

ANALISIS BUTIR SOAL PILIHAN GANDA

Oleh : Sridadi

Dosen Jurusan Pendidikan Olahraga
FIK UNY

Abstrak

Dari serangkaian tugas dan kewajiban guru sebagai pengajar, analisis butir soal merupakan satu langkah penting yang harus dilakukan guru setelah guru memberikan evaluasi tes hasil belajar kepada peserta didik khususnya pada teori pendidikan jasmani dan kesehatan. Dengan analisis tersebut, guru akan dapat mengetahui kelebihan dan kelemahan kualitas soal yang dibuatnya sendiri. Setidak-tidaknya guru: (1) dapat mengetahui taraf kesukaran soal dengan memfungsikan beberapa *distractor*, (2) dapat menentukan daya pembeda soal yang dapat membedakan antara kelompok peserta didik yang pandai dan yang kurang pandai, (3) dapat mengetahui penyebaran pola jawaban soal baik sebagai kunci jawaban maupun sebagai pengecoh, serta (4) kondisi tiap butir soal terhadap skor keseluruhan.

Manfaat lain yang diperoleh dari analisis butir soal, guru dapat mengetahui gambaran kemampuan peserta didik, serta kelebihan dan kelemahan metode mengajarnya sehingga dapat melakukan langkah-langkah positif terhadap perbaikan proses belajar mengajar di waktu-waktu yang akan datang. Dengan demikian jika langkah ini dapat dilaksanakan secara berkala dan berkesinambungan dengan melakukan evaluasi dan perbaikan sepanjang waktu, maka kualitas soal yang dibuatnya makin lama akan menjadi semakin sempurna.

Kata-kata kunci : analisis, dan butir soal

Keberhasilan suatu proses belajar mengajar (PBM) teori pendidikan jasmani dan kesehatan hanya dapat diketahui apabila guru telah melakukan evaluasi tes hasil belajar. Sehingga dengan evaluasi tes hasil belajar ini guru akan dapat mengetahui beberapa kelebihan dan kelemahan komponen-komponen yang terkait pada kegiatan belajar mengajar, diantaranya mengenai keadaan peserta didik, materi pelajaran yang diberikan, dan cara mengajar guru.

Kegiatan evaluasi tidak hanya berhenti setelah guru melakukan evaluasi hasil belajar yang berakhir dengan pemberian skor/penilaian tes

hasil belajar kepada peserta didik. Namun hendaknya dari hasil evaluasi tersebut guru perlu mengolah atau menganalisis komponen-komponen mana saja dari materi yang diberikan dirasa masih lemah baik di dalam penyampaian bahan oleh guru maupun penguasaan materi pelajaran oleh peserta didik.

Menurut Purwanto (1994: 118) pengolahan tes hasil belajar untuk mengetahui kelemahan dan kekurangan proses belajar mengajar dapat dilakukan dengan dua cara, yaitu: pertama dengan analisis soal, dan kedua, dengan menghitung validitas serta reliabilitas tes. Namun untuk masalah ini hanya akan dibahas pada butir pertama yaitu analisis soal.

Menurut Thorndike dan Hagen (1977) analisis terhadap soal-soal tes yang telah dijawab peserta didik mempunyai dua tujuan: (1) Jawaban soal-soal itu merupakan informasi diagnostik untuk meneliti pelajaran dari kelas itu dan kegagalan-kegagalan belajarnya serta selanjutnya untuk membimbing ke arah cara belajar yang lebih baik. (2) Jawaban terhadap soal-soal yang terpisah dan perbaikan (*review*) soal-soal yang didasarkan atas jawaban-jawaban itu merupakan langkah awal bagi penyiapan tes-tes yang lebih baik untuk tahun berikutnya.

Mengapa guru perlu untuk menganalisis setiap butir soal? Tujuan dari analisis butir soal adalah ingin mengetahui butir soal mana yang baik dan butir mana yang kurang baik, selanjutnya guru dapat mencari kemungkinan sebab-sebab mengapa butir soal tersebut kurang baik. Menurut Silverius (1991: 15) baik buruknya (mutu) butir soal ditetapkan dengan melihat taraf kesukarannya, fungsi *stem* (pokok soal), fungsi *distractor* (pengecoh), serta penyebaran jawaban pada pengecoh dalam total kelompok.

