

ANALISIS KESUKARAN SOAL, DAYA PEMBEDA DAN FUNGSI DISTRAKTOR

Laela Umi Fatimah

E-mail: laelaumi93@gmail.com

Mahasiswa Magister Pendidikan Agama Islam
Universitas Islam Negeri Sunan Kalijaga Yogyakarta

Khairuddin Alfath

khairuddin475@gmail.com

Mahasiswa Magister Pendidikan Agama Islam
Universitas Islam Negeri Sunan Kalijaga Yogyakarta

Abstrak

Artikel ini ditulis karena pada dasarnya pendidikan memerlukan sebuah evaluasi. Tak hanya hasil belajar siswa saja yang dievaluasi, melainkan juga alat pengukur yang digunakan untuk mengukur keberhasilan siswa. Yang dimaksud adalah tes hasil belajar, yang mana terdiri dari butir-butir soal atau biasa disebut item. Dan untuk mengetahui apakah butir-butir tersebut sudah menjalankan fungsinya sebagaimana mestinya, maka perlu identifikasi. Dengan harapan akan menghasilkan berbagai informasi berharga, yang menjadi umpan balik guna melakukan perbaikan, pembenahan dan penyempurnaan kembali, hingga nantinya benar-benar dapat menjalankan fungsinya sebagai alat pengukur hasil belajar yang berkualitas tinggi. Rangkaian kegiatan seperti uraian tersebut sering dikenal dengan istilah analisis item. Maka dari itu perlu adanya analisis kesukaran soal, daya pembeda dan fungsi distraktor. Dan ini akan menjadi topik utama yang dipaparkan pemakalah pada artikel ini.

Kata Kunci: Analisis Kesukaran Soal, Daya Pembeda, Fungsi Distraktor

Abstract

This article written because of basic education need the evaluation. Not only student's result of study that need evaluation, but also the instrument of measure used to measure the success of student. That is a test of study result, that composed of numeral questions or we can call item. To know the function of numeral questions, we need an identification. That hope will produce much of cost information, to be feedback for better reparation, ordering and complicating, until can perform the

function to be instrument of measure study's result with high quality. The combination activity like was described familiar called item analysis. Than, necessary to analysis the difficult of question, capacity of distinguishing and the fungction of distractor. And this article will be prominent topic that researched by writer.

Keywords: Analysis Difficult of Question, Capacity of Distinguishing, Fungction of Distractor

A. PENDAHULUAN

Tidak dapat dipungkiri lagi bahwa angka-angka sebagai hasil pengukuran dan penilaian pendidikan mempunyai arti penting, yaitu berfungsi memberi kesaksian tentang orang yang telah berhasil mencapai, dan sebagai kesaksian yang diperlukan dalam banyak peristiwa penting seperti kenaikan kelas, meneruskan ke sekolah yang lebih tinggi, menyelesaikan pendidikan, bahkan dalam memperoleh pekerjaan. Demikianlah pentingnya nilai-nilai yang menjadi tanda simbolik yang harus diperjuangkan dan direbut dengan berbagai usaha, bahkan tipu daya. Dan di sini peran penting pengukuran dan penilaian dalam suatu evaluasi pendidikan diterapkan.²⁵

Berhasilnya suatu pendidikan sangat ditentukan oleh proses pembelajaran. Yang mana guru berperan penting dalam pelaksanaan evaluasi dan proses analisis dari evaluasi untuk mengetahui kekuatan dan kelemahan pembelajaran dalam rangka meningkatkan proses pembelajaran. Menurut Thorndike dan Hagen, analisis terhadap soal-soal (item) tes ada dua tujuan penting. *Pertama*, Jawaban-jawaban soal itu merupakan informasi diagnostic untuk meneliti pelajaran dari kelas dan kegagalan-kegagalan belajarnya, serta melanjutkan untuk membimbing kearah cara belajar baik. *Kedua*, Jawaban-jawaban

²⁵ Raka Joni, *Pengukuran dan Penilaian Pendidikan*, (Malang: YP2LPM, 1984), hlm. 2.

terhadap soal-soal yang terpisah dan perbaikan soal-soal yang didasarkan atas jawaban-jawaban yang basis bagi penyiapan tes-tes yang lebih baik untuk tahun berikutnya.²⁶

Yang dimaksud item atau soal-soal yang baik adalah yang tingkat kesukarannya dapat diketahui tidak terlalu sukar dan tidak terlalu mudah. Sebab tingkat kesukaran item memiliki korelasi dengan daya pembeda. Bilamana item memiliki tingkat kesukaran maksimal, maka daya pembedanya akan rendah, demikian pula bila item itu terlalu mudah juga tidak akan memiliki daya pembeda. Karenanya, sebaiknya tingkat kesukaran soal itu dipertahankan dalam batas yang mampu memberikan daya pembeda.²⁷

Analisis butir soal atau item-item dapat digunakan secara kualitatif dan kuantitatif yang masing-masing mempunyai keunggulan dan kelemahan. Oleh karenanya, teknik terbaik adalah menggabungkan keduanya.

1. Teknik Analisis Secara Kualitatif

Ada beberapa teknik yang digunakan untuk menganalisis secara kualitatif, yaitu dengan teknik moderator dan teknik panel.

Teknik moderator yaitu menganalisis dengan cara berdiskusi yang mana ada satu orang sebagai pencegah. Dengan teknik ini, setiap butir soal didiskusikan secara bersama-sama dengan beberapa ahli seperti guru yang mengajarkan materi, ahli materi, penyusun atau pengembang kurikulum.

