

KUALITAS SOAL FREDIKSI UJIAN NASIONAL HASIL DIKLAT UJIAN NASIONAL GURU BAHASA INGGRIS MA TAHUN 2019

THE QUALITY OF NATIONAL EXAM TEST FREDICTION WRITTEN BY MADRASAH ALIYAH ENGLISH TEACHER 2019

Sholikin

Widyaiswara Balai Diklat Keagamaan Surabaya
sholikin@yahoo.com

ABSTRACT

Evaluation or Exam is an activity that is inseparable from the activity of teaching and learning carried out by the teacher. Evaluation in the learning process has a very important role because from the evaluation, the teacher will get information about the level of competency of the students, the level of effectiveness of a lesson. For this reason teachers need to be given competencies to develop an adequate evaluation tool so that they can provide the right picture. The focus of this research is 1] How is the effectiveness of tests distraction developed by Diklat participants ? 2] How is the level of difficulty index of the test ? 3] How is the quality of the discrimination index of the test ? This research method uses a Descriptive approach, with the subjects of the study were 35 English Teachers training participants. While the data collection techniques used the document, the results of the Prediction Tests developed by the training participants and it was analyzed by Anates Version 04. The results of the analysis showed that the effectiveness of distractions were still not good, beside that there were still too difficult and too easy tests. While the Discrimination Index were more than 50 % of the questions with bad different power.

Key Words : Distraction, Discrimination index, Difficulty index

ABSTRAK

Evaluasi atau Ujian adalah kegiatan yang tidak terpisahkan dari kegiatan belajar mengajar yang dilakukan oleh guru. Evaluasi dalam proses pembelajaran mempunyai peran yang sangat penting karena dengan evaluasi guru akan mendapatkan informasi tentang, tingkat kompetensi peserta didik, tingkat efektifitas sebuah pembelajaran. Untuk itu guru perlu diberikan kompetensi bagaimana menyusun sebuah alat evaluasi yang memadai agar bisa memberikan gambaran yang tepat. Focus penelitian ini adalah : 1] Bagaimana efektifitas pengecoh dari tes yang dikembangkan oleh peserta Diklat ? 2] Bagaimana prosentase penyebaran soal ditinjau dari tingkat kesukarannya ? 3] Bagaimana kualitas soal ditinjau dari sudut daya bedanya ? Metode penelitian ini menggunakan pendekatan Deskriptif Kualitatif, dengan subyek penelitian sejumlah 35 orang guru Bahasa Inggris peserta Diklat. Sedangkan teknik pengumpulan data menggunakan dokumentasi nilai hasil uji coba tes prediksi yang dikembangkan peserta Diklat dan dianalisis dengan Anates Versi 04. Hasil analisis menunjukkan bahwa efektifitas pengecoh masih belum baik, tingkat kesukaran soal juga masih banyak yang terlalu sulit dan terlalu mudah, sedangkan daya beda soal lebih dari 50 % soal dengan daya beda jelek.

Kata kunci : pengecoh, daya beda, tingkat kesukaran

PENDAHULUAN

A. Latar Belakang

Kegiatan evaluasi pembelajaran adalah kegiatan yang sangat penting untuk dilaksanakan karena dengan kegiatan ini apa yang telah

direncanakan untuk dicapai oleh peserta didik akan dapat diketahui oleh seorang guru. Evaluasi dapat memetakan siswa mana yang telah mencapai kompetensi dan siswa mana yang belum. Apakah secara klasikal kompetensi yang telah

ditetapkan telah tercapai atau belum. Dengan demikian siswa yang belum mencapai kompetensi akan mendapat remedial teaching, sehingga seluruh siswa tuntas belajarnya.

Evaluasi pembelajaran merupakan kegiatan yang tidak terpisahkan dengan kegiatan pembelajaran yang dilakukan oleh guru. Karena salah satu tugas guru yang tertuang dalam PP No. 74 Tahun 2008. adalah melaksanakan kegiatan pembelajaran (merencanakan, melaksanakan dan mengevaluasi) sebagai tugas utama disamping membimbing dan melatih sebagai tugas tambahan. Esessment adalah serangkaian kegiatan yang dilakukan untuk mengambil keputusan melalui pemberian tes atau non tes.

Tes adalah suatu pernyataan, tugas atau seperangkat tugas yang direncanakan untuk memperoleh informasi tentang trait atau atribut pendidikan dan psikologi. Setiap butir pertanyaan atau tugas tersebut mempunyai jawaban atau ketentuan yang dianggap benar. Tes dapat diklasifikasikan menurut bentuk, tipe dan ragamnya (Asmawi Zainul, dkk :1997). Sedangkan menurut Lee, (1984) Tes adalah suatu prosedur yang sistematis untuk mengamati dan mendeskripsikan karakteristik seseorang dengan menggunakan skala numerik atau sistem kategori. Maka tes adalah seperangkat tugas yang harus dikerjakan atau direspon oleh *testee*, yang dapat menggambarkan kompetensi atau ketrampilan seseorang atau peserta didik.

