

LAPORAN PENELITIAN

**Pola Kecenderungan Penempatan Kunci
Jawaban Pada Soal Tipe-D
Melengkapi Berganda**

**Oleh:
Drs. Pramono Sidi**

**Fakultas Matematika dan Ilmu Pengetahuan Alam
Mei 1990**

RINGKASAN

Populasi yang diambil untuk penelitian ini adalah soal-soal ujian matakuliah-matakuliah pada program studi Statistika Terapan dan Matematika FMIPA yang mempunyai tipe D - melengkapi berganda. Dugaan semula bahwa ada kecenderungan penulis soal meletakkan kunci jawaban soal di tengah-tengah (menurut Dorothy C. Adkins), setelah melalui penelitian ini kecenderungan tersebut ternyata bahwa penulis lebih tertarik menempatkan kunci jawaban pada option 1, 2 dan 3 benar atau kunci jawaban D.

PENDAHULUAN

Penulisan soal tes/ujian adalah bagian dari fungsi Pendidikan. Intisari fungsi Pendidikan dapat dispesifikasi menjadi fungsi selektif dan fungsi pengembangan. Yang dimaksud dengan fungsi selektif adalah bahwa tugas pendidikan antara lain mengidentifikasi anggota populasi kelas yang diijinkan untuk meneruskan pendidikannya ke jenjang yang lebih tinggi. Dengan fungsi pengembangan dimaksud bahwa tugas pendidikan antara lain untuk mengembangkan efektivitas pengalaman belajarnya.

Soal tes/ujian sebagai salah satu alat pengukuran pendidikan disusun untuk tujuan mengukur sampai seberapa jauh kedua fungsi pendidikan tersebut berhasil dicapai. Bentuk soal ujian yang dipakai oleh Universitas Terbuka pada umumnya pilihan berganda, hanya sebagian saja dalam bentuk essay (uraian).

Dalam hal ini FMIPA Universitas Terbuka sampai dengan masa ujian 90.1 telah mengembangkan sebanyak 61 matakuliah, sebagian besar soal ujian dalam bentuk pilihan berganda kecuali 12 matakuliah yang bentuk soal ujiannya adalah essay (uraian). Universitas Terbuka, dalam penyelenggaraan ujiannya, menetapkan 5 (lima) macam tipe soal ujian yaitu:

1. melengkapi empat pilihan disebut soal tipe A
2. analisis hubungan antar hal disebut soal tipe B
3. analisis kasus disebut soal tipe C
4. melengkapi berganda disebut soal tipe D
5. analisis diagram/grafik disebut soal tipe E.

Adapun banyaknya pilihan setiap butir soal ditetapkan sebanyak empat pilihan.

Dalam membuat soal ujian para penulis soal, khususnya FMIPA biasanya lebih suka membuat soal dengan tipe A, C, D atau E. Namun yang akan dibicarakan di sini adalah mengenai soal tipe D yaitu melengkapi berganda.

TUJUAN PENELITIAN

Seperti sudah disebutkan pada pendahuluan, bahwa penulisan soal tes/ujian adalah bagian dari fungsi selektif dan fungsi pengembangan. Oleh karena itu dalam penulisan soal pilihan berganda perlu diperhatikan beberapa hal antara lain, persyaratan sebagai alat pengukur pendidikan, format sesuai dengan maksud pengukuran, jenjang kemampuan, derajat kesukaran, kejelasan bahasa dan juga apakah soal dapat memisahkan kemampuan menjadi golongan pandai, kurang pandai dan bodoh. Tetapi di lain pihak, meskipun hal-hal di atas sudah kita penuhi, ada yang berpendapat bahwa soal ujian pilihan berganda mempunyai banyak kelemahan. Salah satu kelemahan tersebut adalah jawabannya "dapat diterka". Jika soal ujian pilihan berganda tersebut terdiri dari empat option (pilihan), maka kemungkinan untuk menempatkan jawaban yang benar dengan cara menerka adalah 25%, dengan asumsi bahwa setiap option ataupun stem soal tidak memberikan "petunjuk" ke arah pemilihan suatu jawaban tertentu (fair). Asumsi lain tentunya adalah tidak adanya kecenderungan dari penulis soal untuk meletakkan kunci jawaban soal untuk suatu tipe soal tertentu pada tempat-tempat tertentu. Jika asumsi yang terakhir ini tidak dapat dipenuhi maka kemungkinan untuk menerka jawaban yang benar akan semakin tinggi ($> 25\%$).

Menurut Dorothy C. Atkins (1974):

" Many people would tend to use the third position too often for the answer in five choice items and to under use the first position and, perhaps to a lesser, the fifth position". (p.94).

