 pada soal pilihan ganda Try Out UN pada tahap kedua pada paket 61 sebanyak 40 soal yang dianalisis, dapat diketahui validitas soal yang meliputi indeks tingkat kesukaran soal, daya pembeda,

dan efektivitas pengecoh serta reliabilitas soal. Berdasarkan hasil analisis tersebut, setiap butir dibandingkan dengan kriteria untuk menentukan apakah butir tersebut diterima, direvisi, atau ditolak, baik dilihat dari tingkat kesukaran, daya beda maupun proporsi jawaban. Di samping itu juga dilihat apakah alternatif jawaban berfungsi secara efektif atau tidak. Kriteria suatu soal diterima, direvisi, atau ditolak telah dibahas pada pembahasan sebelumnya.

a. Tingkat Kesukaran

Soal yang diterima terletak pada rentang tingkat kesukaran 0,30 sampai dengan 0,70. Hal ini dilakukan dengan pertimbangan bahwa soal dengan tingkat kesukaran antara 0,30 sampai 0,70 merupakan soal yang homogen dan dapat menghasilkan penyebaran skor yang luas. Berikut data tingkat kesukaran dari analisis keseluruhan sekolah soal pilihan ganda Try Out UN SMA Se-Kota Mataram Paket 61 dapat dilihat pada Tabel 6.

Tabel 6.

Hasil Analisis Tingkat Kesukaran Butir Soal Try Out UN SMA Se-Kota Mataram

Kategori	Jmh	Nomor Butir	Keputusan Butir
Sangat Mudah (TK > 0,90) Sangat Sukar (TK < 0,10)	1	6	Ditolak
Sedang (0,3 < TK < 0,7)	22	1, 2, 3, 5, 7, 8, 10, 11, 13, 14, 15, 18, 20, 22, 23, 26, 27, 28, 31, 33, 38, 39.	Diterima
Mudah (0,70 < TK < 0,90) Sukar (0,10 < TK < 0,29)	1 16	12 4, 9, 16, 17, 19, 21, 24, 25, 29, 30, 32, 34, 35, 36, 37, 40	Direvisi

b. Daya Pembeda

Soal dengan daya beda diatas 0,30 merupakan soal yang dapat membedakan kelompok yang memiliki kemampuan tinggi dan kelompok yang memiliki kemampuan rendah. Berikut data daya beda dari analisis seluruhan sekolah soal pilihan ganda Try Out UN SMA Se-Kota Mataram Paket 61 dapat dilihat pada Tabel 7 berikut.

Tabel 7.

Hasil Analisis Daya Beda Butir Soal Try Out UN SMA Se-Kota Mataram

Kategori	Jmh	Nomor Butir	Keputusan Butir
DB > 0,3	26	1, 2, 3, 4, 5, 7, 8, 10, 11, 12, 13, 14, 15, 16, 20, 22, 23, 26, 27, 29, 30, 31, 33, 37, 38, 39.	Diterima
0,10 < DB < 0,29	14	6, 9, 17, 18, 19, 21, 24, 25, 28, 32, 34, 35, 36, 40.	Direvisi
DB < 0,10	0	-	Ditolak

c. Efektivitas Pengecoh

Berdasarkan hasil analisis tiap butir soal, dapat dilihat apakah masing-masing alternatif jawaban dapat berfungsi sebagai distraktor atau tidak. Alternatif jawaban yang dapat berfungsi dengan baik akan diterima, alternatif jawaban yang berfungsi namun tidak efektif akan direvisi dan alternatif jawaban yang tidak dapat berfungsi sebagai pengecoh ini dapat dilihat pada Tabel 8 berikut.

Tabel 8.

Hasil Analisis Efektivitas Distraktor Soal Try Out UN SMA Se-Kota Mataram pada Paket 61

Kategori	Jumlah	Nomor Butir
Diterima	17	1, 2, 3, 7, 10, 11, 12, 13, 15, 27, 29, 33, 36, 37, 38, 39.

Direvisi	23	4, 5, 6, 8, 9, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 28, 30, 31, 32, 34, 35, 40.
Ditolak	0	-

d. Reliabilitas

Nilai reliabilitas soal *Try Out* UN paket 61 memiliki nilai 0,767 yang berarti memiliki kehandalan atau keajegan tinggi. Berikut ini disajikan pada Tabel 9 tentang analisis kuantitatif dari seluruh sekolah, yaitu hasil analisis reliabilitas soal pilihan ganda *Try Out* UN SMA Se-Kota Mataram.

Tabel 9.
Hasil Analisis Reliabilitas Soal Try Out UN SMA Se-Kota Mataram

Analisis Reliabilitas	Nilai	Kriteria
	0,767	Tinggi

1. Pembahasan

Kegiatan *Try Out* UN merupakan evaluasi untuk mengetahui sejauh mana materi selama tiga tahun dapat diserap siswa sebelum menghadapi UN. Untuk itu perlu dilakukan analisis secara empiris untuk mengetahui mana soal yang dapat diterima dan dimasukkan dalam bank soal.