Untuk mendapatkan manfaat dari pelaksanaan evaluasi pembelajaran tersebut, diperlukan sebuah tes atau alat evaluasi yang baik. Untuk itu guru sebagai penyusun tes harus mempunyai kompetensi yang baik dalam mengembangkan alat evaluasinya. Untuk itu mereka dikirim ke Balai Pendidikan dan Pelatihan Surabaya selama 6 hari mengikuti kegiatan kediklatan. Mereka mendapatkan pelatihan-pelatihan tertentu guna meningkatkan kompetensi mereka dalam mengembangkan sebuah tes atau alat ukur yang baik.

Penelitian ini akan memberikan deskripsi dan analisis tes yang dihasilkan oleh peserta Diklat penyusunan Soal Ujian Nasional. Penelitian

ini sangat bermanfaat karena akan memberikan informasi tentang kualitas pengecoh tes, daya beda dan tingkat kesukaran tes yang dihasilkan oleh peserta diklat. Dari hasil analisis akan memberikan rekomendasi tentang materi mana yang perlu mendapat perhatian lebih dari kegiatan Diklat.

B. Rumusan Masalah

Untuk mempermudah dalam penelitian ini, penulis merumuskan penelitian ini sebagai berikut :

1. Bagaimana efektifitas pengecoh dari tes yang dikembangkan oleh peserta Diklat ?
2. Bagaimana prosentase penyebaran soal ditinjau dari tingkat kesukarannya ?
3. Bagaimana kualitas soal ditinjau dari sudut daya bedanya ?

C. Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah

1. mendeskripsikan dan menganalisis efektifitas pengecoh dari tes yang dikembangkan oleh peserta Diklat
2. mendeskripsikan dan menganalisis prosentase penyebaran soal ditinjau dari tingkat kesukarannya
3. mendeskripsikan dan menganalisis kualitas soal ditinjau dari sudut daya bedanya

D. Manfaat Penelitian

1. Manfaat Teoritis
 - a. Sebagai bahan referensi yang dapat digunakan untuk memperoleh gambaran tentang kemampuan guru dalam mengembangkan tes sebagai alat ukur kompetensi siswa.
 - b. Hasil penelitian digunakan sebagai bahan acuan, informasi, dan perbaikan bagi pelaksanaan Diklat sejenis.
2. Manfaat Praktis
 - a. Bagi guru, hasil penelitian ini diharapkan dapat memberikan informasi tentang kompetensi mereka dalam mengembangkan sebuah tes penilaian.
 - b. Bagi widyaiswara, hasil penelitian ini dapat

digunakan sebagai acuan dalam memberikan tekanan materi yang perlu diperhatikan guna meningkatkan kemampuan peserta diklat.

KAJIAN TEORI

A. Pengecoh Pilihan Jawaban

1. Pengecoh Jawaban

Pengecoh jawaban atau *distractor* dalam sebuah tes berbentuk pilihan ganda mempunyai peran yang sangat penting. Pengecoh jawaban akan menentukan kualitas sebuah tes dalam bentuk pilihan ganda. Pengecoh adalah pilihan jawaban atau *option* yang disediakan oleh penulis soal di luar kunci jawaban yang benar. Menurut Sujiono (2005) pengecoh adalah kemungkinan jawaban diluar kunci yang disediakan oleh penulis soal untuk dipilih oleh *test taker*. Hal senada juga disampaikan oleh Chabib (1991) dalam sebuah tes obyektif selalu disediakan dua unsur jawaban, yaitu jawaban yang benar dan jawaban yang salah. Jawaban yang salah ini disebut dengan pengecoh atau *distractor*. Maka dapat diambil kesimpulan bahwa pengecoh atau *distractor* adalah pilihan jawaban yang disediakan sebagai pilihan alternatif disamping kunci jawaban yang benar dalam sebuah tes pilihan ganda atau *multiple choice*.

2. Efektifitas Pengecoh

Dalam sebuah tes pilihan ganda selalu ada dua pilihan jawaban, yaitu pilihan jawaban benar dan pengecohnya. Pengecoh diberikan dengan tujuan agar peserta tes ada yang tertarik untuk memilih pengecoh itu. Untuk membuat mereka tertarik maka, kualitas pengecoh harus baik agar mereka pengambil mempertimbangkannya sebagai kemungkinan jawaban benar.

Untuk itu, penulis tes harus mempertimbangkan beberapa hal dalam menyiapkan pengecoh antara lain :

1. pengecoh harus homogen. Artinya pengecoh harus setara dengan kunci jawaban yang disediakan, misalnya : untuk bahasa Inggris, kunci jawaban adalah sebuah frase kata benda, maka

pengecoh harus dalam bentuk frase kata benda juga.

2. memperhatikan panjang pendek jawaban. Artinya pandek option atau jawaban harus relatif sama panjang pendeknya.
3. jawaban sama-sama terkait dengan stimulus soal. Jika soal menggunakan stimulus.

