

**ANALISIS BUTIR SOAL UJIAN
MATA PELAJARAN BAHASA INDONESIA SEKOLAH DASAR
TAHUN PELAJARAN 20013/2014
DI KABUPATEN PURBALINGGA**

Dwi Haryanto
Sekolah Dasar Negeri 1 Kutasari, Purbalingga, Jawa Tengah, Indonesia
Sur-el: dwiharyanto1968@yahoo.com

Abstrak: Ujian Sekolah Dasar sebagai salah satu penilaian yang berfungsi untuk (1) pemetaan mutu satuan pendidikan, (2) dasar seleksi masuk jenjang pendidikan berikutnya, (3) penentuan kelulusan peserta didik dari satuan pendidikan, dan (4) dasar pembinaan dan pemberian bantuan kepada satuan pendidikan dalam upayanya meningkatkan mutu satuan pendidikan harus memenuhi kriteria sebagai instrumen penilaian (tes) yang bermutu. Tujuan penelitian ini adalah untuk mengetahui kualitas/ mutu butir soal ujian sekolah secara empiris ditinjau dari kesahihan/ validitas, kepercayaan/ reliabilitas, tingkat kesulitan, daya pembeda dan efektifitas butir pengecoh (distraktor). Data penelitian dikumpulkan melalui teknik dokumentari, yaitu dengan menggunakan lembar jawab ujian yang telah ada. Analisis datanya dengan menggunakan statistik deskriptif dan persentase. Hasil penelitian menunjukkan bahwa butir yang memenuhi kriteria validitas sebanyak 47%, tingkat kesulitan soal 38%, daya beda 42%, dan efektifitas butir pengecoh 94% serta tingkat reliabilitas memenuhi persyaratan. Dari penggabungan keseluruhan kriteria tersebut, diketahui ada 14 butir soal atau 28% memenuhi semua kriteria yang ditetapkan (layak) dan 36 butir soal atau 72% tidak memenuhi satu atau lebih kriteria sehingga dinyatakan tidak layak. Secara keseluruhan soal ujian sekolah tersebut kualitasnya masih rendah sehingga masih perlu diperbaiki untuk menjadi soal ujian.

Kata kunci: analisis butir soal, ujian, Sekolah Dasar, bahasa Indonesia

***ANALYSIS OF TEST ITEMS
USED IN THE ELEMENTARY SCHOOL
INDONESIAN LANGUAGE EXAM
IN THE 20013/2014 ACADEMIC YEAR
IN PURBALINGGA REGENCY***

Abstract: *The elementary school exams as one of the assessments that serve as (1) the mapping of educational unit quality; (2) the basis for selection in a further education entry; (3) the measure of student study completion from the educational unit; and (4) the basis for the upgrading and grant provision to an educational unit in an effort to improve its quality should meet the criteria of a good test instrument. This study was aimed at describing the quality of test items in the Elementary School Indonesian language exam empirically viewed from validity, reliability, item difficulty, discriminating index and effectiveness of distracter. Data were collected through documentary technique, i.e. by using the exam worksheets available. Data were analysed using descriptive statistic and*

percentage. The results have shown that: 47% of the items fulfilled the requirement of item validity, 38% fulfilled the requirement of item difficulty, 42% fulfilled the requirement of discriminating index, 94% fulfilled the requirement of effectiveness of distracters, and the instrument was said to meet the requirement of test reliability. From all those requirements combined, it revealed that only 14 items (28%) fulfilled the criteria, and were considered 'good' items. Overall, the items used in the Elementary School Indonesian language exam were considered poor and, therefore, needed be improved for an exam.

Key words: analysis of test items, exam, elementary school, the indonesia language

PENDAHULUAN

Penilaian pendidikan sebagaimana tercantum dalam Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 adalah proses pengumpulan dan pengolahan informasi untuk menentukan pencapaian hasil belajar peserta didik. Peraturan Menteri Pendidikan dan Kebudayaan Nomor 102 Tahun 2013 Bab II pasal 2 menyatakan bahwa Ujian Sekolah merupakan kegiatan pengukuran pencapaian kompetensi peserta didik pada semua mata pelajaran dan muatan lokal sesuai Standar Nasional. Selanjutnya pada pasal 3 ayat (1) dinyatakan bahwa Ujian Sekolah dapat digunakan sebagai salah satu pertimbangan untuk (a) pemetaan mutu Satuan Pendidikan, (b) dasar seleksi masuk jenjang pendidikan berikutnya, (c) penentuan kelulusan peserta didik dari Satuan Pendidikan, dan (d)

pembinaan dan pemberian bantuan kepada Satuan Pendidikan dalam upayanya untuk meningkatkan mutu pendidikan. Pada Bab IV pasal 5 butir c. dinyatakan bahwa peserta didik dinyatakan lulus dari satuan pendidikan setelah lulus ujian sekolah.

