

LAPORAN KEGIATAN

**SEMINAR PENYUSUNAN BUKU PANDUAN
PENILAIAN HASIL BELAJAR MAHASISWA
IKIP PADANG**

MILIK UPT PERPUSTAKAAN
IKIP PADANG



PERPUSTAKAAN	PERPUSTAKAAN IKIP PADANG
TANGGAL	10-7-95
SUMBER BAHAN	hs
KOLEKSI	KKI
NOMOR	1360 (hs) 195-52(2)
LOKASI	378 ins ①

**PUSAT TESTING DAN EVALUASI PENDIDIKAN
INSTITUT KEGURUAN DAN ILMU PENDIDIKAN
PADANG
1990**

KATA PENGANTAR

Pusat Penilaian dan Evaluasi Hasil Belajar Mahasiswa IKIP Padang adalah suatu lembaga non-struktural pada tingkat institut. Pusat ini berfungsi untuk mengkaji dan mengembangkan sistem penilaian dan evaluasi hasil belajar mahasiswa.

Salah satu prakarsa yang telah dilakukan oleh pusat ini adalah melaksanakan seminar untuk mengumpulkan bahan bagi penyusunan Buku Panduan Penilaian Hasil Belajar Mahasiswa yang diselenggarakan pada tanggal 22 dan 23 Januari 1990, di Aula FPTK IKIP Padang.

Seminar ini menampung semua gagasan, inspirasi, dan perbedaan pandangan yang diakibatkan oleh hakekat bidang studi setiap jurusan dan fakultas yang ada dalam lingkungan IKIP Padang.

Seminar tersebut diikuti oleh wakil-wakil jurusan dan fakultas dalam lingkungan IKIP Padang serta para pakar berbagai disiplin ilmu di IKIP Padang. Enam orang pakar IKIP Padang menyampaikan makalahnya dalam melakukan tinjauan dari berbagai sudut bidang ilmunya tentang evaluasi hasil belajar mahasiswa.

Kesimpulan hasil seminar ini dihimpun dalam bentuk laporan yang diharapkan dapat diharapkan menjadi salah satu masukan bagi suatu panitia adhock yang akan ditugaskan untuk menyusun Buku Panduan Penilaian Hasil Belajar Mahasiswa IKIP Padang.



Padang, 7 April 1990

Rektor,

Prof. H. Jamil Bahar
NIP. 130 428 395

DAFTAR ISI

KATA PENGANTAR	i
DAFTAR ISI	ii
I. PENDAHULUAN	1
A. Latar Belakang	1
B. Tujuan	3
C. Hasil Akhir yang Diharapkan	4
D. Gambaran Pelaksanaan Seminar	4
II. RINGKASAN DISKUSI	7
A. Judul Makalah Penilaian Aspek Kognitif	7
B. Judul Makalah Penilaian Ranah Afektif	11
C. Judul Makalah Penilaian Hasil Belajar di Bengkel Kerja (Psikomotor)	14
D. Judul Makalah Pertimbangan Penentuan Skor dan Nilai Akhir Semester	16
E. Judul Makalah Penilaian Belajar di Laboratorium	18
F. Judul Makalah Penilaian Belajar Bahasa	19
G. Judul Makalah Penilaian Belajar Seni, Musik, dan Tari.	22
H. Judul Makalah Beberapa Saran Tentang Penilaian Hasil Belajar Perkuliahan Praktek Olahraga di FPOK IKIP Padang	23

III. LAMPIRAN	25
A. Susunan Kepanitiaan	25
B. Peserta Seminar	27
C. Makalah-makalah Seminar	30

MILIK UPT PERPUSTAKAAN IKIP PADANG
DITERIMA TGL
SUMBER/HARGA
KOLEKSI
No INVENTARIS
CALL NO

I. PENDAHULUAN

MILIK UPT PERPUSTAKAAN
IKIP PADANG

A. Latar Belakang

Undang-undang No. 2 Tahun 1989, Bab XII, Pasal 43 mengisyaratkan bahwa penilaian terhadap kegiatan dan kemajuan belajar para peserta didik merupakan bagian yang integral dari proses belajar mengajar. Pertimbangan ini telah ditampung dalam Buku Pedoman IKIP Padang 1989 - 1990 yang mencantumkan bahwa sistem penilaian hasil belajar mahasiswa dapat didasarkan kepada berbagai aspek, seperti kehadiran dalam kuliah, keaktifan di dalam kelas, tugas terstruktur, dan melalui ujian (Buku Pedoman IKIP Padang, 1989, hal. 38 - 39). Penentuan "nilai akhir keberhasilan mahasiswa" kemudian ditetapkan berdasarkan dua model, PAN atau PAP. Kedua model ini memberi pengertian yang berbeda-beda bagi setiap dosen. Perbedaan pengertian dan persepsi tentang kedua model ini akan menghasilkan "nilai" yang beragam makna dan arti.

Walaupun disadari bahwa "penyamaan" makna dan arti "nilai" bagi setiap dosen tidak mungkin dilaksanakan, usaha untuk mempersempit perbedaan persepsi tersebut perlu diusahakan. Oleh sebab itu, Rektor melalui PR I mengambil inisiatif dan menugaskan Pusat Testing dan Evaluasi Pendidikan IKIP Padang untuk mengadakan seminar menuju usaha penyusunan Buku Panduan Penilaian Hasil Belajar Mahasiswa IKIP Padang. Buku Panduan ini

diharapkan dapat menampung semua gagasan, inspirasi, dan perbedaan hakekat bidang studi setiap Jurusan dan Fakultas yang ada di IKIP Padang. Gagasan dan inspirasi tersebut dapat dirumuskan dan menjadi masukan dalam melaksanakan evaluasi hasil belajar (dan kegiatan) mahasiswa. Seperti tersebut di atas bahwa setiap Jurusan dan Fakultas memiliki variabilitas pengertian dan hakekat bidang studi masing-masing. Oleh karena itu, perlu dihimpun pendapat atau konsep mereka tentang penilaian hasil belajar mahasiswa untuk menyusun gagasan awal. Gagasan awal ini akan disampaikan dalam makalah yang akan mereka kemukakan pada suatu seminar yang khusus diadakan untuk ini, dan hasil seminar ini kemudian dihimpun dan dirumuskan oleh suatu panitia kecil.

Berdasarkan hal tersebut, ada dua kegiatan yang perlu dilakukan untuk penyusunan Buku Panduan Penilaian Hasil Belajar tersebut, yaitu:

- 1) Seminar, yang melibatkan staf pengajar jurusan dalam jumlah yang cukup representatif (dua wakil tiap jurusan).
- 2) Penyusunan kesimpulan seminar untuk menjadi Buku Panduan Penilaian Hasil Belajar oleh panitia ad-hoc.

Dengan kedua kegiatan ini akan dihimpun semua gagasan dan inspirasi semua jurusan/fakultas dalam penyusunan Buku Panduan tersebut. Sama halnya dengan Buku Panduan Penulisan Tesis, Buku Panduan ini akan disahkan

Senat IKIP Padang sebelum diberlakukan bagi semua fakultas/jurusan di IKIP Padang.

B. Tujuan

a. Seminar bertujuan untuk:

- 1) Mengumpulkan pendapat dan gagasan dari semua jurusan terhadap gagasan yang disampaikan oleh pemakalah (wakil fakultas).
- 2) Melibatkan semua unsur jurusan dan fakultas agar kekhasan masing-masing jurusan dan fakultas dapat dijangkau dan dengan demikian mereka merasa terlibat dalam penyusunan Buku Panduan Penilaian Hasil Belajar tersebut.
- 3) Memungkinkan Buku Panduan yang dirumuskan dapat diaplikasikan dengan partisipasi jurusan, sesuai dengan tanggung jawab mereka sebagai peserta seminar.

b. Panitia ad-hock dibentuk untuk:

- 1) Merumuskan hasil seminar tersebut dan menyusunnya menjadi draft Buku Panduan Penilaian Hasil Belajar Mahasiswa IKIP Padang.
- 2) Menyunting draft Buku Panduan tersebut sehingga dapat dipedomani oleh setiap dosen sesuai dengan keperluan kekhasan mata kuliah yang dibinanya.

- 3) Menyiapkan draft Buku Panduan Penilaian Hasil Belajar yang siap untuk dikirimkan ke Senat IKIP Padang untuk disahkan.

C. Hasil Akhir yang Diharapkan

Dari kedua kegiatan di atas diharapkan tersusun Buku Panduan Penilaian Hasil Belajar Mahasiswa IKIP Padang yang dapat menjadi pedoman penilaian bagi semua dosen IKIP Padang. Buku Panduan ini hendaknya memuat pedoman untuk menilai aspek-aspek pengajaran teori, hasil belajar mahasiswa, praktek, atau efektif (bila diperlukan). Di samping itu, Buku panduan ini juga memberi peluang bagi dosen yang ingin memakai prosedur PAN ataupun PAP sesuai dengan keyakinannya masing-masing, tanpa dirancukan dengan perbedaan konsep (antara PAN dan PAP).

D. Gambaran Pelaksanaan Seminar

Seminar dilaksanakan dua hari penuh berturut-turut mulai dari tanggal 22 Januari 1990 s.d 23 Januari 1990. Dalam waktu dua hari tersebut telah berbicara enam orang pemakalah dengan bahasan yang berbeda. Perinciannya adalah sebagai berikut:

Pemakalah

J u d u l

Drs. A. Muri Yusuf, M.Pd.

Penilaian Aspek Kognitif

Dr. Abizar

Penilaian Ranah Afektif

- Dr. Aljufri B Syarif, M.Sc. Penilaian Hasil Belajar
di Bengkel Kerja
(Psikomotorik)
- Dr. Kumaidi, MA. Pertimbangan Penentuan
Skor dan Nilai Akhir
Semester
- Dr. Ibenzani Usman Penilaian Belajar Seni,
Musik dan Tari
- Drs. Rusdy Rusyd, MA. Penilaian Belajar di Labor
- Dr. Zainil, MA. Penilaian Belajar Bahasa
- Dr. Yanuar Kiram Beberapa Saran Tentang
Penilaian Hasil Belajar
Perkuliahan Praktek
Olahraga di FPOK IKIP
Padang

Seminar dihadiri oleh para pakar pendidikan, wakil/utusan dari fakultas (masing-masing 2 orang), wakil dari jurusan (masing-masing 1 orang), dan mahasiswa program S2 di lingkungan IKIP Padang yang keseluruhannya berjumlah ± 100 orang. Dengan diundang dan hadirnya orang-orang di atas dimaksudkan agar rumusan Buku Panduan tersebut telah dibicarakan oleh banyak orang, sehingga buku tersebut merupakan produk IKIP Padang bukan oleh segelintir orang. Suasana seminar mencerminkan sikap ilmiah. Hal ini terlihat dari setiap tahapan pemakalah selalu mendapat tanggapan dari peserta seminar. Tanggapan

yang diberikan bervariasi, ada yang menyetujui, memodifikasi, bahkan menolak sama sekali model penilaian yang diajukan pemakalah. Semua penanggap memberikan alasannya masing-masing. Karena terbatasnya waktu, maka setiap tahapan hanya memberikan maksimum dua "termin" dengan jumlah penanggap rata-rata lima orang tiap "termin".

Semua pembicaraan seminar direkam yang dimaksudkan agar pembicaraan dan ide-ide yang muncul dapat terjaring untuk dijadikan masukan berharga bagi Panitia Ad-hoc. Peserta seminar berusaha membicarakan semua bentuk penilaian yang mungkin dilaksanakan pada ketiga ranah (kognitif, afektif, dan psikomotorik) di masing-masing fakultas. Pembicaraan diarahkan untuk mencari kesamaan-kesamaan aspek penilaian pada masing-masing fakultas.

Dari tujuan dan analisa sepintas, kelihatannya untuk masing-masing aspek penilaian pada ranah yang sama terdapat kesamaan, sehingga usaha simplifikasi dan standardisasi akan tercapai. Namun terdapat model evaluasi hasil belajar untuk beberapa mata kuliah yang belum dibicarakan dalam seminar dan kelihatannya mempunyai karakteristik yang berbeda. Panitia Ad-hoc akan berusaha memperolehnya melalui pakar mata kuliah bersangkutan.

II. RINGKASAN DISKUSI

A. Judul Makalah : Penilaian Aspek Kognitif
Pembawa makalah : Drs. A. Muri Yusuf, M.Pd.

1. Isi Makalah

Makalah ini membicarakan penilaian hasil belajar aspek kognitif yang mencakup :

- a. Konsep pengukuran dan penilaian
- b. Syarat-syarat tes yang baik
- c. Kisi-kisi tes
- d. Pembuatan tes hasil belajar
- e. Skoring
- f. Pengembangan tes hasil belajar.

Pengukuran berbeda dengan penilaian. Pengukuran adalah deskripsi kuantitatif, sedangkan penilaian adalah deskripsi kuantitatif, kualitatif dan pemberian makna.

Tes adalah kumpulan contoh butiran soal yang jawabnya dapat dinyatakan dengan benar-salah. Syarat tes yang baik adalah:

- a. Valid, yaitu mengukur apa yang mesti diukur
- b. Reliabel, yaitu konsisten dan teliti dalam mengukur apa yang mesti diukur. Reliabilitas dipengaruhi beberapa faktor antara lain:
 - 1) Konstruksi item
 - 2) Panjang/pendeknya item

- 3) Evaluasi
 - 4) Ketidak tetapan waktu yang diberikan
 - 5) Kemampuan yang ada dalam kelompok
 - 6) Luas/tidaknya cakupan tes
 - 7) Kondisi dan situasi pada pengadmnistrsian alat ukur.
- c. Objektivitas, yaitu konsistensi skor apabila dinilai oleh beberapa orang yang berbeda.
- d. Praktikalitas, yaitu:
- 1) Biaya tidak tingi
 - 2) Mudah diadministrasikan
 - 3) Mudah diskor
 - 4) Mudah diinterpretasikan
 - 5) Waktu tidak terlalu lama.
- e. Norma
- Tes penguasaan belajar perlu dirancang secara baik dan mempunyai hubungan erat dengan tujuan pengajaran, menurut Bloom, ada enam kemampuan yang terdapat pada kawasan kognitif, yaitu:
- 1) Pengetahuan
 - 2) Pemahaman
 - 3) Aplikasi
 - 4) Analisis
 - 5) Sintesis
 - 6) Evaluasi
- Agar tujuan pengajaran tertuang pada item butir

soal, perlu dibuat kisi-kisi tes. Hal yang perlu diperhatikan pada pengembangan kisi tes adalah:

a. Informasi umum meliputi:

- 1) Mata kuliah
- 2) Fakultas
- 3) Jurusan
- 4) Waktu pelaksanaan

b. Tipe/bentuk soal yang digunakan

c. Ruang lingkup dan sub pokok bahasan

d. Tingkat kesukaran

e. Jumlah soal tiap pokok bahasan ruang lingkup ujian

f. Jumlah total butir soal

g. Persentase tiap pokok bahasan (ruang lingkup).

Kemampuan bersifat pengetahuan dapat diukur dengan butir soal:

- a. Melengkapi (completion)
- b. Jawaban pendek/melengkapi (supply item)
- c. Menjodohkan (matching item)
- d. Pilihan jamak (multiple choice)
- e. Betul-salah (true-false)

Kemampuan yang bersifat pemahaman lebih baik diukur dengan tipe tes pilihan yang berbentuk pilihan jamak atau kasus.

Kemampuan berbentuk aplikasi, analisis, sintesis dan evaluasi dapat menggunakan pilihan jamak, analisis kasus dan tes essay.

Untuk skoring sebaiknya digunakan rumus:

$$S = B - \frac{W}{n - 1}$$

Untuk menentukan batas lulus dapat digunakan PAP atau PAN.

Pada metoda PAP digunakan persamaan:

$$\frac{\text{Skor Mentah}}{100} \times \text{Skor Maksimal}$$

Pada metode PAN digunakan aturan sebagai berikut:

	> x + 2,00 SD	A	lulus
antara	x + 1,00 dan		
	x + 2,00 SD	B	lulus
antara	x dan x + 1,00 SD	C	lulus
antara	x - 1,00 SD dan x	D	lulus
	< x - 1,00 SD	E	tidak lulus

Untuk menyempurnakan item soal, perlu dilakukan analisis butir soal dengan menggunakan persamaan:

$$TK = \frac{J S B}{J S T}$$

$$DS = \frac{Jp + Jk}{Np + Nk} \times 100 \%$$

$$PP = \frac{Jp}{Np} - \frac{Jk}{Nk} = Pp - Pk$$

2. Pembahasan

Dalam penilaian perlu dilakukan "Self evaluation" sebagai upaya umpan balik.

