ISSN 2088-3021 (media cetak)
ISSN 2598-8077 (media online)

JURNAL BUANA MATEMATIKA. Vol. 7, No. 2, Tahun 2017

ANALISIS BUTIR SOAL PILIHAN GANDA MATEMATIKA SEKOLAH MENENGAH PERTAMA

Hodiyanto
Program Studi Pendidikan Matematika, Fakultas MIPATEK, IKIP PGRI Pontianak
hodiyanto@ikippgriptk.ac.id

Abstrak

Instrumen yang baik memiliki reabilitas yang baik dan setiap butir soal memiliki daya beda yang baik pula. Tetapi fakta di lapangan, kebanyakan guru dalam membuat tes hanya mengambil dari buku. Oleh sebab itu, penulis merasa perlu menganalisis butir-butir soal yang dijadikan instrumen oleh guru-guru, sehingga hasil penelitian bisa memberikan masukan terhadap instrument yang digunakan tersebut. Tujuan penelitian adalah untuk mengetahui baik atau tidaknya instrument ditinjau dari reabilitas soal, baik atau tidaknya butir-butir soal ditinjau dari daya pembeda dan indeks kesukaran, dan befungsi dan tidaknya pengecoh dalam soal tersebut. Metode penelitian ini menggunakan metode diskriptif. Berdasarkan hasil analisis diperoleh bahwa reabilitas istrumen soal tersebut tergolong tidak baik, 12 butir soal memiliki daya beda yang jelek, dan tidak berfungsinya pengecoh.
Kata Kunci: evaluasi pembelajaran, reabilitas, daya pembeda

Abstract

The Good instruments have good realibility and each items have good discrimination power. But the facts, most teachers in making tests just take from a book. Therefore, the author feel the need to analyze the items that are used as an instrument by teachers, so that the results of research can provide input to the instrument used. The purpose of this research is to discribe the instrument viewed from the reability, the items viewed from discrimination power, and test distractors. This research method used is descriptive method. Based on the results of the analysis are obtained that the instrument realibility isn't good, the items have ugly discrimination power, and the distractors are not function.
Keywords: *evaluation of learning, realibility, discrimination power*

PENDAHULUAN

Evaluasi adalah salah satu unsur penting dalam pembelajaran karena tanpa evaluasi seorang pendidik tidak dapat mengukur keberhasilan pembelajaran, sejauh mana peserta didik memahami konsep yang diajarkan oleh gurunya. Guba dan Lincoln (1985: 35), mendefinisikan bahwa evaluasi adalah "*a process for describing an evaluand and judging its merit and worth*" yang artinya suatu proses menggambarkan orang yang dievaluasi dan menimbang makna dan nilainya. Menurut Arifin (2012: 8) evaluasi adalah suatu proses yang sistematis dan berkelanjutan untuk menentukan kualitas (nilai dan arti) daripada sesuatu, berdasarkan pertimbangan dan kriteria tertentu untuk membuat suatu keputusan. Dari dua pendapat ini, dapat disimpulkan bahwa evaluasi dalam pembelajaran adalah suatu proses untuk menentukan kualitas dari pada siswa.

Dalam proses evaluasi tentu dilakuakan pengukuran (*measuremen*t) dan penilaian (*assessmen*t). Nitko (1996 : 4) menjelaskan "*assessment is a broad term defined as a process for obtaining information that is used for making decisions about students, curricula and programs, and educational policy*" yang artinya penilaian adalah penilaian adalah istilah yang didefinisikan sebagai proses untuk mendapatkan informasi yang digunakan untuk membuat keputusan tentang siswa, kurikulum dan program, dan kebijakan pendidikan. Sebelum dilakukan proses penilaian, tentu terlebih dahulu dilakukan pengukuran. Reynolds et al (dalam Budiyono, 2015: 2) mendefinisikan "*measurement is a set of rules for assigning numbers to represent objects, traits, attributes, or behaviors*". Pada definisi ini, pengukuran adalah sekumpulan cara untuk memberikan bilangan kepada objek, kemampuan, atribut, atau perilaku untuk menyatakan objek, kemampuan, atribut, atau perilaku.