²⁶ Ngalim Purwanto, *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*, (Bandung: Remadja Karya, 2002), hlm. 153.

²⁷ M. Chabib Thoha, MA, *Teknik Evaluasi Pendidikan*, (Jakarta: CV. Rajawali, 1991), hlm. 145.

Teknik panel yaitu satu teknik menelaah butir soal berdasarkan kaidah penulisan butir soal. kaidah itu diantaranya materi, konstruksi, bahasa atau budaya, kebenaran kunci jawaban atau pedoman penskoran. Dalam penganalisisan butir soal dengan teknik kualitatif, penggunaan format penelaahan soal akan sangat membantu dan mempermudah prosedur pelaksanaannya.

a. Teknik Analisis Secara Kuantitatif

Analisis soal dengan kuantitatif yaitu penelaahan butir soal didasarkan pada data empirik. Dimana data empirik ini di dapatkan dari soal yang sudah diujikan. Ada 2 pendekatan dalam analisis secara kuantitatif, yaitu pendekatan secara klasik dan modern.

Analisis soal secara klasik adalah proses penelaahan butir soal melalui informasi dari jawaban peserta didik tes guna meningkatkan mutu butir soal yang bersangkutan dengan menggunakan teori tes klasik. Aspek yang sangat penting untuk ditelaah secara klasik ini terlihat dari segi, tingkat kesukaran, daya pembeda butir, fungsi pengecoh pada setiap pilihan jawaban. Reliabilitas dan validitas soal.

Untuk mengetahui lebih dalam apakah soal-soal (item) yang diidentifikasi menjalankan fungsinya sebagai alat pengukur hasil belajar yang memadai, penulis akan membahas lebih dalam tentang menganalisis soal secara klasik yaitu, mengetahui taraf kesukaran soal, daya pembeda serta fungsi distraktor.

B. PEMBAHASAN

1. Analisis Kesukaran Soal

Asumsi yang digunakan untuk memperoleh kualitas soal yang baik, disamping memenuhi validitas dan reliabilitas, adalah adanya keseimbangan dari tingkat kesulitan soal tersebut. Keseimbangan yang dimaksudkan adalah adanya soal-soal yang termasuk mudah, sedang dan sukar secara proporsional. Tingkat kesukaran soal dipandang dari kesanggupan atau kemampuan siswa dalam menjawabnya, bukan dilihat dari sudut guru sebagai pembuat soal.²⁸

Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar kemampuannya.²⁹

Bermutu atau tidaknya butir-butir soal tes hasil belajar pertama-tama dapat diketahui dari derajat kesukaran atau taraf kesulitan yang dimiliki oleh masing-masing butir item tersebut. Butir-butir item dapat dikatakan baik apabila butir-butir tersebut tidak terlalu sukar dan tidak pula terlalu mudah. Maka, apabila seluruh testee tidak dapat menjawab soal dengan betul, (karena terlalu sukar) tidak dapat disebut sebagai item yang baik. pun apabila seluruh testee dapat menjawab dengan betul,

²⁸ Sulistyorini, *Evaluasi Pendidikan Dalam Meningkatkan Mutu Pendidikan*, (Yogyakarta: Teras, 2009), hlm.174.

²⁹ Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, (Yogyakarta: PT. BINA AKSARA, 1984), hlm. 159.

(karena soal terlalu mudah) juga tidak dapat dimasukkan dalam kategori yang baik.³⁰

Di dalam bukunya yang berjudul *Psychological Education*, Witherington mengatakan, bahwa sudah atau belum memadainya derajat kesukaran item tes hasil belajar dapat diketahui dari besar kecilnya angka yang melambangkan tingkat kesulitan dari item tersebut. Angka yang dapat memberikan petunjuk mengenai tingkat kesukaran item tersebut dikenal dengan istilah *difficulty index* (angka indeks kesukaran item), yang dalam dunia evaluasi hasil belajar umumnya dilambangkan dengan huruf P yaitu *Proportion*. Suatu tes tidak boleh terlalu mudah, dan juga tidak boleh terlalu sukar. Sebuah item yang terlalu mudah sehingga seluruh siswa dapat menjawab dengan benar bukanlah item yang baik. Begitupula item yang terlalu sukar sehingga tidak dapat dijawab oleh siswa juga tidak baik. Jadi, item yang baik adalah item yang mempunyai derajat kesukaran tertentu.³¹

Tingkat kesukaran soal adalah peluang untuk menjawab benar suatu soal pada suatu tingkat kemampuan atau bisa dikatakan untuk mengetahui sebuah soal itu tergolong mudah atau sukar.

a. Analisis Kesukaran Soal Pilihan Ganda

Seperti yang diketahui, bahwa jenis soal itu bermacam-macam di antaranya soal pilihan ganda dan essay. Di sini akan dijelaskan analisis kesukaran soal pada soal pilihan ganda.

³⁰ Purwanto, *Evaluasi Hasil Belajar*, (Yogyakarta: Pustaka Pelajar, 2009), hlm. 97.

³¹ Ngalim Purwanto, *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*, (Bandung: Remadja Karya, 2002), hlm. 118.