Hasil ujian harus memberikan informasi tentang peserta didik sesuai dengan keadaan yang mendekati sesungguhnya. Untuk keperluan itu, soal Ujian Sekolah harus memenuhi syarat "baik". Gronlund (dalam Nurgiyantoro, 2013: 150) menyatakan, sebuah alat tes (dalam hal ini soal Ujian Sekolah) yang baik, harus mempunyai kriteria validitas (*validity*), realibilitas (*reliability*), dan kebergunaan (*usability*). Validitas menunjuk pada pengertian *apabila instrumen itu dapat mengukur apa yang hendak diukur* (Widoyoko, 2012: 128). Reliabilitas menunjuk pada

instrumen jika memberikan hasil yang tetap (konsisten) apabila diteskan berkali-kali.

Bertolak dari uraian di atas maka perlu diadakan analisis butir soal Ujian Sekolah Dasar Mata Pelajaran Bahasa Indonesia Tahun Pelajaran 2013/2014 di Kabupaten Purbalingga dengan tujuan untuk mengetahui kriteria soal Ujian Sekolah yaitu valid, reliabel, tingkat kesulitan, daya beda juga efektivitas butir pengecoh dari setiap butir soal. Penelitian ini dilaksanakan dengan tujuan untuk mengetahui: (1) tingkat kesahihan, (2) tingkat keterpercayaan, (3) tingkat kesukaran, (4) daya pembeda, dan (5) efektivitas distraktor (butir pengecoh) setiap butir soal.

Analisis butir soal adalah identifikasi jawaban benar dan salah tiap butir soal yang diujikan kepada peserta didik (Nurgiyantoro, 2013: 190). Melalui kerja analisis ini akan diketahui butir-butir soal mana saja yang dijawab benar oleh peserta tes atau sebaliknya, yaitu butir-butir soal mana saja yang dijawab salah oleh peserta tes. Berdasarkan jawaban benar dan salah itulah akan dihitung validitas, reliabilitas, indeks tingkat kesulitan, daya beda maupun efektivitas butir pengecohnya.

Selanjutnya Thorndike dan Hagen (dalam Purwanto, 1984: 118) menyatakan bahwa analisis terhadap soal-soal (*items*) tes yang telah dijawab oleh peserta didik mempunyai dua tujuan yang penting. Pertama, jawaban-jawaban soal itu merupakan informasi diagnostik untuk meneliti pembelajaran dan kedua adalah perbaikan (*review*) soal-soal didasarkan atas jawaban-jawaban itu merupakan basis bagi penyiapan tes-tes yang lebih baik untuk waktu berikutnya.

Daryanto (2010: 177-178) menyatakan analisis soal adalah suatu prosedur yang sistematis, yang akan memberikan informasi-informasi yang sangat khusus terhadap butir tes. Adapun faedah mengadakan analisis butir soal, yaitu:

1. Membantu guru dalam mengidentifikasi butir-butir soal yang jelek.
2. Memperoleh informasi yang dapat digunakan untuk menyempurnakan soal-soal untuk kepentingan selanjutnya.
3. Memperoleh gambaran keadaan sebuah butir soal secara empiris.

Dari penjelasan di atas, dapat diketahui bahwa analisis butir soal

bertujuan untuk mengidentifikasi soal-soal yang baik, kurang baik dan soal yang jelek.

Ujian sekolah/madrasah adalah kegiatan pengukuran pencapaian kompetensi peserta didik yang dilakukan oleh satuan pendidikan untuk memperoleh pengakuan atas prestasi belajar dan merupakan salah satu persyaratan kelulusan dari satuan pendidikan. Selain itu, ujian sekolah juga digunakan sebagai pemetaan mutu satuan pendidikan, khususnya untuk mata pelajaran Bahasa Indonesia, Matematika, dan IPA pada SD/ MI/ SDLB yang dilakukan oleh Kementerian (Permendikbud, 2013 pasal 3 ayat 2).

Tes (dalam hal ini, soal ujian sekolah) sebagai salah satu alat ukur hasil peserta didik diharapkan mampu memberikan informasi yang dapat dipertanggungjawabkan kebenarannya. Artinya, soal ujian sekolah tersebut dapat memberikan informasi tentang peserta didik sesuai dengan keadaan yang sesungguhnya (Nurgiyantoro, 2013:149). Widoyoko (2009:127) menyatakan bahwa data penilaian yang baik adalah data yang sesuai dengan kenyataan yang sebenarnya dan data tersebut bersifat

tetap, ajeg, dan dapat dipercaya. Arikunto (2009: 57) menyatakan bahwa suatu tes dapat dikatakan baik apabila memenuhi lima persyaratan, yaitu: validitas, reliabilitas, objektivitas, praktikabilitas, dan ekonomis.

Sahih (valid), artinya bahwa alat tersebut (tes) hanya mengukur satu dimensi/aspek saja, misal tes matematika, hanya mengukur pengetahuan matematika saja, tidak mengukur pengetahuan lain (Depdikbud, 1992: 10). Sahih, berarti penilaian didasarkan pada data yang mencerminkan kemampuan yang diukur (Depdiknas, 2007: 1). Selanjutnya Widoyoko (2009: 128) mengemukakan bahwa instrumen (tes) dikatakan valid apabila instrumen tersebut dapat dengan tepat mengukur apa yang hendak diukur atau ketepatan alat ukur.

