

Analisis Soal Modul Laboratorium Fisika Sekolah I Menggunakan *Racsh Model*

Sri Wahyu Widyaningsih^{1,*} dan Irfan Yusuf²

^{1,2} Jurusan Pendidikan Fisika, Universitas Papua

Email: s.widyaningsih@unipa.ac.id

Abstract

This study aims to determine the quality of problems that exist in the school lab module 1 using Racsh Model. The test was conducted on Physics Education Study Program Students of University of Papua Semester Even Semester 2016-2017. This research is descriptive research using quantitative approach. The results showed that 1) there is a question that is too easy (outlier) that is S17 and S38 and need to be revised. 2) There are 3 questions that have low power difference that is about S11, S17, and S38. 3) Reliability questions obtained 0.62 results in the weak category. 4) The results of analysis obtained value Raw variance 30.8% in both category and Unexplned variance value is 9.5%, so it can be concluded that the matter is valid constructively. 5) The results of content validity analysis show that: the degree of problem kesulitas only consists of 2 groups that is easy and difficult problem, the level of accuracy of the problem is in the category enough, there are problems that are not fit ie S50, S41, S44, S40, S37, S45, S31, S27, S28, S31, S42, S12, S20, there is a question that does not meet the Z-standard Output Value (ZSTD) criteria S50, there is a problem that does not meet the standard Point Value Correlation Value (Pt Mean Corr) is a matter of numbers S41, S42, S12, and S20 and there is a problem that has the value of Pt Mean Corr minus is a matter of no 41 which means the matter is confusing. For that reason S11, S17, S38, S50, S41, S44, S40, S37, S45, S36, S47, S43, S32, S27, S28, S31, S42, S12, S20 need to be revised.

Key word: *Questions Analysis, School Laboratory, Racsh Model*

Abstrak

Penelitian ini bertujuan untuk mengetahui kualitas soal yang ada pada modul laboratorium sekolah 1 menggunakan Racsh Model. Pengujian dilakukan pada Mahasiswa Program Studi Pendidikan Fisika Universitas Papua Semester Genap 2016-2017. Penelitian ini merupakan penelitian deskriptif dengan menggunakan pendekatan kuantitatif. Hasil penelitian menunjukkan bahwa 1) terdapat soal yang terlalu mudah (*outlier*) yaitu S17 dan S38 dan perlu direvisi. 2) Terdapat 3 soal yang memiliki daya beda yang rendah yaitu soal S11, S17, dan S38. 3) Reliabilitas soal diperoleh hasil 0,62 yaitu pada kategori lemah. 4) Hasil analisis diperoleh nilai *Raw variance* 30,8% pada kategori baik dan nilai *Unexplned variance* yaitu 9,5%, sehingga dapat disimpulkan bahwa soal valid secara konstruk. 5) Hasil analisis validitas konten menunjukkan bahwa: tingkat kesulitas soal hanya terdiri dari 2 kelompok yaitu soal yang mudah dan sulit, tingkat ketelitian soal berada pada kategori cukup, terdapat soal yang tidak fit yaitu S50, S41, S44, S40, S37, S45, S36, S47, S43, S32, S27, S28, S31, S42, S12, S20, terdapat soal yang tidak memenuhi kriteria Nilai *Output Z-standar* (ZSTD) yaitu S50, terdapat soal yang tidak memenuhi standar Nilai *Point Measure Correlation* (*Pt Mean Corr*) yaitu soal nomor S41, S42, S12, dan S20 dan terdapat soal yang memiliki nilai *Pt Mean Corr* minus yaitu soal no 41 yang artinya soal tersebut membingungkan. Untuk itu maka soal S11, S17, S38, S50, S41, S44, S40, S37, S45, S36, S47, S43, S32, S27, S28, S31, S42, S12, S20, perlu direvisi.

Kata kunci: Analisis Soal, Laboratorium Sekolah, Racsh Model

PENDAHULUAN

Keberhasilan kegiatan pembelajaran dapat dilihat dari hasil evaluasi yang dilakukan. Menurut Suharsimi (2009) evaluasi adalah kegiatan mengumpulkan data-data untuk mengukur sejauh mana tujuan pembelajaran telah tercapai. Terdapat dua teknik untuk melakukan evaluasi yaitu teknik tes dan non tes. Teknik tes merupakan salah satu bentuk evaluasi yang sering digunakan. satu alat ukur yang digunakan dalam evaluasi adalah tes hasil belajar. Sudijono (2013) menyatakan bahwa tes berfungsi untuk mengukur perkembangan yang dicapai oleh peserta didik setelah mengikuti kegiatan pembelajaran dalam rentang waktu tertentu.

Memutuskan hasil belajar dari seorang mahasiswa harus menggunakan instrumen yang dapat dipercaya. Reliabilitas dan validitas instrumen menunjukkan kualitas suatu tes. Tes yang tidak valid dan reliabel akan memberikan hasil yang bias dan bahkan dapat merugikan mahasiswa. Sunyoto D (2012) menyatakan reliabilitas adalah sejauhmana hasil dari evaluasi dapat dipercaya. Ssedangkan, suatu instrumen dikatakan valid jika instrumen tersebut mengukur atribut yang ingin diukur (Sumintono B., & Widhiarso W., 2015).

Mata kuliah Laboratorium Sekolah merupakan mata kuliah yang wajib diambil oleh mahasiswa prodi pendidikan fisika. Mata kuliah tersebut membahas mengenai teori dan praktek pengelolaan laboratorium.

Proses pembelajaran memerlukan sarana dan prasarana pendukung, salah satunya adalah sumber belajar yang digunakan dalam pembelajaran (Firdaus, B, Trapsilasiwi, D, Lestari, N.D.S, 2014). Sumber belajar yang beredar untuk Mata kuliah Laboratorium Sekolah diantaranya adalah modul. Namun soal-soal yang ada didalam modul tersebut belum diketahui tingkat reliabilitas dan validitasnya. Penelitian ini bertujuan untuk mengetahui kualitas soal yang terdapat pada modul Laboratorium Sekolah I.

