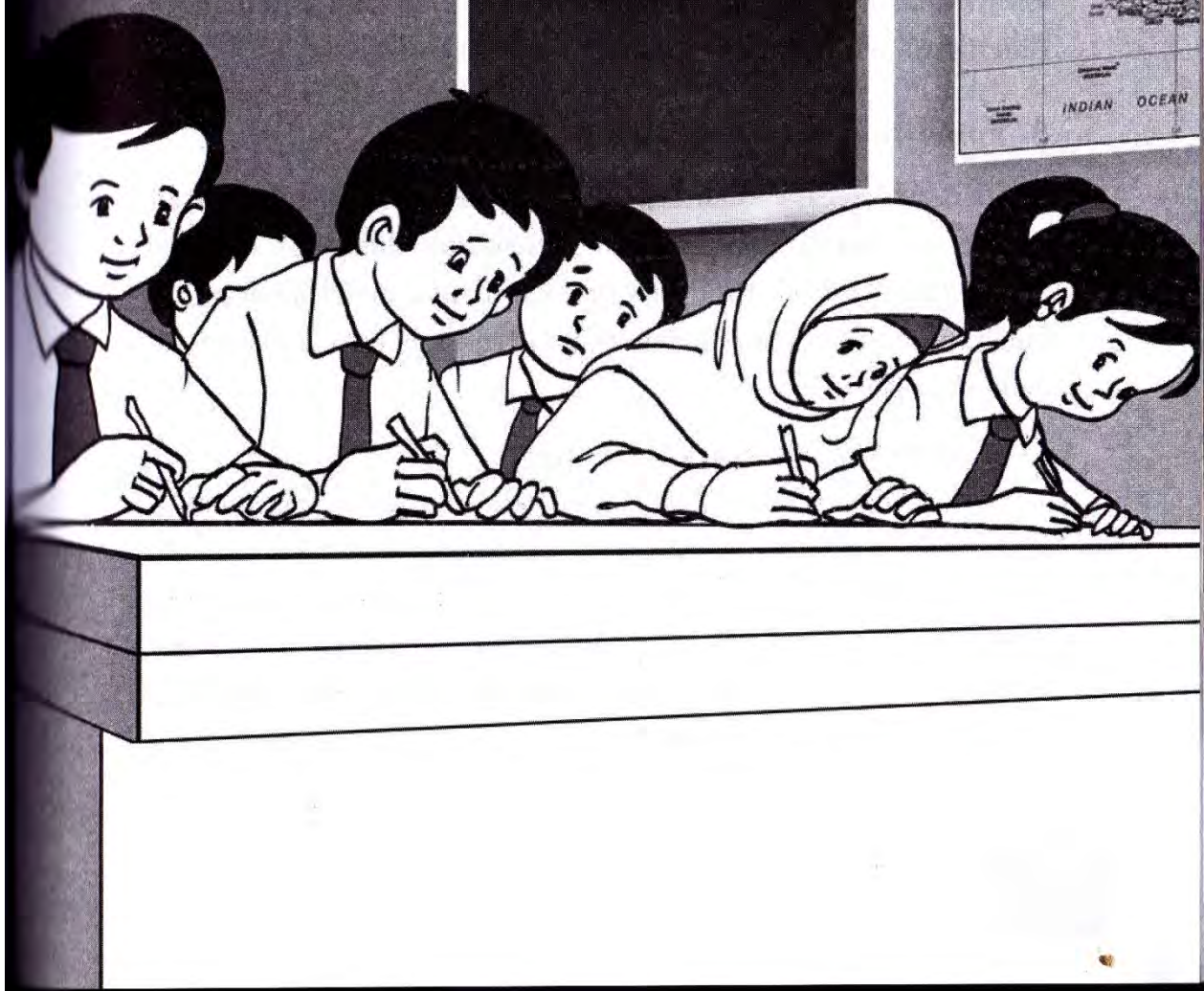


BAB 6

ANALISIS BUTIR SOAL UNTUK TES TULIS



Analisis butir soal pada tes tulis merupakan kegiatan penting dalam pengembangan tes agar diperoleh soal yang bermutu. Soal yang bermutu adalah soal yang dapat memberikan informasi setepat-tepatnya tentang kemampuan siswa. Tujuan kegiatan analisis adalah menelaah setiap soal agar diperoleh soal yang bermutu sebelum digunakan, meningkatkan kualitas butir soal melalui kegiatan revisi soal, dan membuang soal yang tidak efektif.