Reliabel (andal), artinya bahwa alat tes tersebut harus dapat memberikan hasil pengukuran yang tepat, cermat dan ajeg (kurang lebih sama walaupun diberikan pada waktu yang berbeda) (Depdikbud, 1992: 11). Nurkencana (1986: 131) menyatakan bahwa suatu tes dikatakan tes yang reliabilitas apabila tes tersebut menunjukkan hasil-hasil yang mantap. Sedangkan Nurgiyantoro (2013:165)

menyatakan reliabilitas tes adalah jika sebuah tes diujicobakan lebih dari satu kali kepada subjek yang sama dapat menghasilkan data yang kurang lebih sama. Alat tersebut dapat mengukur secara konsisten, secara ajeg.

Nurgiyantoro (2013:194) menyatakan, indeks tingkat kesulitan (ITK) adalah indeks yang menunjukkan seberapa mudah atau sulit suatu butir soal bagi peserta tes yang diuji. Selanjutnya Oller (dalam Nurgiyantoro, 2013:194) mengemukakan bahwa tingkat kesulitan (*item difficulty*) adalah pernyataan tentang seberapa mudah atau sulit sebuah butir soal bagi siswa.

Daya pembeda (*item discriminability*) maksudnya adalah seberapa besar suatu butir soal dapat membedakan antara siswa kelompok tinggi dengan kelompok rendah. Purwanto (2009: 120) menyatakan bahwa daya pembeda suatu soal tes ialah kemampuan soal itu untuk membedakan siswa-siswa yang termasuk kelompok pandai (*upper group*) dengan siswa-siswa yang termasuk kelompok kurang (*lower group*).

Pengecoh (*distracter*) adalah pilihan jawaban yang bukan

merupakan kunci jawaban. Pengecoh diadakan untuk menyesatkan peserta didik agar tidak memilih kunci jawaban (Waridjan, 1984: 387). Daryanto (2010: 192) menyatakan bahwa sebuah pengecoh dapat dikatakan berfungsi dengan baik apabila pengecoh mempunyai daya tarik yang besar bagi pengikut-pengikut tes yang kurang memahami konsep atau kurang menguasai materi.

METODOLOGI PENELITIAN

Penelitian ini mengambil lokasi di Sekolah Dasar di wilayah Kabupaten Purbalingga dan terfokus pada tes hasil belajar dalam ujian sekolah tingkat Sekolah Dasar Mata Pelajaran Bahasa Indonesia yang dilaksanakan pada tanggal 19 Mei 2014, dengan mengambil objek penelitian dan kegiatan analisis dari tanggal 1 sampai dengan 14 Desember 2014. Untuk penelitian ini diambil sampel 5% dari seluruh populasi yang berjumlah 470 sekolah atau 9.572 siswa, sehingga sampelnya ada 24 sekolah atau 479 siswa.

Metode yang digunakan dalam penelitian ini adalah penelitian deskriptif. Penelitian deskriptif tidak

dimaksudkan untuk menguji hipotesis tertentu, tetapi hanya menggambarkan “apa adanya” tentang suatu gejala atau keadaan. Penelitian yang dilakukan diupayakan seobjektif mungkin terhadap hal-hal yang menjadi pusat perhatian dan mendukung penelitian. Peneliti bertugas untuk mengumpulkan data, mendeskripsikannya, menganalisisnya, sampai akhirnya dapat membuat kesimpulan sebagai jawaban terhadap masalah penelitian yang diajukan-nya.

Prosedur atau tahapan penelitian dalam penelitian ini sebagai berikut: (1) menyusun instrumen atau rambu-rambu pengukuran, (2) mengumpulkan data, (3) mendeskripsikan data, (4) menganalisis data, (5) merumuskan simpulan.

HASIL DAN PEMBAHASAN

Analisis uji validitas butir soal ini dilakukan dengan menggunakan korelasi *point biserial* dengan hasil sebagai berikut.

Analisis Validitas Butir Soal Ujian Sekolah Dasar Mata Pelajaran Bahasa Indonesia Tahun Pelajaran 2013/2014 di Kabupaten Purbalingga