Dipertanyakan mengapa nilai huruf yang dipakai rentangan yang tidak sama. Agar dipertimbangkan juga nilai + dan nilai - pada masing-masing nilai huruf yang berlaku saat ini.

Agar butir soal representatif terhadap tujuan pengajaran, kisi-kisi dikelompokkan dengan cara yang benar. Muncul isu selama seminar, bahwa sedikit sekali staf pengajar yang membuat kisi-kisi tes hasil belajar. Hal ini perlu dicarikan pemecahannya.

Pada soal-soal yang diberikan, perlu ada keterangan tentang pembobotan.

B. Judul Makalah : Penilaian Ranah Afektif

Pembawa makalah : Dr. Abizar

1. Isi Makalah

Makalah ini menyatakan bahwa Ranah Afektif merupakan bagian dari taksonomi belajar Bloom. Ranah afektif yang dimaksud pada makalah ini mencakup konsep-konsep nilai dan sikap, moral dan etik, perkembangan diri, emosi dan perasaan. Di IKIP, mata kuliah yang mempunyai proporsi amat besar terhadap sikap adalah mata kuliah DUM (MKDU).

Ada banyak pendapat tentang sikap yang telah

dikeluarkan oleh para ahli, namun pada makalah ini penulis mengambil pendapat Fishien & Ajzen. Menurut Fishien, sikap mempunyai empat komponen, yaitu: keyakinan (belief), sikap (attitude), maksud (intention), dan tingkah laku (behaviour). Keempatnya merupakan sistem.

Hubungan antara sikap dengan keyakinan dinyatakan dengan formula :

$$A = f \Sigma b e$$

Antara sikap dengan maksud berbuat:

$$B - I = f (W1A + W2SN)$$

Antara norma subjektif dengan keyakinan kelompok dan motivasi:

$$SN = F (\Sigma b m)$$

Dalam Pengukuran sikap, TIK merupakan acuan. Taksonomi afektif yang digunakan adalah yang berasal dari Krathwol yang lebih dikenal dengan prinsip internalisasi yaitu: Penerimaan, merespon, penilaian, pengorganisasian dan karakterisasi. Diakui bahwa taksonomi efektif Krathwol bukan tanpa cacat, sebab pengukuran ranah afektif belum semantap pengukuran ranah kognitif, terutama dalam pengembangan alat ukur (hasil belajar)nya.

Untuk pengukuran sikap dapat dilakukan pendekatan sebagai berikut:

- a. Laporan diri (Self Report).
- b. Laporan dari orang lain

Atribut-atribut yang akan diukur perlu diisolasi.

Pada pembuatan butir pertanyaan atau butir pernyataan yang akan direspon murid dapat digunakan matriks dua faktor. Dari matrik ini, butir-butir pernyataan dan pertanyaan dapat dikembangkan.

Pada strategi self Report, skala yang digunakan umumnya adalah: skala Guttman, Thurstone, Likert, dan Semantic Differensial. Jika teknik observasi, digunakan lembaran observasi dan observasi dilakukan berulang-ulang.

2. Pembahasan

Dipermasalahan apakah pengukuran sikap untuk mata kuliah MKDU dapat dilakukan. Sebab, tujuan mata kuliah adalah perubahan sikap, yang diuji biasanya adalah kemampuan kognitifnya.

Pada mata kuliah non MKDU, antara ranah Kognitif dan ranah Afektif terdapat bentuk simultan. Perlu juga dipertimbangkan modifikasi Skala Likert dari ganjil menjadi genap mengingat kondisi sosial-budaya bangsa Indonesia. Patut dipertimbangkan unsur kode etik guru dalam butir-butir pertanyaan/pernyataan yang dibuat.

3. Saran

Untuk melihat apakah perubahan sikap memang diakibatkan proses belajar mengajar, dapat dilakukan analisis pre-postest.

C. Judul Makalah : Penilaian Hasil Belajar di Bengkel Kerja (Psikomotor)

Pembawa Makalah : Dr. Aljufri B Syarif, M.Sc.

1. Isi Makalah

Penilaian hasil belajar di bengkel kerja mencakup tiga aspek, yaitu aspek ranah kognitif, aspek ranah afektif dan aspek ranah psikomotor.

Menurut More, langkah-langkah yang perlu dilakukan dalam pengembangan keterampilan sebagai produk hasil belajar adalah menetapkan:

- a. langkah kerja
- b. kompetensi yang terkait
- c. derajat kesukaran pekerjaan
- d. acuan kegiatan
- e. kemampuan belajar untuk menyesuaikan diri dengan lingkungan kerja
- f. kebutuhan fisik

Makalah ini membagi jenis kegiatan atas tiga jenis, yaitu : bekerja dengan alat, bekerja dengan data, dan bekerja manusia. Sesuai dengan tujuannya, (IKIP) yang menghasilkan guru teknik, produk ini akan

lebih banyak bekerja dengan manusia. Selanjutnya ketiga jenis pekerjaan tersebut dituangkan melalui jabarannya dan diakomodasikan pada bengkel kerja.

Penilaian yang dianjurkan adalah menggunakan Sistem Batas Toleransi.

Untuk menentukan nilai akhir, semua aspek (kognitif, afektif dan psikomotor) harus diperhitungkan. Masing-masing komponen yang dinilai diberi bobot dan selanjutnya digunakan persamaan regresi linier.

2. Pembahasan

Perlu penjabaran "Construct" pada aspek kognitif, afektif, dan psikomotorik pada tiap-tiap "jop" supaya terdapat peniaian yang relatif semua melalui kegiatan "analisis skill" oleh guru.

Di bengkel kerja, penilaian sebaiknya "Go or Not Go" Sedangkan di IKIP dikenal nilai A, B, C, D, E. Hal ini perlu diusahakan penggabungannya.

Dipermasalahan juga mengapa makalah tidak membicarakan taksonomi hasil belajar. Pembobotan pada masing-masing aspek penilaian ditentukan oleh dosen mata kuliah bersangkutan bersama rekan-rekan sejawat. Perlu juga dipertimbangkan mahasiswa melakukan "self evaluation"

3. Saran

Dalam menentukan nilai akhir, digunakan rata-rata nilai yang diperoleh dari "continuous assesment". Penilaian dilakukan dengan menggunakan lembaran observasi.

D. Judul Makalah : Pertimbangan Penentuan Skor dan Nilai Akhir Semester

Pembawa Makalah : Dr. Kumaidi, MA.

1. Isi Makalah

Makalah ini membahas tentang bagaimana caranya menentukan besarnya skor jika perlu menentukan nilai akhir dari beberapa skor yang harus digabungkan.

Untuk menentukan skor komposit (gabungan), cara yang paling sederhana adalah dengan model regresi.

Jika model pembobotan yang digunakan, bobot tersebut ditentukan oleh dosen yang bersangkutan bersama rekan-rekan sejawat dengan pertimbangan yang cukup matang.

Secara teoritis skor komposit selalu dipengaruhi oleh skor komponennya. Dengan demikian skor atau bobot yang diberikan harus mencerminkan tujuan pengukuran dan makna nilai akhir yang dituju secara cermat.

Penetapan nilai akhir perlu memperhatikan makna nilai dan menghubungkannya dengan sistem acuan yang dipakai. Dalam penentuan nilai akhir ada tiga

alternatif yaitu metode PAP dan PAN serta gabungan keduanya. Untuk ranah psikomotor sebaiknya digunakan metode PAP, sedangkan untuk ranah afektif dan kognitif sebaiknya PAN atau gabungan PAP dan PAN.

2. Pembahasan

Jika model acuan penilaian menggunakan gabungan PAP dan PAN, maka butir soal juga ikut terpengaruh. Sebagian butir soal yang pernah diberikan pada mahasiswa tahun sebelumnya, diberikan kembali kepada kelompok baru. Hal ini seakan-akan mengacu kepada bentuk konstan atau adanya usaha standardisasi. Untuk ilmu-ilmu sosial ini menjadi masalah, sebab pada bidang sosial ilmu tersebut berkembang terus.

Apakah penggabungan PAP dan PAN ada formulasinya. Jika ada bagaimana caranya. Atau bukankah yang dikatakan gabungan PAP dan PAN itu hanyalah bentuk PAP saja.

Dalam menentukan sistem penilaian operasional perlu dipikirkan faktor yang akan mengganggu, misalnya faktor "decision maker".

E. Judul Makalah : Penilaian Belajar di Laboratorium

Pembawa Makalah : Drs. Rusdy Rusyd, MA.

1. Isi Makalah

Makalah ini membahas komponen-komponen yang harus dinilai dalam kegiatan belajar di laboratorium. Penilaian belajar di laboratorium mempunyai spesifikasi tersendiri sehingga tidak dapat begitu saja mengikuti pola pada buku pedoman IKIP. Penilaian harus bermakna bagi mahasiswa, dosen dan bagi institut.

Kegiatan belajar di laboratorium terdiri dari komponen.

- a. Membuat persiapan,
- b. Melaksanakan eksperimen,
- c. Menulis laporan.

Walaupun ada 3 komponen kegiatan belajar yang menonjol di laboratorium, tetapi untuk menentukan/menetapkan nilai mahasiswa, diperlukan data yang terdiri dari: persiapan, laporan dan ujian semester. Sedangkan proses tidak dinilai.

Dalam penentuan nilai, masing-masing komponen diberi pembobotan, dengan model persamaan:

$$N = \frac{20 P + Q + 40 R}{100}$$

dimana P nilai persiapan, Q nilai laporan, dan R nilai ujian semester disesuaikan dengan kriteria yang dipakai oleh buku pedoman IKIP Padang.

2. Pembahasan

Persamaan penilaian yang disajikan hanya mengolah nilai akhir. Bagaimana penilaian pada masing-masing topik.

Model penilaian yang disajikan mengabaikan proses praktikum. Mestinya proses dinilai dan diberi bobot tertinggi. Laporan diberi bobot terendah, sebab terdapat kemungkinan mahasiswa mencontoh laporan orang lain. Ujian praktikum tidak perlu dilakukan, cukup dengan melakukan "continuous assesment".

Perlu juga dipertimbangkan "self-evaluation". Pembobotan jangan dimulai dari 0 (nol), sebaiknya dimulai dari 10.

F. Judul Makalah : Penilaian Belajar Bahasa

Pembawa Makalah : Dr. Zainil, MA.

1. Isi Makalah

Makalah ini membahas upaya standarisasi penilaian bahasa. Upaya ini mencakup perencanaan pengembangan tes/instrumen pengukuran keterampilan berbahasa sampai dengan penskoran (scoring), format yang diperlukan dan pertimbangan penetapan skor.

Penilaian skill berbahasa terdiri dari empat bagian, yaitu; Listening (menyimak), Reading (membaca), Writing (menulis), dan Speaking (berbicara).

Penilaian belajar bahasa dapat dilakukan dengan Achievement test, Proficiency test, Diagnostic Language test, dan Aptitude test.

Komponen yang harus diperhatikan adalah komponen keterampilan, komponen informasi dan komponen teori.

Dalam pembuatan test, hendaklah diperhatikan faktor reliabilitas, validitas, dan kepraktisan.

Pada jurusan-jurusan non bahasa, tujuan pengajaran adalah meningkatkan kemampuan membaca sesuatu yang ditulis bahasa Inggris dalam bahasa Inggris. Oleh sebab itu instrumen harus dititikberatkan pada kemampuan membaca. Ada beberapa pendapat tentang tingkatan pertanyaan, diantaranya yang dikemukakan oleh Herber, yaitu literal, interpretiveness dan applied reading comprehension questions.

Panjang wacana yang diberikan pada waktu tes hendaknya berkisar 100 s/d 250 kata. Dalam menentukan besarnya skor dapat dipedomani LAS yang sudah dimodifikasi.

Untuk penentuan nilai akhir dan pemberian label, angka/jumlah skor disesuaikan dengan Buku Pedoman IKIP

Padang. Untuk penentuan nilai akhir tersebut harus juga diperhitungkan kehadiran, pekerjaan terstruktur (home work), partisipasi dan paper.

2. Pembahasan

Kenyataan yang ada antara bahasa dan sastra disatukan, pada hal proses belajar maupun penilaiannya berbeda.

Cara menentukan besarnya pembobotan pada masing-masing komponen menyimak, membaca, menyalin dan berbicara untuk menentukannya perlu dirumuskan oleh suatu tim khusus (tim dosen bahasa).

Pada jurusan non bahasa, fokus tidak hanya pada membaca, tetapi yang paling mendasar adalah frase.

Perlu juga dipikirkan bagaimana menilai pada komponen menulis, khususnya pada Bahasa Indonesia.

3. Saran

Buku pedoman penilaian hasil belajar hendaknya memuat format-format penilaian. Untuk menentukan IP Kumulatif perlu dimodifikasi.

Perlu dicari model (bentuk) penilaian sastra dan bahasa yang jurusan non bahasa.

G. Judul Makalah : Penilaian Belajar Seni, Musik, dan Tari.

Pembawa Makalah : Dr. Ibenzani Usman.

1. Isi Makalah

Makalah ini membahas usaha penyeragaman penilaian dari berbagai jurusan.

Seni merupakan wujud budaya manusia yang mencakup ideal, nilai, norma, aturan' kepercayaan kemudian diproyeksikan pada tingkah laku berpola.

Pada perkembangannya seni terbagi menjadi tiga bagian, yaitu seni tari, seni musik dan seni rupa. Kreatifitas merupakan unsur yang paling penting pada seni.

Pengajaran pada mata kuliah seni ada juga yang disebut praktikum. Pada praktikum mahasiswa melaksanakan pengetahuan terapan dan melakukan eksperimen. Untuk menunjang praktikum ada yang disebut mata kuliah bantu praktikum yang umumnya berisi latihan-latihan.

Penilaian pada belajar seni adalah dengan melihat aktifitas yang telah lewat dengan hasil yang dicapai serta melihat prospek masa depan.

Ada empat kriteria yang harus dinilai:

- a. Perkembangan individu tatal
- b. Kompetensi sosial

- c. Kreativitas
- d. Paham estetika + cakap seni

Proses penilaian dilakukan dengan menggunakan ceklist dan matriks. Ceklist untuk diamati dan matriks untuk diukur. Bentuk matriks yang sesuai adalah matriks dua dimensi dengan membagi bagian atas sel-sel. Penilaian dilakukan dengan menghitung persentase sel yang diperoleh.

Untuk menentukan nilai akhir diperhitungkan juga aktifitas, tugas terstruktur, Mid Semester dan ujian semester dengan pembobotan. Angka yang diperoleh diinterpretasikan pada Buku Pedoman IKIP Padang.

H. Judul Makalah : Beberapa Saran Tentang Penilaian Hasil Belajar Perkuliahan Praktek Olahraga di FPOK IKIP Padang

Pembawa Makalah : Dr. Phil Yanuar Kiram.

1. Isi Makalah

Makalah ini mengatakan bahwa perencanaan, pelaksanaan dan penilaian hasil belajar merupakan suatu sistem terpadu.

Penilaian merupakan suatu hal yang sulit dilakukan. Pada FPOK penilaian dilakukan mencakup ketiga ranah (kognitif, afektif dan psikomotor) dengan penekanan pada psikomotor.

Aspek motorik diutamakan karena guru harus dapat mendemonstrasikan gerakan-gerakan di depan anak didiknya. Pada penilaian, masing-masing aspek diberi pembobotan. Besarnya pembobotan tergantung dari jenjang program yang diambil, dengan mengutamakan aspek motorik.

Acuan penilaian yang umumnya menggunakan PAP dan PAN. Namun kedua metode ini belum seragam dipakai.

Beberapa saran yang diajukan pemakalah:

- a. Untuk mata kuliah dasar atau atau yang diberikan hanya 1 semester, bobot penilaian, kognitif 60 %, motorik 40 % dengan pendekatan PAN.
- b. Untuk mata kuliah pendalaman, kognitif 50 %, motorik 50 % dengan pendekatan PAP.
- c. Untuk mata kuliah spesialisasi, kognitif 40 %, motorik 60 %, dengan pendekatan PAP.

2. Pembahasan

Dalam uraian pembahasan, unsur afektif termasuk bagian yang harus dinilai, tetapi pada model yang diberikan unsur itu tidak ada, hal ini perlu menjadi perhatian.