Analisis soal dapat dilakukan secara kualitatif (berkaitan dengan isi dan bentuknya) dan kuantitatif (berkaitan dengan ciri-ciri statistiknya). Analisis kualitatif mencakup pertimbangan validitas isi dan konstruk. Analisis kuantitatif mencakup pengukuran validitas dan reliabilitas soal, tingkat kesulitan, serta daya pembeda. Kedua teknik masing-masing memiliki keunggulan dan kelemahan. Karenanya, teknik terbaik adalah memadukan keduanya.

A. Analisis Soal secara Kualitatif

Analisis soal secara kualitatif dilaksanakan melalui penelaahan soal berdasarkan kaidah penulisan. Penelaahan dilakukan sebelum soal diujikan. Aspek yang diperhatikan dalam penelaahan secara kualitatif mencakup aspek materi, konstruksi, bahasa atau budaya, dan kunci jawaban.

1) Teknik Analisis secara Kualitatif

Ada beberapa teknik yang dapat digunakan untuk menganalisis soal secara kualitatif, yakni teknik moderator dan teknik panel. Teknik moderator merupakan teknik berdiskusi yang di dalamnya terdapat satu orang sebagai penengah. Berdasarkan teknik ini, setiap soal didiskusikan secara bersama-sama dengan beberapa ahli seperti guru yang mengajarkan materi, ahli materi, penyusun atau

Teknik ini sangat baik karena setiap soal dilihat secara bersama-sama berdasarkan kaidah penulisan. Di samping itu, para penelaah dipersilakan mengomentari berdasarkan kompetensinya masing-masing. Setiap komentar atau masukan dari peserta diskusi dicatat. Masing-masing soal dituntaskan secara bersama-sama, dan perbaikannya seperti apa. Teknik ini memiliki kelemahan karena memerlukan waktu lama untuk mendiskusikan setiap satu soal.

2) Prosedur Analisis secara Kualitatif

103

a. Format Penelaahan Soal Bentuk Uraian

Mata Pelajaran :

Kelas/semester :

Penelaah :

No.	Aspek yang Ditelaah	Nomor Soal							
		1	2	3	4	5	6	7	...
A. Materi									
1.	Soal sesuai dengan indikator (menuntut tes tertulis bentuk uraian)								
2.	Batasan pertanyaan dan jawaban yang diharapkan sudah sesuai								
3.	Materi yang ditanyakan sesuai dengan kompetensi (urgensi, relevansi, kontinuitas, keterpakaian sehari-hari tinggi)								
4.	Isi materi yang ditanyakan sesuai dengan jenjang, jenis sekolah, atau tingkat kelas								
B. Konstruksi									
5.	Menggunakan kata tanya atau perintah yang menuntut jawaban uraian								
6.	Ada petunjuk yang jelas tentang cara mengerjakan soal								
7.	Ada pedoman penskorannya								
8.	Tabel, gambar, grafik, peta, atau yang sejenisnya disajikan dengan jelas dan terbaca								
C. Bahasa/Budaya									
9.	Rumusan kalimat komunikatif								
10.	Soal menggunakan bahasa Indonesia baku								
11.	Tidak menggunakan kata atau ungkapan yang menimbulkan penafsiran ganda atau salah pengertian								
12.	Tidak menggunakan bahasa yang berlaku setempat/tabu								

b. Format Penelaahan Soal Bentuk Pilihan Ganda

Mata Pelajaran :

Kelas/semester :

Penelaah :