Pengecoh dikatakan baik atau berfungsi atau berkualitas jika dipilih oleh sejumlah peserta tes atau *test takers*. Mengutip pendapat Purwanto (2000) bahwa pengecoh dikategorikan baik jika perbandingan kelompok memilih jawaban salah lebih sedikit dibandingkan dengan kelompok bawah yang memilih jawaban salah. Dengan kata lain kelompok bawah memilih jawaban salah lebih banyak dari pada kelompok atas yang mengambil tes. Menurut Sujiono (2005) tujuan penulisan pengecoh adalah dalam rangka memberi pilihan alternatif kepada *test takers* untuk dipilih. Sedangkan menurut Chabib (1991) tujuan pengecoh adalah untuk mengocok peserta test yang kurang atau tidak kompeten pada materi yang diteskan. Sehingga tes benar-benar bisa membedakan siswa yang kompeten dan tidak kompeten.

Pengecoh dikatakan berkualitas dan berfungsi dengan baik apabila dipilih oleh sejumlah *test takers*. Artinya jika *test takers* tidak ada yang memilih berarti pengecoh tidak berfungsi atau tidak efektif dan perlu diperbaiki kualitasnya. Sudijono (2005), Kategori pengecoh dikatakan berfungsi dengan baik jika dipilih oleh sejumlah *test takers*. Penentuan kualitas pengecoh, para ahli penilaian menjelaskan bahwa pengecoh dikategorikan berkualitas jika dipilih oleh 5 % dari jumlah *test takers*. Sedangkan Purwanto tidak memberikan presentase, akan tetapi beliau menjelaskan bahwa pengecoh efektif berfungsi dengan baik jika kelompok bawah banyak yang memilih distractor yang disediakan dari pada kelompok atas. Sedangkan untuk membagi kelompok atas dan bawah Sujiono menjelaskan dengan cara mengurutkan nilai peserta tes dari skor tertinggi ke terendah. Lalu diambil 25 % kelompok atas dihitung dari

skor tertinggi sedangkan kelompok bawah dihitung dari nilai terendah. Jika perta tes jumlahnya banyak. Apabila peserta tes jumlahnya sedikit kelompok atas dan bawah ditentukan dengan membagi 50 % kelompok atas dan 50 % kelompok bawah. Jadi Sebuah pengecoh dikatakan efektif jika mampu mengecoh sejumlah *test takers* dengan mempertimbangkan *test takers* kelompok bawah lebih banyak yang terkecoh dari pada *test takers* kelompok atas dengan tetap memerhatikan pembagian kelompok di atas.

Rumus menghitung efektifitas pengecoh

2. Tingkat Kesukaran Tes

$$D = \left(\frac{\sum \text{Peserta Tes} : 4n}{S \text{ Peserta Tes} : n^2} \right)$$

(Haladyna, Thomas M. : 1994)

n = banyaknya option jawaban

maka untuk hasil analisis pengecoh dikatakan berfungsi dengan hitungan sebagai berikut :

Jika peserta tes : 34

Option : 5

$$D = (34 : 4.5)$$

$$D = 34 : 20$$

$$D = 1.7 \text{ dibulatkan menjadi } 2$$

Jadi pengecoh untuk tes di atas dikatan berfungsi jika pengecoh dipilih minimal 2 peserta tes.

B. Tingkat Kesukaran Tes

1. Definisi

Tes hasil belajar yang baik adalah soal atau tes yang mampu menggambarkan kompetensi dari peserta tes tersebut. Untuk tes yang baik adalah tes yang tidak terlalu mudah atau terlalu sulit. Artinya derajad soal tersebut adalah sedang. (Sudjono, 2009). Sedangkan menurut Aiken (1994) Tingkat kesukaran soal adalah peluang untuk menjawab benar suatu soal padat tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Sedangkan menurut Zainul, dkk (1997) tingkat kesukaran

butir soal adalah proporsi peserta tes menjawab benar terhadap butir soal tersebut. *Item difficulty is simply the percentage of students taking the test who answered the item correctly. The larger the percentage getting an item right, the easier the item. The higher the difficulty index, the easier the item is understood to be* (Wood, 1960). Artinya tingkat kesukaran adalah prosentase dari siswa yang mengikuti tes yang menjawab benar. Semakin besar yang menjawab benar maka semakin mudah ters tersebut. Semakin tinggi skor pada indeks kesulitan, maka semakin mudah tes tesebut.

Bistok Sirait (2009) berpendapat bahwa tinggi rendahnya tingkat kesukaran suatu butir soal dapat disebabkan oleh kerumitan (kompleksitas) pokok soal dan kondisi pilihan jawaban yang disediakan. Tingkat kesukaran butir soal biasanya dilambangkan dengan p. Makin besar nilai p yang berarti makin besar proporsi yang menjawab benar terhadap butir soal tersebut, makin rendah tingkat kesukaran butir soal itu. Hal ini mengandung arti bahwa soal itu makin mudah, demikian pula sebaliknya. Dengan demikian dapat dikatakan bahwa tingkat kesukaran atau *Difficulty Index*, sebuah soal adalah kesempatan untuk menjawab benar soal tersebut, semakin tinggi peluang peserta tes untuk menjawab benar makan semakin mudah tes tersebut, begitu juga sebaliknya.