No Butir	Mp	Mt	SDt	p	Q	$\sqrt{\frac{p}{q}}$	<i>r_{pbi}</i>	<i>r</i> tabel 5%	Interpretasi
1	40.340	40.094	5.116	0.908	0.092	3.144	0.151	0.088	Valid
2	41.257	40.094	5.116	0.658	0.342	1.386	0.315	0.088	Valid
3	40.639	40.094	5.116	0.307	0.693	0.665	0.071	0.088	Tidak valid
4	40.454	40.094	5.116	0.960	0.040	4.92	0.347	0.088	Valid
5	40.981	40.094	5.116	0.781	0.219	1.887	0.327	0.088	Valid
6	41.874	40.094	5.116	0.597	0.403	1.217	0.424	0.088	Valid
7	40.671	40.094	5.116	0.883	0.117	2.748	0.310	0.088	Valid
8	41.593	40.094	5.116	0.672	0.328	1.432	0.420	0.088	Valid
9	40.733	40.094	5.116	0.877	0.123	2.668	0.333	0.088	Valid
10	40.652	40.094	5.116	0.875	0.125	2.643	0.288	0.088	Valid
11	41.273	40.094	5.116	0.649	0.351	1.361	0.314	0.088	Valid
12	40.964	40.094	5.116	0.578	0.422	1.171	0.199	0.088	Valid
13	40.554	40.094	5.116	0.931	0.069	3.676	0.330	0.088	Valid
14	40.347	40.094	5.116	0.937	0.063	3.869	0.192	0.088	Valid
15	40.411	40.094	5.116	0.935	0.065	3.802	0.235	0.088	Valid
16	40.199	40.094	5.116	0.904	0.096	3.068	0.063	0.088	Tidak valid

17	40.935	40.094	5.116	0.772	0.228	1.842	0.303	0.088	Valid
18	40.825	40.094	5.116	0.868	0.132	2.57	0.367	0.088	Valid
19	41.332	40.094	5.116	0.779	0.221	1.876	0.454	0.088	Valid
20	45.082	40.094	5.116	0.889	0.111	2.835	2.764	0.088	Valid
21	40.978	40.094	5.116	0.768	0.232	1.821	0.315	0.088	Valid
22	41.092	40.094	5.116	0.770	0.230	1.832	0.357	0.088	Valid
23	41.210	40.094	5.116	0.814	0.186	2.093	0.457	0.088	Valid
24	40.502	40.094	5.116	0.931	0.069	3.676	0.293	0.088	Valid
25	41.326	40.094	5.116	0.653	0.347	1.373	0.331	0.088	Valid
26	40.628	40.094	5.116	0.948	0.052	4.261	0.445	0.088	Valid
27	40.641	40.094	5.116	0.854	0.146	2.417	0.258	0.088	Valid
28	40.817	40.094	5.116	0.766	0.234	1.81	0.256	0.088	Valid
29	40.457	40.094	5.116	0.923	0.077	3.456	0.245	0.088	Valid
30	40.417	40.094	5.116	0.942	0.058	4.013	0.253	0.088	Valid
31	41.020	40.094	5.116	0.843	0.157	2.321	0.420	0.088	Valid
32	40.560	40.094	5.116	0.925	0.075	3.508	0.319	0.088	Valid
33	40.538	40.094	5.116	0.912	0.088	3.226	0.280	0.088	Valid
34	40.392	40.094	5.116	0.931	0.069	3.676	0.214	0.088	Valid
35	40.817	40.094	5.116	0.831	0.169	2.217	0.313	0.088	Valid
36	40.606	40.094	5.116	0.854	0.146	2.417	0.242	0.088	Valid
37	41.494	40.094	5.116	0.367	0.633	0.762	0.209	0.088	Valid
38	41.467	40.094	5.116	0.630	0.370	1.306	0.351	0.088	Valid
39	40.462	40.094	5.116	0.927	0.073	3.562	0.256	0.088	Valid
40	40.315	40.094	5.116	0.962	0.038	5.061	0.218	0.088	Valid
41	40.779	40.094	5.116	0.823	0.177	2.153	0.288	0.088	Valid
42	40.581	40.094	5.116	0.902	0.098	3.032	0.289	0.088	Valid
43	40.842	40.094	5.116	0.858	0.142	2.458	0.359	0.088	Valid
44	40.524	40.094	5.116	0.948	0.052	4.261	0.358	0.088	Valid
45	41.805	40.094	5.116	0.171	0.829	0.454	0.152	0.088	Valid
46	41.289	40.094	5.116	0.693	0.307	1.503	0.351	0.088	Valid
47	40.456	40.094	5.116	0.714	0.286	1.58	0.112	0.088	Valid
48	41.010	40.094	5.116	0.800	0.200	1.997	0.358	0.088	Valid
49	40.201	40.094	5.116	0.944	0.056	4.092	0.086	0.088	Tidak valid
50	40.209	40.094	5.116	0.927	0.073	3.562	0.080	0.088	Tidak valid

Setelah dikonsultasikan ke tabel harga kritik *r product moment* diketahui signifikan atau tidaknya masing-masing butir soal tersebut.

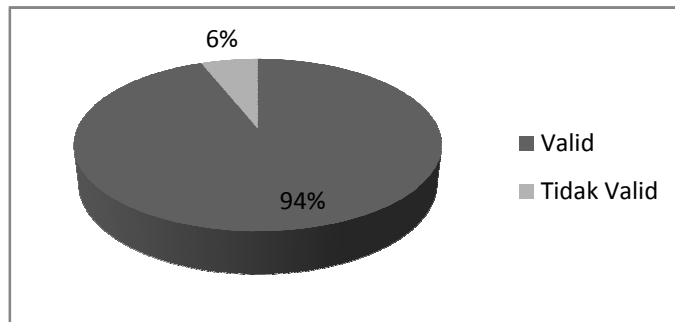
Jika harga r hitung (r_h) lebih kecil dari harga kritik dalam tabel (r_t), maka korelasi tersebut tidak signifikan atau tidak valid (*invalid*). Sebaliknya jika

harga r hitung (r_h) lebih besar dari harga kritik dalam tabel (r_t), maka korelasi butir soal tersebut signifikan atau valid.