No	Aspek yang Ditelaah	Nomor Soal					
		1	2	3	4	5	...
A. Materi							
1.	Soal sesuai dengan indikator (menuntut tes tertulis bentuk pilihan ganda)						
2.	Materi yang ditanyakan sesuai dengan kompetensi (urgensi, relevansi, kontinuitas, keterpakaian sehari-hari tinggi)						
3.	Pilihan jawaban homogen dan logis						
4.	Hanya ada satu kunci jawaban						
B. Konstruksi							
5.	Pokok soal dirumuskan dengan singkat, jelas, dan tegas						
6.	Rumusan pokok soal dan pilihan jawaban merupakan pernyataan yang diperlukan saja						
7.	Pokok soal tidak memberi petunjuk kunci jawaban						
8.	Pokok soal bebas dari pernyataan yang bersifat negatif ganda						
9.	Pilihan jawaban homogen dan logis ditinjau dari segi materi						
10.	Gambar, grafik, tabel, diagram, atau sejenisnya jelas dan berfungsi						
11.	Panjang pilihan jawaban relatif sama						
12.	Pilihan jawaban tidak menggunakan pernyataan "Semua jawaban di atas salah atau benar" dan sejenisnya						
13.	Pilihan jawaban yang berbentuk angka/waktu disusun berdasarkan urutan besar kecilnya angka atau kronologisnya						
14.	Butir soal tidak bergantung pada jawaban soal sebelumnya						
C. Bahasa/Budaya							
15.	Menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia						
16.	Menggunakan bahasa yang komunikatif						
17.	Tidak menggunakan bahasa yang berlaku setempat atau tabu						
18.	Pilihan jawaban tidak mengulang kata/kelompok kata yang sama, kecuali merupakan satu kesatuan pengertian						

B. Penelaahan secara Kuantitatif

Penelaahan soal secara kuantitatif adalah penelaahan yang didasarkan pada data empirik. Data empirik diperoleh dari soal yang telah diujikan. Analisis ini dilakukan dengan cara menelaah soal berdasarkan informasi jawaban peserta tes. Aspek yang diperhatikan di antaranya setiap soal ditelaah dari segi tingkat kesukaran soal, daya pembeda, dan penyebaran pilihan jawaban (untuk soal bentuk pilihan ganda) atau frekuensi jawaban setiap pilihan jawaban.

1) Tingkat Kesukaran (TK)

Tingkat kesukaran soal adalah peluang menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Indeks tingkat kesukaran soal dinyatakan dalam bentuk proporsi yang besarnya berkisar dari 0 sampai 1. Semakin besar indeks tingkat kesukaran yang diperoleh dari hasil perhitungan, semakin mudah soal itu. Suatu soal memiliki TK sebesar 0 berarti tidak ada siswa yang mampu menjawab benar soal tersebut. Bila suatu soal memiliki TK sebesar 1 maka soal itu dipastikan dapat dijawab benar oleh semua siswa.

Perhitungan indeks tingkat kesukaran dilakukan untuk setiap soal. Rumus untuk menghitung tingkat kesukaran soal bentuk pilihan ganda menurut Nitko (1996) adalah:

$$\text{Tingkat Kesukaran (TK)} = \frac{\text{Jumlah siswa yang menjawab benar butir soal}}{\text{Jumlah siswa yang mengikuti tes}}$$

Sementara, untuk menghitung tingkat kesukaran soal uraian digunakan rumus berikut.

$$\text{Tingkat Kesukaran (TK)} = \frac{\text{Mean}}{\text{Skor maksimum yang ditetapkan}}$$

Tabel 6.1
Kriteria Tingkat Kesulitan

2) Daya Pembeda (DP)

Kedua, mengetahui seberapa jauh masing-masing soal dapat mendeteksi atau membedakan kemampuan siswa, yaitu siswa yang pandai atau kurang. Apabila suatu soal tidak dapat membedakan kedua kemampuan siswa itu, maka soal itu kemungkinan: (a) memiliki kunci jawab tidak tepat; (b) memiliki dua atau lebih kunci jawaban yang benar; (c) kompetensi yang diukur tidak jelas; (d) pengecoh tidak berfungsi; (e) materi yang ditanyakan terlalu sulit sehingga banyak siswa yang menebak; dan (f) sebagian besar siswa pandai berpikir ada yang salah informasi dalam soalnya.

“Semakin tinggi indeks daya pembeda soal, semakin tinggi kemampuan soal itu membedakan siswa yang pandai dengan siswa yang kurang.”