2. Teknik Penghitungan Tingkat Kesukaran Tes

Berangkata dari definisi di atas, bagaimanakan cara menganalisis butir soal yang mempunyai tingkat kesukaran yang baik ? Witherington dalam Sudijono (2005) mengatakan bahwa tes dikatakan mempunyai tingkat kesukaran tinggi atau rendah adalah dengan merujuk pada nilai yang ada pada *Difficulty Index*, yang biasa dilambangkan dengan " P ". singkatan dari *Proportion*.

Tingkat kesukaran biasanya dinyatakan dalam bentuk index dengan rentang nilai 0,00 - 100. Semakin tinggi index tingkat kesukaran berarti semakin mudah tersebut begitu juga

sebaliknya. Adapun cara menghitung index kesukaran menurut Gronlund and Linn (1990) adalah menggunakan rumus sebagai berikut :

$$\text{Difficulty Index} = R : T \times 100 \%$$

R : Jumlah siswa yang menjawab benar
T : Jumlah siswa yang ikut tes

Sedangkan menurut Purwanto (2000), langkah menghitung tingkat kesukaran adalah sebagai berikut :

1. Urutlah lembar jawaban dari nilai tertinggi ke nilai terendah
2. Pisahkan kelompok atas dengan mengambil 30 % peserta tes. Dihitung dari nilai tertinggi dan 30 % kelompok bawah dihitung dari nilai terendah.
3. Hitunglah kelompok atas yang menjawab benar dan hitunglah kelompok bawah yang menjawab benar
4. Gunakan rumus berikut untuk menghitung tingkat kesukaran

$$TK = \frac{J_a + J_b}{K_a + K_b} \times 100 \%$$

Keterangan :

- Ja : Kelompok atas yang menjawab benar
Jb : Kelompok bawah yang menjawab benar
Ka : Kelompok atas
Kb : Kelompok bawah

C. Daya Beda Tes

1. Definisi

Salah satu tujuan dari pelaksanaan tes adalah untuk memberikan informasi tentang kompetensi dari peserta tes. Untuk bisa memberikan informasi maka tes tersebut harus mempunyai daya beda yang baik. Daya beda tes adalah kemampuan sebuah tes tersebut dalam membedakan siswa yang *fast learners* (unggul) dan *slow learners* (asor). Menurut Sudijono (2005) adalah kemampuan suatu item tes hasil belajar dalam membedakan antara *tes tee* yang mempunyai kemampuan tinggi dengan *testee* yang mempunyai kemampuan rendah. *The discrimination index, D, is the number of people in the upper group who answered the item correctly minus the number of people in*

the lower group who answered the item correctly, divided by the number of people in the largest of the two groups. Wood (2006). Artinya Indeks Daya Beda, D, adalah sejumlah orang yang dirada di kelompok atas yang menjawab benar dikurnagi sekelompok orang di kelompok bawah yang menjawab benar, dibagi jumlah keompok atas ditambah kelompok bawah.

2. Kualitas Daya Beda

Kualitas daya beda sebuah tes dikatakan baik jika tes tersebut mampu membedakan peserta tes yang mempunyai kemampuan baik dan yang tidak, *A good item discriminates between those who do well on the test and those who do poorly.* (Susan Matlock-Hetzel. 1997) artinya sebuah tes yang baik adalah ts yang dapat membedakan seseorang yang dapat mengerjakan tes dengan baik dan yang tidak dapat mengerjakannya dengan baik. Hal sama juga disampaikan oleh Purwnato (1997), Thoha (1991), Sudijono (2005).

Dalam menghitung indeks diskriminasi D, pertama nilai siswa diranking dari nilai tertinggi ke nilai terenda. Kedua, ambillah 27 % dari total siswa yang mengikuti tes, sebagai kelompok atas dengan cara menghiutng dari ranking teratas ke bawah, sedangkan kelompok bawah 27 % dihitung dari nilai terbawah ke atas. Wiersma and jurs (1990) menyatakan bahwa 27 % digunakan karena hal itu menunjukkan bahwa nilai ini akan memaksimalkan perbedaan pada distribusi normal. Sedangkan Nunnally (1972) menyarankan 25 % dari total peserta tes, untuk menentukan kelompok atas dan bawah.

Adapun rumus menghitung Daya Beda d apat menggunakan Rumus sebagai berikut :

$$DP = \frac{B_a - B_b}{\frac{1}{2}(J_a + J_b)}$$

DP = Daya Pembeda

J_a = banyaknya kelompok atas

J_b = banyaknya kelompok bawah

Kriteria :

- 0,00 ≤ DP < 0,20 Jelek
- 0,20 ≤ DP < 0,40 Cukup
- 0,40 ≤ DP < 0,70 Baik
- 0,70 ≤ DP ≤ 1,00 Baik Sekali

METODE PENELITIAN

A. Pendekatan

Penelitian tergolong penelitian Deskriptif Kualitatif. Data hasil penelitian ini diolah secara kualitatif menggunakan Teknik prosentase dan selanjutnya dideskripsikan dan dianalisis secara kualitatif. (Sugiyono, 2013). Dalam hal ini, peneliti mengelompokkan hasil analisis terhadap efektifitas pengecoh, tingkat kesukaran dan daya beda menggunakan standarnya masing-masing.

