

DOKUMENTASI
UNIVERSITAS TERBUKA

LAPORAN PENELITIAN

HUBUNGAN ANTARA STATUS PENULIS SOAL DAN
KUALITAS SOAL UJIAN YANG DITULIS
PADA FMIPA-UT

oleh

Ir. Tuty Maria Wardiny

Ir. Argadatta Sigit, M.Ed

00172

Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Terbuka

1990

LAPORAN PENELITIAN

HUBUNGAN ANTARA STATUS PENULIS SOAL DAN
KUALITAS SOAL UJIAN YANG DITULIS PADA FMIPA-UT

OLEH :

IR. TUTY MARIA WARDINY
IR. ARGADATTA SIGIT M.Ed.

Mengetahui,
Dekan FMIPA-UT


Dra. Erim Moerwani M.Sc.
NIP. 1300475



Menyetujui,
Pembimbing



Prof. dr. Sigit Muryono
NIP 130098861

DEPARTEMEN PENDIDIKAN DAN KEBUDAYAAN

UNIVERSITAS TERBUKA JAKARTA

RINGKASAN

Terjadinya kesenjangan antara ketersediaan soal-soal yang cukup banyak, berkualitas baik dan tersedia setiap saat untuk digunakan sebagai bahan ujian, bahan tugas mandiri ataupun sebagai bahan dalam pelaksanaan tutorial tertulis.

Sedangkan dilain pihak belum dimanfaatkannya tenaga edukatif yang ada sebagai sumber bahan ujian/soal yang dapat diandalkan dalam ketersediannya dan juga lebih murah dalam proses pengadaannya.

Berdasarkan hal tersebut di atas maka penulis mencoba mencari hubungan keterkaitan antara kualitas soal ujian dengan status dan latar belakang penulis soal.

Dalam menganalisa data yang telah diperoleh digunakan Uji T dengan software STATPACK. Dan tingkat signifikan yang digunakan adalah 0.05.

Dari hasil pengolahan data tersebut dapat ditarik kesimpulan bahwa antara penulis modul dan non-penulis modul dengan kualifikasi yang sama, akan menyebabkan tidak ada perbedaan dalam kualitas soal yang ditulis, sehingga terbuka kesempatan bagi FMIPA-UT khususnya untuk menggunakan tenaga pengajar yang tidak menulis modul untuk menulis soal ujian.

Orang dari dalam Universitas Terbuka dan dari luar Universitas terbuka terdapat perbedaan yang signifikan pada kualitas dan kuantitas soal yang ditulis, sehingga untuk memanfaatkan tenaga edukatif junior ini perlu diadakan peningkatan kemampuan dalam penulisan soal.

KATA PENGANTAR

Penelitian ini terlaksana berkat bantuan dana dari Pusat Penelitian Universitas Terbuka (Puslit-UT). Untuk itu kami ucapkan terima kasih. Dan pada kesempatan ini kami sampaikan terima kasih yang sebesar-besarnya kepada Prof. dr. Sigit Muryono sebagai pembimbing, Dra. Patimah Moerwani, M.Sc. sebagai Dekan FMIPA-UT, dan teman-teman yang telah membantu dalam pelaksanaan penelitian ini yang tidak dapat kami sebutkan satu per-satu.

Kami yakin, bahwa pada laporan ini ada kekurangan-kekurangan yang masih dapat diperbaiki, oleh karena itu kami sangat mengharapkan saran dan masukan dari rekan-rekan sejawat yang membaca laporan ini. Semoga laporan ini ada manfaatnya dalam bidang pembangunan pada umumnya dan pendidikan di UT pada khususnya.

Jakarta, November 1990

Penulis

DAFTAR ISI

KATA PENGANTAR	i
DAFTAR ISI	ii
DAFTAR TABEL	iii
PENDAHULUAN	1
1. Latar belakang	1
2. Perumusan masalah	2
3. Hipotesis	2
TINJAUAN PUSTAKA	3
1. Tingkat kesulitan butir soal	4
2. Daya pembeda butir	5
3. Korelasi tiap butir dan nilai total	8
4. Analisa distractor	8
TUJUAN DAN MANFAAT PENELITIAN	9
METODE PENELITIAN	10
HASIL DAN PEMBAHASAN	11
Hasil	11
Pembahasan	14
KESIMPULAN DAN SARAN	19
Kesimpulan	19
Saran	19
DAFTAR PUSTAKA	21
LAMPIRAN	22

DAFTAR TABEL

Tabel 1.	Persentase Item yang Baik Berdasarkan Data Hasil Item Analisis	13
Tabel 2.	Tingkat Signifikan untuk "Semua", P , R_{bis} dan Mean	15
Tabel 3.	Nilai Tengah untuk "Semua", P , R_{bis} dan Mean	17