METODE

Penelitian ini merupakan penelitian deskriptif kuantitatif untuk mencari informasi dan data yang dapat digunakan untuk mendeskripsikan kualitas soal dalam Modul Laboratorium Sekolah 1. Penelitian dilakukan pada 19 orang mahasiswa semester II yang terdiri dari 15 mahasiswi dan 4 mahasiswa yang mengontrak Mata kuliah Laboratorium Sekolah di Program Studi Pendidikan Fisika Universitas Papua (UNIPA). Pengambilan data dilakukan pada bulan April 2017 yaitu saat Ujian

Tengah Semester (UTS) semester Genap 2016-2017.

Intrumen yang digunakan dalam penelitian ini adalah soal dalam modul Laboratorium Sekolah I yang terdiri dari 50 soal pilihan ganda. Data yang terkumpul dianalisis menggunakan Rasch Model.

Hasil analisis reliabilitas dapat dilihat pada program *Winsteps* dengan menggunakan *output tabels 1. Variabel Map*. Reliabilitas dapat ditentukan melalui nilai *separation*, reliabilitas dan nilai α *Crombach*. *Separation* bernilai baik jika terdapat 4 kelompok soal. Untuk mengetahui *separation* (indikasi pengelompokan) melalui perhitungan menggunakan Persamaan berikut:

$$H = \frac{[(4 \times \textit{separation}) + 1]}{3}$$

Nilai reliabilitas mahasiswa dan butir soal ditentukan menggunakan tabel berikut:

Tabel 1. Nilai Reliabilitas

No	Rentang	Kategori
1	<0,67	Lemah
2	0,67-0,80	Cukup
3	0,80-0,90	Bagus
4	0,90-0,94	Bagus Sekali
5	>0,94	Istimewa

Untuk mengukur reliabilitas yaitu interaksi antara mahasiswa dengan butir soal secara keseluruhan digunakan nilai α *Crombach* dengan ketentuan berikut:

Tabel 2. Nilai α *Crombach*

No	Rentang	Kategori
1	<0,5	Buruk
2	0,5-0,6	Kurang Baik
3	0,6-0,7	Cukup
4	0,7-0,8	Baik
5	>0,8	Baik sekali

Validitas dapat dilihat dari validitas konstruk dan validitas konten. Hasil analisis validitas konstruk dapat dilihat pada program *Winsteps* dengan menggunakan *output tabels 23. Item: Dimensionality*. Validitas konstruk dapat ditentukan dengan melihat *Raw Variance* dan *Unexplned variance*. Nilai *Raw Eariance* menggunakan kriteria berikut:

Tabel 3. Nilai *Raw Eariance*

No	Nilai	Ketegori
1	<20%	Kurang baik
2	20%-40%	Baik
3	40%-60%	Baik Sekali
4	>60%	Istimewa

Sedangkan nilai yang diperbolehkan dalam *Unexplned variance* adalah <15%. Hasil analisis validitas konten dapat dilihat

pada program *Winsteps* dengan menggunakan *output tabels* 10, 13, dan 30.

Konten Validity ditentukan melalui

1. Tingkat kesulitan butir soal (Tabel 13. *Item Measure*) dan ketelitian butir soal. Ketelitian butir soal dilihat pada Model SE. Kriteria Model SE adalah sebagai berikut:

Tabel 4. Nilai Reliabilitas

No	Rentang	Kategori
1	<0,5	Baik
2	0,5-1,0	Cukup
3	>1,0	Kurang Baik

2. Tingkat kesesuaian butir (Tabel 10. *Item Fit Order*) kriteria yang digunakan adalah
 - a. Nilai *Outfit mean square* (MNSQ) yang baik memiliki syarat yaitu $0,5 < \text{MNSQ} < 1,5$.
 - b. Nilai *Output Z-standar* (ZSTD) yang diterima adalah $-2 < \text{ZSTD} < +2$.
 - c. Nilai *Point Measure Correlation* (*Pt Mean Corr*) yang diterima adalah $0,4 < \text{Pt Mean Corr} < 0,85$.
3. Analisis Bias butir soal (Tabel 30. *Item DIF*). Sumintono, B dan Widhiarso, W (2015) menyatakan bahwa suatu butir soal dapat dianggap bias apabila nilai probabilitas butirnya $< 5\%$ ($< 0,05$).