B. Subyek Penelitian

Subyek penelitian ini adalah guru Bahasa Inggris Madrasah Aliyah yang mengikuti Diklat Sunstantif Peningkatan Kompetensi dalam menyusun prediksi Unjian Nasional B. Inggris Madrasah Aliyah pada Balai Diklat Keagamaan Surabaya tahun 2019, sampel penelitian sebanyak 35 guru dari populasi sebesar 35 guru. Jadi dalam penelitian sampel sekaligus populasi penelitian.

C. Metode Pengumpulan data

Data dalam penelitian ini dikumpulkan dengan menggunakan Teknik Dokumentasi. Yaitu, peneliti menganalisis lembar jawaban siswa yang mengerjakan soal prediksi yang dikembangkan oleh peserta DIKLAT di Balai Diklat Keagamaan Surabaya tahun 2019.

D. Analisis Data

Data tentang efektifitas pengecoh, tingkat kesulitan dan daya beda dalam penelitian ini di-analisis menggunakan perangkat analisis Anates Versi 04.

TEMUAN DAN PEMBAHASAN

A. Temuan

Data yang terkumpul dalam penelitian ini adalah hasil dari analisis menggunakan Anates Ver. 4. Adapun hasil analisis dapat disajikan sebagai berikut :

1. Pengecoh

Dari hasil analisis data tentang efektifitas pengecoh dapat ditampilkan data sebagai berikut :

Jumlah Subyek : 34 orang

Jumlah Soal : 35 item soal

Tabel 1.1. analisis Efektifitas Pengecoh

No Butir	a	b	c	d	e	*	Di revisi	Efektif
1	-1	7--	0--	25**	-1	0	3	1
2	31**	0--	0--	0--	3--	0	3	1
3	2+	0--	1+	3--	28**	0	2	2
4	16**	-2	9--	0--	-7	0	1	3
5	4+	20--	1--	0--	9**	0	2	2
6	17**	3+	1--	1--	12--	0	2	2
7	20**	-1	7--	-1	4++	0	2	2
8	2+	2+	1+	1+	28**	0	2	2
9	1++	1++	0--	30**	2--	0	3	1
10	5+	19**	-1	3++	-6	0	1	3
11	21**	4++	4++	3++	2+	0	0	4
12	3--	1+	1+	28**	1+	0	3	1
13	1--	1--	0**	31--	1--	0	3	1
14	-2	13**	12--	4++	3+	0	0	4
15	0--	5--	-1	26**	2++	0	2	2
16	17**	1--	-7	-2	-7	0	1	3
17	4+	23**	-1	6--	0--	0	2	2
18	27**	1+	1+	-3	2++	0	2	2
19	13--	19**	0--	2+	0--	0	2	2
20	3+	12**	-2	0--	17--	0	1	3
21	0--	7**	1--	6++	20--	0	2	2
22	3+	25**	2++	2++	2++	0	0	4
23	5+	-1	-6	3++	19**	0	1	3
24	9++	7++	6++	3**	8++	0	0	4
25	0--	31**	0--	1+	2--	0	3	1
26	-2	3**	-13	11+	4+	1	0	4
27	12--	-2	6++	13**	1--	0	1	3
28	1**	2--	23--	2--	5+	1	0	4
29	9--	18**	-2	1--	3+	1	1	3
30	2++	3+	-1	26**	-1	1	2	2
31	0--	2--	0--	1+	31**	0	3	1
32	22--	11**	0--	1--	0--	0	3	1
33	3+	26**	2++	0--	2++	1	1	3
34	1+	28**	0--	1+	3--	1	3	1
35	9**	11--	5++	7++	-2	0	3	2
Total							60	80
Prosentase							42.8%	57.1%

Keterangan :

1. ** : Kunci Jawaban
2. * : omit

Data temuan diatas menunjukkan bahwa pengecoh yang harus direvisi atau diperbaiki sebesar 41 % atau sebanyak 57 buah dari total pengecoh sejumlah 140 pengecoh. Sedangkan pengecoh yang baik yang harus dipertahankan sebesar 59 % atau sebanyak 70 buah dari 140 pengecoh.

2. Tingkat Kesulitan

Jumlah Subyek : 34 orang

Jumlah Soal : 35 item soal

Tabel 1.2 Hasil Analisis Tingkat Kesulitan

No Butir	Jumlah Betul	Tkt. Kesukaran	Tafsiran				
			Sangat Mudah	Mudah	Sedang	Sukar	Sangat Sukar
1	25	73,53		v			
2	31	91,18	v				
3	28	82,35		v			
4	16	47,06			v		
5	9	26,47				v	
6	17	50			v		
7	20	58,82			v		
8	28	82,35		v			
9	30	88,24	v				
10	19	55,88			v		
11	21	61,76			v		
12	28	82,35		v			
13	0	0					v
14	13	38,24			v		
15	26	76,47		v			
16	17	50			v		
17	23	67,65			v		
18	27	79,41		v			
19	19	55,88			v		
20	12	35,29			v		
21	7	20,59				v	
22	25	73,53		v			
23	19	55,88			v		
24	3	8,82					v
25	31	91,18	v				
26	3	8,82					v
27	13	38,24			v		
28	1	2,94					v
29	18	52,94			v		
30	26	76,47		v			
31	31	91,18	v				
32	11	32,35			v		
33	26	76,47		v			
34	28	82,35		v			
35	9	26,47				v	
Total			4	10	13	4	4
Prosentase			11 %	34 %	37 %	11 %	11 %