UNIVERSITAS TERBUKA

I. PENDAHULUAN

1. Latar Belakang

Lima tahun setelah berdirinya Universitas Terbuka maka pada saat ini hampir seluruh modul yang akan digunakan oleh mahasiswa FMIPA-UT telah tersedia atau tertulis. Di lain pihak saat ini dirasakan adanya kebutuhan akan ketersediaan soal-soal yang cukup banyak, berkualitas baik dan tersedia setiap saat untuk digunakan sebagai bahan uji, bahan tugas mandiri (TM) ataupun sebagai bahan dalam pelaksanaan tutorial tertulis. Mengingat kebutuhan akan bahan soal-soal tersebut yang tinggi, maka diperlukan adanya suatu sumber pemenuhan yang dapat diandalkan ketersediaannya, dan juga lebih murah dalam proses pengadaannya.

Di dalam Universitas Terbuka pada umumnya dan di Fakultas Matematik dan Ilmu Pengetahuan Alam (FMIPA) secara khususnya tersedia cukup banyak tenaga pengajar yang dapat dimanfaatkan sebagai sumber bahan uji/soal tersebut, dan dapat diandalkan dalam ketersediaan waktu dan ketepatan penyelesaian penulisan.

Disamping itu juga perlu dilihat bahwa banyak tenaga edukatif FMIPA-UT belum dimanfaatkan secara maksimal dalam usaha penulisan soal/bahan ujian tersebut. Dengan asumsi bahwa kualifikasi tenaga-tenaga pengajar tersebut dalam hal pengetahuan akan materi ajaran masih diragukan atau belum memenuhi standar, dan pada saat bersamaan tenaga pengajar tersebut membutuhkan latihan dan magang dalam penulisan materi modul, soal-soal, naskah audio-video dan lain-lain.

Beberapa staf FMIPA-UT telah melakukan penulisan tersebut terutama dalam penulisan soal-soal maka rasanya perlu diadakan suatu penelitian mengenai studi kelayakkan seseorang tenaga pengajar junior di Universitas Terbuka dalam menulis soal-soal tersebut. Studi kelayakkan ini dapat mencakup beberapa aspek antara lain pemanaman materi pelajaran, kualitas soal yang telah ditulis dan diujikan.

2. Perumusan Masalah

Dari uraian tersebut di atas maka yang menjadi masalah dalam penelitian ini adalah tersedianya sumber bahan ujian/soal yang dapat diandalkan dalam ketersediannya dan ketepatan penyelesaian penulisan dengan kualitas soal yang baik.

3. Hipotesis

Kualitas soal ujian berkaitan erat dengan status dan latar belakang penulis soal. Latar belakang pendidikan, pengalaman mengajar, dan statusnya terhadap suatu modul pada suatu mata ajaran yang ditawarkan dan diujikan, akan meningkatkan kualitas soal ujian yang ditulisnya.

II. TINJAUAN PUSTAKA

Tes Hasil Belajar atau ujian adalah suatu alat yang paling sering digunakan untuk menentukan keberhasilan seseorang dalam suatu proses belajar-mengajar atau untuk menentukan keberhasilan suatu program pendidikan, makin tinggi kualitas maka ujian tersebut makin baik sebagai alat ukur.

Cronbach (1970) mendefinisikan Tes sebagai suatu prosedur yang sistematis untuk mengobservasi kemampuan seseorang dan menjelaskan hasil dengan bantuan skala atau sistem penilaian lainnya.

Prosedur yang sistematis ini menunjukkan bahwa Tes itu harus di-konstruksi, di-administrasi, dinilai dan dijelaskan sesuai dengan aturan tertentu.

Bentuk ujian yang digunakan di Universitas Terbuka sebagai bahan ujian adalah pilihan berganda. Bentuk ujian pilihan berganda yang digunakan memiliki beberapa ragam :

- a. Ragam melengkapi pilihan.
Ragam ini terdiri atas pokok soal (stem) dan sejumlah alternatif jawaban (options).
- b. Ragam hubungan antar hal.
Ragam ini terdiri atas dua pernyataan yang dihubungkan oleh kata 'sebab'. Kedua pernyataan tersebut dapat benar dan dapat pula salah, dapat mempunyai hubungan dan dapat pula tidak.
- c. Ragam tinjauan kasus.
Ragam ini sangat bagus untuk mengukur kemampuan siswa memecahkan satu masalah (dapat nyata dapat pula fiktif).
- d. Ragam asosiasi pilihan berganda.
Ragam ini memungkinkan jawaban yang betul lebih dari satu. Akan tetapi alternatif jawabannya yang betul hanya tetap satu.

e. Ragam membaca diagram.