Bilangan yang menunjukkan sukar dan mudahnya suatu soal disebut *indeks kesukaran* (difficulty indeks). Besarnya indeks kesukaran antara 0,00 – 1,00. Indeks kesukaran ini menunjukkan taraf kesukaran soal, sehingga soal dengan indeks 0,00 menunjukkan bahwa soal itu terlalu sukar, sebaliknya indeks 1,0 menunjukkan bahwa soalnya terlalu mudah. Dalam evaluasi, indeks kesukaran ini diberi simbol P, singkatan dari “proporsa”. Angka indeks kesukaran item dapat diperoleh dengan rumus yang dikemukakan oleh *Du Bois*, yaitu:³²

$$P = \frac{Np}{N}$$

Di mana :

P : Proporsi atau proporsa atau angka indeks kesukaran item

N_p : Banyaknya testee yang dapat menjawab dengan betul terhadap butir item

N : Jumlah testee yang mengikuti tes hasil belajar.

Rumus lainnya adalah:

$$P = \frac{B}{JS}$$

P : Indeks Kesukaran

B : Banyaknya siswa yang menjawab soal itu dengan benar.

JS : Jumlah seluruh siswa peserta tes.

Sedangkan kriteria yang digunakan adalah, makin kecil indeks yang diperoleh, makin sulit soal tersebut.

³² Anas Sudjiono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Rajawali Press, 2016), hlm.371.

Sebaliknya, makin besar indeks yang diperoleh, makin mudah soal tersebut. Kriteria Indeks kesulitan soal ditafsirkan oleh Robert L. Thorndike dan Elizabeth Hagen sebagai berikut:³³

Besarnya P	Interpretasi
Kurang dari 0,30	Terlalu Sukar
0,30 - 0,70	Cukup (Sedang)
Lebih dari 0,70	Terlalu Mudah

Sedangkan menurut Whiterington dalam bukunya *Psychological Education* sebagai berikut:

Besarnya P	Interpretasi
Kurang dari 0,25	Terlalu Sukar
0,25 - 0,75	Cukup (Sedang)
Lebih dari 0,75	Terlalu Mudah

Contoh analisis kesukaran soal:

Siswa	Nomor Soal					Skor
	1	2	3	4	5	
A	1	1	0	1	1	4
B	0	1	1	0	1	3
C	0	1	0	0	1	2
D	1	0	1	0	0	2
E	1	0	0	0	1	2
Jumlah	3	3	2	1	4	

Ada 5 siswa yang mengikuti suatu ujian yang terdiri dari 5 soal. Dari 5 siswa tersebut terdapat 3 orang yang dapat menjawab soal nomor 1 dengan betul. Maka indeks kesukarannya adalah:

³³ Anas Sudjiono, *Pengantar Evaluasi Pendidikan...*, hlm. 373.

$$P = \frac{B}{JS} = \frac{3}{5}$$

= 0,6 (kriteria sedang)

Dari tabel yang disajikan di atas, dapat ditafsirkan bahwa:

- 1) Soal nomor 1 mempunyai taraf kesukaran $3/5 = 0,6$ (kriteria sedang)
- 2) Soal nomor 3 adalah soal dengan taraf kesukaran $2/5 = 0,4$ (kriteria sedang)
- 3) Soal nomor 4 adalah soal dengan taraf kesukaran $1/5 = 0,2$ (kriteria sukar)
- 4) Soal nomor 5 adalah soal dengan taraf kesukaran $4/5 = 0,8$ (kriteria sedang)

Perlu diketahui bahwa soal-soal yang terlalu mudah atau terlalu sukar, lalu tidak berarti tidak boleh digunakan. Hal ini tergantung dari penggunaannya. Jika dari pengikut yang banyak, kita menghendaki yang lulus hanya sedikit, kita ambil siswa yang paling top. Untuk itu lebih baik diambilkan soal-soal tes yang sukar.

Sebaliknya, jika kekurangan pengikut ujian, kita pilihkan soal-soal yang mudah. Selain itu, soal yang sukar akan menambah semangat belajar bagi siswa yang pandai, sedangkan soal-soal yang terlalu mudah, akan membangkitkan semangat siswa yang tidak pandai.

Maka dari itu, dalam menyusun suatu naskah ujian sebaiknya digunakan butir soal yang mempunyai tingkat kesukaran berimbang. Yaitu: soal berkategori

sukar sebanyak 25%, kategori sedang 50% dan kategori mudah 25%.

b. Analisis Kesukaran Soal Tes Uraian atau Essay³⁴

Tidak hanya pada soal pilihan ganda saja kita dapat menganalisis tingkat kesukaran soal, begitupun pada soal uraian atau essay yang mana indeks tingkat kesukaran ini umumnya dinyatakan dalam bentuk proporsi yang besarnya kisaran 0,00- 1,00. Yang mana jika semakin besar indeks tingkat kesukaran yang diperoleh, maka semakin mudah soal itu. Karena fungsi kesukaran soal biasanya dikaitkan dengan tujuan tes. Misalnya untuk keperluan ujian semester digunakan butir soal yang memiliki tingkat kesukaran sedang, untuk keperluan seleksi digunakan butir soal yang memiliki tingkat kesukaran tinggi, dan untuk keperluan diagnostik biasanya digunakan butir soal yang memiliki tingkat kesukaran rendah. Dan untuk mengetahui tingkat kesukaran soal bentuk uraian digunakan rumus berikut ini:

$$\text{Mean} = \frac{\text{Jumlah skor siswa peserta tes pada butir soal tertentu}}{\text{Banyak siswa yang mengikuti tes}}$$

Hasil perhitungan dengan menggunakan rumus di atas menggambarkan tingkat kesukaran soal itu. Klasifikasi tingkat kesukaran soal dapat dicontohkan seperti berikut:

0,00 – 0,30 : Soal tergolong sukar

0,31 – 0,70 : Soal tergolong sedang

0,71 – 1,00 : Soal tergolong mudah

³⁴ Daryanto, *Evaluasi Pendidikan*, (Jakarta: Rineka Cipta, 2012), hlm. 180-182.