Untuk melakukan pengukuran tersebut tentu pendidik memerlukan instrument berupa tes. Agar pengukuran tersebut memberikan hasil sesuai dengan apa diharapkan tentu tes yang digunakan tersebut harus tepat. Menurut Budiyono (2015: 91) secara garis besar, untuk menyusun tes yang baik, diperlukan langkah-langkah: menginventarisasi bahan yang telah diajarkan, menyusun spesifikasi tes, menyusun butir-butir soal beserta

kuncinya, menelaah butir-butir tes, melakukan uji coba, melakukan analisis tes dan analisis butir soal berdasarkan hasil uji coba, melakukan revisi terhadapt butir-butir soal yang kurang baik jika memungkinkan untuk melakukan uji coba lagi, menetapkan istrumen, melaksanakan pengukuran kepada subjek yang dikehendaki, dan menafsirkan hasil yang diperoleh.

Tetapi, fakta di lapangan masih banyak guru/pendidik yang kurang memperhatikan langkah-langkah maupun aturan dalam penyusunan tes. Kebanyakan guru dalam membuat tes hanya mengambil dari buku padahal soal tersebut belum tentu baik untuk dijadikan alat ukur. Hasil penelitian Amaliya dan Widayati (2012: 4) analisis soal tes kendali mutu baik secara kualitatif maupunkuantitatif belum pernah dilakukan di Kota Yogyakarta sehingga dari tahun ke tahun kualitas soal tes kendali mutu masih belum diketahui. Analisis terhadap soal tes kendali mutu sangatlah penting dilakukan untuk memperbaiki kualitas soal dan peningkatan mutu soal yang akan diujikan pada tahun-tahun selanjutnya. Soal-soal tes kendali mutu dianalisis untuk diketahui soal yang baik dan soal yang tidak baik. Hal tersebut juga terjadi di Kota Pontianak. Artinya, belum ada seorangpun yang melakukan analisis terhadap soal-soal ulangan tengah semester atau ulangan umum yang dibuat oleh guru. Oleh sebab itu, penulis merasa perlu menganalisis butir-butir soal yang dijadikan instrumen oleh guru-guru, sehingga hasil penelitian bisa memberikan masukan terhadap instrumen yang digunakan tersebut. Dengan harapan, semoga dari hasil dari analisis ini bisa memberikan dorongan/masukan kepada guru/pendidik agar dapat membuat tes yang baik. Untuk mempermudah dalam mengnalisis soal maka dalam penelitian ini peneliti menggunakan *software Iteman 3.5*. Iteman adalah salah satu aplikasi yang bisa dimanfaatkan untuk menganalisi soal pilihan ganda.

Adapaun tujuan penelitian ini adalah untuk mendiskripsikan: (1) baik atau tidaknya instrument ditinjau dari reabilitas soal, (2) baik atau tidaknya butir-butir soal ditinjau dari daya pembeda dan indeks kesukaran, (3) befungsi dan tidaknya pengecoh dalam soal tersebut, (4) kesalahan-kesalahan dalam pembuatan stem dan kemungkinan jawaban (options) yang dibuat oleh guru.

METODE

Metode penelitian ini menggunakan metode kualitatif diskriptif. Penelitian ini dilaksanakan di SMP Negeri 11 Pontianak Kota. Adapun tes yang akan dianalisis adalah tes soal MID yang dibuat oleh guru dan sudah diujikan di kelas IX di SMP Negeri 11 Pontianak Kota. Pemilihan kelas dan sekolah ini berdasarkan pada pertimbangan berikut: belum ada penelitian mengenai analisis soal pilihan ganda di SMP Negeri 11 Pontianak, dan peneliti mudah bekerja sama dengan guru di sekolah tersebut. Teknik pengambilan sampel dilakukan dengan *purposive sampling*.

Alat pengumpul data yang digunakan berupa tes. Tes yang akan dianalisis dalam penelitian ini sebanyak 20 butir soal. Teknik analisis data yang digunakan adalah reduksi data (*data reduction*), penyajian data (*data display*), dan penarikan kesimpulan (*conclusion drawing*).

HASIL DAN PEMBAHASAN

Dalam penelitian ini, peneliti akan menganalisis soal MID yang dibuat oleh guru. Berdasarkan hasil analisis *iteman* 3.5 diperoleh bahwa koefisien reabilitas instrumen tersebut adalah 0,546. Ini menunjukkan bahwa soal tersebut memiliki tingkat reabilitas yang kurang baik. Artinya soal tersebut tidak akan memberikan hasil yang sama jika diberikan kepada orang yang sama pada waktu yang berlainan karena berdasarkan Budiono (2015: 49) dan Sudijono (2011: 209) koefisien reabilitas yang baik itu jika lebih dari atau sama dengan 0,7. Tidak baiknya koefisien reabilitas tersebut tentu juga dipengaruhi oleh daya beda soal tersebut. Oleh sebab itu, akan dilanjutkan analisis setiap butir soal ditinjau dari daya beda butir soal. Butir-butir saol yang akan dibahas dibawah ini adalah butir-butir soal yang memiliki daya beda jelek.