Ragam ini terutama digunakan untuk mengukur kemampuan siswa dalam membaca, dalam menafsirkan, diagram-diagram, grafik-grafik, dan sejenisnya.

Terdapat empat karakteristik yang biasanya digunakan untuk mendeterminasi butir soal, yaitu :

1. Tingkat kesulitan butir soal.
2. Daya pembeda butir soal.
3. Korelasi tiap butir dan nilai total.
4. Analisa distractor (pengecoh).

1. Tingkat kesulitan butir soal.

Apabila suatu butir soal dinilai secara dikotomi, lulus dan gagal, maka tingkat kesukarannya dapat diukur dengan tingkat persentase yang lulus. Jadi secara kebalikannya, makin mudah suatu butir soal maka indek kesulitannya makin tinggi.

Jika 50% dari siswa yang diuji lulus dan 50% sisanya gagal, maka keragaman (pxq) dari butir soal tersebut adalah 0,25. Butir soal dengan kesulitan 50% adalah butir soal dengan daya pembeda yang terbaik.

Seandainya bentuk kurva sebaran yang kita dapatkan berbentuk negatif, ini berarti bahwa ujian tersebut terlalu mudah sehingga kita harus menggantinya dengan butir-butir soal yang lebih sulit; karena yang kita inginkan adalah distribusi yang simetris. Demikian juga untuk ujian yang cenderung positif.

2. Daya pembeda butir.

Tujuan dari suatu ujian adalah untuk memberikan informasi tentang perbedaan individu dan konteks yang direncanakan untuk diukur oleh tes itu atau pada kriteria eksternal yang hendak diramalkan oleh skor tes itu. Oleh karenanya dapat didefinisikan bahwa daya pembeda dari suatu ujian adalah kemampuannya dalam membedakan/memisahkan seorang siswa yang baik dari siswa yang kurang baik (buruk).

a. Indeks daya pembeda.

Merupakan salah satu parameter daya pembeda yang sederhana, yang hanya dapat diterapkan pada butir-butir yang diskor secara dikotomi. Perhitungannya menghendaki penentuan satu atau dua titik skor pada distribusi skor kriteria sebagai skor potong kemudian pemisahan siswa ke dalam kelompok bernilai rendah dan bernilai tinggi diatas skor potong.

Apabila kelompok siswa itu besar, perhitungan indeks daya pembeda tidak perlu menggunakan seluruh kelompok bernilai tinggi dan bernilai rendah yang masing-masing sebesar 50%.

Kelley (1939) menunjukkan bahwa dibawah kondisi tertentu indeks diskriminasi yang lebih sensitif dan stabil dapat diperoleh dengan penggunaan 27% kelompok bernilai tinggi dan 27% kelompok bernilai rendah, apabila ukuran sampel cukup besar, sesungguhnya hasil yang sama dapat diperoleh dengan kelompok bernilai tinggi dan bawah 30% atau 50% (Beuchert dan Mendoza, 1979; Englenart 1965).

Rumus Indeks daya pembeda (D) :

$$D = P_u - P_l$$

Keterangan :

p_u = proporsi siswa di dalam kelompok bernilai tinggi menjawab butir soal itu secara benar.

p_i = proporsi siswa di dalam kelompok bernilai rendah menjawab butir soal itu secara benar.

Nilai positif menunjukkan bahwa proporsi siswa di dalam kelompok bernilai tinggi yang menjawab benar lebih besar daripada proporsi siswa di dalam kelompok bernilai rendah yang menjawab benar dan arti sebaliknya untuk nilai negatif.

b. Indeks pembeda butir.

Rumus yang lazim digunakan untuk menaksir tingkat hubungan antara butir dan skor kriteria ialah koefisien korelasi Pearson Product Moment. Pada situasi tertentu rumus ini dapat disederhanakan seperti untuk mengetahui seberapa jauh hubungan antara jawaban pada suatu butir tes yang diskor secara dikotomi dengan skor tes total. Pada situasi ini disebut Korelasi Point Biserial yang dinyatakan sebagai :

$$R_{pbis} = \frac{(\bar{x}_p - \bar{x}_q) \sqrt{p/q}}{S_x}$$

atau

$$R_{pbis} = \frac{(\bar{x}_p - \bar{x}_q) \sqrt{pq}}{S_x}$$

dimana

—
 X_p = rata-rata skor kriteria untuk siswa yang menjawab butir tertentu itu secara benar

—
 X_q = rata-rata skor kriteria untuk siswa yang menjawab salah untuk butir-butir yang sama

—
 X_x = rata-rata skor kriteria untuk seluruh siswa

S_p = simpangan baku skor kriteria

p = kesukaran butir

q = $1 - p$

Koefisien Korelasi Biserial

Rumus Koefisien Korelasi Biserial :

$$R_{bis} = \frac{(\bar{X}_p - \bar{X}_x) (p/y)}{S_x \sqrt{pq}}$$

$$R_{bis} = \frac{\bar{X}_p - \bar{X}_q}{S_x} \times \frac{pq}{Y}$$

dimana

—
 X_p , X_q , X_x , S_p , p , dan q : lihat keterangan atas.