Tingkat kesukaran butir soal dapat mempengaruhi bentuk distribusi soal tes. Untuk tes yang sangat sukar distribusinya berbentuk positif skewed, sedangkan tes yang mudah distribusinya berbentuk negatif skewed.

Tingkat kesukaran butir soal memiliki 2 kegunaan, yaitu bagi guru dan bagi pengujian dan pengajaran. Bagi guru, sebagai pengenalan konsep terhadap pembelajaran ulang dan memberi masukan kepada siswa tentang hasil belajar mereka, memperoleh informasi tentang penekanan kurikulum atau mencurigai terhadap butir soal yang bias. Adapun kegunaannya bagi pengujian dan pengajaran yaitu, sebagai pengenalan konsep yang diperlukan untuk diajarkan ulang, tanda-tanda terhadap kelebihan dan kelemahan kurikulum sekolah, memberi masukan kepada siswa, tanda-tanda kemungkinan ada soal yang bias dan merakit tes yang memiliki ketepatan data soal.

Selain kegunaannya dalam konstruksi tes, tingkat kesukaran soal sangat penting karena dapat mempengaruhi karakteristik distribusi skor yang mana mempengaruhi bentuk dan penyebaran skor tes atau jumlah soal dan korelasi antarsoal, serta berhubungan dengan reliabilitas. Yang mana menurut koefisien alfa dan KR-20, semakin tinggi korelasi antarsoal, semakin tinggi reliabilitasnya.

Dalam kaitannya dengan hasil analisis item dari segi kesukarannya, tindak lanjut yang dilakukan adalah:

- 1) Untuk butir-butir item hasil analisis termasuk kategori baik dalam arti derajat kesukarannya cukup atau sedang, sebaiknya cepat dicatat dalam bank soal. Dan dapat dikeluarkan lagi dalam tes-tes berikutnya.
- 2) Untuk butir-butir item yang termasuk kategori terlalu sukar, ada 3 kemungkinan tindak lanjut yaitu, dibuang atau didrop dan tidak akan dikeluarkan lagi dalam tes berikutnya, diteliti ulang untuk dapat diketahui penyebab banyaknya siswa yang tidak bisa menjawab item tersebut, dan bukannya sola yang sukar itu tidak ada manfaatnya. Dari soal yang sukar bisa nantinya digunakan untuk tes seleksi yang ketat, yang mana membutuhkan orang-orang yang kompeten. Maka soal yang sukar sangat dibutuhkan untuk meluluskan orang-orang yang berkemampuan.
- 3) Untuk butir-butir item yang termasuk kategori terlalu mudah, juga ada kemungkinan tindak lanjutnya yaitu, butir item tersebut dibuang atau didrop dan tidak dikeluarkan lagi pada tes-tes berikutnya, diteliti dan dilacak kenapa item soal sangat mudah sehingga semua testee dapat menjawab soal dengan benar, sehingga adanya perbaikan yang dilakukan, seperti item butir yang sukar, tidak semua item yang mudah tidak ada manfaatnya, item butir soal yang mudah dapat digunakan pada tes-tes terutama tes seleksi yang sifatnya longgar, yang mana kategori soal yang

terlalu mudah hanya sebagai formalitas saja di suatu tes.

Jadi, tidak ada salahnya jika memasukkan butir-butir item yang terlalu sukar dan terlalu mudah, sebab sewaktu-waktu butir-butir seperti itulah yang dibutuhkan.

Tetapi, angka indeks di atas bukanlah angka indeks tanpa cacat. Kelemahan utama yang terdapat pada angka indeks kesukaran rata-rata P ialah, adanya hubungan yang terbalik antara derajat kesukaran item dengan angka indeks itu sendiri. Jadi, semakin rendah angka indeks kesukaran item yang dimiliki oleh sebutir item, akan semakin tinggi derajat kesukaran item tersebut, pun sebaliknya, jika semakin tinggi angka indeks kesukaran item yang dimiliki oleh sebutir item, akan semakin rendah derajat kesukaran item tersebut. Menjadikan keduanya hubungan yang berlawanan arah. Itulah yang menyebabkan banyak orang cenderung mengatakan bahwa istilah angka indeks kesukaran item lebih tepat diganti istilah angka indeks kemudahan item.

Cara kedua untuk menghitung angka indeks kesukaran item bisa dilakukan dengan skala linier. Yang disusun dengan cara mentransformasikan nilai P menjadi nilai z, dimana perubahan itu dilakukan dengan berkonsultasi pada tabel z. Dengan langkah-langkah sebagai berikut:

- 1) Mengoreksi nilai P kotor (P_k) menjadi P nilai bersih (P_b) dengan rumus:

$$P_b = \frac{aP_k - 1}{a - 1}$$

di mana :

P_b = P bersih

P_k = P kotor

a = Alternatif atau option yang disediakan atau dipasang pada butir item yang bersangkutan

1 = Bilangan konstan

- 2) Mentransformasikan nilai P bersih (P_b) menjadi nilai z. Dengan berkonsultasi pada tabel kurva normal. Selain rumus di atas, dikemukakan juga rumus lain untuk mencari P bersih dengan hasil yang sama, yaitu:

$$P_b = \frac{B - \frac{S}{a - 1}}{B + S}$$

Di mana :

P_b : Angka indeks kesukaran item (setelah dikoreksi).