Perlu juga dipikirkan penjabaran "konstruk" bagian-bagian yang harus dinilai pada masing-masing aspek "self evaluation" perlu dipertimbangkan.

III. LAMPIRAN

A. Susunan Kepanitiaan

Susunan dan personalia panitia seminar penyusunan panduan penilain hasil belajar mahasiswa IKIP Padang adalah :

1. Penanggung Jawab : Prof. H. Jamil Bakar (Rektor IKIP Padang)

2. Panitia Pengarah

Ketua : Dr. Mohd. Ansyar

Sekretaris : Dr. Sutjipto

Anggota : Drs. Anas Syafei

Dr. Kumaidi, M.A.

Dr. Aljufri B. Syarif, M.Sc.

3. Panitia Pelaksana

Ketua : Dr. Kumaidi, M.A.

Wakil Ketua : Drs. A. Muri Yusuf, M.Pd.

Sekretaris : Drs. Sufyarma M., M.Pd.

4. Seksi-seksi

Tempat : Drs. Zahri Jas, M.Pd.

Drs. Jasman

Konsumsi : Dra. Nurjani Nurdin

Sekretariat : Ahmad Sabandi, S.Pd.

Yushamdi, S.Pd.

Yusmainal

Nilam Sari

Juin

5. Notulis : Drs. Refdinal

Drs. Waskito

Drs. Syahril

6. Pemakalah : Dr. Kumaidi, M.A.

Dr. Aljufri B. Syarif, M.Sc.

Drs. A. Muri Yusuf, M.Pd.

Dr. Abizar

Dr. Yanuar Kiram

Drs. Rusydi Rusyid, M.A.

Dr. Zainil, M.A.

Dr. Ibenzani Usman

7. Moderator : Dr. Kumaidi, M.A.

Drs. A. Muri Yusuf, M.Pd.

Drs. Zahri Jas, M.Pd.

Drs. Sufyarma M, M.Pd.

8. Penyusunan Laporan : Dr. Kumaidi, M.A.

Drs. A. Muri Yusuf, M.Pd.

Drs. Sufyarma M., M.Pd.

Drs. Waskito

Yushamdi, S.Pd.

Ahmad Sabandi, S.Pd.

Nilam Sari

Yusmainal

B. Peserta Seminar

Peserta seminar penyusunan panduan penilain hasil belajar mahasiswa IKIP Padang

No.	Nama	Jabatan
1	Prof. H. Jamil Bakar	Rektor IKIP Padang
2	Dr. Mohd. Ansyar	PR I IKIP Padang
3	Dr. Sutjipto	PR II IKIP Padang
4	Drs. Anas Syafei	PR III IKIP Padang
5	Prof. Dr. St. Zanti Arbi, M.A.	Guru Besar IKIP Padang
6	Dr. H. Arni Muhammad	Dosen FIP IKIP Padang
7	Dr. Nurtaim	Dosen FIP IKIP Padang
8	Dr. Aliasar, M.Ed.	Dosen FIP IKIP Padang
9	Dr. Amir Hakim Usman	Dosen FPBS IKIP Padang
10	Dr. Hon. Adrin Kahar	Dosen FPBS IKIP Padang
11	Dr. Abizar	Dosen FPIPS IKIP Padang
12	Dr. Azinar Sayuti, M.A.	Dosen FPIPS IKIP Padang
13	Dr. Agusrianto	Dosen FPIPS IKIP Padang
14	Dr. Azmi, M.A.	Dosen FPIPS IKIP Padang
15	Drs. Rafli Kosasi, M.Sc.	Dosen FPMIPA IKIP Padang
16	dr. Armen Arief, MPH	Dosen FPMIPA IKIP Padang
17	Dr. Aljufri B. Syarif, M.Sc.	Dosen FPTK IKIP Padang
18	Dr. Kurnaidi, M.A.	Dosen FPTK IKIP Padang

19	Dr. Nasrūn	Dosen FPTK IKIP Padang
20	Dr. Helmi Suyuthie, M.Ed.	Dosen FPTK IKIP Padang
21	Dr. Nasrullah Aziz.	Dosen FPTK IKIP Padang
22	Dr. Syahron Lubis, M.Ed.	Dosen FPTK IKIP Padang
23	Dr. Agamuddin, M.Ed.	Dosen FPTK IKIP Padang
24	Drs. Djanoeear Sas	Dosen FPOK IKIP Padang
25	Drs. Hanif Badri	Dosen FPOK IKIP Padang
26	Drs. Chalid Marzuki, M.A.	Dosen FPOK IKIP Padang
27	Drs. Abubakar	Dosen FPOK IKIP Padang
28	Drs. Fauzan Hos	Dosen FPOK IKIP Padang
29	Drs. Nurhadi	Dosen FPIPS IKIP Padang
30	Drs. Benyamin Liputo	Dosen FPIPS IKIP Padang
31	Drs. Yasril Yunus	Dosen FPIPS IKIP Padang
32	Drs. Nasrul Sumar	Dosen FPIPS IKIP Padang
33	Drs. Mohd. Nasir B.	Dosen FPIPS IKIP Padang
34	Drs. Thahar Ramli	Dosen FPIPS IKIP Padang
35	Drs. Nur Anas Jamil	Dosen FPIPS IKIP Padang
36	Drs. Syahrūn, M.Pd.	Dosen FPMIPA IKIP Padang
37	Drs. Tahasmin Tamīn	Dosen FPMIPA IKIP Padang
38	Drs. Mawardi	Dosen FPMIPA IKIP Padang
39	Drs. Edwin Musdi, M.Pd.	Dosen FPMIPA IKIP Padang
40	Drs. Syamsul Akmal	Dosen FPMIPA IKIP Padang
41	Drs. Yustini Maaruf	Dosen FPMIPA IKIP Padang
42	Drs. Syafei Said	Dosen FIP IKIP Padang
43	Dra. Syahmiar Syahrūn	Dosen FIP IKIP Padang
44	Dra. Riska Ahmad	Dosen FIP IKIP Padang
45	Drs. Syafruddin Wahid	Dosen FIP IKIP Padang
46	Drs. Erman Amri	Dosen FIP IKIP Padang
47	Drs. Agusfidar Nasution	Dosen FIP IKIP Padang
48	Dra. Maizuar, M.Pd.	Dosen FPTK IKIP Padang

49	Drs. Cecep Syamsuri	Dosen FPTK IKIP Padang
50	Drs. Yusri Abd Hamid	Dosen FPTK IKIP Padang
51	Drs. Samsuri Hasan	Dosen FPTK IKIP Padang
52	Dra. Wildati Zahri	Dosen FPTK IKIP Padang
53	Drs. Daman Suwanto	Dosen FPTK IKIP Padang
54	Drs. Mardi Rasyid, M.Ed.	Dosen FPTK IKIP Padang
55	Drs. Zainuddin Amir	Dosen FPBS IKIP Padang
56	Drs. Nasroel Malano, M.A.	Dosen FPBS IKIP Padang
57	Drs. Yulina Kasim	Dosen FPBS IKIP Padang
58	Drs. Kurnia Ningsih, M.A.	Dosen FPBS IKIP Padang
59	Drs. Ajusril	Dosen FPBS IKIP Padang
60	Drs. M. Zaim	Dosen FPBS IKIP Padang
61	Drs. Rusydi Rusyid, M.A.	Dosen FPMIPA IKIP Padang
62	Yasaratodo Wau	Pasca Sarjana IKIP Padang
63	Nuryanic Hamid	Pasca Sarjana IKIP Padang
64	Rosmaniar	Pasca Sarjana IKIP Padang
65	Rahmulyani	Pasca Sarjana IKIP Padang
66	Robenhart Tamba	Pasca Sarjana IKIP Padang
67	Salfen Hasri	Pasca Sarjana IKIP Padang
68	Aldwin Surya	Pasca Sarjana IKIP Padang
69	Drs. A. Muri Yusuf, M.Pd.	Dosen FIP IKIP Padang
70	Drs. Zahri Jas, M.Pd.	Dosen FPBS IKIP Padang
71	Drs. Sufyarma M, M.Pd.	Dosen FIP IKIP Padang

PENILAIAN ASPEK KOGNITIF

Oleh :

Drs. A. MURI YUSUF, M.Pd.

Disampaikan pada:

Seminar Tanggal 22 dan 23 Januari 1990
dalam Rangka Penyusunan Buku
Panduan Penilaian Hasil Belajar
IKIP Padang

PENILAIAN ASPEK KOGNITIF

Pendahuluan

Memantau kemajuan peserta didik (mahasiswa) dalam belajar merupakan salah satu tugas pendidik (dosen) dalam kegiatan belajar-mengajar. Pemantauan itu hendaklah bersifat menyeluruh (mencakup berbagai aspek) yang terkait dengan tujuan yang diharapkan serta menggunakan instrumen / alat ukur yang valid, reliabel, objektif dan praktis; sebab dalam setiap kegiatan belajar-mengajar ranah/ aspek (domain) yang dikembangkan akan berbeda-beda. Pada kegiatan pertama, yang dikembangkan adalah ranah kognitif, dengan sub kategori pengetahuan dan pemahaman, sedangkan pada kegiatan berikutnya ranah kognitif juga, tetapi dengan sub kategori kemampuan aplikasi, dan evaluasi. Pada pertemuan yang lain yang ditekankan adalah perkembangan sikap dan psikomotor, dengan berbagai sub kategori yang terdapat dalam ranah tersebut.

Sehubungan dengan itu penilaian hasil belajar tidaklah dapat dipisahkan dari tujuan-tujuan yang pernah dirumuskan sebelum belajar-mengajar berlangsung atau dengan konsep utama yang terdapat dalam setiap materi sajian perkuliahan. Konsep-konsep utama itu juga mengacu kepada ranah kognitif, afektif dan psikomotor dalam setiap materi perkuliahan. Keterwakilan semua konsep utama (yang dituangkan dalam proposisi) dalam materi perkuliahan dan kemampuan yang diharapkan dari mahasiswa tercetak dengan baik dalam kisi-kisi ujian (tabel spesifikasi). Sifat "mewakili" (representativeness) dalam materi yang akan diuji akan sangat membantu validitas isi (content validity) dari instrumen yang akan dikembangkan. Kisi-kisi ujian merupakan pedoman dan atau memberi arah dalam menyusun penilaian hasil belajar yang mangkus dan sangkil.

Mengingat luasnya cakupan penilaian hasil belajar itu maka dalam makalah ini hanya akan dibahas penilaian hasil belajar aspek kognitif yang mencakup: (1) konsep pengukuran dan penilaian, (2) syarat-syarat tes yang baik, (3) kisi-kisi tes, (4) pembuatan tes hasil belajar, (5) skorsing dan (6)

pengembangan tes hasil belajar.

Ranah afektif dan psikomotor akan dibahas oleh pemakalah lain dalam seminar ini.

2. Pengukuran dan Penilaian

Pengukuran dan penilaian merupakan dua istilah yang sering digunakan dalam bidang pendidikan dan psikologi. Kedua istilah itu sering pula dikacaukan dalam penggunaannya. Penilaian mempunyai cakupan yang lebih luas dari pengukuran. Hill (1981) menyatakan: "measurement is the assignment of numbers to attribute of objects, events or people according to rules". Sedangkan Campbell menyatakan bahwa: "measurement as the assignment of numerals to object or events according to rules". Hal yang hampir sama dikemukakan oleh Stavens, seperti yang dikutip oleh Rachmias (1981) "measurement may be viewed as a procedure in which one assign numerals, numbers or others symbols to emperical properties (variables) according to rules". Dari kutipan di atas sekurang-kurangnya ada tiga konsep yang perlu diperhatikan, yaitu:

- a. Angka atau simbol yang dapat diubah secara statistis atau dimanipulasi secara matematis.
- b. Aturan yang digunakan sebagai patokan tentang benar/tidaknya tindakan yang dilakukan atau suatu kejadian atau objek yang dikuasai seseorang.
- c. Penerapan atau penetapan angka atau simbol terhadap objek atau kejadian yang dimaksudkan.

Jadi pengukuran adalah suatu prodesur atau suatu usaha penerapan angka atau simbol terhadap objek atau kejadian sesuai dengan aturan. Kalau dikaitkan dengan kegiatan belajar atau hasil belajar, maka pengukuran dapat diartikan sebagai suatu usaha pengumpulan informasi tentang individu (mahasiswa yang belajar) dalam bentuk data kuantitatif. Tahap-tahap yang dilalui dalam pengukuran hasil belajar adalah: (1) pengumpulan data dengan menggunakan tes, (2) jawaban mahasiswa (respon mahasiswa) diskor dengan menggunakan aturan-aturan tertentu (3) Hasilnya berupa angka (data kuantitatif). Angka-angka tersebut

merupakan indikator kemajuan mahasiswa dalam bidang yang diuji.

Penilaian atau evaluasi lebih luas artinya dan cakupannya dari pada pengukuran. Bloom (1981) menyatakan: "evaluation is the systematic collection of evidence to determine whether in fact certain change are taking place in learners as well as to determine the amount or degree of change in individual students". Pada bagian lain ia menemukan pula: "evaluation as a method of acquiring and processing the evidence needed to improve the student's learning in the teaching". Kedua pendapat itu menyatakan bahwa penilaian merupakan suatu proses sistematis untuk mengetahui bukti dalam menentukan peringkat penguasaan mahasiswa atau perubahan tingkah laku sebagai hasil belajar. Sedangkan Gronlund (1981) menyatakan bahwa: "evaluation may be defined as a systematic process of determining the extent to which instructional objectives are achieved by pupils". Jadi jelaslah bahwa penilaian itu merupakan suatu proses pengumpulan data (kuantitatif dan kualitatif) secara sistematis dan pemberian makna (value judgement) terhadap data yang telah dikumpulkan atau dapat juga dikatakan bahwa penilaian adalah pemberian makna/arti terhadap data tentang seseorang, objek atau kejadian. Secara sederhana dapat dikatakan bahwa:

pengukuran = deskripsi kuantitatif

penilaian = deskripsi kuantitatif, kualitatif dan:
pemberian makna (value judgement).

Karena itu penilaian hasil belajar tidak saja melibatkan pengukuran, tetapi juga cara-cara lain sesuai dengan informasi yang diperlukan. Penilaian dapat juga dilakukan tanpa pengukuran.

3. Hakekat Tes Hasil Belajar

Seperti telah disinggung pada uraian terdahulu bahwa prestasi belajar adalah tingkat penguasaan siswa dalam belajar. Penguasaan ini akan tercermin dalam berbagai aspek antara lain (1) kognitif yang meliputi: pengetahuan, pemahaman, aplikasi, analisis, sistesis, dan evaluasi; (2)

afektif (3) psikomotor.

Penguasaan itu bukanlah semata-mata diakhir semester maupun naik kelas melainkan dari awal interaksi belajar sampai akhir belajar. Kegiatan awal ditandai dengan perlunya menguasai kemampuan awal siswa sebagai hasil penguasaan belajar sebelumnya. Tes yang digunakan untuk itu adalah tes penguasaan belajar. Tes dengan maksud yang sama dapat juga digunakan pada bagian demi bagian atau sub bagian dalam keseluruhan proses belajar sebab tes penguasaan (achievement test) adalah tes yang dimaksudkan untuk mengukur penguasaan/pencapaian dalam proses belajar-mengajar yang telah dilaksanakan dan merupakan akumulasi pengetahuan dan keterampilan individual akan bidang tertentu. Jadi tes penguasaan belajar menekankan pada kemajuan yang telah dicapai seseorang terhadap sesuatu bidang yang telah diajarkan, bukan kemampuan untuk masa datang.

Umpama:

- a). Untuk mengetahui kesulitan/kesukaran dalam pelajaran Dasar-dasar Kependidikan tentang Inovasi Pendidikan di Indonesia dapat diketahui dosen dengan tes penguasaan buatan dosen yang di rancang untuk itu. Dari hasil tes itu akan dapat diketahui bagian mana yang belum dikuasai mahasiswa serta bentuk kesukaran yang dialami. Tes penguasaan ini berupa tes diagnostik.
- b). Untuk mengetahui tingkat penguasaan mahasiswa dalam keseluruhan dapat pula diketahui dengan melakukan tes summatif dengan tingkat penguasaan untuk tiap unit satu pelajaran melalui tes penguasaan sub-summatif.
- c). Untuk mengetahui penguasaan mahasiswa dengan tujuan akan menempatkan yang bersangkutan dalam kelompok belajar atau jurusan tertentu dapat digunakan tes penguasaan yang menurut fungsinya adalah tes penempatan (placement test).