Untuk mengetahui daya pembeda soal bentuk pilihan ganda dapat digunakan rumus berikut.

$$DP = \frac{BA - BB}{\frac{1}{2} N} \text{ atau } DP = \frac{2(BA - BB)}{N}$$

N = jumlah siswa yang mengerjakan tes.

$$DP = \frac{\text{Mean kelompok skor atas} - \text{Mean kelompok bawah}}{\text{Skor maksimum soal}}$$

Hasil perhitungan dengan menggunakan rumus di atas dapat menggambarkan tingkat kemampuan soal dalam membedakan antarpeserta tes. Kriteria yang dapat digunakan untuk melihat daya pembeda soal dapat mengacu pada Crocker & Algina (1986) berikut.

Tabel 6.2
Kriteria Daya Pembeda

No.	Range Daya Penilaian	Kategori	Keputusan
1.	0,40 – 1,00	Sangat memuaskan	Diterima
2.	0,30 – 0,39	Memuaskan	Diterima
3.	0,20 – 0,29	Tidak memuaskan	Ditolak/direvisi
4.	0,00 – 0,19	Sangat tidak memuaskan	Direvisi total

3) Distribusi Pilihan Jawaban

Penyebaran pilihan jawaban (pengecoh) juga dijadikan dasar dalam penelaahan soal. Hal ini dimaksudkan untuk mengetahui berfungsi tidaknya pilihan jawaban yang tersedia. Suatu pilihan jawaban dapat dikatakan berfungsi apabila (i) pengecoh paling tidak dipilih oleh 5% siswa, dan (ii) pengecoh lebih banyak dipilih oleh kelompok siswa yang kurang pandai.

C. Analisis Kuantitatif dengan Bantuan Komputer

Analisis dengan menggunakan bantuan komputer sangat tepat, karena tingkat keakuratan lebih tinggi dibandingkan dengan diolah secara manual atau menggunakan kalkulator dan tangan. Terdapat banyak program komputer yang dapat digunakan untuk menganalisis soal, tergantung tujuan dan maksud analisis. Program yang sudah dikenal secara umum adalah EXCEL atau program khusus seperti ITEMAN, BIGSTEPS, RASCAL, ASCAL, BILOG, dan FACETS. Pada buku ini tidak semua program dibahas, hanya

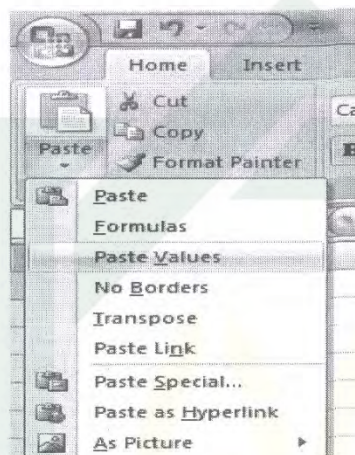
c. Proporsi Jawaban Benar

Jumlah jawaban benar dan proporsi jawaban benar untuk soal nomor 2 sampai dengan nomor 10 dapat diperoleh dengan cara menyalin pada kolom yang sesuai. Dengan demikian, diperoleh hasil sebagai berikut:

112

Untuk menghitung daya pembeda, langkah pertama adalah menyalin isi sel N1 sampai dengan Z24 pada sheet yang baru. Caranya, Sorot sel N1 sampai dengan Z24, kemudian klik kanan → **Copy**. Pindahkan kursor pada sel A1 dalam sheet baru, kemudian klik kanan → **Paste Special** → **Paste Values**.

Untuk menghitung daya pembeda, langkah pertama adalah menyalin isi sel N1 sampai dengan Z24 pada sheet yang baru. Caranya, Sorot sel N1 sampai dengan Z24, kemudian klik kanan → **Copy**. Pindahkan kursor pada sel A1 dalam sheet baru, kemudian klik kanan → **Paste Special** → **Paste Values**.