B : Jumlah testee yang jawabannya betul.

S : Jumlah testee yang jawabannya salah.

a : Alternatif jawaban yang dipasang pada item yang bersangkutan.

Contoh :

Misalkan butir item nomer 5 yang telah diajukan dalam contoh di atas adalah butir item yang dijawab betul oleh 4 orang testee (B= 4). Karena jumlah testee 5 orang, dengan sendirinya dapat diketahui jumlah testee yang jawabannya salah yaitu = 5-4=1

orang ($S=1$). Banyaknya alternatif = 5. Dengan demikian:

$$\begin{aligned}P_b &= \frac{4 - \frac{1}{5-1}}{4+1} \\ &= \frac{4 - 0,25}{5} \\ &= 3,75/5 \\ &= 0,75\end{aligned}$$

- 3) Mencari atau menghitung angka indeks kesukaran item dengan menggunakan angka indeks Davis, yang diberi lambang D dapat diperoleh dengan menggunakan rumus :

$$D = 21,063 z + 50$$

Keuntungan dari rumus ini ialah bahwa kita akan terhindar dari tanda negatif atau tanda minus, seperti yang kemungkinan terjadi apabila kita menggunakan skala kesukaran linier.

2. Analisis Daya Pembeda

Untuk mengetahui intensitas sebuah soal dalam hal kesukaran dibutuhkan sebuah daya pembeda, yaitu kemampuan antara butir soal dapat membedakan antara peserta didik yang menguasai materi yang diujikan dan peserta didik yang belum menguasai materi yang diujikan. Menurut Zainul, daya beda butir soal ialah indeks yang menunjukkan tingkat kemampuan butir soal membedakan kelompok yang berprestasi tinggi dari kelompok yang berprestasi rendah diantara para peserta tes.

Angka yang menunjukkan besarnya daya pembeda disebut indeks diskriminasi. Yang berkisar antara 0,00 sampai 1,00. Pada indeks ini kemungkinan adanya tanda

negatif manakala suatu tes terbalik menunjukkan kualitas tes yaitu anak pandai disebut tidak pandai dan sebaliknya.

Dengan demikian ada 3 titik daya pembeda yaitu:

-1,00	0,00	1,00
Daya pembeda negatif	Daya pembeda rendah	Daya pembeda tinggi (positif)

Bertitik tolak dari titik di atas, terdapat patokan yang dapat digunakan untuk mengetahui sebesar manakah sebuah item butir soal dapat dinyatakan memiliki pembeda yang baik. Patokannya adalah sebagai berikut:

Besarnya angka indeks diskriminasi item (D)	Klasifikasi	Interpretasi
Kurang dari 0,20	Poor	Butir item yang bersangkutan daya pembedanya lemah sekali (jelek), dianggap tidak memiliki daya pembeda yang baik.
0,20 - 0,40	Satisfactor y	Butir item yang bersangkutan telah memiliki daya pembeda yang cukup (sedang)
0,40 - 0,70	Good	Butir item yang bersangkutan telah memiliki daya pembeda yang baik.
0,70 - 1,00	Excellent	Butir item yang bersangkutan telah memiliki daya pembeda yang baik sekali.
Bertanda negatif	-	Butir item yang bersangkutan daya pembedanya negatif (jelek sekali)

Bagi suatu soal yang dapat dijawab benar oleh siswa pandai maupun tidak pandai, maka soal itu tidak baik karena tidak mempunyai daya pembeda. Demikian sebaliknya, jika siswa yang pandai maupun tidak pandai tidak bisa menjawab benar, soal tersebut juga tidak baik karena tidak mempunyai daya pembeda.

Tes yang baik adalah tes yang dapat dijawab dengan benar oleh siswa yang pandai saja. Contohnya, jika suatu kelompok anak yang berprestasi tinggi dapat menjawab dengan benar suatu tes dan seluruh atau hampir suatu kelompok yang berprestasi rendah menjawab salah, dikatakan bahwa soal itu memiliki indeks diskriminasi (D) terbesar. Sebaliknya jika kelompok yang berprestasi rendah seluruhnya menjawab soal dengan benar sedangkan kelompok berprestasi tingginya menjawab dengan salah, maka indeks diskriminasi (D) soal tersebut -1,00. Sedangkan jika antara kedua kelompok sama-sama menjawab dengan benar, berarti indeks diskriminasi (D) soal tersebut 0,00 atau tidak memiliki daya pembeda.³⁵

Untuk mengetahui besar kecilnya angka indeks diskriminasi item dapat digunakan 2 macam rumus berikut ini:³⁶

a. Rumus Pertama,

$$D = P_A - P_B \text{ atau}$$

$$D = P_H - P_L$$

Di mana :

D = Discriminatory power (angka indeks diskriminasi item)

P_A atau P_H = Proporsi testee kelompok atas yang dapat menjawab dengan betul butir item yang bersangkutan.

Dapat diperoleh dengan rumus :

$$P_A = P_H = \frac{B_A}{J_A}$$

³⁵ Syamsudin, *Pengukuran Daya Pembeda, Taraf Kesukaran, dan Pol Jawaban Tes (Analisis Butir Soal)*, Jurnal Ilmu Tarbiyah "At-Tajdid", Vol. 1, No.2, Juli 2012, hlm. 189.

³⁶ Anas Sudjiono, *Pengantar Evaluasi Pendidikan...*, hlm.390-391.

Di mana:

B_A : Banyaknya testee kelompok atas yang menjawab dengan betul butir item bersangkutan.