Soal Nomor 1

Item Statistics

Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.
1	0-1	.87	.04	.02

CHECK THE KEY
D was specified, B works better

Alternative Statistics

Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.	Key
A	.00	.00	.00		
B	.10	.00	.08	.08	?
C	.03	.13	.00	-.18	
D	.87	.88	.92	.02	*
Other	.00	.00	.00		

ISSN 2088-3021 (media cetak)
ISSN 2598-8077 (media online)

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 0,87 yang berarti butir soal tersebut tergolong mudah bagi siswa.
2. Indeks daya pembeda D = 0,2 (dengan point biserial titik). Daya pembeda soal ini tidak baik karena kurnag dari 0,3.
3. Untuk pengecoh A, B dan C tidak berfungsi dengan baik karena pemilih pengecoh A dan C kurang dari 5% dan daya beda pengecoh B positif artinya pengecoh B dipilih lebih banyak kelompok atas daripada kelompok bawah.
4. Perhatikan bahwa Iteman memberikan masukan agar pengembang tes meninjau kembali kunci jawaban. Kunci jawaban yang disebut oleh pengembang adalah D, namun Iteman menyarankan apakah kunci jawabannya bukan B, karena pengecoh B memiliki daya beda yang positif artinya kelompok atas lebih banyak memilih pengecoh B dari pada kunci jawaban D.

Soal Nomor 2

Item Statistics

Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.
2	0-2	.97	.00	.08

Alternative Statistics

Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.	Key
A	.97	1.00	1.00	.08	*
B	.03	.00	.00	-.08	
C	.00	.00	.00		
D	.00	.00	.00		
Other	.00	.00	.00		

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 0,97 yang berarti butir soal tersebut tergolong sangat mudah bagi siswa.
2. Indeks daya pembeda D = 0,08 (dengan point biserial titik). Daya pembeda soal ini tidak baik.
3. Untuk pengecoh B, C, dan D tidak baik karena banyaknya peserta tes yang memilihnya tidak lebih dari 5% bahkan tidak ada peserta tes yang memilih pengecoh B.

Soal Nomor 3

Item Statistics

Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.
3	0-3	.97	.00	.08

Alternative Statistics

Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.	Key
A	.03	.00	.00	-.08	
B	.97	1.00	1.00	.08	*
C	.00	.00	.00		
D	.00	.00	.00		
Other	.00	.00	.00		

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 0,97 yang berarti butir soal tersebut tergolong sangat mudah bagi siswa.
2. Indeks daya pembeda D = 0,08 (dengan point biserial titik). Daya pembeda soal ini tidak baik.
3. Untuk pengecoh B, C, dan D tidak baik karena banyaknya peserta tes yang memilihnya tidak lebih dari 5%. Soal no 3 sama hasilnya dengan soal no 2.

Soal Nomor 4

Item Statistics

Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.
4	0-4	1.00	.00	

Alternative Statistics

Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.	Key
A	1.00	1.00	1.00		*
B	.00	.00	.00		
C	.00	.00	.00		
D	.00	.00	.00		
Other	.00	.00	.00		

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 1,00 yang berarti butir soal tersebut tergolong sangat mudah sekali karena seluruh siswa daat menjawab soal dengan benar.
2. Indeks daya pembeda D = 0,00 (dengan point biserial titik). Daya pembeda soal ini tergolong jelek.
3. Untuk pengecoh B, C, dan D tidak baik karena banyaknya peserta tes yang memilihnya tidak lebih dari 5% bahkan tidak ada peserta tes yang memilih pengecoh B.

Soal Nomor 6

Item Statistics

Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.
6	0-6	.32	.29	.27

Alternative Statistics

Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.	Key
A	.10	.00	.25	.25	
B	.32	.13	.42	.27	*
C	.52	.63	.33	-.24	
D	.06	.25	.00	-.33	
Other	.00	.00	.00		

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 0,32 yang berarti butir soal tersebut tergolong sedang.
2. Indeks daya pembeda D = 0,27 (dengan point biserial titik). Daya pembeda soal ini tergolong cukup.
3. Untuk pengecoh A tidak berfungsi dengan baik karena memiliki daya beda positif artinya lebih banyak dipilih oleh kelompok atas dari pada kelompok bawah.
4. Untuk pengecoh C dan D berfungsi dengan baik karena peserta tes yang memilihnya lebih dari 5% dan memiliki daya beda negatif artinya lebih banyak dipilih oleh kelmpok bawah dari pada kelompok atas.