Dengan demikian dapat pula dikatakan bahwa tes penguasaan atau pencapaian belajar mahasiswa tidaklah semata-mata untuk menentukan angka atau naik kelas saja,

tetapi jauh lebih luas lagi. Apabila dikaitkan dengan fungsi tes, maka tes penguasaan belajar dapat berfungsi untuk mengetahui kesiapan mahasiswa dalam belajar; kesulitan/kesukaran mahasiswa dalam belajar; perbaikan pengajaran maupun penentuan angka atau nilai mahasiswa dalam belajar.

Konsep tersebut didukung oleh pendapat Gronlund dan Hill

sebagai berikut:

Gronlund (1982) mengemukakan:

Achievement tests should support and reinforce others aspect of the instructional process. They are can aid both the teacher and the student in asesing learning readiness monitoring learning process diagnosing learning outcomes

(Tes penguasaan akan membantu dan memperkuat aspek lain dalam proses pengajaran. Tes ini dapat membantu guru dalam menilai kesiapan belajar... memantau proses belajar ... mendiagnosis kesulitan belajar ... dan menilai hasil belajar).

Hill (1981) menyatakan: tes penguasaan belajar dirancang untuk mengukur apa yang telah diajarkan/dipelajari dalam bidang studi yang bersifat formal. Justru karena itu keberhasilan dalam penilaian hasil belajar sangat terkait dengan baik tidaknya tes yang digunakan, pengadministrasian tes dan pengolahan data yang telah dikumpulkan.

4. Syarat-syarat Tes yang baik

Secara sederhana dapat didefinisikan bahwa tes adalah kumpulan butir soal yang jawabannya dapat dinyatakan dengan benar-salah. Definisi ini dipakai untuk membedakan instrumen penilaian lain yang mungkin dipakai dalam penilaian hasil belajar, seperti angket, skala, atau checklist.

Penilaian hasil belajar akan dapat mencapai tujuan yang

diinginkan apabila menggunakan tes hasil belajar yang baik; dengan memiliki karakteristik yaitu: valid, reliabel, praktis, objektif, dan norma. Tiap-tiap karakteristik di atas akan diuraikan lebih lanjut.

a. Validitas

1). Pengertian

Apabila suatu tes atau alat ukur akan dipergunakan, pertanyaan yang mula-mula timbul dan perlu mendapat perhatian adalah: apakah tes atau alat ukur itu valid atau sah? Betulkah ia dapat mengukur suatu aspek atau perilaku yang ingin kita ukur? Anastasi (1968), menyatakan: "the validity of a test concern what does the test measure and how well it does so". Pendapat yang hampir bersamaan dengan itu dikemukakan pula oleh Nachmias (1981) : Validity is concerned with the question : Is one measuring what one thinks one is measuring? Cronbach menyatakan bahwa validity adalah : How well a test or evaluative technique does the job that it is employed to do" (Thorndike, 1970). Sedangkan Anderson dkk (1975) menulis : "A test is valid if it measures what it purpose to measure"

Dari pendapat-pendapat tersebut dapat ditarik suatu kesimpulan bahwa suatu tes atau alat ukur dikatakan valid apabila tes/alat ukur tersebut benar-benar mengukur (objek) yang hendak diukurnya. Umpama : apabila kita ingin mengukur kemampuan mahasiswa dalam ilmu pemerintahan, maka materi yang diujikan hendaklah terfokus pada materi ilmu pemerintahan. Jangan terjadi salah arah dengan memberikan sebanyak mungkin istilah asing, sehingga berubah menjadi ujian-bahasa asing, bukan ilmu pemerintahan.

Makin tinggi validitas suatu alat ukur makin baik alat ukur itu untuk digunakan. Tetapi perlu diingat

bahwa validitas alat ukur itu tidaklah dapat dilepaskan dari kelompok yang dikenai alat ukur itu. Oleh karena itu berlakunya validitas hanya terbatas pada kelompok tersebut atau kelompok lain yang kondisinya hampir sama dengan kelompok itu.

Gronlund (1981) menyatakan bahwa :

- a. Validity menunjuk pada hasil suatu tes atau instrumen evaluasi untuk kelompok atau individual, tidak untuk instrumen itu sendiri.
- b. Validitas merupakan "degree" (derajat) seperti : tinggi, sedang dan kurang.
- c. Validitas itu selalu spesifik penggunaannya. Umpama : Suatu tes Bahasa Indonesia dalam tipe objektif adalah valid untuk kemampuan/pengetahuan siap tetapi tes yang sama, tak valid untuk kemampuan melahirkan/mengemukakan pendapat.

Menurut Test Standard 1985, seperti yang dikemukakan Kumaedi (1989), validitas didefinisikan sebagai: "refers to appropriateness, meaningfulness and usefulness of the specific inferences made from test scores". Sedangkan Tuckman lebih menekankan bahwa "validity or appropriateness address themselves to the question of whether the test measures what it is supposed to measure".

Oleh karena itu validitas akan terkait dengan (1) appropriateness yaitu apakah sampel butir soal betul-betul mewakili suatu populasi/universe dari isi materi perkuliahan tertentu, (2) meaningfulness yaitu keberartian/kebermaknaan butir-butir soal yang diberikan untuk mewujudkan kemampuan yang ingin diukur, (3) usefulness, menyangkut kegunaan suatu instrumen pengukuran untuk mengungkapkan tujuan tertentu.

b. Reliabilitas

1. Pengertian

Dalam pembicaraan tentang reliabilitas tidak lagi dipertanyakan tentang ketepatan pengukuran, tetapi tentang konsistensi dan ketelitian dalam mengukur apa yang seharusnya diukur. Apakah skor-skor yang didapat melalui pengukuran yang dilakukan konsisten? Tidakkah hasilnya akan berubah kalau dilakukan pengukuran ulang terhadap objek yang sama pada waktu yang berlainan?

Wrightstone (1959) mengemukakan bahwa :

"Reliability is commonly defined as an estimate of degree of consistency or constancy among repeated measurements of individuals with the same instrument".

Pendapat yang cukup lama ini tetap berlaku. Gronlund (1981) menyatakan: "Reliability refers to the consistency of measurement results from one measurement to another", (p. 94). Lebih lanjut pembicaraan tentang reliabilitas mengandung arti bahwa :

- a). Reliabilitas merupakan hasil yang diperoleh dengan instrument evaluasi, bukan kepada instrumen itu sendiri.
- b). Perkiraan reliabilitas itu menunjuk kepada konsistensi dari skor tes.
- c). Reliabilitas itu penting tetapi tidak cukup untuk menjamin validitas suatu tes. Ia hanya menyediakan kenyataan tentang konsistensi, bukan mengukur isi instrumen.
- d). Reliabilitas dinyatakan dalam "coefficient reliability" dan/atau dengan "standard error measurement".

Jadi suatu alat ukur dikatakan reliabil, apabila alat ukur itu dicobakan kepada objek atau subjek yang sama secara berulang-ulang, maka hasilnya akan tetap sama atau relatif sama.

2. Faktor-faktor yang Mempengaruhi Reliabilitas

Banyak faktor yang mempengaruhi reliabilitas alat ukur. Di antara faktor itu yang menonjol adalah:

a). Kontruksi item yang tidak tepat, sehingga tidak dapat mempunyai daya pembeda yang kuat. Misalnya seseorang murid yang berinteligensi cukup tinggi tetapi karena kontruksi item yang salah, murid tersebut tidak dapat menjawab tes yang dimaksud, sehingga ia menjadi salah.

b. Panjangnya/pendeknya instrumen

Instrumen yang panjang dengan mengkaji beberapa tujuan akan lebih reliabil dibandingkan instrumen yang pendek, karena menyediakan sampel penampilan siswa yang lebih representative.

c. Evaluasi yang subjektif akan menurunkan realibilitas Kelelahan dan kebosanan seorang guru dalam memeriksa hasil suatu tes juga akan mempengaruhi reliabilitas tersebut.

Pengenalan diri secara pribadi oleh guru yang menjurus kepada setuju dan tidak setuju pada siswa, hubungan kekeluargaan, tulisan yang terlalu buruk, kecantikan dan lain-lain dapat mempengaruhi kemurnian hasil belajar siswa. Yang dinilai bukan lagi hasil kerja siswa, tetapi ada faktor lain yang menentukan. Dengan tes essay; evaluasi subjektif dari guru lebih mudah timbul dibandingkan dengan tes objektif.

Standar skor yang tidak sama atau tidak sesuai antara test satu dan dua; antara bagian awal dengan bagian akhir akan menyebabkan reliabilitas akan terganggu.

d. Ketidak tepatan waktu yang diberikan.

Tiap tes mempunyai waktu tersendiri sesuai dengan

tingkat kesukaran, jumlah serta bentuk soal yang digunakan. Tes yang tidak seimbang antara beban dan waktu, menyebabkan hasil yang didapat tidak memuaskan. Reliabilitas tes tinggi seandainya tes itu diberikan sesuai dengan kemampuan siswa dengan perkiraan waktu yang digunakan relatif tepat sehingga siswa dapat berpikir untuk menjawab suatu item (butir^a soal) dengan benar, tetapi kalau waktu terlalu pendek maka ia salah menjawabnya. Kekurang tepatan pengalokasian waktu akan mengganggu reliabilitas alat ukur.

e. Kemampuan yang ada dalam kelompok

Makin berbeda kemampuan siswa dalam suatu kelas yang mengikuti tes makin tinggi reliabilitas tes. Apabila suatu tes diberikan pada kelas yang sama, maka reliabilitasnya lebih rendah dibandingkan apabila tes itu diberikan pada kelas yang berbeda.

f. Luas/tidaknya sample yang diambil

Makin luas sample yang diambil makin tinggi reliabilitas suatu tes; sebab makin luas cakupan responden dan makin bervariasi individu yang terlibat. Koefisien reliabilitas suatu tes dalam suatu kelompok, bervariasi, menurut rentangan (range) bakat yang ada dalam kelompok itu. Apabila jarak rentangan itu melebar, maka reliabilitas tes akan tinggi, tetapi apabila rentangannya kecil maka reliabilitas tes akan rendah.

g. Kondisi dan situasi pada pengadministrasian alat ukur

Makin baik penilai mengontrol situasi dan kondisi pada pengadministrasian tes makin reliabel tes yang diberikan. Apabila kondisi tersebut sangat berbeda

maka soal yang dicapai adalah refleksi dari kejadian tersebut dan akan mengurangi reliabil tes.

c. Objektivitas

Objektivitas suatu tes ditentukan oleh kesamaan skor-skor yang diperoleh siswa meskipun tes itu dinilai oleh beberapa orang penilai. Objektivitas suatu tes akan bervariasi :

1). Tinggi

2). Sedang

3). Berubah-ubah

Tinggi/rendahnya objektivitas suatu tes itu akan berkaitan erat dengan tersedianya kunci jawaban yang baik dan benar serta standard interpretasi yang baku.

Oleh karena itu setiap tes dan instrument hendaklah objektif; bebas dari pengaruh-pengaruh yang mengganggu, yaitu subjektivitas orang yang menilai dan melakukan tes. Siapapun yang melakukan tes kesimpulan yang diperdapat hendaklah sama.

d. Praktikalitas

Suatu tes yang mempunyai validitas yang tinggi adalah suatu tes yang reliabel, tetapi suatu tes yang reliabel belum tentu valid. Apabila suatu tes telah valid dan reliabel, biasanya tes itu telah dianggap memenuhi syarat pokok. Di samping itu perlu pula diperhatikan kepraktisan dari alat ukur yang digunakan. Sifat praktis dari suatu tes hendaklah tercermin dalam hal : (1) biaya yang digunakan tidak terlalu tinggi, (2) mudah diadministrasikan, (3) mudah diskor, (4) mudah diinterpretasikan, (5) waktu yang dipakai tepat dan tidak

terlalu lama.

1. Biaya yang digunakan tidak terlalu tinggi

Faktor biaya merupakan faktor yang tampaknya tidak penting tetapi perlu diperhatikan. Testing adalah sesuatu yang tidak mahal, namun kalau digunakan dengan cara yang tidak tepat akan mempengaruhi pelaksanaan tes itu. Walaupun fasilitas yang dipakai tidak terlalu banyak, namun kalau jumlah siswa yang ikut ujian cukup besar, maka hal itu akan membawa pengaruh pada jumlah kertas yang dibutuhkan dan fasilitas lainnya. Sehubungan dengan itu akan lebih baik bila dirancang suatu tes yang dapat dipakai secara berulang-ulang sehingga akan menimbulkan penghematan dalam biaya.

2. Mudah diadministrasikan

Suatu alat ukur yang baik ditandai oleh mudahnya alat itu diadministrasikan. Ini berarti:

- a). Alat ukur itu mudah diberikan kepada mahasiswa, dengan petunjuk yang jelas bagaimana cara mengerjakannya dan mudah dimengerti, sehingga dosen tidak perlu lagi memberikan penjelasan-penjelasan.
- b). Alat ukur itu mudah dilaksanakan dan waktu yang disediakan cukup dibandingkan dengan tingkat kesukaran alat ukur itu.
- c). Mudah dikumpulkan kembali setelah waktu yang tersedia untuk mengerjakan habis.

3). Mudah diskor

Salah satu kesukaran yang sering ditemui dalam

ujian-ujian di masa lampau ada sulitnya tes itu diskor, sehingga terdapat perbedaan hasil dari beberapa orang pemeriksa yang memeriksa satu jawaban tes yang sama atau bahkan terdapat perbedaan hasil dari satu orang pemeriksa yang memeriksa ulang satu hasil tes yang sama. Alat ukur dapat dianggap baik apabila mudah diskor. Ini berarti :

- a). Ada standar yang dapat digunakan sehingga hasilnya lebih homogen.
- b). Siapa saja yang memeriksa kertas jawaban ujian dalam waktu yang berlainan, hasil/skornya tidak akan berubah.
- c). Waktu yang digunakan untuk memeriksa hasil ujian itu tidak terlalu lama.
- d). Pemeriksa hasil ujian itu tidak perlu orang yang ahli betul dalam bidang yang di tes itu.

Untuk memudahkan pekerjaan skorsing biasanya untuk suatu alat ukur perlu adanya pemisahan antara buku tes dengan lembaran jawabannya, lembaran jawaban yang terpisah yang diskor, baik dengan menggunakan mesin ataupun secara manual.

4. Mudah diinterpretasikan

Skor yang didapat sebagai hasil dari pengukuran belum mempunyai arti kalau skor itu tidak diterjemahkan atau diinterpretasikan. Untuk itu seorang pengajar hendaklah memiliki kemampuan tentang bagaimana menginterpretasikan hasil tes. Hal itu dapat dilakukan dengan bantuan statistik, sehingga angka-angka itu dapat diolah menjadi sesuatu yang berarti.

5. Waktu yang tepat dan tidak terlalu lama

Tes yang pengerjaannya memakan waktu terlalu lama

akan membosankan dan sebaliknya tes yang terlalu cepat juga merugikan, walaupun tes itu mungkin power tes atau speed test. Waktu yang disediakan untuk setiap tes atau alat ukur lainnya hendaklah mempertimbangkan keadaan mahasiswa maupun lingkungannya. Berikanlah waktu itu sesuai dengan tingkat kesukaran tes yang diberikan, tetapi jangan terlalu lama. Di samping itu berikanlah ujian itu pada waktu yang baik, atau tidak pada udara yang sangat panas atau dilingkungan yang sangat terganggu.

Contoh: Untuk tes bentuk objektif (pilihan jamak), biasanya untuk tiap soal akan diberikan waktu antara 45 detik sampai dengan 1 (satu) menit.

e. Norma

Penilaian pendidikan bukanlah semata suatu proses sistematis untuk mengumpulkan bukti-bukti (data dan informasi), tetapi juga melakukan "judgment" tentang sesuatu yang dinilai.

Untuk mendapat informasi dan pengambilan keputusan yang tepat maka alat ukur yang baik haruslah mempunyai norma sebagai patokan, sehingga akan memberikan kesimpulan yang tepat.

5. Prinsip-prinsip Penilaian Hasil Belajar

a. Prinsip-prinsip umum penilaian

Seperti telah disinggung pada uraian terdahulu evaluasi pengajaran dapat disinggung pada uraian terdahulu evaluasi pengajaran dapat dikembangkan sesuai dengan fungsi dan tujuan yang diharapkan.