Menurut Purwanto (1994) dengan membuat analisis soal, sedikitnya guru dapat mengetahui tiga hal penting yang dapat diperoleh dari tiap soal. Pertama, sampai dimana tingkat atau taraf kesukaran soal. Kedua, apakah soal tersebut mempunyai daya pembeda (*discriminating power*) sehingga dapat membedakan kelompok siswa yang pandai dengan kelompok siswa yang bodoh (kurang). Ketiga, apakah semua alternatif jawaban (*option*) menarik jawaban-jawaban, ataukah ada yang demikian tidak menarik sehingga tidak perlu dimasukkan ke dalam soal.

TARAF KESUKARAN (TK) SOAL

Dalam setiap soal pilihan ganda selalu digunakan alternatif jawaban yang mengandung dua kemungkinan, yaitu jawaban yang tepat/benar dan jawaban yang salah sebagai penyesat/pengecoh (*distractor*). Tujuan penggunaan pengecoh ini adalah untuk menyesatkan mereka yang kurang pandai agar dapat dibedakan dengan yang pandai. Oleh karena itu pengecoh yang baik adalah yang dapat dihindari oleh peserta didik yang pandai dan terpilih oleh peserta didik yang kurang pandai, jangan sampai terjadi sebaliknya.

Butir soal dikatakan memiliki tingkat kesukaran yang tinggi apabila *distractor* (pengecoh) dapat berfungsi dengan baik, sehingga alternatif jawaban yang benar dapat dijawab tepat oleh beberapa orang pada kelompok peserta didik yang pandai. Sedangkan kelompok peserta didik yang bodoh maupun kelompok tengah sebagian besar terkecoh untuk memilih dan menyebar pada beberapa *distractor* yang ada. Jika pengecoh berfungsi, maka jawaban soal yang benar dan jawaban soal yang salah hampir tidak dapat dibedakan satu sama lain. Hanya kelompok peserta didik yang pandai saja yang pada umumnya dapat mengetahui alternatif jawaban yang paling tepat.

Namun untuk mengetahui bahwa soal tersebut memiliki taraf kesukaran yang tinggi, maka guru perlu menganalisis setiap butir soal yang ada. Menurut Suharsimi (1995: 209) analisis soal merupakan prosedur sistematis yang akan memberikan informasi-informasi yang sangat khusus terhadap butir tes yang akan disusun. Untuk dapat menganalisis tingkat kesukaran setiap butir soal, menurut Purwanto (1994) langkah awal yang harus dilakukan guru adalah mengelompokkan hasil tes menjadi tiga kelompok berdasarkan peringkat dari keseluruhan skor atau nilai yang diperoleh masing-masing peserta didik. Ketiga kelompok yang dimaksud adalah:

1. Kelompok peserta didik yang pandai (*Upper Group*) yang diambil 25% atau 27% dari jumlah peserta didik yang berada pada peringkat paling atas berdasarkan skor nilai yang diperoleh.
2. Kelompok peserta didik yang bodoh/kurang (*Lower Group*), yang diambil 25% atau 27% dari jumlah peserta didik yang berada pada peringkat paling bawah berdasarkan skor nilai yang diperoleh.

3. Kelompok peserta didik yang berada di tengah (*Middle Group*), yaitu selain kelompok pandai maupun kelompok bodoh.

Dari ketiga kelompok yang ada, dua kelompok (*Upper dan Lower*) dapat digunakan untuk membantu menganalisis setiap butir soal. Untuk dapat mengetahui taraf kesukaran setiap butir soal, dapat digunakan rumus sbb:

$$TK = \frac{U + L}{T}$$

Keterangan :

- TK : Indeks taraf kesukaran yang dicari
U : Jumlah peserta didik yang termasuk kelompok pandai (*Upper Group*) yang menjawab benar untuk tiap soal.
L : Jumlah peserta didik yang termasuk kelompok bodoh/kurang (*Lower Group*) yang menjawab benar untuk tiap soal.
T : Jumlah peserta didik dari kelompok pandai dan kelompok kurang/bodoh

Misalnya suatu tes yang terdiri dari beberapa soal diberikan kepada 40 orang peserta didik. Dari hasil tes tersebut guru dapat menganalisis taraf kesukaran setiap butir soal yang ada. Awalnya hasil tes yang berupa skor/nilai itu disusun berdasarkan peringkat dari tertinggi hingga terendah. Langkah selanjutnya adalah mengambil 25% (= 10 lembar jawaban) dari kelompok pandai (atas), dan 25% (= 10 lembar jawaban) dari kelompok kurang (bawah).