HASIL DAN PEMBAHASAN

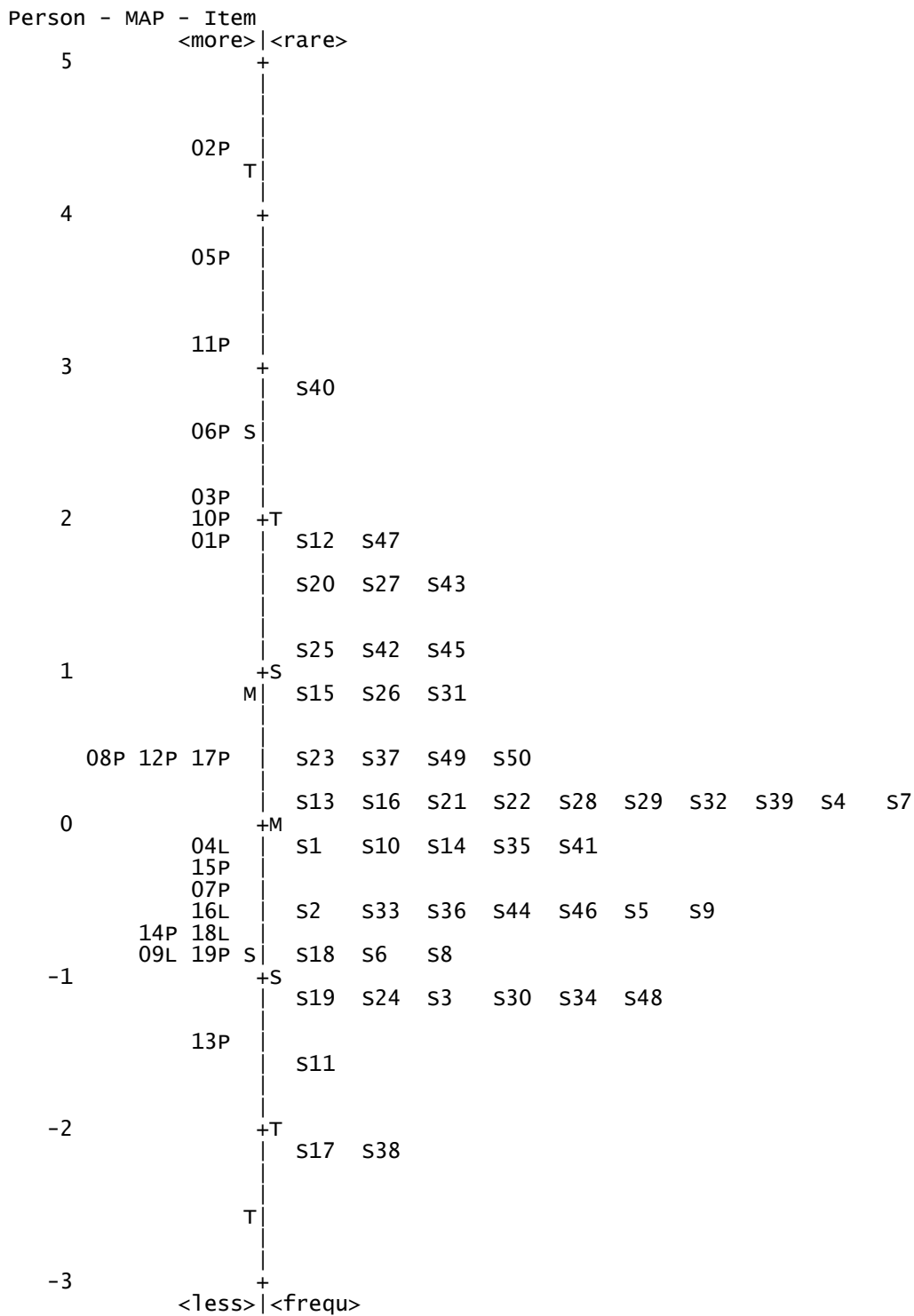
Gambaran umum butir soal yang dianalisis dapat dilihat pada Gambar 1. Gambar tersebut menggambarkan sebaran kemampuan dari 19 orang mahasiswa dan sebaran tingkat kesukaran soal dengan skala yang sama. Dari Gambar tersebut terlihat bahwa ada seorang mahasiswa yang memiliki kemampuan yang tinggi yaitu mahasiswa dengan kode 02P. Nilai logit dari mahasiswa ini adalah lebih dari +4 logit.

Mahasiswa tersebut juga berada di luar batas dua Standar Deviasi (SD) yang ditandai dengan T. Hal ini menunjukkan bahwa mahasiswa ini memiliki kecerdasan tinggi yang berbeda (*outlier*) dari kelompok yang diuji. Mahasiswa yang memiliki kemampuan paling rendah yaitu dengan kode 13P dengan nilai logit lebih kecil dari -1 logit, namun masih berada pada daerah antara 1 SD dan T.

Gambar tersebut juga menunjukkan bahwa soal nomor 40 atau S40 memiliki tingkat kesukaran paling tinggi yang ditunjukkan dengan nilai logit diatas +2 logit. Soal ini menunjukkan mahasiswa yang dapat mengerjakan dengan benar kemungkinan sebanyak 3 orang mahasiswa yaitu 02P, 05P dan 11P.

Soal nomor 17 (S17) dan 38 (S38) memiliki tingkat kesulitan paling rendah yaitu memiliki nilai logit dibawah -2 logit

dan kemungkinan semua mahasiswa dapat mengerjakan soal ini dengan benar.



Gambar 1. Gambaran Umum Butir Soal yang Dianalisis

Dari hasil tersebut terlihat bahwa ada dua soal yang *outlier* yaitu soal S17 dan S38 keduanya posisi dibawah 2SD atau T. Kedua soal ini merupakan soal yang terlalu mudah sehingga perlu direvisi. Menurut Arikunto (2013) soal yang terlalu mudah tidak merangsang keterampilan berfikir peserta didik.

Ada soal-soal yang memiliki tingkat kesukaran yang sama, salah satunya adalah S12 dan S47 yang memiliki nilai logit di bawah +2 logit. Soal-soal yang memiliki tingkat kesukaran yang sama sebaiknya direvisi kembali. Revisi yang harus dilakukan adalah memperbaiki S47 agar memiliki tingkat kesulitan diantara S12 dan S40.

Selain tingkat kesukaran, Gambar 1 juga menunjukkan ada 3 soal yang memiliki daya beda yang rendah yaitu soal S11, S17, dan S38. Soal-soal ini dapat dijawab dengan benar oleh seluruh mahasiswa, artinya soal ini tidak bisa membedakan mahasiswa yang pandai dan

kurang pandai. Arikunto (2013) menyatakan bahwa suatu soal yang dapat dijawab benar oleh siswa pandai dan siswa yang kurang pandai adalah soal yang tidak baik karena tidak mempunyai daya pembeda. Boopathiraj and Chellamani (2013) mengatakan bahwa soal yang memiliki daya beda tinggi adalah soal dimana peserta didik yang memiliki skor tes tinggi dapat menjawab dengan benar sedangkan peserta didik yang memiliki skor tes rendah tidak dapat menjawab soal tersebut dengan benar.

Gambar 1 juga terlihat bahwa rata-rata logit kemampuan mahasiswa berada di atas 0,0 logit. Hal ini menunjukkan bahwa prestasi rata-rata mahasiswa pendidikan fisika angkatan 2016 di atas rata-rata tingkat kesukaran standar soal, beda letak M satu logit menunjukkan perbedaan kemampuan yang cukup besar.

1. Reliabilitas

Reliabilitas butir soal dapat dilihat pada Tabel 5.