Dari seluruh 50 butir soal,

yang memenuhi kriteria (signifikan/valid) sebanyak 47 butir soal (94%) sedangkan untuk butir soal yang tidak memenuhi kriteria sebanyak 3 butir soal atau 6%.

Porsi Butir Soal Valid



Penghitungan reliabilitas butir soal keseluruhan, dilakukan dengan teknik *split half* lalu dengan rumus korelasi *product moment*. Untuk

memperoleh indeks reliabilitas dihitung dengan rumus Spearman Brown, dengan hasil sebagai berikut.

Reliabilitas Keseluruhan Soal/Instrumen

No	Jumlah Soal	r_{hitung}	r_{tabel}	Keterangan
1	50	0,795	0,279	reliabel

Untuk r_{tabel} yang diminta dengan jumlah soal (N) sebanyak 50 butir adalah dengan taraf signifikansi 5% pada tabel r *product moment* adalah 0,279. Dari hasil penghitungan di atas harga r hitung (r_h) yaitu 0,795 ternyata lebih besar dari r tabel (r_t) yaitu 0,279 yang dapat diartikan ada korelasi yang signifikan sehingga dapat disimpulkan bahwa soal reliabel.

Untuk menghitung indeks kesulitan butir soal (ITK) dilaksanakan melalui beberapa tahapan, di antaranya menentukan kelompok rendah dan tinggi dengan pedoman 27,5% dari seluruh sampel, maka jumlah subjek untuk masing-masing kelompok sebanyak $27,5\% \times 479 = 132$ responden/ subjek. Setelah itu dilakukan pengurutan skor perolehan siswa untuk menentukan responden

mana saja yang masuk kelompok tersebut.

Untuk tingkat kesulitan yang memenuhi persyaratan (indeks yang dapat ditoleransi), berkisar antara 0,20

sampai dengan 0,80 sedangkan ITK yang kurang 0,20 atau lebih dari 0,80

berarti soal tidak layak (Nurgiyantoro, 2013: 195).

Indeks Tingkat Kesulitan (ITK) Butir Soal

No. Soal	Σ FKT	Σ FKR	N	ITK	Keterangan
1	129	110	264	0.905	Tidak Layak
2	114	69	264	0.693	Layak
3	49	40	264	0.337	Layak
4	132	119	264	0.951	Tidak Layak
5	118	82	264	0.758	Layak
6	116	51	264	0.633	Layak
7	126	105	264	0.875	Tidak Layak
8	122	55	264	0.670	Layak
9	130	98	264	0.864	Tidak Layak
10	127	100	264	0.860	Tidak Layak
11	100	59	264	0.602	Layak
12	105	64	264	0.640	Layak
13	132	107	264	0.905	Tidak Layak
14	129	118	264	0.936	Tidak Layak
15	130	112	264	0.917	Tidak Layak
16	125	118	264	0.920	Tidak Layak
17	119	83	264	0.765	Layak
18	130	99	264	0.867	Tidak Layak
19	127	71	264	0.750	Layak
20	130	94	264	0.848	Tidak Layak
21	116	84	264	0.758	Layak
22	126	74	264	0.758	Layak
23	131	78	264	0.792	Layak
24	130	110	264	0.909	Tidak Layak
25	108	59	264	0.633	Layak
26	132	112	264	0.924	Tidak Layak
27	126	95	264	0.837	Tidak Layak
28	110	86	264	0.742	Layak
29	129	111	264	0.909	Tidak Layak
30	129	116	264	0.928	Tidak Layak
31	130	82	264	0.803	Tidak Layak
32	132	110	264	0.917	Tidak Layak

33	126	108	264	0.886	Tidak Layak
34	125	114	264	0.905	Tidak Layak
35	127	87	264	0.811	Tidak Layak
36	127	102	264	0.867	Tidak Layak
37	70	33	264	0.390	Layak
38	119	55	264	0.659	Layak
39	127	116	264	0.920	Tidak Layak
40	132	120	264	0.955	Tidak Layak
41	126	93	264	0.830	Tidak Layak
42	130	108	264	0.902	Tidak Layak
43	124	95	264	0.830	Tidak Layak
44	130	114	264	0.924	Tidak Layak
45	32	11	264	0.163	Tidak Layak
46	111	64	264	0.663	Layak
47	104	84	264	0.712	Layak
48	126	80	264	0.780	Layak
49	128	119	264	0.936	Tidak Layak
50	128	117	264	0.928	Tidak Layak

Keterangan :

$\sum FKT$ = Jumlah skor kelompok tinggi

$\sum FKR$ = Jumlah skor kelompok rendah

N = Jumlah siswa kelompok tinggi dan rendah

ITK = Indeks tingkat kesulitan

Dari penghitungan indeks tingkat kesulitan maka didapat hasil bahwa soal yang memenuhi kriteria ada 19 butir soal atau 38%, sedang 31 butir soal atau 62% adalah butir soal yang tidak memenuhi kriteria (tidak layak).