J_A : Jumlah testee yang termasuk dalam kelompok atas

P_B atau P_L = Proporsi testee kelompok bawah yang dapat menjawab dengan betul butir item yang bersangkutan.

Dapat diperoleh dengan rumus:

$$P_B = P_L = \frac{B_B}{J_B}$$

Di mana:

B_B = Banyaknya testee kelompok bawah yang dapat menjawab dengan benar butir item yang bersangkutan

J_B = Jumlah testee yang termasuk dalam kelompok bawah.

- b. Rumus Kedua, angka indeks diskriminasi item diperoleh dengan menggunakan teknik korelasi Phi (ϕ) dengan rumus,

$$\phi = \frac{P_H - P_L}{2\sqrt{(P)(q)}}$$

Di mana:

ϕ : Angka indeks korelasi Phi, yang dalam hal ini dianggap sebagai angka indeks diskriminasi item.

P_H : Proportion of the higher group

P_L : Proportion of the lower group

2 : Bilangan konstan

P : Proporsi seluruh testee yang jawabannya betul

q : Proporsi seluruh testee yang jawabannya salah, di mana $q = (1 - p)$

Cara Menentukan Daya Pembeda³⁷

Membedakan menjadi kelompok kecil (kurang dari 100) dan kelompok besar (100 ke atas).

a) Kelompok kecil (kurang 100)

Seluruh kelompok tes terbagi 2 sama besar, separuh kelompok atas dan separuh kelompok bawah sebagai berikut:

Siswa	Skor		Siswa	Skor	
A	10	Kelompok atas (JA)	E	5	Kelompok Bawah (JB)
B	9		F	4	
C	8		G	3	
D	7		H	2	

b) Kelompok besar (100 ke atas)

Untuk memudahkan analisis cukup diambil kedua kutub atas dan bawahnya saja, masing-masing 27% sebagai JA dan JB nya. Contoh sebagai berikut:

9
 9 27% sebagai JA
 8
 8
 2 27% sebagai JB
 1
 1
 0

Dari tabel kelompok atas dan bawah itu dicari menggunakan rumus:

$$D = (B_a/J_A)-(B_b/J_B)=P_a-P_b$$

D : Daya Pembeda

³⁷ Syamsudin, *Pengukuran Daya Pembeda, Taraf Kesukaran, dan Pola Jawaban Tes (Analisis Butir Soal)*, hlm. 190.

J : Jumlah Peserta

J_A : Jumlah Peserta Atas

J_B : Jumlah Peserta Bawah

B_b : Jumlah Peserta Kelompok bawah menjawab benar

B_a : Jumlah peserta kelompok atas menjawab benar

$PB - BB/JB$: Proporsi peserta kelompok bawah yang menjawab benar

$PA - BA/JA$: Proporsi peserta kelompok atas yang menjawab benar.

Contoh perhitungan sebagai berikut:

Tabel Analisa 10 butir soal, 10 siswa

Siswa	Kelompok	1	2	3	4	5	6	7	8	9	10	Skor
A	A	1	1	1	1	1	1	1	1	1	0	9
B	B	0	1	0	1	1	1	0	1	0	0	5
C	B	1	0	0	0	0	0	0	1	1	1	4
D	B	0	1	0	1	0	1	1	1	0	0	5
E	B	1	0	0	1	1	0	1	0	1	1	6
F	A	1	1	1	1	1	1	1	1	0	0	8
G	A	1	1	1	1	1	1	1	1	1	1	10
H	A	1	0	0	0	1	1	1	1	1	1	7
I	B	0	1	0	1	1	0	0	0	0	0	3
J	A	1	1	1	1	0	0	1	1	1	1	8
10=N		7	7	4	8	7	6	7	8	6	5	65

Berdasarkan nama-nama siswa dapat diperoleh skor-skor sebagai berikut:

A = 9, B = 5, C = 4, D = 5, E = 6, F = 8, G = 10, H = 7, I = 3, J = 8

Dari angka-angka yang belum teratur lalu disusun menjadi array (urutan penyebaran), dari skor yang paling tinggi ke skor yang paling rendah.

Kelompok Atas	Kelompok Bawah
10	6
9	5
8	5
8	4
7	3
J _A : 5 orang	J _b : 5 orang

Array ini sekaligus menunjukkan adanya kelompok atas (JA) dan kelompok bawah (JB) dengan pemilikinya sebagai berikut:

Kelompok atas JA	Kelompok bawah JB
G = 10	E = 6
A = 9	B = 5
F = 8	D = 5
J = 8	C = 4
H = 7	I = 3

Selanjutnya dilihat tabel analisa lagi butir-butir soal sehingga diketahui hasilnya :

Nomor Butir Item	Ba	Bb	Ja	Jb	Pa = Ba/Ja	Pb = Bb/Jb	D = Pa - Pb
1	5	2	5	5	1,00	0,4	0,6
2	4	3	5	5	0,8	0,6	0,2
3	4	0	5	5	0,8	0	0,8

4	4	4	5	5	0,8	0,8	0
5	4	3	5	5	0,8	0,6	0,2
6	4	2	5	5	0,8	0,4	0,4
7	5	2	5	5	1,00	0,4	0,6
8	5	3	5	5	1,00	0,6	0,4
9	4	2	5	5	0,8	0,4	0,4
10	3	2	5	5	0,6	0,4	0,2

Kemudian diinterpretasikan terhadap D menjadi:

Nomor Butir Item	Besarnya D	Klasifikasi	Interpretasi
3	0,8	Excellent	Daya pembeda itemnya baik sekali
1 dan 7	0,6	Good	Daya pembedanya baik
6,8 dan 9	0,4	Satisfactory	Daya pembedanya cukup (sedang)
5	0,2	Poor	Daya pembedanya lemah sekali (jelek)
4	0,0	Poor	tidak memiliki daya pembeda sama sekali (jelek)

Sebagai tindak lanjut atas hasil analisis daya pembeda adalah:³⁸

- a. Butir-butir item yang sudah memiliki daya pembeda baik (satisfactory, good, dan excellent) hendaknya dimasukkan dalam bank soal. Dan dapat dikeluarkan lagi pada tes berikutnya karena kualitasnya sudah cukup memadai.
- b. Butir-butir item yang daya pembedanya masih rendah (poor), ada 2 kemungkinan:

³⁸ Anas Sudjiono, *Pengantar Evaluasi Pendidikan...*, hlm. 409.

- 1) Ditelusuri untuk kemudian diperbaiki, dan setelah diperbaiki dapat diajukan lagi dalam tes hasil belajar, yang kemudian dianalisis lagi apakah meningkat atau tidak.
 - 2) Dibuang atau didrop dan tidak dikeluarkan lagi untuk tes hasil belajar.
- c. Khusus butir-butir item yang angka indeks diskriminasi itemnya bertanda negatif, sebaiknya pada tes hasil belajar tidak usah dikeluarkan lagi, sebab butir yang demikian kualitasnya sangat jelek (testee yang pandai lebih banyak menjawab salah ketimbang testee yang tidak pandai, justru hanya sedikit yang menjawab salah)
3. Analisis Fungsi Distraktor

Pengertian distraktor yaitu, "*Distractor are classified as the incorrect answer in a multiple-choice question.*".³⁹ Dalam setiap tes obyektif selalu digunakan alternatif jawaban yang mengandung 2 unsur sekaligus, yaitu jawaban tepat dan jawaban yang salah sebagai penyesat (distraktor). Tujuan pemakaian distraktor ini adalah mengecohkan mereka yang kurang mampu atau tidak tahu untuk dapat dibedakan dengan yang mampu. Oleh karena itu, distraktor yang baik adalah yang dapat dihindari oleh anak-anak yang pandai dan terpilih oleh anak-anak yang kurang pandai.⁴⁰ Dan apabila terpilih minimal 5% dari jumlah peserta.

³⁹ Shafizan Sabri, *Item Analysis of student Comprehensive Test for Reasearch Teaching Begiliner String Ensemble Using Model Based Teaching Among Music Students In Public Universities*", International Journal of Rducation and Reasearch, 1 (12), hlm. 1.

⁴⁰ M. Chabib Thoha, M.A, *Teknik Evaluasi Pendidikan...*, hlm. 149.

Option atau alternatif yaitu beberapa kemungkinan jawaban berjumlah kisaran antara 3 sampai dengan 5 buah, dan dari kemungkinan-kemungkinan jawaban yang terpasang pada setiap soal, salah satu diantaranya adalah merupakan jawaban betul, sedangkan sisanya jawaban salah.⁴¹

Jawaban di setiap soal mempunyai pola jawaban. Yang dimaksud pola jawaban di sini adalah distribusi testee dalam hal menentukan pilihan jawaban pada soal bentuk pilihan ganda. Dan untuk memperolehnya dengan menghitung banyaknya testee yang memilih pilihan jawaban a,b,c, atau d yang tidak memilih pilihan manapun. Dari pola jawaban soal dapat ditentukan apakah pengecoh (distractor) berfungsi sebagai pengecoh dengan baik atau tidak. Pengecoh yang tidak dipilih sama sekali oleh testee berarti bahwa pengecoh itu jelek, terlalu menyolok menyesatkan. Sebaliknya sebuah distraktor dikatakan berfungsi dengan baik apabila distraktor tersebut mempunyai daya tarik yang besar bagi pengikut tes yang kurang memahami konsep atau kurang menguasai materi.⁴²

Suatu distraktor dapat diperlakukan dengan 3 cara:⁴³

- a. Diterima, karena sudah baik.
- b. Ditolak, karena tidak baik.
- c. Ditulis kembali, karena kurang baik.

⁴¹ Meita Fitriawanati, *Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru dan Hasil Belajar Peserta Didik*, Tulisan Artikel di Seminar Nasional Pendidikan PGSD UMS & HDPGSDI Wilayah Jawa, hlm. 287.

⁴² Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan...*, hlm. 173

⁴³ *Ibid.*, hlm. 174

Contoh menganalisis fungsi distraktor. Misalkan tes hasil belajar bidang studi Aqidah Akhlak diikuti oleh 50 orang siswa Madrasah Tsanawiyah. Bentuk soalnya adalah *multiple choice* item dengan item sebanyak 40 butir, yang dilengkapi lima alternatif yaitu, A,B,C,D dan E. Dari 40 butir item, khusus butir item 1,2 dan 3 diperoleh pola penyebaran jawaban item sebagai berikut:⁴⁴

Nomor Butir Item	Alternatif (= option)					Keterangan
	A	B	C	D	E	
1	4	6	5	(30)	5	() : Kunci Jawaban
2	1	(44)	2	1	2	
3	1	1	(10)	1	37	

Dengan pola penyebaran jawaban item seperti tabel di atas, dengan mudah mengetahui berapa persen testee yang telah “terkecoh” untuk memilih distraktor yang dipasangkan pada item 1, 2 dan 3, yaitu:

- 1) Item nomer 1, dengan jawaban D, dan distraktornya A,B,C dan E.