Dengan demikian, pada sheet baru tersebut diperoleh:

	A	B	C	D	E	F	G	H	I	J	K	L	M
1			1	2	3	4	5	6	7	8	9	10	Jumlah
2			Skor per Item										Skor
3	Siswa-01		1	1	1	0	1	0	0	1	0	1	6
4	Siswa-02		0	0	1	1	0	0	0	1	0	0	3
5	Siswa-03		0	1	1	0	1	0	0	1	1	1	6
6	Siswa-04		1	1	1	1	1	1	1	0	1	1	9
7	Siswa-05		1	1	1	0	1	1	1	1	1	1	9
8	Siswa-06		0	0	1	0	0	0	1	1	0	1	4
9	Siswa-07		1	1	1	1	1	0	0	1	1	1	8
10	Siswa-08		0	0	0	0	0	0	0	0	0	0	0
11	Siswa-09		0	1	1	0	1	0	1	1	1	1	7
12	Siswa-10		1	0	1	1	0	0	0	0	0	1	4
13	Siswa-11		0	1	1	0	1	0	1	1	0	1	6
14	Siswa-12		0	1	1	0	1	1	1	1	1	1	8
15	Siswa-13		0	1	1	0	1	0	0	1	1	1	6
16	Siswa-14		1	1	1	1	1	1	1	0	1	1	9
17	Siswa-15		1	1	1	1	1	0	0	1	1	1	8
18	Siswa-16		1	0	1	0	1	0	1	0	1	1	6
19	Siswa-17		0	1	1	1	1	0	1	1	0	0	6
20	Siswa-18		1	1	1	0	0	0	1	1	1	1	7
21	Siswa-19		1	1	1	1	1	0	1	1	1	0	8
22	Siswa-20		0	1	1	1	0	0	1	1	1	1	7

ke skor terendah. Caranya, Sorot sel A1 sampai dengan M22, klik **Data** → **Sort** → Sort by: **Colomn M**, Order: **Largest to smallest**.



Diperoleh hasil sortir berdasarkan skor tertinggi ke terendah seperti tampilan berikut:

[illegible]

	A	B	C	D	E	F	G	H	I	J	K	L	M
1			1	2	3	4	5	6	7	8	9	10	Jumlah
2			Skor per Item										Skor
3	Siswa-04		1	1	1	1	1	1	1	0	1	1	9
4	Siswa-05		1	1	1	0	1	1	1	1	1	1	9
5	Siswa-14		1	1	1	1	1	1	1	0	1	1	9
6	Siswa-07		1	1	1	1	1	0	0	1	1	1	8
7	Siswa-12		0	1	1	0	1	1	1	1	1	1	8
8	Siswa-15		1	1	1	1	1	0	0	1	1	1	8
9	Siswa-19		1	1	1	1	1	0	1	1	1	0	8
10	Siswa-09		0	1	1	0	1	0	1	1	1	1	7
11	Siswa-18		1	1	1	0	0	0	1	1	1	1	7
12	Siswa-20		0	1	1	1	0	0	1	1	1	1	7
13	Siswa-01		1	1	1	0	1	0	0	1	0	1	6
14	Siswa-03		0	1	1	0	1	0	0	1	1	1	6
15	Siswa-11		0	1	1	0	1	0	1	1	0	1	6
16	Siswa-13		0	1	1	0	1	0	0	1	1	1	6
17	Siswa-16		1	0	1	0	1	0	1	0	1	1	6
18	Siswa-17		0	1	1	1	1	0	1	1	0	0	6
19	Siswa-06		0	0	1	0	0	0	1	1	0	1	4
20	Siswa-10		1	0	1	1	0	0	0	0	0	1	4
21	Siswa-02		0	0	1	1	0	0	0	1	0	0	3
22	Siswa-08		0	0	0	0	0	0	0	0	0	0	0
23	Atas	⬆	7	10	10	6	8	4	8	8	10	9	
24	Bawah		3	5	9	3	6	0	4	7	3	7	

115

dengan cara menyalin formula tersebut pada sel yang sesuai. Dengan demikian, diperoleh:

[illegible]

e. Penyebaran Jawaban

Untuk menganalisis penyebaran jawaban, perlu dihitung jumlah setiap pilihan pada setiap butir soal. Penghitungan ini menggunakan *sheet* pertama yang berisi jawaban sebelum diskor. Untuk menghitung pilihan jawaban A pada nomor 1 digunakan **=COUNTIF(C3:C22,"A")** yang ditempatkan pada sel C23. Untuk menghitung pilihan jawaban B pada nomor 1 digunakan **=COUNTIF(C3:C22,"B")** yang ditempatkan pada sel C24. Demikian seterusnya sampai dengan pilihan jawaban E. Cara yang sama dilakukan pada butir soal ke-2 sampai dengan ke-10, sehingga diperoleh:

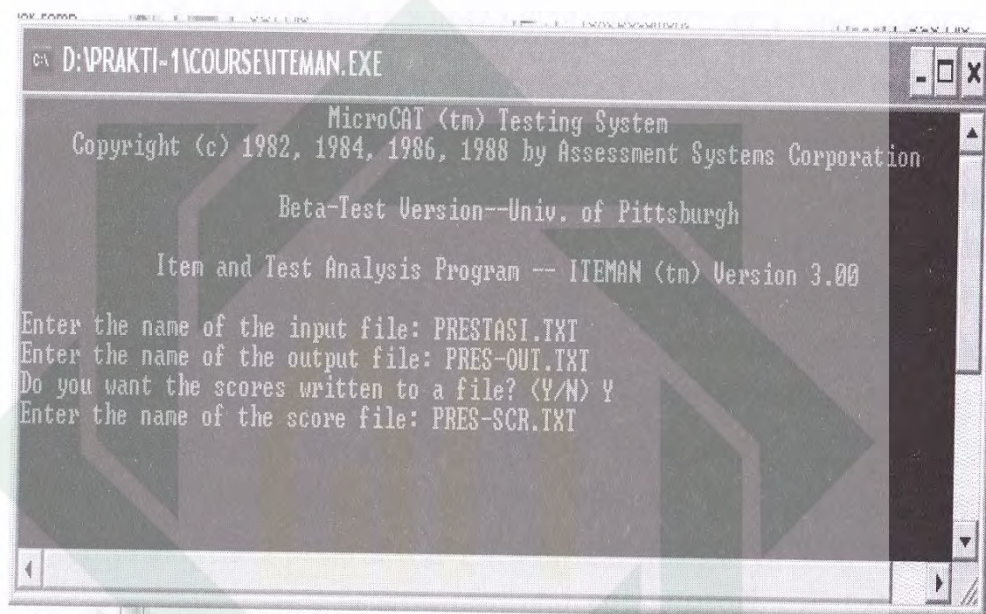
[illegible]

5. *Point Biser* adalah indeks daya pembeda soal dan pilihan jawaban (alternatif) dengan menggunakan koefisien korelasi point biserial. Penafsirannya sama dengan statistik biserial. Statistik pilihan jawaban (alternatif) memberikan informasi yang sama dengan statistik butir soal. Perbedaannya adalah statistik pilihan jawaban dihitung secara terpisah, untuk setiap pilihan jawaban dan didasarkan pada dipilih tidaknya alternatif tersebut, bukan pada benarnya jawaban. Tanda bintang (*) yang muncul di sebelah kanan hasil analisis menunjukkan kunci jawaban.
6. *Prop. Endorsing* menunjukkan proporsi alternatif jawaban yang dijawab oleh peserta tes. Diharapkan harganya lebih dari 0,05 agar tiap distraktor alternatif jawaban mempunyai fungsi yang baik (Mardapi, 2008).

Sebuah tes prestasi belajar terdiri atas 30 butir soal, masing-masing memiliki 5 pilihan jawaban, diujikan pada 300 peserta tes. Misalkan kunci jawabannya adalah: CDCDEDAABCBBAAABCAEEBAACAADDAC. Berdasarkan hasil ujicoba, dilakukan analisis dengan menggunakan program ITEMAN.

Buat file data (misalkan: PRESTASI.TXT) yang berisi:

Perintah di atas, bila dilihat di layar komputer akan tampak seperti di bawah ini.