Data diatas menunjukkan bahwa tingkat kesulitan untuk katagori sangat mudah sebesar 11 % atau sebesar 4 buah item soal dari 35 item soal. Sedangkan untuk soal dengan katagori mudah sebesar 34 % atau 10 item soal dari 35. Soal dengan katagori sedang sebesar 37 % atau sebesar 13 item soal dari 35. Untuk soal dengan katagori sukar sebesar 9 % atau sebesar 3 soal dan soal dengan katagori sangat sukar sebesar 11 % atau 4 soal dari 35.

3. Daya Beda Tes

Jumlah Subyek : 34 orang

Klp atas/bawah (n) : 9

Butir Soal : 35

Tabel 1.3 Hasil Analisis Daya Beda Tes

No	Kelp atas	Kel. Bawah	Beda	Indeks DP (%)	Jelek	Cukup	Baik	Baik Sekali
1	8	4	4	44,44			V	
2	9	8	1	11,11	v			
3	8	6	2	22,22		v		
4	7	2	5	55,56			V	
5	3	2	1	11,11	v			
6	8	2	6	66,67			V	
7	6	1	5	55,56			V	
8	9	5	4	44,44			V	
9	9	6	3	33,33		v		
10	6	4	2	22,22		v		
11	7	2	5	55,56			V	
12	9	7	2	22,22		v		
13	0	0	0	0	v			
14	3	3	0	0	v			
15	7	5	2	22,22		v		
16	8	2	6	66,67			V	
17	7	5	2	22,22		v		
18	9	5	4	44,44			V	
19	5	4	1	11,11	v			
20	5	1	4	44,44			V	
21	3	1	2	22,22	v			
22	8	5	3	33,33		v		
23	4	7	-3	-33,33		v		
24	2	0	2	22,22		v		
25	8	8	0	0	v			
26	1	0	1	11,11	v			
27	5	2	3	33,33		v		
28	0	0	0	0	v			
29	5	5	0	0	v			
30	8	7	1	11,11	v			
31	9	9	0	0	v			
32	4	1	3	33,33		v		
33	9	6	3	33,33		v		
34	8	7	1	11,11	v			
35	9	6	3	33,33		v		
Total					13	13	9	0
Prosentase					38 %	38 %	26 %	0 %

Dari data yang ditampilkan diatas bahwa daya beda soal dikategorikan pada empat katagori, yaitu : jelek, cukup, baik dan baik sekali. Untuk penjelasan setiap katagori dapat dilihat pada tabel berikut :

No	Hasil Penghitungan	Katagori Daya Beda
1	$0,00 \leq DP < 0,20$	Jelek
2	$0,20 \leq DP < 0,40$	Cukup
3	$0,40 \leq DP < 0,70$	Baik
4	$0,70 \leq DP \leq 1,00$	Baik Sekali

Hasil analisis menunjukkan bahwa tidak semua soal yang disusun oleh peserta Diklat mempunyai katagori baik, banyak item sola yang berkatagori jelek atupun cukup. Untuk lebih lengkapnya penulis paparkan sebagai berikut. Soal dengan kagori jeleks sebesar 38 % atau sebesar 13 item soal, yang katagori cukup 12 % atau sebesar 12 item soal, sedangkan item soal dengan daya beda baik, sebesar 26 % atau sebesar 9 item soal dan tidak soal dengan katagori sangat baik.

B. Pembahasan

1. Pengecoh Pilihan Jawaban

Mengamati data yang ditampilkan pada analisis hasil penelitian untuk pengecoh pada item soal yang telah *ditry out* pada siswa Madrasah Aliyah Negeri Bangil jelek. Sebesar 41 % atau 57 pengecoh atau distractor masuk pada katagori buruk. Artinya pegecoh tersebut tidak dipilih oleh *test takers* atau dipilih oleh *test takers* kurang dari 2. Artinya pengecoh ini harus direvisi atau diganti. Karena tujuan dilaksanakan tes antara lain adalah seperti apa yang disampaikan oleh Ground Lund (1991), ia memandang evaluasi sebagai proses sistematis untuk mengumpulkan, menganalisis serta menafsirkan informasi guna menentukan keberhasilan siswa dalam mencapai kompetensi. Dengan kata lain bahwa pengecoh yang jelek harus diperbaiki agar mampu menggoda *test takers* untuk memilihnya.