Y = ordinat dari kurva normal standar pada skor-z yang berhubungan dengan nilai p untuk butir ini.

Secara matematis hubungan antara korelasi biserial dan point biserial adalah :

$$R_{bis} = \frac{V_{pq}}{Y} R_{pbis}$$

Apabila skor dari butir diskor secara dikotomi dikorelasikan dengan skor dari suatu kriteria dikotomi maka dapat digunakan koefisien phi.

3. Korelasi tiap butir dan nilai total.

Apabila butir diskor secara dikotomi, pengukur yang terbaik dari korelasi adalah Point Biserial yang sangat berbeda dengan Korelasi Product Moment. Ada beberapa pengukuran lain untuk hubungan item-total. Kegunaan dari Korelasi Point Moment dengan butir yang dikotomi adalah Point Biserial memungkinkan digunakan beberapa metoda dari analisa multivarian.

4. Analisa distractor (Pengecoh).

Aspek keempat dari analisis butir soal difokuskan pada bentuk dari bermacam-macam jawaban dari pengecoh. Kita tidak menginginkan pilihan yang menarik perhatian. Ini semacam ketentuan yang tak tertulis bahwa setiap pilihan harus dihilangkan atau direvisi, oleh karena itu harus diitung nilai effisiensinya masing-masing.

Hal ini memerlukan para pengotahusan dari murid dalam memilih setiap pengecoh dari kelompok bernilai tinggi ke kelompok bernilai rendah. Jika kita menemukan pengecoh yang kebalikan, yang menarik murid dari kelompok bernilai tinggi untuk menjawab sebagai jawaban yang benar, kita harus mempelajari pengecoh tersebut secara hati-hati.

III. TUJUAN DAN MANFAAT PENELITIAN

Tujuan

Penelitian ini bertujuan untuk mengetahui adanya hubungan antara status penulis soal/bahan, yang dalam penelitian ini dikelompokkan dalam 4 kelompok, dengan kualitas soal yang ditulisnya. Yang diamati dalam penelitian ini adalah karakteristik soal yaitu : P, Rbis, Mean dan "Semua".

Manfaat

Temuan-temuan dari hasil penelitian ini diharapkan dapat memberikan pertimbangan bagi para Dekan khususnya FMIPA-UT untuk memanfaatkan tenaga edukatif yang ada secara maksimal didalam penulisan soal.

IV. METODE PENELITIAN

1. Teknik Pengumpulan Data

Penelitian ini mengambil data sekunder yaitu data item analisis dari matakuliah yang ada di FMIPA-UT.

2. Populasi dan Sampel

Penentuan populasi dan sampel yang diambil yakni populasinya adalah hasil item analisa matakuliah di FMIPA-UT.

Sampelnya adalah 12 item analisa matakuliah di FMIPA-UT yaitu :

1. Fisika
2. Program Linear
3. Matematika Berhingga
4. Kimia II
5. Kalkulus I
6. Struktur Data Analisa
7. Analisa Data Statistik
8. Struktur Aljabar
9. Fisika I
10. Kalkulus Lanjutan II
11. Pengantar Psikologi
12. Teori Survei Sampel

3. Teknik Analisa Data

Metode analisis yang digunakan adalah Uji F dengan menggunakan software STATPACK. Tingkat signifikan ditetapkan sebesar 0.05.

V. HASIL DAN PEMBAHASAN

HASIL

Dalam penelitian ini dilakukan pengamatan terhadap beberapa matakuliah yang terdapat pada FMIPA di lingkungan Universitas Terbuka .