Beberapa persiapan umum dalam evaluasi pengajaran adalah:

1). Evaluasi yang baik bersifat komprehensif dan internal

Prinsip ini menunjukkan pada betapa pentingnya cakupan yang luas dari alat ukur yang digunakan, sesuai dengan materi pelajaran. Cakupan itu bukan semata-mata dilihat dari luas materi yang di ujikan, tetapi juga domain (aspek) yang diukur. Melalui tes objektif lebih banyak informasi dapat diukur, tetapi sangat sedikit sekali yang berkaitan dengan minat, keterampilan maupun sikap. Sedangkan dengan menggunakan tes essay, sedikit informasi yang dikumpulkan, tetapi kemampuan menalar, dan mengemukakan pendapat dapat dijangkau dengan cara demikian. Selanjutnya keterampilan perlu pula diketahui, kalau materi pelajaran memang terkait dengan aspek itu. Justru karena itu evaluasi harus luas cakupannya baik dilihat dari isi maupun aspek-aspek yang diukur maupun dinilai.

2). Evaluasi hendaklah kontinu

Evaluasi yang baik bukanlah dilakukan pada awal, dan akhir suatu kegiatan saja, atau sesuatu yang bersifat sewaktu atau momentum, melainkan hendaklah dilakukan secara terus-menerus. Mulai pada saat program pengajaran dirancang seharusnya sudah ada evaluasi, yaitu untuk mengetahui seberapa jauh mahasiswa telah memiliki materi yang akan diberikan. Dengan cara demikian dapat dipilih materi dan strategi mengajar yang tepat, organisasi kelas yang baik, dan waktu yang sesuai, serta sumber belajar yang mendukung kegiatan belajar-mengajar. Pada saat kegiatan mulai dilaksanakan, evaluasi proses sudah harus berjalan, sehingga dapat diketahui kesulitan-kesulitan, dan hambatan mahasiswa dalam belajar.

Demikian juga kesukaran-kesukaran yang dialami dosen perlu diketahui, sehingga dapat dilakukan penyempurnaan pada kegiatan-kegiatan berikutnya. Evaluasi summatif dapat dilakukan pada akhir kegiatan untuk mengetahui

tingkat pencapaian mahasiswa maupun epektipitas pengajaran.

Evaluasi yang dilakukan secara tidak kontinu kurang dapat merekam semua keadaan dalam proses belajar-mengajar, sehingga hasil evaluasi itu belum dapat menggambarkan hasil belajar secara keseluruhan.

3. Evaluasi yang baik bersifat objektif

Hasil belajar yang terkumpul dengan menggunakan alat ukur selanjutnya ditafsirkan dengan jelas dan tegas, serta tidak memihak. Artinya, gambaran hasil belajar itu tidak dipengaruhi oleh faktor lain diluar hasil yang dicapai mahasiswa. Hendaknya ada patokan atau norma yang jelas dengan klasifikasi yang tegas, sehingga gambaran apa yang didapat mahasiswa itu akan menjamin ketetapan kecakapan mahasiswa yang sebenarnya.

4. Evaluasi yang baik berpijak pada tujuan yang jelas

Perumusan tujuan yang jelas dan mendapat prioritas, akan dapat membantu terwujudnya evaluasi kegiatan belajar yang baik. Tujuan itu hendaklah terjabar dengan baik, jelas dan mudah diukur atau dinilai, sehingga dapat memberi bimbingan dalam menyusun alat ukur yang tepat.

5. Evaluasi yang baik menggunakan alat ukur yang ganda dan sah

Tidak ada alat penilaian tunggal yang dapat menilai semua kemajuan mahasiswa dalam belajar. Untuk menilai pengetahuan siap dapat digunakan tes dalam bentuk true false; tetapi bentuk ini tidak baik digunakan untuk mengetahui tingkat pemahaman, keterampilan berfikir atau perubahan sikap mahasiswa. Untuk yang tersebut terakhir itu, guru hendaklah mencari atau menyusun alat ukur lain yang lebih cocok. Oleh karena itu gunakan bermacam-macam tipe alat ukur atau alat penilaian, sehingga dapat merangkum semua yang dibutuhkan sesuai dengan keadaan mahasiswa yang sesungguhnya.

6. Evaluasi yang baik hendaknya dilakukan oleh suatu tim

Cara ini dapat mengurangi beberapa subjektivitas yang mungkin timbul dibandingkan dengan apabila penilaian itu dilakukan oleh suatu orang saja.

7. Evaluasi bukanlah tujuan, melainkan adalah cara untuk mencapai suatu tujuan.

Banyak "kesalahan" yang mungkin terjadi pada alat penilaian yang dipakai. Kesalahan pertama akan ada pada waktu menyusun instrumen. Apakah instrumen itu telah dirakit sedemikian rupa menurut cara yang sebenarnya? Apakah tujuan yang dirumuskan sudah benar? Kesalahan lain terletak pada betulkah yang di uji telah mencakup semua aspek ataukah aspek-aspek yang dimunculkan itu telah mewakili keadaan yang sebenarnya? Mengingat kelemahan-kelemahan mungkin terjadi, baik pada alat ukur maupun pada aspek yang dinilai, maka hendaklah di pandang bahwa penilaian itu adalah untuk menyediakan informasi tentang mahasiswa yang digunakan sebagai dasar untuk mengambil keputusan.

b. Prinsip-prinsip khusus evaluasi penguasaan belajar.

Penguasaan belajar akan dapat diungkapkan menurut fungsi yang diharapkan apabila instrument yang dipakai memenuhi syarat untuk hal tersebut. Tes hendaklah valid dan reliable di samping syarat-syarat yang lain, seperti telah diutarakan dalam uraian terdahulu.

Sehubungan dengan itu beberapa prinsip yang perlu di perhatikan dalam penyusunan tes penguasaan belajar adalah:

1. Tes yang disusun hendaklah betul mengukur tujuan pengajaran.

Seseorang dosen di P.T.N pada waktu penyusunan suatu pelajaran hendaklah membayangkan dengan tuntas hubungan

tujuan yang disusun dengan tuntas hubungan tujuan yang disusun dengan bentuk evaluasi pengajaran sesuai dengan fungsinya. Justru karena itu, langkah awal seseorang dosen adalah mengidentifikasi dengan jelas tujuan menurut kawasan kognitif, efektif atau psikomotor. Pada kegiatan berikutnya merinci tiap-tiap kawasan menjadi lebih spesifik dan operasional sehingga dapat dilihat dalam penampilannya.

Umpama :

Guru ingin menilai hasil belajar.

- 1. Ia perlu menetapkan terlebih dahulu apa yang mau diukur ; pengetahuan , pemahaman ataukah kemampuan mendemonstrasikan.
- 2. Merumuskan tujuan yang telah ditetapkan itu dalam bentuk umum
- 3. Merumuskan tujuan khusus.

Umpama :

- Siswa dapat mendemonstrasikan
- Siswa dapat menyebutkan

Tes yang disusun hendaklah mampu menjawab tujuan tersebut. Ini berarti tes itu harus dirancang berdasarkan tujuan tersebut.

2. Tes yang disusun merupakan sampel yang representatif dari semua materi pengajaran.

Dalam hal ini guru perlu terlebih dahulu membuat kisi-kisi ujian atau table of specification yang akan memberikan gambaran yang lebih terinci tentang kawasan dan materi pengajaran.

Dengan adanya kisi-kisi tersebut bobot materi yang diberikan dapat dirinci menjadi lebih proporsional dan

tiap kelompok terwakili dalam instrument/tes yang disusun.

3. Bentuk/format tes yang dipilih hendaklah sesuai dengan tujuan yang akan dicapai.

Banyak bentuk tes yang dapat digunakan, namun tiap-tiap bentuk tes mempunyai keterbatasan dan keuntungan sendiri-sendiri. Baik tes essay maupun tes objektif mempunyai peran yang berbeda dalam mengungkapkan sesuatu. Tes essay lebih diharapkan untuk mengungkapkan kemampuan menjelaskan menerangkan dan mengorganisasikan pendapat sendiri terhadap sesuatu yang ditanyakan. Sedangkan untuk "jawaban pendek" maupun isian" siswa juga diharapkan men "supply" jawabannya, tetapi dalam hal ini lebih bersifat fakta.

Tes objektif lebih terarah pada pengumpulan informasi lebih banyak dari essey tetapi dalam pelaksanaannya mahasiswa hanya mengambil dari jawaban yang lebih disediakan:

Justru karena itu tipe penguasaan belajar hendaklah dirancang sesuai dengan tujuan yang diinginkan.

4. Tes penguasaan belajar hendaklah seterandal mungkin. Dalam hubungan ini penyusun tes penguasaan belajar :

- a. Memperhatikan panjangnya tes yang disusun
Tes yang panjang lebih reliable dari tes yang pendek

- b. Memperbaiki butir demi butir soal sehingga kualitas tes menjadi lebih baik.

- c. Menyediakan butir soal yang dapat mewakili tingkah laku yang diukur.

- d. Menguji cobakan tes yang telah disusun itu untuk mengetahui kesalahan (error) dan kemudian menyempurnakannya.

Melalui cara yang demikian penyusun tes akan mendapatkan tes yang lebih reliabel.

6. Perencanaan Tes Hasil Belajar

Tes yang baik bukanlah hasil sekali jadi, tetapi membutuhkan waktu yang cukup lama dengan tenaga yang ahli di bidangnya serta memahami konsep-konsep tes hasil belajar. Ketekunan dan ketelitian sangat diperlukan dalam mendisiplinkan diri pada waktu penyusunan tes.

Tes penguasaan belajar perlu dirancang secara baik dan mempunyai hubungan erat dengan tujuan pengajaran. Tujuan merupakan titik awal dari perencanaan dan penyusunan tes penguasaan belajar.

Sehubungan dengan itu maka perencanaan tes dalam setiap aktivitas pengajaran akan sangat penting sebab dengan tersedianya tes berdasarkan perencanaan yang matang akan dapat diketahui :

- a. Apakah tes itu betul-betul relevan dengan tujuan pengajaran.
- b. Apakah tes itu betul-betul mewakili keseluruhan bahan dan menguji penguasaan belajar yang diinginkan.

Perencanaan tes yang benar akan membantu memudahkan dalam penyusunan tes penguasaan belajar yang sesungguhnya.

a. Tingkat Penguasaan.

Kawasan kognitif (Cognitiv domain) merupakan kemampuan intelektual pada diri siswa dalam mengenal lingkungannya. Kawasan ini terdiri dari 6 macam kemampuan yang disusun dari yang rendah/sederhana kepada yang tinggi/kompleks.

1. Pengetahuan (knowledge)
yaitu kemampuan untuk mengingat kembali atau mengenal hal-hal yang telah dipelajari.
2. Pemahaman (Comprehension) merupakan kemampuan memahami hubungan atau menangkap arti dan makna diantara konsep dan fakta-fakta tentang suatu hal.
3. Aplikasi (Application).
Merupakan kemampuan untuk memilih konsep, fakta, dalil, aturan, hukum dan sebagainya serta menerapkan/menggunakan hal itu secara tepat dan benar dalam situasi baru, atau kehidupan sehari-hari.

Kemampuan ini dapat diukur dengan mengajukan suatu keadaan (kondisi) harus terlebih dahulu kemudian siswa menerapkan sesuatu yang dipelajari yang terkait dengan kasus itu.

4. Analisis (Analysis)
Merupakan kemampuan menganalisis kembali atau menjabarkan suatu yang kompleks menjadi bagian/hal yang lebih sederhana dan mudah dipahami. Berhubung karena kemampuan analisis bukan lagi menerapkan/mengaplikasikan pengetahuan, tetapi lebih menguraikan, maka dalam evaluasi belajar siswa harus dihadapkan pada kasus yang dirangkum sendiri oleh penulis soal dan kurangi mengambil uraian yang mungkin sudah diketahui mahasiswa lebih dahulu.

5. Sintesis (Synthesis)
Adalah kemampuan menyusun kembali atau memadukan bagian menjadi keseluruhan yang lebih berarti. Oleh karena itu mahasiswa diajak untuk menformulasikan, mengorganisasikan, merencana-

kan, membulatkan dengan mengatur, merencanakan, mengkombinasikan, menyusun kembali tentang sesuatu.

6. Evaluasi (evaluation)

Merupakan kemampuan sejauh mana siswa dapat menerapkan konsep aturan atau pengetahuan yang ada untuk menilai sesuatu yang lain. Atau dapat juga dikatakan kemampuan untuk membuat keputusan (judgment) tentang sesuatu berdasarkan kriteria/standar yang telah ditetapkan.

Apa yang akan diukur pada prinsipnya adalah apa yang ingin dicapai pada tujuan yang dijabarkan. Tujuan-tujuan haruslah tertuang dalam penguasaan belajar atau hasil belajar yang dikuasai mahasiswa.

b. Kisi-kisi Ujian (Tabel Spasifikasi)

Kisi-kisi ujian yang baik dan benar akan memberikan bimbingan yang terarah kepada penyusun tes disuatu pihak, namun pada pihak lain akan dapat memberikan bantuan untuk menyiapkan tes sesuai dengan atau mewakili apa yang pernah diberikan baik secara langsung maupun tidak langsung atau kemampuan yang diharapkan seharusnya dimiliki oleh seorang mahasiswa dalam bidang tertentu (yang diujikan). Sifat mewakili (representativeness) adalah sangat perlu, sebab kita menyadari bahwa tes itu hanya merupakan sample dari tingkah laku yang akan diukur dalam waktu yang terbatas. Dalam penyusunan kisi-kisi ujian, seorang penyusun tes betul-betul hendaknya menghayati apa yang seharusnya dimiliki mahasiswa berdasarkan silabus yang telah diberikan dan kemudian menyusunnya dalam bentuk lebih sederhana dan

merupakan suatu matrik dua jalur.

Pada jalur mendatar diisi dengan jenjang kemampuan dan bentuk item yang akan digunakan, sedangkan jalur ke bawah diisi dengan ruang lingkup mata ujian. Banyak soal pada masing-masing ruang lingkup tergantung pada silabus dan pentingnya beban itu bagi mahasiswa serta kegunaannya di dalam masyarakat setelah mereka menyelesaikan studinya. Secara terinci beberapa hal yang perlu ada dalam kisi-kisi ujian adalah :

- a. Mata kuliah
- b. Fakultas
- c. Jurusan
- d. Waktu
- e. Aspek yang diukur
- f. Tipe/bentuk soal yang digunakan
- g. Ruang lingkup dan sub pokok bahasan
- h. Tingkat kesukaran
- i. Jumlah Soal tiap pokok bahasan ruang lingkup ujian
- j. Jumlah
- k. Persentase tiap pokok bahasan (ruang lingkup)

Selanjutnya perhatikan contoh 1, 2 dan 3. Ketiga contoh tersebut dari yang paling sederhana sampai kepada yang kompleks (terlampir)

7. Teknik Penilaian untuk Pengetahuan dan Pemahaman

Pengetahuan (knowlegde) merupakan tingkatan kemampuan yang paling rendah dalam ranah kognitif. Mahasiswa diharapkan mengingat kembali apa yang telah disampaikan kepadanya berupa : fakta, definisi, terminologi dan pengetahuan siap lainnya. Sedangkan dalam pemahaman mahasiswa telah mengerti/menguasai pengetahuan yang ada padanya. Telah terjadi secara sistematis

tis proses internalisasi terhadap pengetahuan yang disampaikan padanya. Mahasiswa telah mampu menterjemahkan konsep-konsep yang disampaikan. Mereka juga telah mampu menginterpretasikan ;membedakan yang penting dan tidak penting serta meng ekstrapolasikan sesuatu dengan jelas.

Kemampuan pengetahuan dapat diukur dengan tipe tes antara lain:

pilihan v: (1) melengkapi (complation item), (2) jawaban pendek/ melengkapi (supply item), (3) menjodohkan (matching item), (4) pilihan jamak (multiple choce), (5) Betul Salah (True-false). Sedangkan kemampuan pemahaman lebih baik diukur dengan tipe tes pilihan yang berbentuk pilihan jamak atau kasus.

a. Aturan umum tes pilihan

1. Tiap-tiap bentuk soal hendaklah didahului oleh petunjuk atau penjelasan bagaimana mengerjakan soal itu. Petunjuk itu harus jelas, ringkas dan sederhana. Apabila dalam tes itu akan digunakan formula tertentu dalam men skornya, maka perlu dinyatakan dalam petunjuk atau penjelasan.
2. Jadganlah mengutip secara langsung kalimat yang sama dengan materi yang terdapat dalam buku pegangan, kecuali untuk menangkap isi bacaan itu.
3. Usahakan agar mahasiswa tidak mengalami kesukaran dalam membaca butir soal. Bahasa yang dipakai disesuaikan dengan tingkat kematangan peserta didik yang mengikuti ujian.
4. Kalau pendapat yang ditanyakan bersumber dari orang tertentu, maka sebaiknya dinyatakan sumber tersebut.