Tabel 5. Reliabel Butir Soal

Variabel	Rata-rata logit (SD)	Separation	Reliabilitas	α Crombach
Mahasiswa	0,80 (1,71)	3,50	0,92	0,95
Soal	0,00 (1,03)	1,27	0,62	

Dari hasil analisis diperoleh bahwa rata-rata nilai seluruh mahasiswa dalam mengerjakan soal yang diberikan adalah

0,80 logit. Nilai tersebut lebih besar dari nilai rata-rata kesulitan soal. Artinya terdapat kecenderungan kemampuan siswa

yang lebih tinggi daripada tingkat kesulitan soal. Sebaran mahasiswa diperoleh nilai 3,50 dan sebaran soal diperoleh nilai 1,27. Strata (indikasi pengelompokan) diperoleh melalui perhitungan berikut:

$$H = \frac{[(4 \times \text{separation}) + 1]}{3}$$

$$H_{\text{mahasiswa}} = \frac{[(4 \times 3,50) + 1]}{3} = 5$$

$$H_{\text{soal}} = \frac{[(4 \times 1,27) + 1]}{3} = 2,03$$

Melalui perhitungan berikut ini diperoleh $H_{\text{mahasiswa}}$ ada 5 sedangkan H_{Soal} ada 2 maka pengelompokan yang bisa kita lakukan maksimal hanya dua karena ini indikasi separasi yang ada. Reliabilitas mahasiswa diperoleh hasil 0,92 berada pada kategori bagus sekali. Reliabilitas soal diperoleh hasil 0,62 dimana lebih kecil dari pada 0,67 yaitu pada kategori lemah sehingga dapat disimpulkan bahwa soal yang digunakan dalam Ujian Tengah Semester (UTS) memiliki reliabilitas yang lemah. Hal ini terjadi karena sebaran tingkat kesulitan soal yang terlalu banyak mengumpul di tengah. Maka soal yang

memiliki tingkat kesulitan sama harus direvisi atau menghapusnya sehingga jumlah soal tidak terlalu banyak. Selain itu juga dikarenakan jumlah mahasiswa yang di uji masih kurang. Miller et.al (2009) mengatakan bahwa reliabilitas dipengaruhi oleh beberapa faktor yaitu jumlah *teste*, sebaran skor dan objektivitas.

Untuk nilai α Crombach diperoleh nilai 0,95 pada kategori baik sekali yang menunjukkan bahwa reliabilitas yaitu interaksi antara mahasiswa dan butir-butir soal secara keseluruhan baik sekali.

2. Validitas

2.1 Validitas Konstruk

Dari hasil analisis diperoleh nilai *Raw variance* 30,8% pada kategori baik dan nilai *Unexplned variance* berada pada <15% yaitu 9,5%. Sehingga dapat disimpulkan bahwa soal UTS valid secara konstruk.

2.2 Conten Validity

2.2.1 Tingkat Kesulitan Butir Soal.

Tingkat Kesulitan Soal dapat dilihat Pada Tabel 6.

Tabel 6. Tingkat Kesulitan Butir Soal

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S. E.	INFIT		OUTFIT		PT-MEASURE		EXACT MATCH		Item
					MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.	OBS%	EXP%	
40	4	19	2.82	.71	.88	-.2	2.42	1.4	.54	.59	94.7	85.5	S40
12	6	19	1.91	.65	.41	-1.9	.27	-1.6	.88	.64	100.0	82.1	S12
47	6	19	1.91	.65	1.61	1.5	1.79	1.3	.38	.64	68.4	82.1	S47
20	7	19	1.51	.62	.33	-2.4	.25	-1.9	.91	.64	100.0	81.9	S20
27	7	19	1.51	.62	.64	-1.0	.45	-1.2	.80	.64	89.5	81.9	S27
43	7	19	1.51	.62	1.75	1.8	1.58	1.1	.36	.64	57.9	81.9	S43
25	8	19	1.14	.60	.98	.1	1.10	.4	.62	.63	84.2	80.9	S25
42	8	19	1.14	.60	.43	-2.0	.33	-1.7	.87	.63	94.7	80.9	S42

45	8	19	1.14	.60	1.95	2.3	1.76	1.4	.26	.63	52.6	80.9	S45
15	9	19	.79	.58	.71	-.9	.80	-.2	.72	.61	89.5	78.9	S15
26	9	19	.79	.58	.75	-.8	.67	-.6	.72	.61	89.5	78.9	S26
31	9	19	.79	.58	.49	-1.9	.37	-1.5	.83	.61	89.5	78.9	S31
23	10	19	.45	.57	1.08	.4	.90	.0	.56	.58	73.7	76.3	S23
37	10	19	.45	.57	1.57	1.9	2.08	1.6	.29	.58	63.2	76.3	S37
49	10	19	.45	.57	1.02	.2	.90	.0	.58	.58	73.7	76.3	S49
50	10	19	.45	.57	1.64	2.0	4.98	3.9	.14	.58	73.7	76.3	S50
4	11	19	.13	.56	1.07	.4	.92	.1	.53	.55	73.7	75.4	S4
7	11	19	.13	.56	.93	-.2	.71	-.3	.60	.55	73.7	75.4	S7
13	11	19	.13	.56	.97	-.1	.74	-.2	.59	.55	73.7	75.4	S13
16	11	19	.13	.56	.75	-1.0	.58	-.5	.67	.55	84.2	75.4	S16
21	11	19	.13	.56	.73	-1.2	.55	-.6	.69	.55	84.2	75.4	S21
22	11	19	.13	.56	.92	-.3	.71	-.3	.60	.55	73.7	75.4	S22
28	11	19	.13	.56	.54	-2.2	.41	-.9	.76	.55	94.7	75.4	S28
29	11	19	.13	.56	.81	-.8	.59	-.5	.66	.55	73.7	75.4	S29
32	11	19	.13	.56	.65	-1.5	.49	-.8	.72	.55	84.2	75.4	S32
39	11	19	.13	.56	.74	-1.1	.56	-.6	.68	.55	84.2	75.4	S39
1	12	19	-.18	.56	1.25	1.1	1.15	.5	.40	.52	68.4	74.0	S1
10	12	19	-.18	.56	.73	-1.3	.53	-.5	.65	.52	78.9	74.0	S10
14	12	19	-.18	.56	1.16	.8	.88	.1	.47	.52	68.4	74.0	S14
35	12	19	-.18	.56	1.06	.4	.81	.0	.51	.52	68.4	74.0	S35
41	12	19	-.18	.56	1.98	3.6	2.91	1.9	-.03	.52	47.4	74.0	S41
2	13	19	-.50	.57	.92	-.3	.66	-.1	.53	.48	68.4	73.6	S2
5	13	19	-.50	.57	.78	-1.1	.54	-.3	.59	.48	78.9	73.6	S5
9	13	19	-.50	.57	.83	-.8	.60	-.2	.57	.48	78.9	73.6	S9
33	13	19	-.50	.57	.88	-.5	.62	-.2	.55	.48	78.9	73.6	S33
36	13	19	-.50	.57	1.33	1.5	1.82	1.0	.26	.48	68.4	73.6	S36
44	13	19	-.50	.57	1.79	3.1	2.59	1.5	.02	.48	57.9	73.6	S44
46	13	19	-.50	.57	.85	-.7	.60	-.2	.56	.48	78.9	73.6	S46
6	14	19	-.84	.59	.96	-.1	.68	.0	.47	.43	84.2	75.3	S6
8	14	19	-.84	.59	.82	-.8	.54	-.2	.53	.43	73.7	75.3	S8
18	14	19	-.84	.59	1.11	.5	.78	.1	.40	.43	63.2	75.3	S18
3	15	19	-1.20	.62	.98	.0	.62	.1	.42	.38	73.7	79.4	S3
19	15	19	-1.20	.62	1.08	.4	.78	.3	.36	.38	73.7	79.4	S19
24	15	19	-1.20	.62	1.05	.3	.76	.2	.37	.38	73.7	79.4	S24
30	15	19	-1.20	.62	.92	-.2	.68	.2	.43	.38	84.2	79.4	S30
34	15	19	-1.20	.62	.96	-.1	.61	.1	.43	.38	73.7	79.4	S34
48	15	19	-1.20	.62	.96	-.1	.71	.2	.41	.38	84.2	79.4	S48
11	16	19	-1.61	.68	.86	-.3	.51	.1	.41	.33	84.2	84.2	S11
17	17	19	-2.14	.78	1.14	.4	1.06	.5	.18	.27	89.5	89.5	S17
38	17	19	-2.14	.78	.94	.1	.50	.0	.32	.27	89.5	89.5	S38
MEAN	11.4	19.0	.00	.60	.99	-.1	.98	.1			77.7	77.7	
S.D.	2.9	.0	1.03	.05	.37	1.3	.82	1.0			11.5	4.0	