Adapun matakuliah yang diamati adalah :

- | | |
|----------------------------|---------------|
| 1. Fisika | - PM dan UT |
| 2. Program Linear | - PM dan NUT |
| 3. Matematika Berhingga | - PM dan NUT |
| 4. Kimia II | - PM dan NUT |
| 5. Kalkulus I | - PM dan NUT |
| 6. Struktur Data Statistik | - PM dan NUT |
| 7. Analisa Data Statistik | - PM dan NUT |
| 8. Struktur Aljabar | - PM dan NUT |
| 9. Fisika I | - NPM dan UT |
| 10. Kalkulus Lanjutan II | - NPM dan UT |
| 11. Pengantar Psikologi | - NPM dan UT |
| 12. Teori Survei Sampel | - NPM dan NUT |

Keterangan :

1. PM = Penulis Modul
2. NPM = Non Penulis Modul
3. UT = Dari dalam Universitas Terbuka
4. NUT = Dari luar Universitas Terbuka

Data yang digunakan adalah merupakan hasil dari item analisis pada paket soal yang ditulis oleh penulis butir soal. Penulis butir soal tersebut diberi kode angka 1 s/d 12 sesuai dengan matakuliah yang dituliskannya. Tiap-tiap penulis soal menulis satu paket/set soal ujian berdasarkan bidang keahliannya dan berisi 30 butir soal.

Butir-butir soal tersebut dianalisa berdasarkan :

1. P
2. R_{bis}
3. Mean
4. Semua (P, R_{bis} dan Mean secara bersamaan)

Apabila suatu item memiliki nilai P yang lebih kecil dari 0.2 maka item tersebut dinilai terlampau sulit bagi mahasiswa dan apabila nilai P lebih besar dari 0.8 dinilai item tersebut terlalu mudah. Jadi suatu item dinilai baik apabila nilai P nya berada antara 0.2 dan 0.8.

Data besarnya persentase soal yang baik berdasarkan nilai P dari masing-masing penulis soal tampak tertera pada Tabel 1 kolom 3.

R_{bis} dari masing-masing matakuliah dicari dengan menggunakan rumus. Apabila suatu item memiliki R_{bis} yang sama atau lebih besar dari R_{bis} butir soal maka item tersebut dinilai baik.

Data besarnya persentase soal yang baik berdasarkan R_{bis} nya dari masing-masing penulis soal tampak tertera pada Tabel 1 kolom 4.

Mean suatu item dikatakan baik apabila nilai mean kunci lebih daripada nilai mean jawaban yang lainnya.

Data besarnya persentase soal yang baik berdasarkan mean dari masing-masing penulis soal tampak tertera pada Tabel 1 kolom 5.

Yang dimaksud dengan "semua" adalah nilai P, R_{bis} dan Mean dilihat secara bersamaan. "Semua itu dikatakan baik apabila nilai P, R_{bis} dan Mean memenuhi syarat seperti tersebut diatas.

Data besarnya persentase soal yang baik berdasarkan "semua" dari masing-masing penulis soal tampak tertera pada Tabel 1 kolom 2.

Tabel 1. Persentase item yang baik berdasarkan data hasil item analisis.

No.	Semua	P	R_{bis}	Mean
1.	0.43	0.867	0.70	0.80
2.	0.76	0.967	0.96	0.83
3.	0.56	0.83	0.70	0.667
4.	0.76	0.967	0.73	0.867
5.	0.76	0.93	0.86	0.50
6.	0.50	0.767	0.60	0.767
7.	0.76	0.93	0.93	0.73
8.	0.4	0.63	0.50	0.83
9.	0.3	0.867	0.70	0.53
10.	0.30	0.767	0.33	0.50
11.	0.26	0.80	0.43	0.567
12.	0.73	0.70	0.93	0.83

Data di atas dapat dikelompokkan berdasarkan status penulis soal, yaitu :

1. Penulis Modul (PM)

- a. dari dalam Universitas Terbuka (UT)
- b. dari luar Universitas Terbuka (NUT)

2. Non Penulis Modul (NPM)
 - a. dari dalam Universitas Terbuka (UT)
 - b. dari luar Universitas Terbuka (NUT)

Setelah itu data dianalisa dengan menggunakan Statistik Uji T dengan menggunakan software STATPACK.

PEMBAHASAN

Tingkat Signifikansi

Untuk menguji apakah suatu variabel menunjukkan hubungan yang signifikan dengan variabel yang lain, maka nilai T yang didapat melalui Uji T-statistic perlu dibandingkan dengan nilai T pada tabel. Dalam penelitian ini tingkat signifikan telah ditetapkan sebesar 0.05 maka dengan mengacu pada tabel T diperoleh nilai T sebesar 3.182.

Dengan demikian maka hubungan signifikan dapat dilihat dengan membandingkan antara nilai T hitung dan nilai T tabel. Jika nilai T lebih kecil atau sama dengan T tabel maka berarti tidak terdapat hubungan signifikan antara variabel tersebut. Sebaliknya jika nilai T lebih besar daripada nilai T tabel maka berarti terdapat hubungan signifikan antara variabel tersebut.