Kemudian dari masing-masing soal pada lembar jawaban peserta didik yang diperoleh itu ditabulasikan. Misalnya dari hasil tabulasi soal no.1 diperoleh hasil sebagai berikut: yang menjawab benar pada soal nomor satu dari kelompok pandai = 9 orang, dan yang menjawab benar dari kelompok kurang = 4 orang. Dengan menggunakan rumus tersebut di atas, maka tingkat kesulitan (TK) dari soal no.1 adalah:

$$TK-1 = \frac{U+L}{T} = \frac{9+4}{20} = 0,65 \text{ atau } 65\%$$

Untuk menafsirkan baik buruknya satu butir soal berdasarkan hasil penghitungan tingkat kesulitan, maka sebelumnya perlu mengetahui daya pembeda butir soal. Namun besar kecilnya bilangan yang ditunjukkan dari hasil perhitungan taraf kesukaran setiap butir soal merupakan indeks kesukaran yang menggambarkan sulit/sukar dan mudahnya suatu soal. Dan besarnya indeks kesukaran ditunjukkan dengan bilangan dari 0,00 sampai dengan 1,0. Artinya semakin kecil bilangan yang dihasilkan dari penghitungan rumus di atas, menunjukkan bahwa taraf kesukaran butir soal semakin tinggi. Sebaliknya, jika bilangan yang dihasilkan semakin besar, maka taraf kesukaran butir soal semakin mudah.

Sukar 0,0 ----- 1,0 Mudah

Contoh Format Tabulasi Jawaban Soal Kelompok *Upper* dan *Lower*
Mata Pelajaran : Pendidikan Jasmani dan Kesehatan
Kelas II, Semester I,
SMU Ma'arif 2 Sleman

Kelompok Upper & Lower	NOMOR SOAL DAN JAWABAN KELOMPOK UPPER/LOWER																				Jumlah				
	1	2	3	4	5	6	7	8	9	10	1	2	3	4	5	6	7	8	9	10					
U	L	U	L	U	L	U	L	U	L	U	L	U	L	U	L	U	L	U	L	U	L	U	L	U	L
U1	L1	1			1	1	1	1		1		1	1	1		1	1	1		1		1		8	4
U2	L2	1		1	1			1	1	1		1	1	1		1	1	1		1		1		8	5
U3	L3	1	1	1	1	1		1	1	1		1	1	1	1		1	1	1		1		9	5	
U4	L4	1		1	1			1	1	1	1		1	1	1		1	1	1		1		8	5	
U5	L5	1	1	1			1			1		1	1	1	1		1	1	1		1		8	3	
U6	L6	1		1	1			1	1	1		1	1	1	1		1	1	1		1		8	5	
U7	L7	1		1	1			1	1	1	1		1		1		1	1	1		1		7	4	
U8	L8	1	1	1	1	1		1	1	1		1		1		1	1	1		1		9	4		
U9	L9	1		1	1			1				1	1	1	1		1		1		1		7	5	
U10	L10	1	1	1	1	1		1	1	1	1		1		1	1	1	1	1		1		9	6	

Daya Pembeda (DP) Soal

Menurut Purwanto (1994: 120) dan Suharsimi (1991: 215) yang dimaksud dengan daya pembeda suatu soal tes adalah bagaimana kemampuan soal itu untuk membedakan siswa-siswa yang termasuk kelompok pandai (*upper group*) dengan siswa-siswa yang termasuk kelompok kurang (*lower group*). Daya pembeda suatu soal dapat dihitung dengan menggunakan rumus (Purwanto, 1994: 120, Thoha, 1994: 147) sebagai berikut:

$$DP = \frac{U - L}{\frac{1}{2} T}$$

Keterangan :

- DP : Indeks Daya Pembeda soal yang dicari
U, L, dan T : sama dengan keterangan yang diberikan pada rumus "tingkat kesulitan"

Sedang angka yang diperoleh dari hasil perhitungan menggunakan rumus di atas menunjukkan daya pembeda soal atau indeks diskriminasi soal, yang besarnya antara -1,00 sampai dengan +1,00.