Analisis daya pembeda artinya mengkaji butir-butir soal tes dari segi kesanggupan tes tersebut membedakan

peserta didik yang masuk dalam katagori peserta didik yang rendah dan tinggi prestasinya. Nurgiyantoro (2013: 198) menyatakan bahwa untuk indeks daya pembeda yang diper-syaratkan minimal 0,25. Apabila sebuah butir soal tidak memenuhi kriteria tersebut maka dikelompokkan dalam kriteria tidak layak.

Daya Pembeda Butir Soal

No. Soal	$\sum FKT$	$\sum FKR$	n	IDB	Keterangan
1	129	110	132	0.144	Tidak Layak
2	114	69	132	0.341	Layak

3	49	40	132	0.068	Tidak Layak
4	132	119	132	0.098	Tidak Layak
5	118	82	132	0.273	Layak
6	116	51	132	0.492	Layak
7	126	105	132	0.159	Tidak Layak
8	122	55	132	0.508	Layak
9	130	98	132	0.242	Layak
10	127	100	132	0.205	Layak
11	100	59	132	0.311	Layak
12	105	64	132	0.311	Layak
13	132	107	132	0.189	Tidak Layak
14	129	118	132	0.083	Tidak Layak
15	130	112	132	0.136	Tidak Layak
16	125	118	132	0.053	Tidak Layak
17	119	83	132	0.273	Layak
18	130	99	132	0.235	Tidak Layak
19	127	71	132	0.424	Layak
20	130	94	132	0.273	Layak
21	116	84	132	0.242	Tidak Layak
22	126	74	132	0.394	Layak
23	131	78	132	0.402	Layak
24	130	110	132	0.152	Tidak Layak
25	108	59	132	0.371	Layak
26	132	112	132	0.152	Tidak Layak
27	126	95	132	0.235	Tidak Layak
28	110	86	132	0.182	Tidak Layak
29	129	111	132	0.136	Tidak Layak
30	129	116	132	0.098	Tidak Layak
31	130	82	132	0.364	Layak
32	132	110	132	0.167	Tidak Layak
33	126	108	132	0.136	Tidak Layak
34	125	114	132	0.083	Tidak Layak
35	127	87	132	0.303	Layak
36	127	102	132	0.189	Tidak Layak
37	70	33	132	0.280	Layak
38	119	55	132	0.485	Layak
39	127	116	132	0.083	Tidak Layak
40	132	120	132	0.091	Tidak Layak
41	126	93	132	0.250	Layak
42	130	108	132	0.167	Tidak Layak
43	124	95	132	0.220	Tidak Layak

44	130	114	132	0.121	Tidak Layak
45	32	11	132	0.159	Tidak Layak
46	111	64	132	0.356	Layak
47	104	84	132	0.152	Tidak Layak
48	126	80	132	0.348	Layak
49	128	119	132	0.068	Tidak Layak
50	128	117	132	0.083	Tidak Layak

Keterangan :

\sum FKT = Jumlah skor kelompok tinggi

\sum FKR = Jumlah skor kelompok rendah

n = Jumlah siswa kelompok tinggi atau rendah

IDB = Indeks daya beda

Dari data tabel di atas dapat dilihat bahwa 21 butir soal memenuhi kriteria yang ditentukan atau sebanyak 42% disebut layak. Sedangkan sebanyak 29 butir soal atau 58% ternyata indeks daya bedanya tidak memenuhi kriteria yang ditentukan dan dikatakan tidak layak.

Untuk mengetahui efektivitas tiap alternatif jawaban (*option*) atau sebaliknya, yaitu adanya penyimpangan alternatif jawaban, perlu dilakukan juga kegiatan analisis

butir pengecoh (*distracter*). Dari kegiatan ini akan diketahui efektif tidaknya tiap alternatif pilihan jawaban dan sebaran frekuensi jawaban.

Berdasarkan data dari kegiatan analisis sebaran frekuensi jawaban peserta didik maka dapat diketahui efektivitas butir pengecoh sehingga jika ada penyimpangan pada alternatif jawaban (*option*) perlu juga dilakukan revisi. Deskripsi dan analisis sebaran frekuensi jawaban masing-masing butir soal dapat dilihat pada tabel berikut.