Soal Nomor 10

Item Statistics

Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.
10	0-10	1.00	.00	

Alternative Statistics

Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.	Key
A	.00	.00	.00		
B	1.00	1.00	1.00		*
C	.00	.00	.00		
D	.00	.00	.00		
Other	.00	.00	.00		

ISSN 2088-3021 (media cetak)
ISSN 2598-8077 (media online)

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 1,00 yang berarti butir soal tersebut tergolong sangat mudah sekali karena seluruh siswa daat menjawab soal dengan benar.
2. Indeks daya pembeda D = 0,00 (dengan point biserial titik). Daya pembeda soal ini tergolong jelek.
3. Untuk pengecoh A, C, dan D tidak baik karena banyaknya peserta tes yang memilihnya tidak lebih dari 5% bahkan tidak ada peserta tes yang memilih pengecoh B.

   Hasil analisis soal no 10 sama dengan soal no 4.

Soal Nomor 13

Item Statistics

Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.
13	0-13	.00	.00	

CHECK THE KEY
B was specified, C works better

Alternative Statistics

Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.	Key
A	.77	.63	.75	.17	
B	.00	.00	.00		*
C	.10	.00	.25	.25	?
D	.13	.38	.00	-.43	
Other	.00	.00	.00		

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 0,00 yang berarti butir soal tersebut tergolong sangat sukar.
2. Indeks daya pembeda D = 0,00 (dengan point biserial titik). Daya pembeda soal ini tergolong jelek.
3. Pengecoh D berfungsi dengan baik karena dipilih oleh peserta tes lebih dari 5% dan memiliki daya beda yang negatif. Pengecoh C dan A tidak berfungsi dengan baik karena memiliki daya beda negatif.
4. Perhatikan bahwa Iteman memberikan masukan agar pengembang tes meninjau kembali kunci jawaban. Kunci jawaban yang disebut oleh pengembang adalah B, namun Iteman menyarankan apakah kunci jawabannya bukan C, karena pengecoh A memiliki daya beda positif yaitu 0,25 dan lebih besar dari pada yang lain.

Soal Nomor 14

Item Statistics

Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.
14	0-14	.03	.00	.01

CHECK THE KEY
B was specified, C works better

Alternative Statistics

Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.	Key
A	.13	.25	.00	-.28	
B	.03	.00	.00	.01	*
C	.42	.25	.83	.55	?
D	.42	.50	.17	-.36	
Other	.00	.00	.00		

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 0,03 yang berarti butir soal tersebut tergolong sukar.
2. Indeks daya pembeda D = 0,01 (dengan point biserial titik). Daya pembeda soal ini tergolong jelek.
3. Pengecoh A dan D berfungsi dengan baik karena dipilih oleh peserta tes lebih dari 5% dan memiliki daya beda negatif.
4. Pengecoh C tidak berfungsi dengan baik karena dipilih lebih banyak kelompok atas dari pada kelompok bawah.
5. Perhatikan bahwa Iteman memberikan masukan agar pengembang tes meninjau kembali kunci jawaban.

Kunci jawaban yang disebut oleh pengembang adalah B, namun Iteman menyarankan apakah kunci jawabannya bukan C, karena pengecoh A memiliki daya beda positif yaitu 0,55 dan lebih besar dari pada yang lain.

Soal Nomor 16

Item Statistics

Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.
16	0-16	.00	.00	

CHECK THE KEY
A was specified, C works better

Alternative Statistics

Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.	Key
A	.00	.00	.00		*
B	.00	.00	.00		
C	.74	.63	.75	.23	?
D	.26	.38	.25	-.23	
Other	.00	.00	.00		

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 0,00 yang berarti butir soal tersebut tergolong sangat sukar.
2. Indeks daya pembeda D = 0,00 (dengan point biserial titik). Daya pembeda soal ini tergolong jelek.
3. Pengecoh A dab B tidak berfungsi dengan baik karena tidak ada peserta tes yang memilih pengecoh tersebut. Pengecoh C juga tidak baik karena memiliki daya beda yang positif.
4. Perhatikan bahwa Iteman memberikan masukan agar pengembang tes meninjau kembali kunci jawaban. Kunci jawaban yang disebut oleh pengembang adalah A, namun Iteman menyarankan apakah kunci jawabannya bukan C, karena pengecoh C memiliki daya beda positif yaitu 0,23 dan lebih besar dari pada yang lain bahkan tidak ada peserta tes yang memilih kunci jawaban A.