-1,00 ----- 0,00 ----- +1,00
daya pembeda negatif daya pembeda rendah daya pembeda tinggi

(Suharsimi, 1991: 215)

Perhitungan daya pembeda dapat mencapai angka +1,00 apabila seluruh kelompok pada *Upper Group* dapat menjawab benar, dan seluruh kelompok pada *Lower Group* menjawab salah. Artinya, bahwa soal tersebut memiliki indeks diskriminasi yang tinggi, yaitu dapat membedakan antara kelompok yang pandai dengan kelompok yang bodoh.

Selanjutnya perhitungan daya pembeda dapat memiliki koefisien 0,00, apabila kelompok pandai maupun kelompok kurang sama-sama menjawab benar atau sama-sama menjawab salah. Artinya soal tersebut tidak memiliki daya pembeda, sehingga tidak dapat membedakan antara

kelompok atas dengan kelompok bawah. Sebaiknya soal tersebut direvisi atau diganti.

Sebaliknya apabila seluruh kelompok pandai tidak dapat menjawab benar, namun seluruh kelompok bodoh justru dapat menjawab benar, maka akan diperoleh hasil perhitungan indeks diskriminasi soal = -1,00. Artinya perbedaan yang ditunjukkan menunjukkan koefisien negatif sehingga kelompok pandai dapat dikatakan bodoh sedangkan kelompok bodoh dapat dikatakan pandai. Soal seperti ini sebaiknya diganti.

Contoh hasil penghitungan Taraf Kesukaran dan Daya Pembeda setiap butir soal pada format Tabulasi Soal di atas.

No. Soal	U	L	TK	DP	KETERANGAN
1	9	4	$\frac{9+4}{20} = 0,65 = 65\%$	$\frac{9-4}{10} = 0,5 = 50\%$	-
2	9	6	$\frac{9+6}{20} = 0,75 = 75\%$	$\frac{9-6}{10} = 0,3 = 30\%$	-
3	4	2	$\frac{4+2}{20} = 0,3 = 30\%$	$\frac{4-2}{10} = 0,2 = 20\%$	Revisi
4	7	8	$\frac{7+8}{20} = 0,75 = 75\%$	$\frac{7-8}{10} = -0,1 = -10\%$	Revisi
5	9	3	$\frac{9+3}{20} = 0,6 = 60\%$	$\frac{9-3}{10} = 0,6 = 60\%$	-
6	8	4	$\frac{8+6}{20} = 0,7 = 70\%$	$\frac{8-6}{10} = 0,2 = 20\%$	Revisi
7	9	5	$\frac{9+5}{20} = 0,7 = 70\%$	$\frac{9-5}{10} = 0,4 = 40\%$	-
8	9	5	$\frac{9+5}{20} = 0,7 = 70\%$	$\frac{9-5}{10} = 0,4 = 40\%$	-
9	9	2	$\frac{9+2}{20} = 0,55 = 55\%$	$\frac{9-2}{10} = 0,7 = 70\%$	-
10	8	2	$\frac{8+2}{20} = 0,5 = 50\%$	$\frac{8-2}{10} = 0,6 = 60\%$	-

KRITERIA MENENTUKAN KUALITAS SOAL

Soal yang dibuat untuk melakukan evaluasi tes hasil belajar mengajar, umumnya sebelum diberikan kepada peserta didik sebaiknya dilakukan uji coba. Manfaat uji coba disini tidak lain untuk mengetahui apakah soal yang akan diberikan ini benar-benar sudah baik atau belum. Dari uji coba tersebut guru dapat melakukan analisis setiap butir soal yang ada, sehingga dapat diketahui mana soal yang baik dan mana soal yang jelek sehingga perlu direvisi. Beberapa ketentuan dalam menentukan baik buruknya soal diantaranya :