Pengecoh A dipilih 4 orang, berarti: $4/50 \times 100\% = 8\%$
(telah berfungsi dengan baik, karena angka persentasenya sudah lebih dari 5%)

Pengecoh B dipilih 6 orang, berarti: $6/50 \times 100\% = 12\%$
(telah berfungsi dengan baik)

Pengecoh C dipilih 5 orang, berarti: $5/50 \times 100\% = 10\%$
(telah berfungsi dengan baik)

Pengecoh E dipilih oleh 5 orang, berarti: $5/50 \times 100\% = 10\%$
(telah berfungsi dengan baik)

⁴⁴ Anas Sudjiono, *Pengantar Evaluasi Pendidikan...*, hlm. 413

Jadi, keempat distraktor tersebut sudah menjalankan fungsinya dengan baik.

- 2) Item nomer 2, dengan jawaban B, dan distraktornya A,C,D dan E.

Pengecoh A dipilih 1 orang, berarti: $1/50 \times 100\% = 2\%$
(belum berfungsi)

Pengecoh C dipilih 2 orang, berarti: $2/50 \times 100\% = 4\%$
(belum berfungsi)

Pengecoh D dipilih 1 orang, berarti: $1/50 \times 100\% = 2\%$
(belum berfungsi)

Pengecoh E dipilih 2 orang, berarti: $2/50\% \times 100\% = 4\%$
(belum berfungsi)

Jadi, keempat distraktor tersebut belum menjalankan fungsinya dengan baik seperti yang diharapkan.

- 3) Item nomer 3, dengan jawaban C, dan distaktornya A,B,D dan E.

Pengecoh A,B dan D masing-masing dipilih 1 orang, berarti: $1/50 \times 100\% = 2\%$ (belum berfungsi)

Pengecoh E dipilih 37 orang, berarti: $37/50 \times 100\% = 74\%$ (telah berfungsi baik)

Jadi, pada item ini, hanya 1 distraktor saja yang sudah dapat menjalankan fungsinya dengan baik seperti yang diharapkan.

C. KESIMPULAN

Dalam suatu evaluasi pembelajaran dibutuhkan juga evaluasi proses penilaian dan pengukuran siswa. Salah satunya dengan menganalisis kesukaran soal, daya pembeda dan fungsi distraktor atau pengecoh. Yang mana dalam analisis kesukaran soal sangat dibutuhkan sekali. Karena untuk mengetahui

apakah testee bisa mengerjakan soal yang diberikan atau justru soal yang diberikan terlalu mudah bahkan terlalu sukar. Karena proporsi dalam kesukaran soal paling tidak terdiri dari 25% sukar, 50% sedang dan 25% mudah. Yang mana dengan porsi yang sudah diatur bagi testee pandai maupun tidak pandai mudah mengerjakan soal atau tidak merasa kesulitan. Begitupula dibutuhkan daya pembeda khususnya dalam soal *multiple choice*, sangat berarti yang mana dari sini bisa diketahui mana anak yang pandai dan tidak. Dan ditambah dengan distraktor atau pengecoh jawaban pada soal *multiple choice*. Yang berfungsi sebagai alat yang dapat menggambarkan apakah butir soal yang dibuat baik atau gagal. Dalam keseluruhannya, analisis soal-soal pada sebuah tes sangat diperlukan guna untuk mengadakan identifikasi soal-soal yang baik, kurang baik dan soal yang jelek. Dengan menganalisis soal akan diperoleh informasi tentang kejelasan sebuah soal dan petunjuk untuk mengadakan perbaikan dalam pembelajaran. Dan sebagai acuan para pendidik untuk terus menilai dan mengukur hasil belajar para siswa dengan baik dan adil.

DAFTAR PUSTAKA

- Arikunto, Suharsimi, *Dasar-Dasar Evaluasi Pendidikan*, Yogyakarta: PT. BINA AKSARA, 1984.
- Daryanto, *Evaluasi Pendidikan*, Jakarta: Rineka Cipta, 2012.
- Fitrianawati, Meita, *Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru dan Hasil Belajar Peserta Didik*, Tulisan Artikel di Seminar Nasional Pendidikan PGSD UMS & HDPGSDI Wilayah Jawa.
- Joni, Raka, *Pengukuran dan Penilaian Pendidikan*, Malang: YP2LPM, 1984.
- Purwanto, Ngalim, *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*, Bandung: Remadja Karya, 2002.
- Purwanto, *Evaluasi Hasil Belajar*, Yogyakarta: Pustaka Pelajar, 2009.
- Sabri, Shafizan, *Item Analysis of student Comprehensive Test for Reasearch Teaching Begüliner String Ensemble Using Model Based Teaching Among Music Students In Public Universities*”, International Journal of Rducation and Reasearch, 1 (12)
- Sudjiono, Anas, *Pengantar Evaluasi Pendidikan*, Jakarta: Rajawali Press, 2016.
- Sulistyorini, *Evaluasi Pendidikan Dalam Meningkatkan Mutu Pendidikan*, Yogyakarta: Teras, 2009.
- Syamsudin, *Pengukuran Daya Pembeda, Taraf Kesukaran, dan Pol Jawaban Tes (Analisis Butir Soal)*, Jurnal Ilmu Tarbiyah “At-Tajdid”, Vol. 1, No.2, Juli 2012.
- Thoha, M.A, M. Chabib, *Teknik Evaluasi Pendidikan*, Jakarta: PT. Raja Grafindo Persada, 1994