Tabel 6 diurutkan berdasarkan tingkat kesukaran soal yaitu berdasarkan *measure*. Nilai logit yang tinggi menunjukkan tingkat kesulitan soal yang lebih tinggi. Hal ini berhubungan dengan kolom total skor yang menunjukkan

jumlah mahasiswa yang menjawab benar. Dari tabel juga dapat dilihat nilai rata-rata soal yaitu 0,0 logit dan nilai SD yaitu 1,03. Jika terdapat dua kelompok soal maka tingkat kesulitan soal dapat dikelompokkan sebagai berikut:

Tabel 7. Kelompok Soal

No	Rentang	Keterangan	No Soal
1	> 0,0 logit	Sulit	40, 12, 47, 20, 27, 43, 25, 42, 45, 15, 26, 31, 23, 37, 49, 50, 4, 7, 13, 16, 21, 22, 28, 29, 32, 39
2	< 0,0 logit	Mudah	1, 10, 14, 35, 41, 2, 5, 9, 33, 36, 44, 46, 6, 8, 18, 3, 19, 24, 30, 34, 48, 11, 17, 38

Tabel ini secara tepat menggambarkan tentang pengelompokkan tingkat kesulitan soal. Dari Tabel 6. Terlihat bahwa soal yang sulit terdiri dari 26 soal dan soal mudah terdiri dari 24 soal. Dari proporsinya sudah baik, hanya saja perlu ditambah dengan soal yang sangat sulit dan soal yang sangat mudah. Tabel 5. Juga dapat menunjukkan tingkat ketelitian

soal dengan melihat Model SE (*Standar Error*). Seluruh soal yang dianalisis memiliki ketelitian yang berada pada kategori Cukup yaitu antara 0,5 sampai 1.

2.2.2 Tingkat Kesesuaian Butir Soal

Data tingkat kesesuaian butir soal dapat dilihat pada Tabel 8.