Ketidak mampuan pengecoh mengecoh *test takers* dikarenakan kualitasnya tidak baik. Dengan kata lain agar pengecoh dapat berfungsi

dengan baik, maka perlu analisis kualitatif dilaksanakan dengan sebaik-baiknya. Analisis kualitatif yang meliputi analisis konstruksi, isi dan bahasa. Telaah kualitatif atau analisis teoritik dilakukan sebelum butir-butir soal diuji-cobakan dan di analisis secara empirik. Aspek-aspek yang diperhatikan dalam telaah kualitatif adalah aspek materi, konstruksi, dan bahasa/ budaya ditelaah berdasarkan kaidah-kaidah yang telah ditentukan (1997/ 1998).

2. Tingkat Kesukaran Tes

Tingkat kesukaran adalah porsi kesempatan test taker untuk menjawab tes dengan benar. Menurut Zainul, dkk (1997) tingkat kesukaran butir soal adalah proporsi peserta tes menjawab benar terhadap butir soal tersebut. Saifudin Azwar (2006 : 129) mengatakan bahwa tingkat kesukaran butir soal adalah proporsi antara banyaknya peserta tes yang menjawab butir soal dengan benar dengan banyaknya peserta tes. Untuk menganalisis tingkat kesukaran dari sebuah tes harus dilakukan sebuah analisis. Analisis kuantitatif adalah analisis yang dilakukan setelah tes dilaksanakan.

Pada kesempatan ini analisis tingkat kesukaran Soal prediksi UN bahasa Inggris MA menunjukkan bahwa tingkat kesukaran belum seluruh soal baik. Tingkat kesukaran yang baik menunjukkan diagram kurva normal. Artinya jumlah soal dengan katagori mudah mempunyai porsi yang tidak banyak, begitu juga porsi soal dengan katagori sulit dan sangat sulit. Data menunjukkan (11 % sangat mudah, 34 % mudah, 37 % sedang, 9 % sulit dan 11 % sangat sulit). Kebanyakan soal berada dalam katagori sedang. Menurut Dawson (1972) butir soal yang memiliki tingkat kesukaran 0,25 – 0,75 dikatakan baik.

Wayan Nurkencana (2002 : 161) mengemukakan tentang klasifikasi tingkat kesukaran soal yaitu jika nilai p (prosentase tingkat kesukaran soal) sebagai berikut :

Jika Nilai $p = 0,81 - 1,00$ butir soal sangat mudah,
 $p = 0,61 - 0,80$ butir soal mudah,
 $p = 0,41 - 0,60$ butir soal sedang,
 $p = 0,21 - 0,40$ butir soal sukar,
 $p = 0,00 - 0,20$ butir soal sangat sukar.

Merujuk pada pendapat dari Wayan Nurkuncara bahwa soal yang dikategorikan baik adalah soal yang mempunyai nilai p . antara 0.25 – 0.75, dengan kata lain soal yang dalam katagori mudah – sukar. Artinya soal yang terlalu mudah dan terlalu sukar sebaiknya direvisi.

Data dari analisis menunjukkan bahwa jumlah soal sangat mudah sebesar 4 buah atau 11 %. Begitu juga soal sangat sukar sebesar 4 buah atau 11 % harus direvisi. Soal-soal sangat mudah tersebut terdapat pada soal no, 2, 9, 25 dan 31. Sedangkan soal-soal dengan katagori sangat sukar terdapat pada no. 13, 24, 26 dan 28.

3. Daya Beda Tes

Menurut Ebel (1972) butir yang memiliki daya pembeda 0,41 atau lebih besar dikatakan baik, sedangkan Fernandes mengatakan butir yang memiliki daya pembeda lebih besar dari 0,2 dikatakan baik. Merujuk hasil analisis untuk daya beda tes menunjukkan bahwa 13 item soal atau 38 % katagori jelek. Sedangkan soal katagori cukup sejumlah 12 item soal atau 35 %. Artinya soal yang dengan katagori jelek harus direvisi atau diperbaiki. Merujuk pendapat Ebel, bahwa soal dapat diterima apabila soal tersebut diatas 0,41. Berdasarkan itu maka, ada 24 soal yang harus diperbaiki. Dan ada 10 soal dengan katagori baik dan baik sekali, dan soal ini yang bisa diterima.

Soal yang dengan katagori jelek dan harus direvisi terdapat pada soal no, 2, 5, 13, 14, 19, 21, 25, dan 26, soal dengan katagori cukup, dianjurkan untuk direvisi terdapat pada soal no, 3, 9, 10, 12, 15, 17, 22, 23, 24, 27, 32, 33, dan 35. Sedangkan soal dengan katagori baik, sudah tidak perlu direvisi terdapat pada soal no, 1, 4, 6, 7, 8, 11, 16, 18, dan 20.

Mencermati hasil analisis dan temuan dari soal try out di atas, analisis kualitatif perlu diperkuat terutama pada pemilihan atau penyediaan jawaban, baik itu kunci jawaban maupun *destraction*. Disamping itu tingkat kesulitan dengan katagori sangat mudah dan sangat mudah, seperti soal dengan nomor 2, 9 dan 25. Begitu juga soal dengan katagori sangat sulit

cenderung daya bedanya kurang baik.