5. Usahakan setiap butir soal, tidak memberi petunjuk pada butir soal yang lain.
6. Dalam menyusun kunci jawaban janganlah dilakukan secara sistematis.
7. Hindarkan pertanyaan-pertanyaan yang meragukan.
8. Hindarkan membuat soal yang tidak ada gunanya dan

b. Jawaban Pendek/ Melengkapi.

Butir soal bentuk melengkapi ini dapat diklasifikasikan menjadi

1. Jawaban pendek (short answer)
2. Isian (Completion)

Kedua jenis ini hanya berbeda dalam bentuk persoalan yang disajikan. Jika persoalan disajikan dalam bentuk pertanyaan, maka ia adalah jawaban pendek, dan apabila disajikan dalam bentuk pertanyaan yang belum selesai, maka ia adalah isian.

Butir-butir bentuk melengkapi dapat digunakan untuk mengukur pembedaharaan kata-kata, nama-nama tanggal dan tahun serta kemampuan memecahkan masalah atau persoalan yang berhubungan dengan angka.

Contoh : 1. Perang Paderi dipimpin oleh

Kunci Tuanku Imam Bonjol

2. Siapakah yang memimpin Perang paderi ?

Jawab Tuanku Imam Bonjol

1. Beberapa saran dalam menyusun soal bentuk Melengkapi

a. Kata-kata untuk jawaban yang disediakan

- hendaklah pendek dan khusus
- b. Jangan mengambil pertanyaan langsung dari buku (textbook)
 - c. Pertanyaan langsung biasanya lebih disukai dari pada pertanyaan tidak lengkap
 - d. Apabila pertanyaan dinyatakan dalam unit satuan maka menyatakan unit kesatuan itu, seperti meter, kilogram dan sebagainya.
 - e. Pengosongan tempat yang disediakan hendaklah diusahakan sama panjangnya
 - f. Jangan terlalu banyak bagian yang dikosongkan Andai kata satu, sebaiknya ditempatkan di akhir atau pada permulaan kalimat
 - g. Hindari pemakaian kata-kata atau petunjuk yang mengarah pada jawaban yang benar

Kurang baik

Pengarang dari pada tes inteligensi yang mula-mula adalah

.....

Lebih baik

Tes inteligensi verbal yang mula-mula dipersiapkan oleh...

.....

Siapakah yang mempersiapkan tes intelegensi verbal yang mula-mula ?

d. Pilihan Jamak

Tes bentuk ini terdiri dari dua bahagian. Bagian pertama disebut dengan stem, yaitu : pertanyaan/pokok/batang yang menampilkan masalah tertentu.

Pertanyaan itu dapat berupa pertanyaan yang lengkap atau dapat pula berupa pertanyaan yang belum selesai.

Bagian kedua disebut dengan option, yaitu kemungkinan jawaban yang disediakan.

Contoh :

1. Pernyataan yang belum selesai

Validitas suatu tes menyatakan :

- (A). tingkat penghargaan orang pada tes itu
- (B). tepat-tidaknya tes itu sebagai alat pengukur sesuatu yang hendak diukurnya
- (C). tingkat kepercayaan yang dapat diberikan kepada tes itu sebagai alat pengukur
- (D). tingkat ketetapan yang dimiliki oleh tes itu

Kunci (B)

2. Bentuk pertanyaan

Siapakah pelopor gerakan Budi Utomo tahun 1908 ?

- A. Ir. Soekarno
- b. Mohamad Hatta
- C. Wahidin Sudirohusodo
- D. Tjipto Mangun Kusumo
- E. Sutomo

Kunci C

Bagi dosen yang belum berpengalaman dalam menulis soal bentuk pertanyaan jauh lebih mudah dari kalimat yang tidak lengkap.

Jumlah "option" (alternatif) jawaban berbeda-beda. Makin sedikit optionnya, makin besar pula faktor terkaan

option yang wajar 4 atau 5 dan setiap alternatif hendaklah berfungsi dengan baik.

1. Beberapa pedoman dalam menyusun soal bentuk pilihan Jamak:

a. Stem hendaklah dirumuskan atau dinyatakan dengan jelas sebagai suatu masalah atau pernyataan yang perlu dijawab.

Contoh :

Kurang baik

Didalam penelitian kualitatif banyak faktor yang perlu dipertimbangkan dalam menentukan validitas internal, maka faktor itu adalah :

- A. History
- B. Conference
- C. Maturity
- D. Testing

Kunci : B

Lebih baik

Banyak faktor yang mempengaruhi validitas internal. Tiga diantaranya, seperti tersebut dibawah ini, kecuali :

- A. History
- B. Conference
- C. Maturity
- D. Testing

Kunci B

b. Masukkan sebanyak mungkin hal kedalam stem (Pernyataan)

Contoh :

Kurang baik

Salah satu faktor pendukung munculnya konsep pendidikan seumur hidup adalah :

- A. Bertambahnya negara barat memberikan bantuan pendidikan
- B. Bertambahnya ilmu dan teknologi
- C. Bertambahnya jumlah penduduk yang tak bersekolah
- D. Bertambahnya permintaan akan pendidikan

Kunci B

Lebih baik

Salah satu faktor pendukung munculnya konsep pendidikan seumur hidup adalah :

- A. Bertambahnya negara barat memberikan bantuan pendidikan
- B. banyaknya fasilitas sekolah yang belum dimanfaatkan
- C. berkembangnya dengan cepat ilmu dan teknologi
- D. meningkatnya kebutuhan pendidikan di dunia

Kunci C :

c. Jangan dibebani stem dengan pernyataan yang tidak berarti atau artinya tidak jelas.

Contoh :

Kurang baik :

Teori belajar menurut psikologi daya, sebagai mana yang telah diutarakan oleh beberapa ahli di barat, memandang bahwa :

- A. Mengajar itu merupakan usaha melatih daya daya jiwa
- B. Mengajar itu adalah mengembangkan struktur kognitif
- C. Mengajar adalah menanamkan pengertian
- D. Mengajar merupakan pembentukan hubungan stimulus respon

Kunci A :

Lebih baik

Menurut Psikologi daya, mengajar merupakan usaha

- A. melatih daya-daya jiwa
- B. mengembangkan struktur kognitif
- C. menanamkan pengertian
- D. membentuk hubungan Stimulus-respon

Kunci A :

- d. Masalah yang ditampilkan dalam stem harus sedemikian rupa sehingga benar-benar hanya ada satu jawaban yang benar.

Contoh :

Kurang baik

Salah satu masalah pendidikan yang dihadapi pemerintah pada permulaan Orde Baru adalah :

- A. Mutu pendidikan yang rendah
- B. Ketimpangan horizontal dan vertikal
- C. Peledakan jumlah penduduk usia sekolah
- D. Fasilitas pendidikan yang terbatas

Lebih baik

Salah satu masalah yang dihadapi pemerintah pada tahun 1966-1968 adalah :

- A. jumlah penduduk usia sekolah dasar meningkat dengan cepat
- B. banyaknya kelas VI yang tidak lulus
- C. terbatasnya kemampuan guru SD
- D. beratnya biaya pendidikan yang dipikul oleh masyarakat

Kunci A

- e. Stem hendaklah dinyatakan dengan bahasa yang jelas
- f. Hindari memakai kalimat yang panjang, karena cenderung memberi petunjuk. Usahakan semua alternatif tidak berbeda panjangnya
- g. Gunakan pernyataan yang bersifat positif, kecuali untuk maksud tertentu
- h. Semua kemungkinan jawaban yang diberikan hendaklah masuk akal
- i. Hendaklah dihindari adanya hubungan asosiasi verbal antara stem dengan kemungkinan jawaban

Contoh :

Kurang baik

Badan yang memberikan saran dan pertimbangan

kepada Presiden dalam melaksanakan pemerintahan adalah :

- A. Dewan Perwakilan Rakyat
- B. Majelis Permusyawaratan Rakyat
- C. Dewan Pertimbangan Agung
- D. Bapenas

Lebih baik

Menurut UUD 1945 badan yang dapat memberikan saran kepada

Presiden dalam melaksanakan pemerintahan sehari-hari adalah:

- A. DPR
- B. MPR
- C. DPA
- D. Bappenas

- j). Pilihan jawaban yang diberikan janganlah menunjuk kunci jawaban. Usahakan semua alternatif jawaban hampir sama dan hanya satulah jawaban yang baik.
- k). Pilihan jawaban yang benar janganlah diletakkan secara sistematis, tetapi letakkanlah secara acak (random) dalam keseluruhan soal.
- l). Jangan menggunakan pilihan jamak kalau bentuk soal lain dapat digunakan dan lebih baik hasilnya.
- m). Hindarkan menggunakan susunan kalimat seperti yang terdapat dalam buku pelajaran.
- n). Usahakan alternatif jawaban yang disajikan agak homogen, baik yang ditinjau yang dari isi maupun bentuknya.
- o). Jangan menggunakan kata-kata petunjuk seperti selalu, kadang-kadang, jarang dan pada umumnya.
- p). Bahasa yang dipakai hendaklah sederhana dan mudah dipahami siswa.
- q). Petunjuk yang diberikan harus jelas, sehingga siswa tidak ragu-ragu dalam mengerjakannya.
- r). Kalimat-kalimat pokok dalam satu soal tidak tergantung pada butir soal lain sehingga memberi petunjuk.

Soal-soal pilihan jamak juga sering disebut orang dengan pilihan ganda atau pilihan "multiple respon test".

f. Bentuk Kecuali

Bentuk lain yang dapat kita pakai untuk kurangi faktor terkaan ialah membuat salah satu pilihan jawaban menjadi "bukan salah satu di atas", atau merubah menjadi bentuk KEQUALI:

Contoh:

1. Reliabilitas suatu test menyatakan:

- (A). tingkat ketetapan (consistency) yang dimiliki tes itu secara keseluruhan
- (B). tingkat ketetapan yang dimiliki oleh sebagian oleh tes itu
- (C). tingkat kebaikan tes itu untuk suatu maksud tertentu
- (D). Semua pernyataan di atas benar.

Kunci : (A)

2. Kebaikan tes objektif dibandingkan tes essay adalah seperti tersebut di bawah ini, KEQUALI:

- (A). skoring dapat dilakukan dengan cepat dan tepat
- (B). dapat diskor siapa saja
- (C). dapat mengetahui fakta dalam ilmu dengan efisien
- (D). bebas dari kemungkinan terkaan

Kunci : (C)

g. Menjodohkan

Bentuk soal ini seringkali lebih efisien apabila digunakan untuk mengetahui tentang fakta, seperti tanggal kejadian, hubungan pengarang dengan buku-buku atau suatu karangan, hubungan simbol kimia dengan campuran kimia, hubungan tahun kejadian dan tempat kejadian, dan sebagainya. Soal bentuk ini terdiri dari dua bagian, yaitu pokok persoalan dan kemungkinan jawaban.

Biasanya kedua bagian itu ditetapkan sejajar, pokok persoalan disebelah kanan dan bagian jawaban disebelah kiri.

Dengan demikian, orang yang dites dapat melihat kedua bagian itu sekali gus.

Seperti juga dalam soal pilihan jamak, maka dalam soal menjodohkan ini juga perlu diperhatikan faktor keseragaman, sehingga dapat dikurangi faktor terkaan. Disamping itu kemungkinan jawaban yang disediakan hendaknya lebih banyak dari pada jumlah persoalan yang dikemukakan.

Contoh:

1. Spranger A. Yang mula-mula menyusun tes kecerdasan untuk membedakan anak-anak yang dungu dari anak yang normal
2. Heyman B. Yang membagi manusia atas kelompok-kelompok: manusia ekonomi, sosial estetis, politis dan teoritis, berdasarkan nilai budaya yang dijunjungnya.
3. Binet C. Penganjur ilmu jiwa analisa dan ilmu jiwa dalam
D. Yang membagi manusia atas dua golongan: introvert dan extrovert.
4. Povlov E. Yang melakukan eksperimen dengan anjing, untuk mendasari pendapatnya bahwa perbuatan belajar itu adalah proses kondisioning.
F. Yang mengemukakan delapan jenis watak (karakter) berdasarkan tiga buah sifat dasar:
 - a. perasaan (emotionality)
 - b. keinginan (activity)
 - c. proses pengiring

Kunci :

1 - B

2 - F

3 - A

4 - E

1). Aturan-aturan dalam menyusun soal bentuk menjodohkan.

a). Persoalan yang dikemukakan hendaklah homogen. Misalnya semua tentang nama orang yang ahli dalam bidang pendidikan atau tentang nama tahun berdirinya pergerakan kemerdekaan.

Contoh :

Kurang baik

1. Tuanku Imam Bonjol

A. Usaha sadar dalam membimbing anak agar berkembang seoptimal mungkin

2. Pendidikan

B. Tahun 1821 - 1937

3. Teuku Umar

C. Pahlawan dalam perang Aceh melawan Belanda

D. Tahun 1874 - 1904

Lebih baik

1. Tuanku Imam Bonjol

A. Pemimpin perang Paderi

B. Pemimpin perang Aceh

2. Teuku Umar

C. Pemimpin perang Diponegoro

3. Wahidin Sudirohusodo

D. Pelopor pergerakan Budi Utomo

b). Jumlah jawaban hendaklah lebih banyak disediakan dari masalah atau persoalan (lihat contoh nomor satu).

Hal ini dimaksudkan agar dapat mengurangi faktor terkaan sebab siswa harus memilih dari jumlah yang lebih banyak.

c). Persoalan yang disajikan hendaklah pendek dan jelas (lihat contoh no.1)

Pernyataan pendek dan jelas akan memudahkan homogenitas dalam menyusun butir soal. Disamping itu akan membantu siswa dalam memilih jawaban yang tepat.

d). Respon atau jawaban hendaklah diatur secara logis.

Umpamanya: Menurut alpabet atau secara kronologis.

Hal ini akan membantu siswa dalam menentukan jawabannya.

e). Petunjuk hendaklah jelas. Dalam petunjuk hendaklah dinyatakan apakah satu alternatif respon hanya digunakan untuk satu kali atau lebih dari satu persoalan.

f). Tempatkan semua persoalan dan jawaban pada satu halaman, kalau semuanya merupakan satu kesatuan soal.

Hal ini akan membantu siswa dalam menjawab persoalan dengan cepat. Sebab ia tidak perlu bolak balik dan memahami kesatuan soal secara utuh.

h. Pilihan Berganda

Pada prinsipnya bentuk ini sama dengan pilihan jamak dengan lima pilihan. Yang membendakannya adalah yang benar dapat satu, dua, tiga atau empat.

Petunjuk:

Pilihlah:

(A). Jika (1), (2) dan (3) benar

(B). Jika (1) dan (3) benar

(C). Jika (2) dan (4) benar

(D). Jika (4) yang benar

(E). Jika semuanya benar

Contoh:

1. Pendidikan adalah tanggung jawab dari pada
 1. orang tua
 2. masyarakat
 3. pemerintah
 4. guru

2. Tingkah laku seorang anak unik artinya ia memiliki sifat-sifat yang hanya dimilikinya sendiri. Hal ini disebabkan perbedaan unsur-unsur/faktor-faktor sebagai berikut.
 1. perasaan
 2. pikiran
 3. kemauan
 4. pola hidup

8. Teknik penilaian untuk :Aplikasi, Analisis, Sintesis dan Evaluasi.

Seperti telah disinggung pada uraian di muka, kemampuan Aplikasi, Analisis, Sintesis dan Evaluasi merupakan kemampuan kognitif yang lebih tinggi tingkatannya. Kemampuan aplikasi lebih terkait dengan kemampuan mahasiswa mengaplikasikan prinsip-prinsip yang sudah dipelajari terhadap masalah baru atau situasi baru. Disamping itu juga kemampuan menyusun pernyataan yang bersifat umum yang dapat diaplikasikan dalam situasi baru.