Dalam pembahasan ini peneliti akan membahas lebih mendalam tentang hasil empiris tiap butir soal yang telah dihitung dengan menggunakan ITEMAN, baik statistik butir soal (Item Statistik) dan hasil statistik tes serta interpretasi setiap butir soal.

a. Statistik Butir soal

Item Statistics				
Seq No	Scale item	Prop Correct	Point Biser	Biser
1	0-1	0.322	0.528	0.405

Soal 1 memiliki tingkat kesukaran (*proportion correct*) 0.322 (sedang). Sebanyak 32% siswa peserta *Try Out* dapat menjawab soal dengan benar, karena alternatif jawaban D merupakan kunci, maka tanda positif pada validitas soal (biserial) 0.528 menunjukkan bahwa kunci jawaban sudah berfungsi sebagaimana mestinya. Hal ini berarti bahwa peserta *Try Out* yang berkemampuan tinggi

menjawab benar, sedangkan peserta tes yang berkemampuan rendah menjawab salah.

Alternative Statistics				
Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
A	0.299	-0.267	-0.202	
B	0.076	-0.386	-0.209	
C	0.223	-0.021	-0.015	
D	0.322	0.528	0.405	*
E	0.080	-0.235	-0.129	
Other	0.000	-9.000	-9.000	

Ditinjau dari distribusi jawaban, yaitu prosentase peserta tes merespon alternatif jawaban, semuanya berfungsi. Sebanyak 29% peserta *Try Out* merespon alternatif A, 7% peserta *Try Out* merespon alternatif B, 22% peserta *Try Out* merespon alternatif C, dan 8% peserta *Try Out* merespon alternatif E.

Ditinjau dari segi validitas yaitu biserial alternatif jawaban, masing-masing pengecoh A = -0,267, B = -0,386, C = -0,021, dan E = -0,235 cukup sempurna. Tanda negatif pada validitas pengecoh A, B, C, dan E menunjukkan bahwa pengecoh sudah berfungsi dengan baik dimana peserta *Try Out* yang skornya rendah memilih pengecoh sebagai jawaban yang benar. Artinya, peserta *Try Out* yang rendah kemampuannya memilih pengecoh A, B, C, dan E sebagai kunci jawaban.

Terlihat dari distribusi jawaban dan validitas biserial alternatif jawaban bahwa alternatif jawaban A, B, C, E, telah berfungsi sebagaimana mestinya karena yang merespon lebih dari 5% dan biser bertanda negatif karena pengecoh dikatakan berfungsi apabila dipilih lebih dari 5% dan biser bertanda negatif. Secara umum soal ini baik dan dapat diterima.

Proses menginterpretasi butir soal tersebut dilakukan terus sebanyak 40 butir soal yang ada. Proses ini dilakukan untuk menentukan mana soal yang memiliki karakteristik baik, direvisi maupun ditolak berdasarkan kriteria pada Tabel 3. Adapun hasil rangkuman butir soal yang memiliki karakteristik baik untuk digunakan, yaitu butir soal nomor 1, 2, 3, 7, 10, 11, 12, 13, 15, 27, 29, 33, 37, 38, dan 39. Butir soal ini diterima berdasarkan kategori tingkat kesukaran yang berada pada interval $0,3 < TK < 0,7$, daya pembeda $DB > 0,3$ dan pengecoh

yang berfungsi yakni dipilih lebih dari 5% dan validitasnya negatif. Berbeda dengan butir soal yang lain, butir soal nomor 29 dan 37 tergolong sangat mudah dan sukar, tetapi keduanya memiliki daya beda dalam kategori diterima dan semua pengecoh berfungsi. Oleh karena itu butir soal ini masih dapat diterima untuk digunakan.

Butir soal yang diterima dengan revisi adalah butir soal nomor 4, 5, 8, 9, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 28, 30, 31, 32, 34, 35, 36, 40. Pada butir-butir soal tersebut, perbaikan perlu dilakukan pada pengecoh yang kurang berfungsi. Butir soal nomor 6 dan 25 merupakan butir soal yang ditolak. Hal ini dikarenakan butir soal nomor 6 memiliki kategori yang sangat sulit dan daya beda yang tidak berfungsi dengan baik, sedangkan nomor 25 memiliki kategori tingkat kesukaran sukar, daya beda yang kurang baik dan semua pengecoh tidak berfungsi.

b. Statistik Tes

Secara umum hasil analisis statistik perangkat Soal Try Out UN SMA pada mata pelajaran matematika Se-Kota Mataram dengan menggunakan program ITEMAN paket 61 adalah sebagai berikut.