Analisa data dengan menggunakan Uji T-statistic didapat hasil seperti dibawah ini.

Tabel 2. Tingkat Signifikan untuk "Semua", P, R_{bis} dan Mean

No	T	Ttabel	Hasil
1. Semua			
- PM dan NPM	1.644	3.182	TS
- UT dan NUT	6.265*	3.182	S
2. P			
- PM dan NPM	2.554	3.182	TS
- UT dan NUT	3.607*	3.182	S
3. R _{bis}			
- PM dan NPM	1.225	3.182	TS
- UT dan NUT	2.533	3.182	TS
4. Mean			
- PM dan NPM	2.057	3.182	TS
- UT dan NUT	3.249*	3.182	S

Keterangan :

TS = tidak signifikan

S = signifikan

Pada faktor-faktor pengamatan "Semua", P, R_{bis} dan Mean, dari Tabel 2. Lampak bahwa pada faktor-faktor tersebut terdapat perbedaan antara penulis modul (PM) dan non-penulis modul (NPM) dari kualitas soal yang ditulis. Hal ini dapat disebabkan karena penulis modul (PM) ataupun non-penulis

modul (NPM) yang ditugaskan oleh FMIPA-UT untuk menulis soal adalah orang-orang yang benar-benar telah berpengalaman dalam bidangnya masing-masing.

Sedangkan untuk faktor "Semua", P, dan Mean dari Tabel 2. tampak bahwa pada faktor-faktor tersebut ada perbedaan antara orang dari dalam Universitas Terbuka (UT) dan dari luar Universitas Terbuka (NUT) dari kualitas soal yang ditulis.

Hal ini mungkin dapat disebabkan oleh :

- P untuk butir soal yang ditulis oleh orang dalam Universitas Terbuka banyak yang terletak di bawah 0.20 dan jarang yang berada pada/di atas 0.80, sedangkan untuk orang dari luar Universitas Terbuka P-nya sebagian besar berada pada $0.20 < P < 0.80$; sehingga butir soal yang ditulis oleh orang dalam Universitas Terbuka dapat dikategorikan sebagai soal yang sulit/susah.
- Mean kunci lebih kecil jika dibandingkan dengan Mean pada jawaban yang lainnya, hal ini dapat disebabkan karena pengecoh yang kurang baik atau salah kunci jawaban.
- Dari kedua hal tersebut di atas maka "Semua" untuk orang dari dalam Universitas Terbuka juga berbeda dengan orang dari luar Universitas Terbuka.

Tabel 3. Nilai Tengah untuk "Semua", P, R_{bis} dan Mean

No	Nilai Tengah
1. Semua	
- PM	0.62
- NPM	0.40
2. P	
- PM	0.91
- NPM	0.78
3. R_{bis}	
- PM	0.78
- NPM	0.60
4. Mean	
- PM	0.78
- NPM	0.61

Dari Tabel 3. tampak bahwa nilai tengah untuk "Semua", P, R_{bis} dan Mean dari penulis modul lebih besar dari non-penulis modul.

Hal ini mungkin dapat disebabkan oleh :

- pengecoh yang kurang baik
- kunci jawaban salah
- matakuliah yang diuji itu kurang dikuasai atau sulit

Sedangkan menurut Sentot (1987), sumber-sumber kekeliruan (error) dalam tes pada umumnya ada dua, yaitu :

1. Faktor Ekstrinsik (diluar tes)
2. Faktor Intrinsik (didalam tes sendiri)

Sumber kekeliruan intrinsik ini biasanya berhubungan dengan masalah penentuan soal dalam tes, misalnya tes hanya berisi sebagian kecil saja dari seluruh butir soal yang mungkin disusun dan dilaksanakan. Padahal seharusnya butir-butir soal dalam tes harus dapat mewakili semua bahan pelajaran yang telah diajarkan.

Dan hal ini dapat juga terjadi bagi penulis soal yang masih junior terutama bagi staf pengajar FMIPA-UT, karena sebagian besar dari mereka masih kurang ilmu pengetahuannya dalam hal penyusunan dan pengembangan soal, serta juga kebanyakan dari staf FMIPA-UT ini berasal dari sarjana non-kependidikan.

Analisis soal dengan memperhatikan jawaban pada soal dan setiap alternatif jawaban dapat digunakan untuk menyempurnakan soal itu sendiri atau memperbaiki soal pada waktu yang akan datang.

Analisis soal dapat dilakukan dengan cara dalam mengadakan untuk merevisi butir soal yang ada atau menambahkan kuantitas soal yang ada.