Sehubungan dengan itu berapa bentuk tes yang dapat digunakan: pilihan jamak, analisis kasus dan tes essay.

Kemampuan analisis bukan sekedar mengerti ide atau masalah dan bukan pula hanya mengaplikasikan ide, konsep maupun konstruk terhadap masalah dan situasi baru, tetapi lebih komplek dari itu. Mahasiswa diharapkan mampu menganalisis, memisah-misah masalah ide atau dokumen menjadi unsur dan mampu melihat hubungan diantara unsur-unsur itu. Bentuk tes yang dapat digunakan untuk mengukur

kemampuan ini antara lain adalah: analisis kasus, pilihan jamak, analisis hubungan dan pilihan berganda serta tes essay.

Untuk mengukur kemampuan sintesis dan evaluasi dapat pula digunakan pilihan jamak, analisis kasus dan tes essay.

Contoh: Aplikasi pemogokan terhadap situasi baru

Di bawah ini dikemukakan berita-berita surat kabar tentang tindakan yang dilakukan oleh para pegawai. Untuk setiap berita itu tuliskanlah pada titik di-depannya huruf sebagai

berikut:

- A. jika hal itu merupakan suatu pemogokan biasa
- B. jika hal itu merupakan suatu mogok duduk
- C. jika hal itu merupakan suatu pemogokan yang berdasarkan undang-undang
- D. jika hal itu merupakan suatu pemogokan yang berdasarkan perasaan simpati
- E. jika hal itu tidak diperkenankan undang-undang perburuhan
- F. jika tidak satupun diantara kelima di atas yang cocok.

menghentikan kegiatannya. Mereka menolak untuk meninggalkan

kantor itu.

..D... 2. Para sopir bis kota hari ini menolak untuk membawa penumpang dari terminal X ketika sopir becak sedang mogok.

..F... 3. Kemaren jam 12.30 seluruh pegawai di negara Y menghentikan pekerjaannya selama 24 jam.

Tingkat penguasaan: Aplikasi

Dalam suatu proyek pembangunan gedung bertingkat, tukang-tukang tembok memasang palang-palang kayu untuk menahan temboknya. Kemudian tukang-tukang kayu tidak mau bekerja lagi dan menuntut bahwa pemasangan palang-palang kayu merupakan bagian-nya juga.

Peristiwa itu adalah contoh dari suatu:

- A. pemogokan umum
- B. pemogokan berdasarkan undang-undang
- C. mogok duduk
- D. pemboikotan
- E. pemogokan berdasarkan rasa simpati.

Tingkat penguasaan: Analisis

Bentuk tes yang digunakan Analisis Hubungan.

Contoh: PILIHAN:

- (A). Jika pernyataan dan alasan benar, dan alasan merupakan sebab dari pernyataan
- (B). Jika pernyataan dan alasan keduanya benar, tetapi alasan bukan sebab pernyataan tersebut
- (C). Jika pernyataan benar tetapi alasan salah
- (D). Jika pernyataan salah, tetapi alasan benar
- (E). Jika pernyataan pernyataan dan alasan merupakan jalan

pikiran yang tidak benar.

Pernyataan:

Alasan

- | | | |
|--|-------|---|
| (1). Dalam menyusun kurikulum suatu sekolah kita perlu mempertimbangkan tujuan pendidikan, faktor anak dan organisasi kurikulum. | sebab | Faktor-faktor itu akan menentukan dan mempengaruhi proses belajar mengajar di sekolah, dan mutu pemakai hasil pendidikan. |
|--|-------|---|

Kunci : (A)

- | | | |
|-----------------------------|-------|--------------------|
| (2). Sistim klasikal sangat | sebab | Interaksi belajar- |
|-----------------------------|-------|--------------------|

membantu dalam mengem-
bangkan murid menurut
polanya masing-masing

mengajar yang baik
akan membantu anak
dalam pertumbuhannya.

Tingkat penguasaan: sintesis dan evaluasi dengan menggunakan tes essay.

Contoh: Ceritakanlah dengan jelas bagaimana hubungan antara perencanaan pendidikan dengan pembangunan ekonomi.

Jawaban: Pada kertas terpisah

Contoh: Jelaskanlah 2 cara yang dapat digunakan dalam menentukan realibilitas suatu tes.

Bentuk lain yang sering digunakan adalah dengan menyusun suatu cerita terlebih dahulu dan kemudian mahasiswa yang diuji menjawab pertanyaan yang bersifat evaluatif.

9. Skoring dan Penentuan Lulus

Skoring tes pilihan dan tes essei berbeda-beda menurut bentuk dan tipe tes yang digunakan. Untuk tes pilihan ada yang menggunakan rumus, seperti

$$S = \frac{B - W}{n-1}$$

tetapi ada pula yang hanya menghitung yang benar saja.

Menurut hemat penulis, kalau dengan menggunakan formula itu akan didapatkan hasil yang lebih cermat lagi maka dianjurkan untuk menggunakannya, tetapi seandainya akan merugikan hasil yang akan dicapai, maka sebaiknya tidak akan menggunakan rumus, kecuali kalau yang digunakan adalah tes yang telah baku (standard ized test).

Seorang mahasiswa mengikuti suatu tes pilihan lima option (100 butir) jawaban yang benar 70 buah dan yang salah 30 buah, maka skor mentah yang bersangkutan adalah:

Tanpa rumus 70, tetapi dengan menggunakan rumus adalah

$$\begin{aligned} S &= 70 - \frac{30}{5} \\ &= 64. \end{aligned}$$

Langkah-langkah yang perlu diperhatikan dalam menskor tes essei adalah:-

- a. Persiapan dengan jelas dan terinci jawaban yang diharapkan untuk setiap pertanyaan.
- b. Pertimbangan bobot skor untuk setiap butir soal tersebut.
- c. Skor menurut butir soal untuk semua mahasiswa. Nomor 1, sampai selesai semua mahasiswa, no 2, dstnya. Ini dimaksudkan untuk menghilangkan "Halo effect", yaitu pengaruh nilai butir sebelumnya terhadap nomor berikutnya.
- d. Jumlahkan skor mentah setiap butir soal masing-masing mahasiswa.

Dalam menentukan batas lulus dapat digunakan PAP atau PAN dan kemudian dapat pula ditentukan nilai mahasiswa.

a. Penilaian Acuan Patokan. (PAP)

Sejalan dengan konsep tuntas, dalam penilaian dapat digunakan PAP. Konsep ini menekankan, seseorang dinyatakan telah menguasai suatu materi apabila ia telah mencapai standar yang telah ditetapkan. Umpamanya, seseorang mahasiswa baru dinyatakan lulus kalau ia menguasai 60% dari bahan ujian, apa bila penguasaannya dibawah itu maka ia akan dinyatakan gagal.

Selanjutnya, skor-skor yang berada di atas 60% tingkat penguasaannya dari keseluruhan bahan ujian disusun dan diklasifikasikan menurut sistem tertentu dan masing-masing tingkat penguasaan diberi nilai tertentu. Misalnya patokan yang dimaksudkan itu ditentukan sebagai berikut:

Derajat Penguasaan

Nilai akhir

90 % - 100 %

A

80 % - 89 %

B

65 % - 79 %

C

60 % - 64 %

D

(60%

F (tidak lulus)

Contoh :

No	Nama	Skor mentah
1.	Ali	70
2.	Corry	90
3.	Sujono	60
4.	Ely	75
5.	Martinus	40
6.	Budi	55
7.	Hasbullah	78
8.	Jiyono	80
9.	Munarfi	85

Skor maksimal 110.

Batas lulus adalah apabila siswa r 60% dari bahan yang sukar dalam meteri/soal ujian. Ini berarti batas lulus adalah:

60

$$X \ 110 = 66$$

100

Mhs yang mendapatkan skor 66 ke atas dinyatakan lulus, sedangkan kecil dari 66, dinyatakan gagal.

Dengan pembatasan berdasarkan tingkat penguasaan itu dapat ditentukan nilai masing-masing siswa.

Ali	70	----- X 100 = 64 %	D
	110		
Corry	90	----- X 100 = 82 %	B
	110		
Sujono	60	----- X 100 = 55 %	Gagal
	110		
Ely	75	----- X 100 = 68 %	C
	110		
Martinus	40	----- X 100 = 36 %	Gagal
	110		
Budi	55	----- X 100 = 50 %	Gagal
	110		
Hasbullah	80	----- X 100 = 71 %	C
	110		
Jiyono	80	----- X 100 = 73 %	C
	110		
Munarfi	85	----- X 100 = 77 %	C
	110		

b. Acuan norma kelompok

Selanjutnya kalau ingin di gunakan penilaian berdasarkan norma kelas, maka dosen hendaklah mencari rata-rata kelompok (X) dan standar deviasinya. Misalnya jenjang nilai itu ditentukan sebagai berikut:

Batas lulus $X - 1,00$ SD (ditentukan sesudah penyebaran skor siswa diketahui).

Batas daerah dalam kurva	Nilai	Keterangan
Lebih dari $X + 2,00$ SD	A	Lulus
antara $X + 1,00$ SD dan $X + 2,00$ SD	B	Lulus
antara X dan $X + 1,00$ SD	C	Lulus
antara $X - 1,00$ SD dan X	D	Lulus
antara dari $X - 1,00$ SD	E	Gagal

Contoh di atas dapat dimodifikasi, tetapi perlu diingat bahwa penyebaran skor dalam kelompok itu hendaklah merupakan kurva normal.

Dengan menggunakan data pada Penilaian Acuan Patokan, dapat di cari:

$$X = \frac{633}{9} = 70,3$$

$$SD = 15,04$$

$$\begin{aligned} \text{Batas lulus} &= X - 1,00 \text{ SD} \\ &= 70,3 - 15,04 \\ &= 55,26 \end{aligned}$$

Letakkan skor batas lulus sebagai batas bawah lulus interval terendah bagi yang lulus, dan lanjutkan dengan kelas interval di atas dan di bawahnya, dengan mempedomani tabel konversi sehingga didapat tabel berikut:

skor	nilai huruf
> 100	A
86 - 100	B
70 - 85	C
55 - 70	D
< 55	E

Berdasarkan acuan kelompok/kelas tersebut dapatlah ditentukan nilai masing-masing siswa:

Nama	Skor mentah	Nilai huruf
Ali	70	D
Corry	90	B
Suyono	60	D
Ely	75	C
Martinus	40	Gagal
Budi	55	D
Hasbullah	78	C
Jiyono	80	C
Munarfi	85	C

Di samping cara di atas skor mentah dapat pula di robah menjadi persentile dan selanjutnya baru dirobah menjadi nilai huruf. dengan menggunakan PAP dan PAN.

10. Indeks Kesukaran dan daya pembeda butir soal.

Pengembangan alat ukur penilaian hasil belajar merupakan salah satu upaya untuk mendapatkan alat ukur yang valid dan reliable. Untuk itu perlu dilakukan analisis butir soal dan daya pembeda butir soal.

Indek Kesukaran Butir Soal

Beberapa ahli psikometrik telah mengembangkan beberapa cara untuk menentukan indeks butir soal. Pada prinsipnya indeks kesukaran butir soal dapat didefinisikan sebagai proporsi (persentase) mahasiswa yang menjawab benar suatu butir soal. Indeks tersebut dapat ditentukan dengan menggunakan rumus:

$$TK = \frac{JSB}{JST}$$

JST

Tk = Tingkat kesukaran

JSB = jumlah siswa yang menjawab benar

JST = jumlah siswa total

Umpama: jumlah mahasiswa yang mengikuti tes 270 orang jumlah yang menjawab benar 56 orang

$$Tk = \frac{56}{270}$$

$$= 0.21 \text{ atau } 21 \%$$

270

Selanjutnya membandingkan indeks tingkat kesukaran tersebut dengan klasifikasi Indeks Kesukaran Butir soal dibawah ini.

Klasifikasi Indeks Kesukaran Butir

Tingkat kesukaran butir soal	Klasifikasi butir soal
> 0,90	sangat mudah
0.70 - 0.90	mudah
0.30 - 0.70	sedang
< 0.30	sukar

Soal tersebut di atas (butir 1) adalah sukar.

Cara lain yang dapat digunakan ialah apabila distribusi skor normal , dan jumlah peserta ujian banyak serta analisis akan dilakukan secara manual, adalah dengan mengambil 27 % yang berskor tinggi dan 27 % , yang berskor terendah, dengan langkah-langkah sebagai berikut:

- Susunlah lembaran jawaban yang telah di skor dengan menempatkan skor paling tinggi di atas sekali dan kemudian diikuti lembaran jawaban berikutnya, dan paling bawah adalah skor yang paling rendah.
- Ambillah 27 % jawaban dari kelompok yang mendapatkan skor tinggi (kelompok pintar) dan 27 % dari kelompok yang mendapatkan skor kurang atau paling rendah. Pengambilan 27 % dari kelompok kemudian dari atas sekali kelompok ini merupakan kelompok pintar. Pengambilan 27 % dari bawah; mulai dari paling bawah, sehingga merupakan kelompok kurang.
- Ambil lembaran jawaban mereka yang termasuk 27 % pintar dan 27 % kurang. Masukkanlah isi jawaban mereka kedalam lembaran jawaban analisis soal.
- Tentukanlah siswa yang menjawab benar dalam setiap item, baik dalam kelompok pintar maupun kelompok kurang.

Contoh:

=====

No. :	Kunci:	NP :	Pilihan jawaban					:%JP:	D :	r :
soal	soal	NK :	A :	B :	C :	D :	E :	00 :%JK:	S	Delta

bis

1	:	B	:	135:	63:	37:	15:	15:	7:	-	:	27:	21:	18:	15,3
---	---	---	---	------	-----	-----	-----	-----	----	---	---	-----	-----	-----	------

	:		:	135:	67:	19:	25:	19:	5:		:	14:			
--	---	--	---	------	-----	-----	-----	-----	----	--	---	-----	--	--	--

2	:	C	:	135:	105:	15:	4:	5:	6:	-	:	78:			
---	---	---	---	------	------	-----	----	----	----	---	---	-----	--	--	--

	:		:	135:	28:	50:	18:	14:	25:	-	:	21:			
--	---	--	---	------	-----	-----	-----	-----	-----	---	---	-----	--	--	--

=====

e. Masukkan ke dalam rumus sebagai berikut:

$$JP + JK$$

$$DS = \frac{\quad}{NP + NK} \times 100 \%$$

$$NP + NK$$

dimana :

DS = derajat kesukaran item (dalam %)

JP = jumlah kelompok pintar yang menjawab benar

JK = jumlah kelompok lemah/kurang yang menjawab benar

NP = jumlah peserta tes yang termasuk 27 % kelompok pintar

NK = jumlah peserta tes yang termasuk 27 % kelompok kurang

Contoh : dengan menggunakan data nomor 1

$$DS = \frac{37 + 19}{135 + 135}$$

$$= \frac{56}{270} \times 100$$

$$= 21 \%$$

$$= 21 \%$$

f. Langkah berikutnya ialah mencari dalam tabel "Composite Graph for Determining Student ABILITY OR DELTA ITEM" (terlampir). Caranya adalah sebagai berikut: Lihat pada tabel itu angka 21 %. Kemudian tarik garis horizontal sehingga memotong kurva yang berbentuk Ogive. Pada potongan itu tarik garis vertikal sehingga menunjukkan besarnya delta doal tersebut. Soal nomor 1 deltanya adalah 15,5. Selanjutnya membandingkan harga delta yang diperdapat dengan tingkat kesukaran yang baik dalam suatu tes. Kategori itu adalah sebagai berikut:

Delta (d)	Kategori:
> 14,1	sukar
10,9 - 14,00	sedang
< 10.8	mudah

Dari contoh di atas ternyata derajat kesukaran yang didapat lebih tinggi dari 14,1. Ini berarti bahwa item nomor 1 adalah sukar. Karena itu perlu direvisi.

Butir soal yang baik adalah jika butir soal itu tidak terlalu sukar dan tidak pula terlalu mudah. Di samping itu butir soal hendaklah mampu membedakan kelompok yang pintar dengan kelompok yang tidak pintar. Untuk mengetahui yang terakhir ini, perlu dilakukan analisis butir untuk mengetahui daya pembeda tiap butir soal.

Apabila distribusi skor yang akan dianalisis berdistribusi normal, maka lembaran jawaban yang akan dianalisis cukup 27 % dari kelompok pintar dan 27 % dari kelompok kurang.