Sebaran Frekuensi Jawaban Siswa Terhadap Alternatif-alternatif Jawaban

Nomor Soal	Kelompok Tinggi					Kelompok Rendah				Opsis yang perlu direvisi
	A	B	C	D		A	B	C	D	
1	0	2	129*	1	3	5	110*	14	-	
2	9	1	8	114*	33	2	28	69*	-	
3	49*	69	1	14	40*	66	6	20	-	
4	1	2	1	130*	2	11	0	119*	C	

5	118*	5	6	5	82*	14	8	28	-
6	6	1	9	116*	17	2	62	51*	-
7	0	1	126*	5	5	1	105*	21	-
8	1	122*	1	6	4	55*	6	67	-
9	130*	1	1	1	98*	20	9	5	-
10	1	127*	2	2	0	100*	8	24	A
11	2	6	100*	24	23	21	59*	29	-
12	20	1	6	105*	56	3	9	64*	-
13	132*	0	0	0	107*	11	5	9	-
14	1	129*	1	1	7	118*	2	5	-
15	1	1	130*	0	14	4	2	112*	-
16	125*	2	5	0	118*	6	7	1	-
17	0	119*	10	3	11	83*	22	16	-
18	130*	1	1	1	99*	27	4	2	-
19	3	1	127*	1	28	14	71*	19	-
20	0	130*	0	2	13	94*	15	10	-
21	0	116*	5	1	14	84*	32	2	-
22	126*	6	0	0	74*	45	11	2	-
23	0	1	131*	0	33	4	78*	17	-
24	130*	1	0	1	110*	14	4	4	-
25	20	108*	2	2	22	59*	17	34	-
26	132*	0	0	0	112*	8	1	11	-
27	126*	0	1	5	95*	6	13	18	-
28	0	110*	1	21	16	86*	6	24	-
29	1	1	1	129*	17	2	2	111*	-
30	2	129*	1	0	13	116*	1	2	-
31	130*	0	1	1	82*	17	17	16	-
32	0	0	132*	0	11	12	110*	9	-
33	4	126*	2	0	14	108*	7	3	-
34	125*	5	1	1	114*	10	4	4	-
35	0	127*	5	0	7	87*	34	4	-
36	0	3	127*	2	8	5	102*	17	-
37	0	132*	0	0	33	33*	19	47	-
38	3	6	4	119*	31	17	29	55*	-
39	5	0	0	127*	3	10	3	116*	A
40	132*	0	0	0	120*	8	2	2	-
41	1	126*	4	1	16	93*	22	1	-
42	1	1	130*	0	8	6	108*	10	-
43	8	0	124*	0	27	8	95*	2	-
44	130*	1	1	0	114*	7	10	1	-
45	5	95	0	32*	11	93	17	11*	-
46	111*	1	3	7	64*	16	13	39	-
47	1	1	104*	26	4	11	84*	33	-
48	1	3	2	126*	16	20	16	80*	-

49	128*	4	0	0		119*	10	1	2	-
50	3	1	0	128*		6	3	6	117*	-

Keterangan :

A, B, C, D = alternatif jawaban (*option*) yang disediakan

* = kunci jawaban

Dari data pada tabel tersebut di atas maka diketahui ada 47 butir soal yang telah memenuhi kriteria atau tidak perlu ada revisi pada alternatif jawaban atau 94%, sedangkan butir soal yang perlu direvisi optionnya sejumlah 3 butir soal atau hanya 6%.

Selanjutnya dari beberapa penghi-tungan kriteria tersebut dapat digabung ke dalam satu tabel, sehingga sebuah butir soal dapat dikatakan layak (memenuhi semua kriteria) atau tidak layak (tidak memenuhi kriteria/ perlu revisi) berdasarkan kriteria yang telah ditetapkan. Sebab dimungkinkan sebuah butir soal memenuhi satu atau dua kriteria yang telah ditetapkan, tetapi belum tentu memenuhi semua kriteria yang telah ditentukan.

Secara keseluruhan berdasarkan analisis butir soal yang telah dilakukan maka dapat dilihat bahwa:

1. Tingkat kesahihan butir soal pada perangkat tes (soal ujian sekolah dasar tersebut) sebanyak 47 butir (94%) memenuhi kriteria dari jumlah seluruh butir soal yaitu 50 butir soal. Sedangkan yang tidak

memenuhi sebanyak 3 butir soal atau 6%.

2. Tingkat reliabilitas terpenuhi karena r hitung (r_h) lebih besar dari r tabel (r_t), yaitu $r_h (0,795) > r_t (0,279)$.
3. Tingkat kesukaran, ada 19 butir soal (38%) yang layak atau memenuhi kriteria yang dipersyaratkan, dan 31 butir soal (62%) yang tidak memenuhi kriteria atau tidak layak.
4. Tingkat daya beda, ada 21 butir soal yang memenuhi kriteria atau layak dan ada 29 butir soal yang tidak memenuhi kriteria atau tidak layak.
5. Butir pengecoh, sebanyak 47 butir soal atau 94% butir pengecohnya berfungsi dengan baik dan ada 3 butir soal atau 6% butir pengecohnya perlu direvisi.

Dari analisis yang telah dilakukan, maka secara keseluruhan dapat diketahui dari keseluruhan butir soal yang memenuhi semua kriteria (layak) ada 14 butir soal atau 28%, dan ada 36 atau 72 % butir soal yang tidak memenuhi satu atau lebih kriteria yang ditentukan termasuk butir soal yang tidak layak.