1. Butir soal yang baik apabila memiliki indeks diskriminasi (DP) antara 0,4 – 0,7
2. Klasifikasi Daya Pembeda Soal: (Nitko: 1983 dalam Suharsimi, 1995:223)
 - a. 0,00 – 0,20 : *poor* (jelek)
 - b. 0,20 – 0,40 : *satisfactory* (cukup)
 - c. 0,40 – 0,70 : *good* (baik)
 - d. 0,70 – 1,00 : *excellent* (baik sekali)
 - e. - 1,00 semuanya tidak baik
3. Untuk pilihan ganda dengan *option* 4, jika tingkat kesulitan di bawah 0,30 dikategorikan soal yang sukar. Sedang jika tingkat kesulitan soal sama atau lebih besar dari 0,70 dikategorikan soal yang mudah. (Siverius, 1991:169)
4. Jika daya pembeda soal = 0 atau = -1 (negatif) maka soal perlu direvisi atau diganti (Purwanto, 1994: 132)

POLA JAWABAN DAN ANALISIS SOAL

Pola jawaban soal diperoleh dengan menghitung berapa peserta didik yang memilih *option* A, berapa yang memilih *option* B, berapa yang memilih *option* C, dan berapa yang memilih *option* D. Dari hasil penghitungan itu akan menunjukkan sejauh mana suatu *option* (pilihan) dapat berfungsi, baik sebagai kunci jawaban maupun sebagai pengecoh (*distractor*).

Apabila penyebaran jawaban peserta didik telah menunjukkan fungsinya beberapa pengecoh yang ada, dan dapat menunjukkan perbedaan antara kelompok pandai dan kurang pandai dengan melihat

pada kunci jawaban, dimana kelompok pandai memilih kunci jawaban lebih banyak dari pada kelompok yang kurang pandai, maka butir soal semacam itu dapat dikatakan sudah baik.

Adakalanya peserta didik lebih banyak memilih kunci jawaban, tetapi hanya sedikit yang memilih pengecoh sehingga terdapat kemungkinan penyebaran jawaban atas pengecoh tidak seimbang dan bahkan ada yang tidak memilih pengecoh sama sekali, maka hal ini perlu dicari penyebabnya. Mengapa sebagian dari mereka menghindari pengecoh?. Apakah pengecoh itu tidak menarik perhatian peserta didik?. Jika kurang menarik perhatian peserta didik, maka butir soal semacam itu perlu ditelaah kembali, bila masih memungkinkan diadakan perbaikan beberapa pengecoh yang tidak berfungsi. Selanjutnya apabila dirasa terlalu sukar untuk melakukan perbaikan atas soal tersebut, maka sebaiknya dihilangkan saja.

Terdapat kemungkinan terjadi banyak peserta didik memilih salah satu pengecoh dan hanya sedikit yang memilih kunci jawaban. Jika seperti ini terjadi, maka terdapat kemungkinan salah meletakkan kunci jawaban, atau pengecoh tersebut sesungguhnya adalah kunci jawaban. Namun mungkin pula kuncinya sudah benar, tetapi pengecoh terlalu menarik untuk dipilih. Jika demikian perlu pertimbangan untuk mengubah pengecoh tersebut, atau butir tersebut dibuang saja.

Suatu cara sederhana untuk menyiapkan pencatatan jawaban tiap butir soal dalam menganalisis dapat dibuat dalam bentuk tabel seperti contoh di bawah ini.

No. Soal & kunci Jawaban	Kelompok	Alternatif Jawaban				Jumlah Jawaban	Tidak Menjawab
		A	B	C	D		
1 (B)	Upper	1	9*	-	-	10	0
	Lower	3	4*	2	1	10	0
2 (C)	Upper	-	-	9*	1	10	0
	Lower	1	-	6*	3	10	0
3 (A)	Upper	4*	-	-	6	10	0
	Lower	2*	1	3	4	10	0
4 (D) dst.	Upper	-	2	-	7*	9	1
	Lower	-	2	-	8*	10	0

Keterangan : Yang diberi tanda * adalah kunci jawaban

Dari contoh penyebaran pola jawaban 4 soal seperti yang tercantum dalam tabel di atas, dapat diambil suatu gambaran bahwa untuk soal no. 1 dapat dikatakan cukup baik karena penyebaran peserta didik dalam memilih *option* yang ada cukup merata. Lebih-lebih kalau dilihat pada kunci jawaban, tampak ada perbedaan jumlah antara kelompok peserta didik yang pandai dengan yang kurang pandai, dimana kelompok yang pandai memilih lebih banyak pilihan benar dari pada yang kurang pandai.