Dengan demikian perlu dilihat bagaimana kecenderungan seorang penulis soal (dalam hal ini soal tipe D) membuat jawaban yang benar. Jika kunci jawaban soal tipe D ini mempunyai pola tertentu, maka dikhawatirkan bahwa seorang mahasiswa mengerjakan soal akan menjawab soal-soal dengan jawaban sesuai pola tersebut. Untuk itulah, apabila penelitian ini menemukan suatu pola tertentu mengenai penempatan kunci jawaban pada soal tipe D, maka kita harus memberitahu agar penulis sebaiknya tidak menggunakan pola tertentu dalam menempatkan kunci jawaban.

DATA DAN METODE

DATA

Penelitian ini menggunakan sampel yang diambil dari soal-soal ujian FMIPA-UT yang sudah diujikan pada masa ujian 89.1 dan 89.2. Soal-soal ujian yang diambil adalah soal-soal ujian yang belum pernah direvisi oleh penulis lain, sehingga kunci jawabannya pun belum dirubah.

Matakuliah yang diujikan pada masa ujian 89.1 adalah 57 matakuliah sedang pada masa ujian 89.2 adalah 61 matakuliah. Sedang jumlah soal-soal yang diujikan pada dua masa ujian tersebut adalah 3540 butir dan 232 butir atau 6,55% di antaranya adalah butir-butir soal bertipe D. Sisanya sejumlah 3308 butir atau 93,5% adalah butir-butir soal bertipe A, B, C, dan E.

METODE

Sebelum kita sampai pada analisis data, perlu dijelaskan lebih dulu apa yang disebut soal dengan tipe D tersebut. Bentuk soal tipe D ini hampir sama dengan bentuk soal tipe A (melengkapi empat pilihan), yaitu satu pernyataan yang tidak lengkap yang diikuti dengan beberapa kemungkinan. Perbedaannya adalah jawaban yang benar bisa satu, dua atau tiga, sehingga jawaban akan mempunyai kategori sebagai berikut:

jawaban

- A jika 1 dan 2 benar
- B jika 1 dan 3 benar
- C jika 2 dan 3 benar
- D jika 1, 2 dan 3 benar

Analisis yang digunakan adalah membuat distribusi frekuensi dari pilihan kunci jawaban A, B, C dan D atau distribusi frekuensi dari yang menjawab 1 dan 2 benar, 1 dan 3 benar, 2 dan 3 benar dan 1, 2 dan 3 benar.

Kemudian kita uji hipotesis yang menyatakan bahwa tidak ada kecenderungan penulis soal dalam menempatkan kunci jawaban pada pilihan tertentu, artinya kita uji seberapa cocok frekuensi kunci jawaban yang diobservasi dari sampel dengan frekuensi yang diharapkan.

Pada penelitian ini, ukuran sampel yang diambil adalah 232 butir soal, sehingga frekuensi yang diharapkan untuk setiap kunci jawaban adalah 58 butir soal. Untuk menguji kecocokan antara frekuensi yang diobservasi dengan frekuensi yang diharapkan, digunakan uji kebaikan saat (=Goodness of fit), dengan statistik pengujinya adalah

$$W = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

di mana

O_i = frekuensi jawaban ke-i yang diobservasi

E_i = frekuensi jawaban ke-i yang diharapkan

W mempunyai distribusi χ^2 dengan derajat kebebasan $k-1$; $k = 4$

Jika frekuensi yang diobservasi cukup dekat dengan frekuensi yang diharapkan, maka nilai W akan kecil, dan $W < \chi^2(k-1; \alpha)$, sehingga hipotesis diterima artinya kecocokan antara frekuensi yang diobservasi dengan

frekuensi yang diharapkan cukup bagus. Sedangkan kalau frekuensi yang diobservasi jauh berbeda dengan frekuensi yang diharapkan, maka nilai W akan besar dan $W > \chi^2(k-1; \alpha)$, sehingga hipotesis ditolak, artinya tidak ada kecocokan sama sekali antara frekuensi yang diobservasi dengan frekuensi yang diharapkan.

HASIL DAN PEMBAHASAN

Dari hasil pengamatan dari sampel berukuran 232 butir soal diperoleh distribusi frekuensi kunci jawaban sebagai berikut:

Tabel 1. Distribusi Frekuensi Kunci Jawaban

Kunci Jawaban Frekuensi	A	B	C	D
Diobservasi	56	53	32	91
Diharapkan	58	58	58	58

Tabel 2. Proporsi Kunci Jawaban

Kunci Jawaban Frekuensi	A	B	C	D
Diobservasi	0,241	0,228	0,138	0,392
Diharapkan	0,25	0,25	0,25	0,25

Dari tabel 1 dan 2 terlihat bahwa kunci jawaban D mempunyai proporsi paling besar yaitu 39,2% (= 91 butir) diikuti kunci jawaban A sebesar 24,1% (= 56 butir), kunci jawaban B sebesar 22,8% (= 53 butir) terakhir kunci jawaban C sebesar 13,8% (= 32 butir).