Soal Nomor 17

Item Statistics

Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.
17	0-17	.03	.08	.21

Alternative Statistics

Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.	Key
A	.00	.00	.00		
B	.97	1.00	.92	-.21	
C	.03	.00	.08	.21	*
D	.00	.00	.00		
Other	.00	.00	.00		

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 0,03 yang berarti butir soal tersebut tergolong sedang.
2. Indeks daya pembeda D = 0,21 (dengan point biserial titik). Daya pembeda soal ini tergolong jelek.
3. Untuk pengecoh A dan D tidak berfungsi dengan baik karena pemilih pengecoh D kurang dari 5%.

Soal Nomor 19

Item Statistics

Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.
19	0-19	.00	.00	

CHECK THE KEY
B was specified, D works better

Alternative Statistics

Alt.	Prop. Total	Endorsing Low	Endorsing High	Point Biser.	Key
A	.87	1.00	.75	-.23	
B	.00	.00	.00		*
C	.00	.00	.00		
D	.13	.00	.25	.23	?
Other	.00	.00	.00	.00	

ISSN 2088-3021 (media cetak)
ISSN 2598-8077 (media online)

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 0,00 yang berarti butir soal tersebut tergolong sangat sukar.
2. Indeks daya pembeda D = 0,00 (dengan point biserial titik). Daya pembeda soal ini tergolong jelek.
3. Untuk pengecoh C dan D tidak berfungsi dengan baik karena tidak ada peserta tes yang memilih pengecoh C. Pengecoh D juga tidak berfungsi dengan baik karena memiliki daya beda positif.
4. Perhatikan bahwa Iteman memberikan masukan agar pengembang tes meninjau kembali kunci jawaban. Kunci jawaban yang disebut oleh pengembang adalah B, namun Iteman menyarankan apakah kunci jawabannya bukan D, karena pengecoh D memiliki daya beda positif yaitu 0,23 dan lebih besar dari pada yang lain bahkan tidak ada peserta tes yang memilih kunci jawaban B.

Soal Nomor 20

Item Statistics					Alternative Statistics				
Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	Point High	Biser.	Key
20	0-20	.94	.13	.19	A	.03	.13	.00	-.18	
					B	.00	.00	.00		
					C	.94	.88	1.00	.19	*
					D	.03	.00	.00	-.08	
					Other	.00	.00	.00		

Berdasarkan keluaran tersebut diperoleh, diproleh hal–hal sebagai berikut:

1. Indeks tingkat kesulitan P = 0,94 yang berarti butir soal tersebut tergolong sangat mudah.
2. Indeks daya pembeda D = 0,19 (dengan point biserial titik). Daya pembeda soal ini tergolong jelek.
3. Untuk pengecoh A, B dan D tidak berfungsi dengan baik karena pemilih pengecoh D kurang dari 5%.

Item Statistics					Alternative Statistics				
Seq. No.	Scale -Item	Prop. Correct	Disc. Index	Point Biser.	Alt.	Prop. Total	Endorsing Low	Point High	Biser.	Key
18	0-18	.58	.75	.60	A	.58	.25	1.00	.60	*
					B	.00	.00	.00		
					C	.42	.75	.00	-.60	
					D	.00	.00	.00		
					Other	.00	.00	.00	00	

Pengecoh B dan D tidak berfungsi dengan baik karena tidak ada peserta tes yang memilih pengecoh tesebut walaupun analisis dari butir soal menunjukkan daya beda yang baik yaitu 0,6. Hasil ini ini menunjukkan bahwa

Hasil analisis menunjukkan bahwa terdapat 12 soal dari 20 soal yang memiliki daya beda kurang dari 0,3, artinya memiliki daya beda jelek sehingga dari 12 soal tersebut perlu adanya revisi dan perbaikan. Selain itu, pengecoh yang dibuat oleh guru lebih banyak yang tidak berfungsi dengan baik. Soal yang baik ditinjau dari daya pembeda juga memiliki pengecoh yang tidak baik. Seperti soal no 18 di bawah ini.

soal yang dibuat oleh guru belum cukup baik walapun memiliki daya pembeda butir soal yang baik yaitu 0,6. Berikut akan disajikan secara tererinci pada Tabel 1 hasil analisis butir soal pilihan ganda.