Scale Statistics

Scale:	0
N of Items	40
N of Examinees	314
Mean	13.471
Variance	31.848
Std. Dev.	5.643
Skew	0.238
Kurtosis	-0.818
Minimum	3.000
Maximum	28.000
Median	13.000
Alpha	0.767
SEM	2.725
Mean P	0.337
Mean Item-Tot.	0.305
Mean Biserial	0.402

Berdasarkan analisis statistik Soal Try Out UN SMA pada mata pelajaran matematika Se-Kota Mataram dengan

menggunakan program ITEMAN, bahwa pada paket 61 dapat disimpulkan, yaitu:

1. Tingkat kesukaran yang dimiliki soal tersebut tergolong sedang tampak dari $mean p = 0,337$.
2. Validitas soal tes tersebut dapat dikategorikan baik, hal ini tampak dari besarnya indeks daya pembeda yang ditunjukkan dalam nilai $mean biserial$ lebih besar dari 0,30 ($mean biserial = 0,402$).
3. Reliabilitas soal tes dapat dikatakan cukup tinggi, hal ini dapat dilihat dari nilai $alpha$ sebesar 0,767.
4. Kesalahan pengukuran yang ditimbulkan karena perangkat soal tes ini sebesar 2,725.

5 KESIMPULAN DAN SARAN

Berdasarkan analisis statistik Soal Try Out UN pada Mata Pelajaran Matematika Se-Kota Mataram dengan menggunakan program ITEMAN, bahwa pada paket 61 disimpulkan:

1. Terdapat 23 soal yang diterima dengan perbaikan, yaitu pada butir soal nomor 4, 5, 8, 9, 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 28, 30, 31, 32, 34, 35, 36, dan 40. Perbaikan perlu dilakukan pada pengecoh yang kurang berfungsi. Terdapat 15 soal yang diterima, yaitu 1, 2, 3, 7, 10, 11, 12, 13, 15, 27, 29, 33, 37, 38, dan 39. Serta terdapat 25 soal yang ditolak, yaitu pada butir soal nomor 6 dan 25.
2. Soal Try Out UN SMA pada mata pelajaran matematika Se-Kota Mataram dengan menggunakan program ITEMAN, paket 61 memiliki reliabilitas dengan kategori tinggi.
3. Soal Try Out UN SMA pada mata pelajaran matematika Se-Kota Mataram dengan menggunakan program ITEMAN, paket 61 memiliki kategori tingkat kesukaran sedang.
4. Soal Try Out UN SMA pada mata pelajaran matematika Se-Kota Mataram dengan menggunakan program ITEMAN, paket 61 memiliki daya beda baik.

Setelah dilakukan analisis Soal Try Out UN SMA pada mata pelajaran matematika Se-Kota Mataram Se-Kota Mataram dengan menggunakan program ITEMAN pada paket 61, maka disarankan agar melakukan uji coba terlebih dahulu dan juga analisis soal, dan bagi peneliti lain bisa melakukan analisis butir secara kualitatif dan kuantitatif pada populasi yang lebih luas untuk mengetahui kualitas guru dalam menyusun soal secara luas.

6 DAFTAR PUSTAKA

- [1] Budi Durachman. 2006. *“Himpunan Peraturan Perundang-Undangan Guru dan Dosen”*, Bandung: Fokus Media.
- [2] Wika Sevi Oktanin dan Sukirno, *“Analisis Butir Soal Ujian Akhir Mata Pelajaran Ekonomi Akuntansi”* Jurnal Pendidikan Akuntansi Indonesia, Vol. XIII, No.1, Tahun 2015.
- [3] Muslikah Purwanti, *“Analisis Butir Soal Ujian Akhir Mata Pelajaran Akuntansi Keuangan Menggunakan Microsoft Office Excel 2010”* Jurnal Pendidikan Akuntansi Indonesia, Vol. XII, No.1, Tahun 2014.
- [4] Djemari Mardapi. 2009. *“Teknik Penyusunan Instrumen Tes dan Non Tes”*, Yogyakarta: Mitra Medika.
- [5] Nana Sudjana. 2008. *“Penilaian Hasil Proses Belajar Mengajar”*, Bandung: Remaja Rosda Karya.
- [6] Saifuddin Azwar. 2009. *“Tes Prestasi”*, Yogyakarta: Pustaka Pelajar.
- [7] Sumarna Surapranata. 2005. *“Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes”*, Bandung: Remaja Rosdakarya.
- [8] Anas Sudijono. 2011. *“Pengantar Evaluasi Pendidikan”*, Jakarta: Rajawali Pres.
- [9] Asmawati Zainul. 2005. *“Penilaian Hasil Belajar”*, Jakarta: PAU-PPAI, Universitas Terbuka.
- [10] Bahtiar, *“Analisis Butir Soal Obyektif”* Jurnal Beta, IAIN Mataram, Mataram, 2009.
- [11] Moh. Nazir. 1999. *“Metode Penelitian”*, Jakarta: Ghalia Indonesia.
- [12] Suharsimi Arikunto. 1999. *“Dasar-dasar Evaluasi Pendidikan Edisi Revisi”*, Jakarta: Bumi Angkasa.
- [13] Sugiyono. 2010. *“Metode Penelitian Pendidikan Pendekatan Kuantitatif, Kualitatif dan R&D”*, Bandung: Alfabeta.