PENUTUP

Dari hasil kegiatan analisis butir soal yang telah dilakukan, maka dapat ditarik simpulan sebagai berikut: (1) tingkat kesahihan sebanyak 47 butir soal (94%) memenuhi kriteria (valid), sedangkan 3 butir soal atau 6% tidak valid, (2) tingkat reliabilitas, terpenuhi karena r hitung (r_h) lebih besar dari r tabel (r_t), yaitu $r_h = 0,795$ sedangkan $r_t = 0,279$, (3) tingkat kesukaran, ada 19 atau 38% yang layak dan 31 butir soal atau 62% tidak layak, (4) tingkat daya beda, ada 21 (42%) butir soal layak dan ada 29 (58%) butir soal yang tidak layak, (5) butir pengecoh, sebanyak 47 butir soal atau 94% butir pengecohnya berfungsi dengan baik dan ada 3 butir soal atau 6% butir pengecohnya perlu direvisi.

Dari penggabungan beberapa kriteria tersebut di atas, dapat diketahui bahwa butir soal yang memenuhi semua kriteria (layak) ada 14 butir soal atau 28% dan ada 36 butir soal atau 72 % yang tidak memenuhi satu atau lebih kriteria yang ditentukan termasuk butir soal yang tidak layak.

Penyusun soal dalam menyusun soal hendaknya memperhatikan kriteria-kriteria penyusunan soal yang ada baik dari segi materi, konstruksi, maupun bahasa. Selain hal tersebut

soal perlu juga diujicobakan sehingga dapat dilakukan analisis secara empiris yang selanjutnya dapat diketahui tingkat validitas, reliabilitas, tingkat kesukaran, daya beda serta efektivitas butir pengecohnya.

DAFTAR PUSTAKA

- Arikunto, Suharsimi. 2009. *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, Suharsimi. 2009. *Manajemen Pendidikan*. Jakarta: Rineka Cipta.
- Arikunto, Suharsimi. 2010. *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta: Rineka Cipta.
- Azwar, Saifudin. 1997. *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Pelajar.
- Balibang Kemendikbud. 2014. *Peraturan Ka. Balibang Kemendikbud Nomor 001/H/HK/2014 tentang POS Penyelenggaraan Ujian Sekolah/ Madrasah Tahun Pelajaran 2013/2014*. Jakarta: Balitbang Kemendikbud.
- Daryanto. 2010. *Evaluasi Pendidikan*. Jakarta: Rineka Cipta.
- Depdikbud. 1994. *Pedoman Penyusunan Perangkat Soal SD dan SMP*. Jakarta: Proyek Peningkatan Mutu Pendidikan Dasar.
- Heryadi, Dedi. 2010. *Metode Penelitian Pendidikan Bahasa*. Bandung: Pustaka Billah.

- Heryadi, Dedi. 2014. *Statistika Praktis untuk Penelitian Pendidikan*. Tasikmalaya: Universitas Siliwangi.
- Nasution, Noehi dan Adi Suryanto. 2005. *Evaluasi Pengajaran*. Jakarta: Universitas Terbuka.
- Nurgiyantoro, Burhan. 2013. *Penilaian dalam Pengajaran Bahasa Berbasis Kompetensi*. Yogyakarta: BPFE.
- Nurkencana dan Sumartana. 1983. *Evaluasi Pendidikan*. Surabaya: Usaha Nasional.
- Kemendikbud. 2013. *Permendikbud Nomor 102 Tahun 2013 tentang Penyelenggaraan Ujian Sekolah Dasar/ Madrasah Ibtidaiyah, SDLB dan Program Paket A/ ULA*. Jakarta: Kemendikbud.
- Mendiknas. 2007. *Permendiknas Nomor 20 Tahun 2007 tentang Standar Penilaian di Sekolah*. Jakarta: Kemendiknas.
- Purwanto. 2010. *Metodologi Penelitian Kuantitatif*. Yogyakarta: Pustaka Pelajar.
- Purwanto, M. Ngalm. 1984. *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT Remaja Rosdakarya.
- Sudijono, Anas. 2010. *Pengantar Evaluasi Pendidikan*, Jakarta: PT Raja Grafindo Persada.
- Sudjana, Nana. 1984. *Dasar-dasar Penilaian Hasil Belajar*. Jakarta: CV Serajaya.
- Sugiyono. 2010. *Statistika untuk Penelitian*. Jakarta: CV Serajaya.
- Sukardi. 2013. *Metodologi Penelitian Pendidikan*. Surabaya: PT Bina Ilmu.
- Sutomo. 1985. *Teknik Penilaian Pendidikan*. Surabaya: PT Bina Ilmu.
- Suwartono. 2014. *Dasar-dasar Metodologi Penelitian*. Yogyakarta: CV Andi Offset.
- Thoha, M. Chabib. 1994. *Teknik Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada.
- Waridjan. 1991. *Tes Hasil Belajar Gaya Objektif*. Semarang: IKIP Semarang Press.
- Widoyoko, Eko Putro. 2012. *Evaluasi Program Pembelajaran*. Yogyakarta: Pustaka Pelajar.