Tabel 8. Tingkat Kesesuaian Butir Soal

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT		OUTFIT		PT-MEASURE		EXACT OBS%	MATCH EXP%	Item
					MNSQ	ZSTD	MNSQ	ZSTD	CORR.	EXP.			
50	10	19	.45	.57	1.64	2.0	4.98	3.9	A .14	.58	73.7	76.3	S50
41	12	19	-.18	.56	1.98	3.6	2.91	1.9	B-.03	.52	47.4	74.0	S41
44	13	19	-.50	.57	1.79	3.1	2.59	1.5	C .02	.48	57.9	73.6	S44
40	4	19	2.82	.71	.88	-.2	2.42	1.4	D .54	.59	94.7	85.5	S40
37	10	19	.45	.57	1.57	1.9	2.08	1.6	E .29	.58	63.2	76.3	S37
45	8	19	1.14	.60	1.95	2.3	1.76	1.4	F .26	.63	52.6	80.9	S45
36	13	19	-.50	.57	1.33	1.5	1.82	1.0	G .26	.48	68.4	73.6	S36
47	6	19	1.91	.65	1.61	1.5	1.79	1.3	H .38	.64	68.4	82.1	S47
43	7	19	1.51	.62	1.75	1.8	1.58	1.1	I .36	.64	57.9	81.9	S43
1	12	19	-.18	.56	1.25	1.1	1.15	.5	J .40	.52	68.4	74.0	S1
14	12	19	-.18	.56	1.16	.8	.88	.1	K .47	.52	68.4	74.0	S14
17	17	19	-2.14	.78	1.14	.4	1.06	.5	L .18	.27	89.5	89.5	S17
18	14	19	-.84	.59	1.11	.5	.78	.1	M .40	.43	63.2	75.3	S18
25	8	19	1.14	.60	.98	.1	1.10	.4	N .62	.63	84.2	80.9	S25
23	10	19	.45	.57	1.08	.4	.90	.0	O .56	.58	73.7	76.3	S23
19	15	19	-1.20	.62	1.08	.4	.78	.3	P .36	.38	73.7	79.4	S19
4	11	19	.13	.56	1.07	.4	.92	.1	Q .53	.55	73.7	75.4	S4
35	12	19	-.18	.56	1.06	.4	.81	.0	R .51	.52	68.4	74.0	S35
24	15	19	-1.20	.62	1.05	.3	.76	.2	S .37	.38	73.7	79.4	S24
49	10	19	.45	.57	1.02	.2	.90	.0	T .58	.58	73.7	76.3	S49
3	15	19	-1.20	.62	.98	.0	.62	.1	U .42	.38	73.7	79.4	S3
13	11	19	.13	.56	.97	-.1	.74	-.2	V .59	.55	73.7	75.4	S13
6	14	19	-.84	.59	.96	-.1	.68	.0	W .47	.43	84.2	75.3	S6
34	15	19	-1.20	.62	.96	-.1	.61	.1	X .43	.38	73.7	79.4	S34
48	15	19	-1.20	.62	.96	-.1	.71	-.2	Y .41	.38	84.2	79.4	S48
38	17	19	-2.14	.78	.94	.1	.50	.0	Y .32	.27	89.5	89.5	S38
7	11	19	.13	.56	.93	-.2	.71	-.3	X .60	.55	73.7	75.4	S7
2	13	19	-.50	.57	.92	-.3	.66	-.1	W .53	.48	68.4	73.6	S2
30	15	19	-1.20	.62	.92	-.2	.68	.2	V .43	.38	84.2	79.4	S30
22	11	19	.13	.56	.92	-.3	.71	-.3	U .60	.55	73.7	75.4	S22
33	13	19	-.50	.57	.88	-.5	.62	-.2	T .55	.48	78.9	73.6	S33
11	16	19	-1.61	.68	.86	-.3	.51	.1	S .41	.33	84.2	84.2	S11
46	13	19	-.50	.57	.85	-.7	.60	-.2	R .56	.48	78.9	73.6	S46
9	13	19	-.50	.57	.83	-.8	.60	-.2	Q .57	.48	78.9	73.6	S9
8	14	19	-.84	.59	.82	-.8	.54	-.2	P .53	.43	73.7	75.3	S8
29	11	19	.13	.56	.81	-.8	.59	-.5	O .66	.55	73.7	75.4	S29
15	9	19	.79	.58	.71	-.9	.80	-.2	N .72	.61	89.5	78.9	S15
5	13	19	-.50	.57	.78	-1.1	.54	-.3	M .59	.48	78.9	73.6	S5
16	11	19	.13	.56	.75	-1.0	.58	-.5	L .67	.55	84.2	75.4	S16
26	9	19	.79	.58	.75	-.8	.67	-.6	K .72	.61	89.5	78.9	S26
39	11	19	.13	.56	.74	-1.1	.56	-.6	J .68	.55	84.2	75.4	S39
10	12	19	-.18	.56	.73	-1.3	.53	-.5	I .65	.52	78.9	74.0	S10
21	11	19	.13	.56	.73	-1.2	.55	-.6	H .69	.55	84.2	75.4	S21
32	11	19	.13	.56	.65	-1.5	.49	-.8	G .72	.55	84.2	75.4	S32
27	7	19	1.51	.62	.64	-1.0	.45	-1.2	F .80	.64	89.5	81.9	S27
28	11	19	.13	.56	.54	-2.2	.41	-.9	E .76	.55	94.7	75.4	S28
31	9	19	.79	.58	.49	-1.9	.37	-1.5	D .83	.61	89.5	78.9	S31
42	8	19	1.14	.60	.43	-2.0	.33	-1.7	C .87	.63	94.7	80.9	S42
12	6	19	1.91	.65	.41	-1.9	.27	-1.6	B .88	.64	100.0	82.1	S12
20	7	19	1.51	.62	.33	-2.4	.25	-1.9	A .91	.64	100.0	81.9	S20
MEAN	11.4	19.0	.00	.60	.99	-.1	.98	.1			77.7	77.7	
S.D.	2.9	.0	1.03	.05	.37	1.3	.82	1.0			11.5	4.0	

Tabel berikut menjelaskan apakah butir soal berfungsi normal melakukan pengukuran atau tidak. Nilai Outfit mean square (MNSQ) yang baik memiliki syarat yaitu $0,5 < \text{MNSQ} < 1,5$. Berdasarkan Nilai tersebut ada beberapa soal yang tidak fit yaitu memiliki indikasi terjadi miskonsepsi dalam memahami dan mengerjakan soal tersebut. Soal yang perlu di hapuskan yaitu soal dengan nomor 50, 41, 44, 40, 37, 45, 36, 47, 43, 32, 27, 28, 31, 42, 12, 20.

Nilai Output Z-standar (ZSTD) yang diterima adalah $-2 < \text{ZSTD} < +2$. Soal yang tidak memenuhi kriteria adalah soal no 50. Nilai Point Measure Correlation (Pt Mean Corr) yang diterima adalah $0,4 < \text{Pt Mean Corr} < 0,85$. Soal yang tidak memenuhi standar adalah soal nomor 41, 42, 12, dan 20. Dari nilai Pt Mean Corr juga dapat dilihat bahwa soal no 41 memperoleh nilai minus yaitu $-0,03$ yang artinya soal tersebut membingungkan.

Kualitas suatu soal juga dapat dilihat dari distraktornya. Distraktor yang baik adalah distraktor yang mampu menarik mahasiswa untuk memilihnya. Sudijono (2013) mengatakan bahwa distraktor dikatakan berfungsi dengan baik apabila

distraktor mempunyai daya rangsang, sehingga mahasiswa yang diuji ragu-ragu sehingga terkecoh untuk memilih distraktor sebagai jawaban betul. Dalam penelitian ini terdapat penempatan distraktor yang kurang tepat. Soal-soal yang penempatan distraktornya tidak tepat yaitu soal nomor 50, 41, 44, 37, 45, 36 dan 23.