PENUTUP

A. Simpulan

Kemampuan para guru yang telah mengikuti Diklat Teknis Substantif Ujian Nasional B. Inggris Aliyah dalam menyusun Tes Prediksi Ujian Nasional 2019 dapat disimpulkan sebagai berikut :

1. Soal yang baik adalah soal yang pengecohnya berfungsi. Dalam penelitian ini, pengecoh dikategorikan berfungsi dengan baik apabila pengecoh tersebut dipilih peserta tes sejumlah 2 orang. Hasil dari analisis penelitian menunjukkan bahwa pengecoh yang harus direvisi atau diperbaiki sebesar 41 % atau sebanyak 57 buah dari total pengecoh sejumlah 140 pengecoh. Sedangkan pengecoh yang baik yang harus dipertahankan sebesar 59 % atau sebanyak 70 buah dari 140 pengecoh. Dengan kata lain masih banyak pengecoh yang tidak berfungsi dengan baik dan perlu direvisi.
2. Tingkat Kesukaran Tes yang baik adalah antara 0.25 – 0.75. Dari analisis tes prediksi Ujian Nasional di atas mengindikasikan bahwa soal dengan katagori sangat mudah sebesar 4 buah atau 11 %. Sedangkan soal sangat sukar sebesar 4 buah atau 11 % harus direvisi. Soal-soal dengan katagori sangat mudah dan sangat sukar harus direvisi. Sedangkan yang dapat diterima adalah soal dengan katagori sukar, sedang dan mudah.
3. Sedangkan untuk daya beda soal yang dapat diterima adalah soal dengan skor 0.41 atau lebih. Dari data di atas bahwa 13 item soal atau 38 % katagori jelek. Sedangkan soal katagori cukup sejumlah 12 item soal atau 35 %. Artinya soal yang dengan katagori jelek harus direvisi atau diperbaiki. Merujuk pendapat Ebel, bahwa soal dapat diterima apabila soal tersebut diatas 0,41. Berdasarkan itu maka, ada 24 soal yang harus diperbaiki. Dan ada 10 soal dengan katagori baik dan baik sekali, dan soal ini yang bisa diterima.

B. Rekomendasi

Melihat paparan data dan analisis hasil penelitian diatas penulis memberikan rekomendasi sebagai berikut :

1. Widyaiswara

- a. Kemampuan guru dalam menyusun soal standar masih perlu ditingkatkan melalui banyak latihan megembangkan Indikator soal dan mengembangkan indikator soal menjadi soal.

- b. Kemampuan guru dalam menganalisis soal baik itu analisis kullitatif maupun kuantitatif perlu diasah lagi dengan memperbanyak latihan.

2. Institusi

- a. Mata Diklat Pengembangan Kisi-Kisi, Penyusunan Soal, dan Analisis Soal, perlu mendapat waktu yang memadai, agar peserta Diklat mempunyai kesempatan untuk berlatih mengembangkan Kisi-kisi, penyusunan Soal dan analisisnya dengan baik. [α]

DAFTAR PUSTAKA

- Aiken, Lewis R. 1994. *Psychological Testing and Assessment*. Eight Edition, Allyn and Bacon, Boston
- Asmawi Zainul dan Noehi Nasoetion. 1997. *Penilaian Hasil Belajar*. Pusat Antar Universitas, Direktorat Jenderal Pendidikan Tinggi : Departemen Pendidikan Dank ebudayaan.
- Azwar, Saifuddin. Tes Prestasi. Yogyakarta : Pustaka Pelajar, 2006..
- Gronlund, N.E., & Linn, R.L. (1990). *Measurement and evaluation in teaching* (6th ed.). New York: MacMillan.
- Lee J.Cronbach, *Essential sof Psychological Testing*, (NewYork : Harper and Row
- Nunnally, J.C. (1972). *Educational measurement and evaluation* (2nd ed.). New York: McGraw-Hill.
- Nur kancana, Wayan dan Supartana. Evaluasi Hasil Belajar. Surabaya: Usaha Nasional, 2002
- PP no. 74 tahun 2008, pasal 52 ayat 1 dan pasal 52 ayat 1 hurug (e). Publisher, 1984), hal. 26.
- Purwanto, Ngalim. (2000). *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*, PT. Remaja Rosdakarya. Jl. Ibu Inggit Garnasih no. 40, Bandung, 40252
- Sirait, Bistok. Bahan Pengajaran Untuk Mata Kuliah Evaluasi Hasil Belajar Siswa. Jakarta: Depdikbud, 2009.
- Sujiono, Anas. (2005). *Pengantar Evaluasi Pendidikan*. PT. Raja Grafindo Persada. Jl. Pelepah Hijau IV TNI. No. 14-15 Kelapa Gading Permai, Jakarta.
- Susan Matlock-Hetze 1997. *I* Paper presented at the annual meeting of the Southwest Educational Research Association, Austin. USA
- Thoha, Chabib M. (1991). *Teknik Evaluasi Pendidikan*, CV, Rajawali, Jl. Pelepah Hijau IV TNI. No. 14-15 Kelapa Gading Permai, Jakarta.
- Wood, D.A. (1960). *Test construction : Development and interpretation of achievement tests*. Columbus, OH: Charles E. Merrill Books, Inc.