VI. KESIMPULAN DAN SARAN

Kesimpulan

Atas dasar pembahasan dan hasil yang diperoleh seperti tersebut di atas maka dapatlah disimpulkan sebagai berikut :

1. penulis modul dan non-penulis modul dengan kualifikasi yang sama, akan menyebabkan tidak ada perbedaan dalam kualitas soal yang ditulis, sehingga terbuka kesempatan bagi FMIPA-UT khususnya untuk menggunakan tenaga pengajar yang tidak menulis modul untuk menulis soal ujian.
2. orang dari dalam Universitas Terbuka dan dari luar Universitas Terbuka terdapat perbedaan yang signifikan pada "Semua", P, dan Mean dalam kualitas soal yang ditulis, sehingga untuk memanfaatkan tenaga pengajar junior ini perlu diadakan peningkatan kemampuan dalam penulisan soal.

Saran-saran

Setelah melakukan, menggunakan dan membahas penelitian ini untuk langkah lebih lanjut ada beberapa hal yang dapat disarankan berkaitan dengan pemanfaatan tenaga edukatif di dalam penulisan soal yakni :

1. Tenaga edukatif lebih dahulu diberikan kesempatan untuk magang atau mengikuti kuliah singkat mengenai pengembangan dan penulisan soal. Dengan demikian mereka sedikit banyaknya mengetahui bagaimana suatu soal itu dapat dikatakan baik dan benar.

2. Untuk tingkat permulaan tenaga odukaurif ditugaskan terlebih dahulu untuk merakit/menulis soal tugas mandiri (TM) kemudian kalau sudah terbiasa atau soal yang ditulis itu sudah dapat dikatakan baik, barulah tenaga odukaurif itu diberikan kepercayaan untuk merakit bahkan kalau perlu untuk menulis soal ujian.

UNIVERSITAS TERBUKA

DAFTAR PUSTAKA

- Furhanda, H.D.K. 1984. "Testing And Measurement". Jakarta: National Education Planning, Evaluation And Curriculum Development.
- Gronlund, Norman, E. 1981 "Measurement And Evaluation in Teaching" (4th Edition). New York: McMillan.
- Nasection, Nochi. 1989. "Mengembangkan Alat Evaluasi Hasil Belajar". Jakarta: Makalah Dalam Lokakarya Pemanfaatan Kisi-Kisi Soal, Guna Percepatan Ujian Negara; Dalam Rangka Operasional Pendidikan Dan Pengajaran, Di Lingkungan Koperlis Wilayah III.
- Noer, Mohammad. 1987. "Pengantar Teori Tes". Jakarta: Proyek Pengembangan Lembaga Pendidikan Tenaga Kependidikan, Ditjen DIKTI, DEPDIKBUD.
- Subino. 1987. "Konstruksi Dan Analisis Tes". Jakarta: Proyek Pengembangan Lembaga Pendidikan Tenaga Kependidikan, Ditjen DIKTI, DEPDIKBUD.

LAMPIRAN

UNIVERSITAS TERBUKA

ANALISA "SEMUA" SOAL UJIAN

PENULIS DAN NON PENULIS MODUL

pen modul (PM)

Mean	= 0.62
Variance	= 0.02
Standard deviation	= 0.13
Standard error of the mean	= 0.05

non pen modul (NPM)

Mean	= 0.40
Variance	= 0.04
Standard deviation	= 0.19
Standard error of the mean	= 0.11

T-Test statistics

Difference (Mean X - Mean Y)	= 0.222
Standard error of the difference	= 0.135
t - statistic	= 1.644
Degrees of freedom	= 3
Probability of t (One tailed test)	= 0.099
Probability of t (Two tailed test)	= 0.198

Correlation coefficient = 0.483

Valid cases	= 4
Missing cases	= 4
Response percent	= 50 %

ANALISA "SEMUA" SOAL UJIAN

DALAM DAN LUAR UNIVERSITAS TERBUKA

dalam UT (UT)

Mean	= 0.32
Variance	= 0.00
Standard deviation	= 0.00
Standard error of the mean	= 0.04

luar UT (NUT)

Mean	= 0.70
Variance	= 0.01
Standard deviation	= 0.08
Standard error of the mean	= 0.05

T-Test statistics

Difference (Mean X - Mean Y)	= -0.380
Standard error of the difference	= 0.061
t - statistic	= 6.265
Degrees of freedom	= 3
Probability of t (One tailed test)	= 0.003
Probability of t (Two tailed test)	= 0.007

Correlation coefficient = 0.242

Valid cases	= 4
Missing cases	= 4
Response percent	= 50 %

ANALISA P SOAL UJIAN

PENULIS DAN NON PENULIS MODUL

pen modul (PM)