Hasil analisis menunjukkan bahwa soal nomor 1 memiliki tingkat kesulitan 0,61, indeks daya pembeda yang dihitung dengan menggunakan biserial sebesar 0,458 dan point biserial sebesar 0,361. Dengan demikian, soal ini dilihat dari tingkat kesukaran dan daya pembeda tergolong soal yang sedang, baik, dan dapat diterima. Selanjutnya, pengecoh (*distractor*) A dipilih sebanyak 5,3% peserta tes, B dipilih 4,3% peserta tes, D dipilih 15% peserta tes, dan E dipilih 14,3. Bila dicermati, dari keempat pengecoh yang disediakan, pengecoh B tampaknya perlu ditinjau kembali karena hanya dipilih oleh 4,3% peserta tes. Artinya, pengecoh B perlu direvisi

MicroCAT (tm) Testing System									
Copyright (c) 1982, 1984, 1986, 1988 by Assessment Systems Corporation									
Item and Test Analysis Program -- ITEMAN (tm) Version 3.00									
Item analysis for data from file PRESTASI.TXT					Page 1				
Item Statistics					Alternative Statistics				
Seq. No.	Scale -Item	Prop. Correct	Biser. Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser. Biser.	Point Biser.	Key
1	0-1	0.610	0.458	0.361	A	0.053	-0.431	-0.208	
					B	0.043	-0.193	-0.087	
					C	0.610	0.458	0.361	*
					D	0.150	-0.190	-0.124	
					E	0.143	-0.296	-0.191	
					Other	0.000	-9.000	-9.000	
2	0-2	0.737	0.642	0.476	A	0.017	-0.122	-0.040	
					B	0.030	-0.365	-0.145	
					C	0.110	-0.548	-0.329	
					D	0.737	0.642	0.476	*
					E	0.107	-0.418	-0.249	
					Other	0.000	-9.000	-9.000	

Skala statistik merupakan gambaran umum mengenai mutu soal yang telah diujikan. Ringkasan statistik dari hasil analisis butir soal disajikan seperti gambar di bawah ini.

121

D. LATIHAN

- | | A | B | C* | D |
|----------------|------|------|------|------|
| Kelompok Atas | 0,25 | 0,35 | 0,40 | 0,00 |
| Kelompok Bawah | 0,40 | 0,10 | 0,50 | 0,00 |
| Keseluruhan | 0,30 | 0,20 | 0,45 | 0,05 |
-
- | | A* | B | C | D |
|----------------|------|------|------|------|
| Kelompok Atas | 0,75 | 0,05 | 0,15 | 0,05 |
| Kelompok Bawah | 0,73 | 0,00 | 0,20 | 0,07 |
| Keseluruhan | 0,74 | 0,02 | 0,18 | 0,06 |

122

Sesi Murni A	Pilihan Jawaban			
	A	B	C	D*
Kelompok Atas	0,02	0,00	0,05	0,93
Kelompok Bawah	0,10	0,29	0,45	0,16
Keseluruhan	0,05	0,15	0,30	0,15

2. Carilah dokumen penilaian yang dikembangkan guru (bisa ulangan harian, UAS, ataupun UNAS). Analisislah tingkat kesukaran, daya pembeda, dan pola distribusi pilihan jawaban dari dokumen penilaian tersebut. Apa yang dapat Saudara sarankan untuk membantu guru tersebut agar instrumen penilaiannya menjadi lebih baik.
3. Ibu Salamah mengajar matematika pada salah satu sekolah di Mojokerto. Ia memiliki pengalaman mengajar kelas VIII lebih 10 tahun. Beberapa kali, ia merevisi tes yang digunakan. Soal-soal yang digunakan diambilkan dari buku. Tahun ini, Bu Salamah mencoba membuat tes sendiri. Tes yang dibuat terdiri dari soal pilihan ganda dan uraian. Namun, para siswanya banyak yang mengalami masalah dengan hasil tesnya, karena banyak di antara mereka mendapat nilai rendah termasuk siswa yang dianggap pandai. Ini berdasarkan hasil tes sebelumnya. Padahal, berdasarkan pengamatan Bu Salamah, siswa-siswinya antusias berpartisipasi dalam diskusi kelas dan tampak memahami materi yang diajarkan.

Ketika Bu Salamah menggunakan berbagai macam penilaian yang digunakan di kelas, para siswa tampak memahami materi. Namun, mereka memiliki hasil yang rendah pada bab-bab yang diujikan. Bu Salamah akhirnya curiga dengan soal yang diujikan, dan dianggap sebagai sumber masalah. Dapatkah Saudara memberi saran pada Bu Salamah terkait langkah-langkah yang harus diambil untuk mengevaluasi tes dan soal yang diujikan?