Tabel 1. Hasil Analisis Butir Soal Pilihan Ganda

No Soal	Indeks Kesulitan	Daya pembeda	Pengecoh	Kesimpulan
1	0,87	0,2	Semua Tidak Berfungsi	Revisi
2	0,97	0,08	Semua Tidak Berfungsi	Revisi
3	0,97	0,08	Semua Tidak Berfungsi	Revisi
4	1,00	0	Semua Tidak Berfungsi	Revisi
5	0,90	0,50	Dua Tidak Berfungsi	Revisi
6	0,32	0,27	Satu Tidak Berfungsi	Revisi

No Soal	Indeks Kesulitan	Daya pembeda	Pengecoh	Kesimpulan
7	0,97	0,38	Semua Tidak Berfungsi	Revisi
8	0,90	0,50	Dua Tidak Berfungsi	Revisi
9	0,90	0,38	Dua Tidak Berfungsi	Revisi
10	1,00	0	Semua Tidak Berfungsi	Revisi
11	0,29	0,80	Satu Tidak Berfungsi	Revisi
12	0,74	0,31	Dua Tidak Berfungsi	Revisi
13	0,00	0,00	Dua Tidak Berfungsi	Revisi
14	0,03	0,01	Satu Tidak Berfungsi	Revisi
15	0,45	0,72	Satu Tidak Berfungsi	Revisi
16	0,00	0,00	Semua Tidak Berfungsi	Revisi
17	0,03	0,21	Dua Tidak Berfungsi	Revisi
18	0,58	0,60	Dua Tidak Berfungsi	Revisi
19	0,00	0,00	Dua Tidak Berfungsi	Revisi
20	0,94	0,19	Semua Tidak Berfungsi	Revisi

Berdasarkan Tabel 1 diperoleh bahwa tidak ada soal yang memiliki kriterium sesuai dengan yang diharapkan, artinya semua soal harus direvisi. Ada beberapa soal yang sebenarnya tidak perlu revisi soal, tetapi hanya perlu revisi pengecoh jawaban seperti butir soal nomor 5, 7, 8, 9, 11, 12, 15, dan 18 karena daya pembeda butir soal tersebut tergolong tegolong baik.

PENUTUP

Simpulan

Berdasarkan kajian teori dan pembahasan maka dalam penelitian ini dapat disimpulkan secara umum bahwa soal yang dibuat oleh guru tergolong tidak baik. Adapun kesimulan secara terperinci adalah sebagai berikut: (1) realibilitas instrumen yang dibuat oleh guru tergolong kurang baik. (2) jika dilihat dari daya pembeda butir soal, butir soal yang memiliki kriteria baik hanya 8 soal dari 20 butir soal. (3) pengecoh tidak berfungsi dengan baik dari 20 butir soal.

Saran

Hasil penelitian ini menunjukkan bahwa soal yang dibuat guru untuk dijadikan alat ukur tes hasil belajar tergolong tidak baik, sehingga tentu disarankan kepada pihak pendidik guru maupun dosen agar menelaah soal atau tes yang akan digunakan untuk divalidasi dan diujicobakan agar tes yang digunakan tapat sesuai dengan apa yang hendak diukur. Selain itu, perlu adanya bimbingan, pelatihan, maupun *workshop* terkait pembuatan soal/tes yang baik dan benar, serta pemanfaatan aplikasi-aplikasi atau *software* terkait analisis butir soal.

DAFTAR PUSTAKA

Amalia, A. N. & Widayati, A. 2012. Analisis Butir Soal Tes Kendali Mutu Kelas XII SMA Mata Pelajaran Ekonomi Akuntansi di Kota Yogyakarta Tahun 2012. Jurnal Pendidikan Akuntansi Indonesia. Vol. X, No. 1. Hal. 1-26.

Arifin, Z. 2012. Evaluasi Pembelajaran. Jakarta: Direktorat Jenderal Pendidikan Islam, Kementerian Agama RI.

Budiyono. 2015. Pengantar Penilaian Prestasi belajar. Surakarta: UPT UNS Press.

Guba, E. G. and Lincoln, Y. S. 1985. Effective Evaluation. San Francisco: Jossey – Bass Pub.

Nitko, A. J. 1996. Educational Assessment of Students, Second Edition. New Jersey: Englewood Cliffs.

Sudijono, A. 2011. Pengantar Evaluasi Pendidikan. Jakarta: PT Rajagrafindo Persada.