2.2.3 Analisis Bias Butir Soal

Suatu soal dapat dikatakan bias apabila diperoleh bahwa salah satu mahasiswa dengan karakteristik tertentu lebih diuntungkan dibandingkan dengan mahasiswa dengan karakteristik lain. Soal UTS diberikan pada mahasiswa dengan melihat jenis kelamin yaitu laki-laki (L) dan perempuan (P). Variabel jenis kelamin ini dapat digunakan untuk mendeteksi bias butir soal.

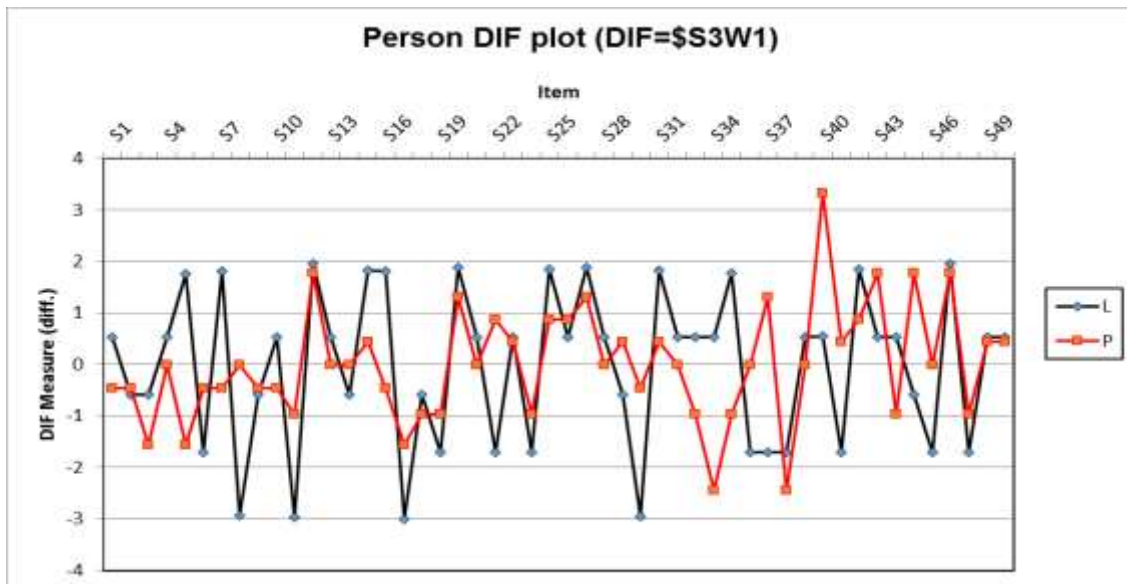
Sumintono, B dan Widhiarso, W (2015) menyatakan bahwa suatu butir soal dapat dianggap bias apabila nilai probabilitas butirnya $< 5\%$ ($< 0,05$). Tabel 9 menunjukkan tidak terdapat soal yang memiliki probabilitas kurang dari 5%.

Tabel 9. Bias Butir Soal

Person CLASSES	SUMMARY DIF		PROB.	BETWEEN-CLASS		Item	
	CHI-SQUARE	D.F.		MEAN-SQUARE	t=ZSTD	Number	Name
2	.4589	1	.4981	.2361	-.3387	1	S1
2	.0085	1	.9267	.0041	-1.3091	2	S2
2	.4821	1	.4875	.2627	-.2912	3	S3
2	.1394	1	.7088	.0655	-.7948	4	S4
1	.0000	0	1.0000	.0000	.0000	5	S5
2	.7029	1	.4018	.3908	-.0989	6	S6
1	.0000	0	1.0000	.0000	.0000	7	S7
1	.0000	0	1.0000	.0000	.0000	8	S8
2	.0085	1	.9267	.0041	-1.3091	9	S9
2	.4589	1	.4981	.2361	-.3387	10	S10
1	.0000	0	1.0000	.0000	.0000	11	S11
1	.0000	0	1.0000	.0000	.0000	12	S12
2	.1394	1	.7088	.0655	-.7948	13	S13
2	.1958	1	.6582	.0964	-.6774	14	S14
1	.0000	0	1.0000	.0000	.0000	15	S15
1	.0000	0	1.0000	.0000	.0000	16	S16
1	.0000	0	1.0000	.0000	.0000	17	S17
2	.0817	1	.7750	.0419	-.9133	18	S18
2	.2478	1	.6186	.1312	-.5720	19	S19
1	.0000	0	1.0000	.0000	.0000	20	S20
2	.1394	1	.7088	.0655	-.7948	21	S21
2	2.6572	1	.1031	1.7957	.9285	22	S22
2	.0048	1	.9448	.0020	-1.3816	23	S23
2	.2478	1	.6186	.1312	-.5720	24	S24
1	.0000	0	1.0000	.0000	.0000	25	S25
2	.0548	1	.8149	.0225	-1.0510	26	S26
1	.0000	0	1.0000	.0000	.0000	27	S27
2	.1394	1	.7088	.0655	-.7948	28	S28
2	.5874	1	.4434	.2938	-.2397	29	S29
1	.0000	0	1.0000	.0000	.0000	30	S30
1	.0000	0	1.0000	.0000	.0000	31	S31
2	.1394	1	.7088	.0655	-.7948	32	S32
2	.9681	1	.3252	.5497	.0878	33	S33
2	2.6056	1	.1065	1.9249	.9889	34	S34
1	.0000	0	1.0000	.0000	.0000	35	S35
2	1.2818	1	.2576	.7572	.2836	36	S36
2	3.4161	1	.0646	2.4724	1.2185	37	S37
2	.1914	1	.6618	.1025	-.6570	38	S38
2	.1394	1	.7088	.0655	-.7948	39	S39
2	2.3990	1	.1214	.7845	.3066	40	S40
2	1.9409	1	.1636	1.2249	.6198	41	S41
1	.0000	0	1.0000	.0000	.0000	42	S42
2	.6683	1	.4136	.2415	-.3289	43	S43
2	.9681	1	.3252	.5497	.0878	44	S44
2	2.5122	1	.1130	1.3924	.7189	45	S45
2	1.2818	1	.2576	.7572	.2836	46	S46
1	.0000	0	1.0000	.0000	.0000	47	S47
2	.2478	1	.6186	.1312	-.5720	48	S48
2	.0048	1	.9448	.0020	-1.3816	49	S49
2	.0048	1	.9448	.0020	-1.3816	50	S50

Jika dilihat dari jenis kelamin, keadaan soal dapat dilihat pada Gambar 2. Gambar 2 menunjukkan bahwa S40 memiliki tingkat kesukaran yang tinggi dan S34 dan S37 memiliki tingkat

kesulitan yang rendah bagi mahasiswa perempuan sedangkan S7, S10, S16, dan S31 memiliki tingkat kesulitan yang rendah bagi mahasiswa laki-laki.