Mean = 0.91
Variance = 0.00
Standard deviation = 0.06
Standard error of the mean = 0.03

non pen modul (NPM)

Mean = 0.78
Variance = 0.00
Standard deviation = 0.06
Standard error of the mean = 0.03

T-Test statistics

Difference (Mean X - Mean Y) = 0.125
Standard error of the difference = 0.049
t - statistic = 2.554
Degrees of freedom = 3
Probability of t (One tailed test) = 0.041
Probability of t (Two tailed test) = 0.083

Correlation coefficient = -.735

Valid cases = 4
Missing cases = 4
Response percent = 50 %

ANALISA P SOAL UJIAN

DALAM DAN LUAR UNIVERSITAS TERBUKA

dalam UT (UT)

Mean	= 0.71
Variance	= 0.00
Standard deviation	= 0.00
Standard error of the mean	= 0.00

Tuar UT (NUT)

Mean	= 0.91
Variance	= 0.00
Standard deviation	= 0.00
Standard error of the mean	= 0.00

T-Test statistics

Difference (Mean X - Mean Y)	= -0.184
Standard error of the difference	= 0.051
t - statistic	= 3.607
Degrees of freedom	= 3
Probability of t (One tailed test)	= 0.017
Probability of t (Two tailed test)	= 0.035

Correlation coefficient = -.066

Valid cases	= 4
Missing cases	= 4
Response percent	= 50 %

ANALISA R_{bis} SOAL UJIAN

PENULIS DAN NOI. PENULIS MODUL

pen modul (PM)

Mean	= 0.78
Variance	= 0.01
Standard deviation	= 0.11
Standard error of the mean	= 0.06

non pen modul (NPM)

Mean	= 0.60
Variance	= 0.06
Standard deviation	= 0.23
Standard error of the mean	= 0.14

T-Test statistics

Difference (Mean X - Mean Y)	= 0.183
Standard error of the difference	= 0.149
t - statistic	= 1.225
Degrees of freedom	= 3
Probability of t (One tailed test)	= 0.154
Probability of t (Two tailed test)	= 0.308

Correlation coefficient = -.494

Valid cases	= 4
Missing cases	= 4
Response percent	= 50 %

ANALISA R_{bis} SOAL UJIAN

DALAM DAN LUAR UNIVERSITAS TERBUKA

dalam UT (UT)

Mean	= 0.54
Variance	= 0.03
Standard deviation	= 0.16
Standard error of the mean	= 0.09

luar UT (LUT)

Mean	= 0.82
Variance	= 0.01
Standard deviation	= 0.10
Standard error of the mean	= 0.06

T-Test statistics

Difference (Mean X - Mean Y)	= -0.280
Standard error of the difference	= 0.111
t - statistic	= 2.533
Degrees of freedom	= 3
Probability of t (One tailed test)	= 0.042
Probability of t (Two tailed test)	= 0.084

Correlation coefficient = 0.176

Valid cases = 4

Missing cases = 4

Response percent = 50 %

ANALISA MEAN SOAL UJIAN

PENULIS DAN NON PENULIS MODUL

pen modul (PM)

Mean	= 0.78
Variance	= 0.00
Standard deviation	= 0.07
Standard error of the mean	= 0.04

non pen modul (NPM)

Mean	= 0.61
Variance	= 0.02
Standard deviation	= 0.13
Standard error of the mean	= 0.08

T-Test statistics

Difference (Mean X - Mean Y)	= 0.175
Standard error of the difference	= 0.085
t - statistic	= 2.057
Degrees of freedom	= 3
Probability of t (One tailed test)	= 0.066
Probability of t (Two tailed test)	= 0.131

Correlation coefficient = 0.249

Valid cases	= 4
Missing cases	= 4
Response percent	= 50 %

ANALISA MEAN SOAL UJIAN

DALAM DAN LUAR UNIVERSITAS TERBUKA

dalam UT (UT)

Mean	= 0.60
Variance	= 0.01
Standard deviation	= 0.10
Standard error of the mean	= 0.07

luar UT (LUT)

Mean	= 0.82
Variance	= 0.00
Standard deviation	= 0.01
Standard error of the mean	= 0.01

T-Test statistics

Difference (Mean X - Mean Y)	= -0.223
Standard error of the difference	= 0.069
t - statistic	= 3.249
Degrees of freedom	= 3
Probability of t (One tailed test)	= 0.023
Probability of t (Two tailed test)	= 0.046

Correlation coefficient = 0.338

Valid cases = 4

Missing cases = 4

Response percent = 50 %



UNIVERSITAS TERBUKA