Ini berarti bahwa penulis soal lebih cenderung membuat soal dengan option jawabannya 1, 2 dan 3 benar, kemudian diikuti dengan option 1 dan 2 benar, lalu 1 dan 3 benar dan terakhir option 2 dan 3 benar. Karena soal tipe D - melengkapi berganda, kemungkinan option yang benar bisa satu, dua atau tiga, maka perlu diperhatikan hal-hal seperti berikut ini:

- jika pertanyaan terdapat pada stem soal, maka pilihan (option) yang terdiri dari 3 (tiga) jawaban, dua atau tiga pernyataannya harus merupakan kesatuan atau jawaban yang utuh.
- Jika pertanyaan terdapat pada pilihan (option) jawaban 1, 2 dan 3, maka 2 atau 3 pilihan (option) tersebut boleh tidak merupakan kesatuan, tetapi harus berasal dari satu TIK.

Kembali kepada kesimpulan deskriptif seperti yang disebutkan di atas, walaupun sudah diketahui, tetapi untuk memperkuatnya harus dilakukan uji secara statistik, apakah kesimpulan tersebut betul-betul signifikan.

Uji Hipotesis

Seperti diketahui, bahwa tujuan penelitian ini adalah untuk melihat kecenderungan penulis soal dalam membuat kunci jawaban, apakah mereka cenderung memilih/meletakkan jawaban pada pilihan A, B, C ataupun D. Oleh karena itu Hipotesis Ho-nya adalah tidak ada kecenderungan penulis soal dalam meletakkan jawaban pada pilihan (option) tertentu, dengan perkataan lain distribusi frekuensi kunci jawaban yang diobservasi cocok dengan distribusi frekuensi jawaban yang Diharapkan.

Untuk ini harus dihitung dulu statistik penguji

$$W = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

dimana:

$$O_1 = O_A = 56$$

$$O_2 = O_B = 53$$

$$O_3 = O_C = 32$$

$$O_4 = O_D = 91$$

$$E_1 = E_A = E_2 = E_B = E_3 = E_C = E_4 = E_D = 58$$

$$= \frac{4 + 25 + 676 + 1089}{58} = \frac{1794}{58} = 30,93$$

Statistik W berdistribusi X^2 (Chi Kuadrat) dengan dk = n - 1 jika dipakai taraf kepercayaan $\alpha = 5\%$ (signifikan), maka $X^2 (n - 1; \alpha) = X^2 (4 - 1; 0,05) = 7,815$ (lihat tabel disribusi X^2 dengan $\alpha = 0,05$ dan d.k = v = 4 - 1 = 3. Di sini terlihat bahwa statistik $W > X^2 (3 ; 0,05)$, berarti Hipotesis Ho ditolak secara signifikan, dengan perkataan lain secara signifikan tidak ada kecocokan antara frekuensi jawaban yang diobservasi dengan frekuensi jawaban yang diharapkan, atau dapat pula disimpulkan bahwa ada kecenderungan penulis soal untuk menempatkan jawaban pada pilihan tertentu. Artinya pola yang dipakai penulis, menurut tabel 1 dan 2 serta uji statistik adalah

- bahwa penulis soal cenderung membuat soal di mana option 1, 2 dan 3 benar, yang kunci jawabannya D.
- bahwa penulis soal tidak tertarik membuat soal di mana option 2 dan 3 benar yang kunci jawabannya C.

KESIMPULAN DAN SARAN

1. Pola kecenderungan penempatan kunci jawaban soal tipe D melingkupi berganda adalah memilih option 1, 2 dan 3 benar, dengan kata lain cenderung memilih kunci D.
2. Pola kecenderungan bahwa penulis tidak tertarik membuat soal dengan kunci jawaban C, yaitu memilih option 2 dan 3 benar
3. Hasil penelitian ini bertentangan dengan apa yang dikemukakan oleh Dorothy C. Adkins (1974) yang menyatakan bahwa penulis cenderung menempatkan kunci jawaban di tengah (kunci B atau C).
4. Kesimpulan penelitian ini sebaiknya perlu diinformasikan kepada penulis.

DAFTAR PUSTAKA

ADKINS, C.D., "*Test Construction development interpretation of achievement tests*" 2nd ed. Columbus, Ohio 43216, Charles E. Herrill Publishing Co., 1974

BHATTACHARYYA, GK and JOHNSON, RA, "*Statistical Concepts and Methods*", J. Wiley 1977.