Gambar 2. Tingkat Kesukaran Soal Dilihat Dari Jenis Kelamin

Seorang pendidik harus menguasai teknik penyusunan soal agar soal-soal yang dihasilkan berupa soal yang berkualitas. Hal ini dapat ditempuh dengan cara mengikuti pelatihan atau dengan membaca buku panduan penyusunan soal sehingga dapat membantu tugas guru dalam menganalisis butir soal (Purwanti M, 2014). Jika pendidik ingin menggunakan soal-soal yang ada di buku-buku, maka hendaknya dilakukan analisis terlebih dahulu untuk memastikan kualitas soal yang ada. Damopolii, I (2016) menyatakan bahwa analisis butir soal bertujuan untuk menentukan peran setiap butir soal yang berhubungan dengan seluruh tes.

SIMPULAN DAN SARAN

Berdasarkan hasil analisis butir soal terhadap kualitas butir soal dapat disimpulkan bahwa: terdapat soal yang

sangat mudah (*outlier*) yaitu soal S17 dan S38 dan perlu direvisi. Terdapat 3 soal yang memiliki daya beda yang rendah yaitu soal S11, S17, dan S38. Reliabilitas soal diperoleh hasil 0,62 dimana lebih kecil dari pada 0,67 yaitu pada kategori lemah. Dari hasil analisis diperoleh nilai *Raw variance* 30,8% pada kategori baik dan nilai *Unexplned variance* berada pada <15% yaitu 9,5%. Sehingga dapat disimpulkan bahwa soal valid secara konstruk. Tingkat kesulitas soal hanya dikelompokkan menjadi 2 kelompok yaitu soal yang mudah dan sulit. Terdapat beberapa soal yang tidak fit yaitu memiliki indikasi terjadi miskonsepsi dalam memahami dan mengerjakan soal tersebut yaitu soal S50, S41, S44, S40, S37, S45, S36, S47, S43, S32, S27, S28, S31, S42, S12, S20. Soal yang tidak memenuhi kriteria Nilai *Output Z-standar* (ZSTD)

adalah S50. Soal yang tidak memenuhi standar Nilai *Point Measure Correlation* (*Pt Mean Corr*) adalah soal nomor S41, S42, S12, dan S20. Dari nilai *Pt Mean Corr* juga dapat dilihat bahwa S41 memperoleh nilai minus yaitu -0,03 yang artinya soal tersebut membingungkan. Untuk itu maka soal S11, S17, S38, S50, S41, S44, S40, S37, S45, S36, S47, S43, S32, S27, S28, S31, S42, S12, S20, perlu direvisi.

Berdasarkan hasil analisis terhadap kualitas butir soal maka saran yang dapat diajukan yaitu sebagai berikut: Soal yang baik dapat digunakan kembali dengan tetap menjaga kerahasiaan soal tersebut. Soal yang belum baik perlu direvisi sehingga menjadi soal yang baik. Cara yang dapat dilakukan untuk merevisi antara lain dengan menyamakan isi soal dengan indikator yang sesuai dengan materi atau dengan mengganti pengecoh agar dapat menimbulkan daya tarik bagi peserta tes untuk memilihnya. Proporsi soal yang memiliki daya pembeda yang sama sebaiknya dikurangi atau direvisi agar lebih bisa membedakan siswa yang berkemampuan tinggi dan berkemampuan rendah, yang akan lebih menuntut siswa untuk berpikir. Guru perlu meningkatkan pemahaman serta kemampuannya dalam menyusun soal dan menganalisis butir soal karena dapat membantu dalam mengetahui

kualitas butir soal sehingga soal yang diujikan terdiri dari soal-soal berkualitas.

DAFTAR PUSTAKA

- A. Sudijono. 2013. *Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada.
- Arikunto, S. 2013. *Dasar-dasar Evaluasi Pendidikan*, Edisi 2. Jakarta: Bumi Aksara.
- C. Boopathiraj and K. Chellamani. 2013. Analysis of Test Items On Difficulty Level and Discrimination Index In The Test For Research In Education, *International Journal of Social Science & Interdisciplinary Research*, 2 (2), pp 189-193.
- Damopolii, I. 2016. Analisis Butir Soal Ujian Tengah Semester Mata Kuliah Biologi Umum Di Universitas Papua. *Prosiding Seminar Nasional*, 2 (1), 145-153.
- Firdaus B, Trapsilasiwi D, Lestari N.D.S. 2014. Analisis Soal Dalam Buku Matematika Kelas VII Kurikulum 2013 Berdasarkan *Mathematical Literacy Assessment Taxonomy*. *Jurnal Edukasi UNEJ*, I (3): 42-46.
- Miller M. D., Linn R. L., and N. E. Gronlund. 2009. *Measurement and Assessment in Teaching, Tenth Edition*. United State: Pearson Education Inc.

- Purwanti, M. 2014. Analisis Butir Soal Ujian Akhir Mata Pelajaran Akuntansi Keuangan Menggunakan *Microsoft Office Excel 2010*. *Jurnal Pendidikan Akutansi Indonesia*, 7 (1), 81-94.
- Suharsimi, A. (2009). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Sumintono B. & Widhiarso W. 2015. *Aplikasi Model Rasch untuk Penelitian Ilmu-ilmu Sosial*. Cimahi: Trim Komunikata.
- Sumintono B. & Widhiarso W. 2015. *Aplikasi Pemodelan Rasch pada Assessment Pendidikan*. Cimahi: Trim Komunikata.
- Sunyoto, D. 2012. *Validitas dan Reliabilitas*. Yogyakarta: Nuha Medika.