Soal no. 2 juga cukup baik karena banyak peserta didik yang memilih kunci jawaban. Namun demikian disini pengecoh **B** tampak tidak berfungsi karena tidak ada seorangpun yang memilihnya. Dengan demikian pengecoh ini perlu diubah/direvisi dengan yang lebih baik agar dapat berfungsi seperti dua pengecoh yang lain.

Pada soal no. 3 dapat dikatakan kurang bagus, karena hanya 6 dari 20 orang peserta didik pada kelompok pandai dan kurang pandai yang dapat menjawab dengan benar. Selebihnya yang 14 orang terkecoh dengan beberapa alternatif jawaban yang lain. Dari 14 orang tersebut sebagian besar banyak yang terkecoh pada *option D*, yang berarti soal tersebut perlu dipertanyakan kebenaran kunci jawabannya. Dalam hal ini ada dua kemungkinan yang dapat terjadi. Pertama soal tersebut mungkin bersifat *ambiguous* (mempunyai dua arti), terutama bagi peserta didik yang mengetahui banyak tentang hal yang ditanyakan. Kedua, peserta didik belum pernah sama sekali mempelajari hal-hal seperti yang ditanyakan pada soal.

Soal no. 4 dapat dikatakan kurang baik, karena penghitungan DP-nya menunjukkan angka negatif. hal ini disebabkan karena jumlah kelompok bawah (kurang pandai) yang memilih kunci jawaban ternyata lebih banyak dari pada kelompok atas (pandai). Kedua soal ini tidak memiliki daya pembeda soal, sehingga tidak dapat membedakan antara kelompok pandai dengan yang kurang pandai.

KESIMPULAN

Sebagai tindak lanjut pembuatan butir soal, analisis butir soal sangat bermanfaat bagi perbaikan penyusunan kembali butir-butir soal yang dirasa masih kurang sempurna. Apabila butir soal memiliki taraf kesukaran yang memadai, dapat membedakan antara kelompok yang

pandai dengan yang kurang pandai, dan memiliki pola jawaban yang menyebar pada semua pilihan, maka soal tersebut sebaiknya dipertahankan. Sebaliknya bila butir soal tidak dapat membedakan antara kelompok yang pandai dengan yang kurang pandai maka sebaiknya direvisi atau diganti dengan yang baru.

Selanjutnya analisis butir soal pilihan ganda dapat digunakan untuk melihat/meneliti materi-materi mana yang belum dikuasai peserta didik (lihat pada soal yang sukar) sehingga guru dapat mengulang kembali atau memperbaiki proses belajar mengajarnya. Untuk butir soal yang sukar bagi keseluruhan kelas dapat berguna untuk pembimbingan ke arah eksplorasi yang lebih luas dengan cara mendiskusikan bersama peserta didik, sehingga dapat memperluas pengetahuan dan menghilangkan salah pengertian.

Selain itu, analisis butir soal juga sangat bermanfaat bagi guru dalam mencari kelemahan-kelemahan yang ada dalam cara mengajar teori penjasokes serta keterampilan guru dalam penulisan soal.

Daftar Pustaka

- Arikunto, Suharsimi. (1995). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Nitko, A.J. (1983). *Educational Test and Measurement, An Introduction*. New York: Harcourt Brace Jovanovich. Inc.
- Purwanto, M. Ngalim. (1994). *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT. Remaja Rosdakarya.
- Silverius, Suke. (1991). *Evaluasi Hasil Belajar dan Umpan Balik*. Jakarta: PT. Gramedia Widiasarana Indonesia.
- Thoha, M. Chabib. (1994). *Teknik Evaluasi Pendidikan*. Jakarta: PT. Raja Grafindo Persada.

Thorndike, R.I., dan Hagen, E.P., (1977). *Measurement and Evaluation in Psychology and Education, fourth edition*. New York: John Wiley and Sons.