Daya pembeda dari suatu item dinyatakan dalam bilangan desimal dan berkisar antara 0.00 sampai 1.00. Apabila suatu soal dapat dijawab dengan benar oleh kelompok pintar dan kelompok kurang maka soal itu dikatakan tidak baik sebab tidak dapat membedakan kelompok pintar dari kelompok kurang.

Seandainya suatu soal dapat dijawab oleh kelompok kurang, sedangkan kelompok pintar tidak menjawab sama sekali, atau tidak dapat dijawab oleh kelompok pintar maupun kelom-

pok kurang maka soal itu dikatakan tidak baik karena tidak mempunyai daya pembeda atau mempunyai daya pembeda negatif.

Soal yang baik adalah apabila soal itu hanya dapat dijawab oleh kelompok pintar saja. Ini berarti bahwa daya pembeda soal itu adalah 1.00 daya pembeda tiap butir soal adalah 0.00 (nol), apabila jumlah siswa yang menjawab benar dari kelompok pintar dan kurang adalah sama.

Langkah-langkah dalam menentukan daya pembeda adalah sebagai berikut:

- a. Tentukan skor masing-masing siswa sesuai dengan pedoman evaluasi (manual) yang telah ditetapkan.
- b. Bagi kelompok itu menjadi dua kelompok berdasarkan skor yang mereka per dapat. Kedua kelompok adalah kelompok pintar dan kelompok kurang.
- c. Apabila jumlah kelompok besar, ambillah 27% dari kelompok pintar dan dimuali dengan nilai tertinggi, selanjutnya ambil pula 27% dari kelompok kurang dan dimulai dari nilai terendah.

```
-----+
Contoh: 90 |
        87 |
        86 | - kelompok pintar
        .. |           27%
        .. |
        -----+
```

```
-----+
        60 |
        56 | - kelompok kurang
        55 |           27%
        .. |
        .. |
        -----+
```

- d. Masukkanlah masing-masing lembaran jawaban pada kartu analisis soal menurut kelompoknya, seperti mencari tingkat kesukaran soal
- e. Langkah berikutnya adalah menghitung persentase kelompok pintar yang menjawab benar dan persentase kelompok kurang yang menjawab benar, dengan rumus:

$$PH = \frac{JP}{135} \times 100 \% = 27,1$$

$$PL = \frac{19}{135} \times 100 \% = 14,07$$

- f. Mencari r_{bis} korelasi, dengan jalan menggunakan tabel "normalized biserial coefficients of coorelation" (terlampir).

Cara mencari r_{bis} pada tabel itu adalah sebagai berikut:

Pada lajur atas (horizontal) carilah angka 27 atau mendekati 27 (dalam hal ini angka 26). Kemudian untuk lajur vertikal (kiri atau kanan) carilah angka 14. Tarik garis vertikal dari angka 26 dan garis horizontal dari angka 14.

Kedua garis itu akan berpotongan pada angka 18. Jadi untuk soal nomor satu r_{bis} adalah 0.18.

- g. Langkah terakhir adalah membandingkan r_{bis} yang didapat dengan tabel daya pembeda sebagai berikut:

Daya pembeda	Kategori
0.40 - 1.00	baik sekali
0.30 - 0.39	baik
0.20 - 0.29	sedang
0.01 - 0.19	kurang

(Chung Teh Fan; 1952)

Dengan daya pembeda 0.18 (yang didapat), berarti daya pembeda itu jatuh pada kategori "kurang". Ini berarti bahwa item

nomor satu tidak dapat membedakan antara kelompok yang pintar dengan kelompok yang kurang.

Cara lain yang dapat digunakan untuk mencari daya pembeda dengan menggunakan rumus:

$$\text{Daya pembeda (DP)} = \frac{Jp}{Np} - \frac{Jk}{Nk} = Pp - Pk$$

dalam mana: Jp = jumlah siswa kelompok pintar yang menjawab benar

Jk = jumlah siswa kelompok kurang yang menjawab

Np = jumlah siswa kelompok pintar

Nk = jumlah siswa kelompok kurang

Pp = proporsi kelompok pintar yang menjawab benar

Pk = proporsi kelompok kurang yang menjawab

benar.

Contoh :

Jumlah kelompok pintar 135 orang

Jumlah kelompok kurang 135 orang

Penyebaran jawaban :	A	B	C	D	E	0
Kelompok pintar	8	15	21	38	51	2
Kelompok kurang	27	22	17	47	21	1

Kunci E

$$\text{DP} = \frac{51}{135} - \frac{21}{135} = 0,38 - 0,16 = 0,22$$

Ini berarti daya pembeda soal tersebut adalah sedang, dan dapat digunakan selanjutnya

Daftar Bacaan

- ~~Anastasi, A., (1968), Psychological Testing, London: Mc Millan.~~
- Ahmann J., Stanley, Glock Marvin D. (1981), Evaluating Student Progress, Principles of Tests and Measurements Boston: Allyn and Bacon, Inc.
- Bloom, B., S., Madaus, G.S., (1981), Evaluation to Improve Learning, New York: Mc Graw Hill Book Company.
- Brown, Frederick G., (1970), Principles of Educational and Psychological Testing, Illinois: The Eryden Press Inc.
- Bloom B., S., Hastings J.T., & Madaus G/F. (1971) Handbook on formative and Summative Evaluation of Student Learning, New York: Mr Graw Hill, Inc.
- Cronbach, L.J. (1970), Essential of Psychological Testing. (1970), New York: Harper & Row.
- Ebel R.L. (1979) Essentials of Educational Measurement. (1979). New York: Prentice Hall Inc.
- Gronlund, N.E. (1981). Measurement and Evaluation in Teaching. (1981). New York: Macmillan Publishing.
- (1973), Preparing referenced Test for Classroom Instruction, New Yersey: Prentice Hall Inc.
- Constructing Achivement Tests, (1982). New Yersey: Prentice Hall Inc.
- Gulford, J.P. (1956). Psychometric Method, New York: Mc Graw Hill Book Company Inc.
- Hill, R., (1981) Measurement and Evaluation in The Classroom,

Ohio: Charles & Merrill Publishing Co.

Kumaidi, (1989) Konsep Dasar Tentang Penilaian Hasil Belajar, Padang.

Popham, W. J. (1975) Educational Evaluation, N. Y. Prentice Hall.

Thorndike, R. L., Hagen, E., (1960) Measurement and Evaluation in Psychology and Education, New York: John Wiley and Sons, Inc.

Tuckman B. W., (1975) Measuring Educational Outcomes, Fundamentals of Testing, New York: American Book Company.

Wrightstone, J. W., Justman J., R., Irving, (1956) Evaluation in Modern Education, New York: American Book Company.

Yusuf, A. Muri, (1988) Dasar-dasar dan Teknik: Evaluasi Pendidikan, Padang.

1. Format tabulasi untuk menentukan Kelompok pintar dan kurang

Mata pelajaran : Semester

No. Urut	Nomor Siswa	Nomor Soal							Jumlah soal yang betul
		1	2	3	4	5	6	7	
1	0001	0	0	1	1	0	0	1	3
2	0003	0	1	0	1	1	1	0	3
3	0004	1	0	0	0	1	0	0	2

Berdasarkan jumlah jawaban yang betul dari tiap siswa maka disusunlah suatu penyebaran data (skor) dari yang tinggi kepada yang rendah, dan kemudian ditentukanlah kelompok pintar dan kelompok kurang. Apabila penyusun soal ingin mengetahui penyebaran jawaban menurut masing-masing alternatif, maka sebaiknya digunakan format 2, . . . kalau tidak gunakan format 3,4.

3. Format tabulasi jawaban tes
dari kelompok pandai (H)

Semester :

No.	Nomor	Nomor Soal							Jumlah
Urut.	Siswa	1	2	3	4	5	6	7	Jawaban
1	: 0012	: 1	: 0	: 1	: 1	: 0	: 1	: 1	: 5
2	: 0014	: 1	: 1	: 1	: 1	: 1	: 1	: 1	: 7
3	: 0017	: 1	: 1	: 0	: 1	: 1	: 1	: 1	: 6
4	: 0018	: 1	: 1	: 1	: 1	: 0	: 1	: 1	: 6
5	: 0020	: 0	: 1	: 1	: 1	: 1	: 0	: 1	: 5
6	: 0023	: 1	: 1	: 0	: 0	: 1	: 1	: 1	: 5
7	: 0024	: 0	: 1	: 1	: 1	: 1	: 1	: 1	: 6
8	: 0025	: 1	: 1	: 1	: 1	: 1	: 1	: 0	: 6
9	: 0028	: 1	: 0	: 1	: 1	: 1	: 1	: 0	: 5
10	: 0029	: 0	: 1	: 1	: 1	: 1	: 0	: 1	: 5
11	: 0030	: 1	: 1	: 0	: 1	: 0	: 1	: 1	: 5
12	: 0031	: 1	: 1	: 0	: 1	: 1	: 1	: 1	: 6
13	: 0032	: 1	: 1	: 1	: 0	: 0	: 1	: 1	: 5
Jumlah jawaban benar betul									: 10 : 11 : 9 : 11 : 9 : 11 : 11

4. Format tabulasi jawaban tes
dari kelompok kurang (L)

Semester :

No.	Nomor	Nomor Soal							Jumlah
Urut.	Siswa	1	2	3	4	5	6	7	Jawaban
1	: 0001	: 0	: 0	: 1	: 1	: 0	: 0	: 1	: 3
2	: 0003	: 0	: 1	: 0	: 1	: 1	: 1	: 0	: 3
3	: 0004	: 1	: 0	: 0	: 0	: 1	: 0	: 0	: 2
4	: 0005	: 1	: 0	: 0	: 0	: 0	: 1	: 0	: 2
5	: 0006	: 0	: 0	: 0	: 0	: 1	: 1	: 0	: 2
6	: 0007	: 0	: 0	: 0	: 0	: 1	: 1	: 0	: 2
7	: 0008	: 0	: 0	: 0	: 0	: 0	: 1	: 0	: 1
8	: 0009	: 0	: 0	: 0	: 0	: 0	: 1	: 1	: 2
9	: 0010	: 0	: 0	: 0	: 1	: 0	: 0	: 0	: 1
10	: 0011	: 1	: 0	: 0	: 0	: 0	: 1	: 1	: 3
11	: 0012	: 0	: 0	: 1	: 1	: 0	: 0	: 0	: 2
12	: 0013	: 0	: 0	: 0	: 1	: 0	: 1	: 0	: 2
13	: 0015	: 0	: 0	: 1	: 0	: 1	: 0	: 0	: 2
Jumlah jawaban benar betul									: 3 : 0 : 3 : 5 : 5 : 8 : 3

Kisi -kisi Ujian Ranah Kognitif

Bidang Studi : Penilaian Hasil Belajar

Waktu :

Materi	Tujuan	Mengetahui			Pemahaman prinsip	Aplikasi prinsip	Jlh
		Istilah	Fakta	Prosedur			
1. Pengertian & fungsi tes 10%		2	2	1			5
2. Prinsip-prinsip Evaluasi 5%					1	1	2
3. Jenis-jenis Alat Evaluasi 10%		2	1	-	1	1	5
4. Perencanaan test 10%		2	1	1	-	1	5
5. Penyusunan tes 30%		2	2	5	3	3	15
6. Pengadministrasian tes 5%		1	1	1	-	-	3
7. Pengolahan nilai 20%		1	3	4	2	-	10
8. Pengembangan tes 10%		1	-	2	2	-	5
Jumlah		10	11	14	9	5	50

Aspek yang diukur Materi	Ingatan			Pemahaman					Aplikasi				Analisis				Sintesis				Evaluasi				Jlh. soal	%
	A	B	CDE	A	B	C	D	E	A	B	CDE	A	B	CDE	A	B	CDE	A	B	CDE						
1. Pengertian dan fungsi Evaluasi	5																							5	10	
2. Prinsip-prinsip					1						1													2	5	
3. Jenis-jenis Evaluasi			2		1						1													3	5	
4. Perencanaan test			4								1													5	10	
5. Penyusunan test			9		3						3													15	30	
6. Pengadministrasian tes			3																					3	5	
7. Pengolahan nilai			8		2																			10	20	
8. Pengembangan tes			3		2																			5	10	
Jumlah soal			35		9						6													50		
% aspek intelektual.			70		18						12														100	

Bentuk Soal:

A = Betul salah B = Pilihan ganda C = Analisis Hubungan

D = Menjodohkan. E = essey

Persentase materi (bobot tiap bahan) disesuaikan dengan silabus GBPP
 Persentase aspek intelektual disesuaikan dengan tujuan pengajaran

SEKOLAH :
 KELAS :
 WAKTU :

ASPEK INTELEKTUAL YANG DIUKUR DAN BENTUK SOAL	INGATAN PEMAHAMAN			APLIKASI			ANALISIS			SINTESIS			EVALUASI			JUMLAH SOAL				Persentase menurut po- kok bahasan			
	C ₁			C ₂			C ₃			C ₄			C ₅			C ₆			SK		SD	MD	Total
	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C					
Pokok Bahasan 1 Sub pokok bahasan 1.1, 1.2, 1.3, 1.4				1	2		1	1		1	2			1		1		1	4	9	3	16	23
Pokok Bahasan 2 Sub pokok bahasan 1.2, 2.2, 3.2, 4		4		3			3	4		2	1				1				5	13	6	24	34
Pokok Bahasan 3 Sub pokok bahasan 1.3, 2.3, 3.3, 4		2		3			2					2							9	17	4	30	43
Jumlah Bentuk Soal Aspek yg diukur	9	10	3	7	8	2	9	7		5	4		3		2	2		1					
	22			17			16			9			3			3							
Persentase Aspek Inte- lektual yang diukur	32			24			23			13			4			4			18	39	13	70	100

Bentuk Soal A = Pilihan Ganda. B = Analisis Hubungan. C = Essey
 Bentuk soal dapat disesuaikan dengan tujuan.
 Pokok bahasan disesuaikan dengan silabi

PENILAIAN RANAH AFEKTIF

oleh: ~~Abdi~~

PANTIA PERSIAPAN PENYUSUNAN
BUKU PANDUAN PENILAIAN HASIL BELAJAR
INSTITUT KEGURUAN DAN ILMU PENDIDIKAN
PADANG - 1990

PENDAHULUAN

Telah sejak beberapa dekade yang lalu para ahli psikologi merasa cukup membagi kebelajaran dan hasil belajar manusia ke dalam tiga kategori: ranah kognitif, ranah afektif, dan ranah psikomotorik. Lalu masing-masing kategori ini diklasifikasikan lagi ke dalam hierarki, yang secara keseluruhannya disebut taksonomi tujuan-tujuan pendidikan.

Pada umumnya, jika orang mendengar kategori afektif, orang membayangkan konsep-konsep nilai dan sikap. Sesungguhnya ranah ini mempunyai cakupan yang lebih luas dari itu. Dimensi lain seperti perilaku moral dan etik, perkembangan diri, emosi dan perasaan, minat dan motivasi. Dinamika kelompok, atribusi juga termasuk ke dalam ranah afektif ini. dengan demikian bila ketiga ranah secara keseluruhan kita misalkan = X, maka ranah afektif = X (kognitif + psikomotor).

Bahwa pembicaraan mengenai nilai dan sikap telah menguasai pembicaraan mengenai ranah afektif, tentu ada sebabnya. Barangkali, banyaknya bidang kajian (misalnya Sosiologi, Antropologi, Politik, Ekonomi) berkepentingan dengan masalah-masalah nilai dan sikap melebihi masalah-masalah lain yang tercakup dalam ranah afektif ini, menyebabkan unsur nilai dan sikap menduduki tempat yang lebih dominan.

Tanpa mengurangi arti unsur-unsur pembicaraan mengenai ranah afektif ini dibatasi pada unsur-unsur nilai dan sikap. Sejumlah mata kuliah dalam kelompok MKDU (misalnya PMP, PSPB, Agama) secara intra-kurikuler dalam proporsi yang amat besar bersangkutan dengan tujuan-tujuan yang bersifat nilai dan sikap ini. Ini adalah alasan kita memusatkan diri pada unsur nilai dan sikap pada kesempatan ini.

SEGI TEORITIS DAN NILAI SIKAP

Dalam Kehidupan sehari-hari orang sering tidak begitu memisahkan pemakaian konsep nilai dan sikap dalam penggunaannya untuk kawasan afektif. Menurut Koentjaraningrat (1976:33), istilah nilai banyak dipakai dalam ilmu-ilmu sosial yang terutama menfokuskan pada kebudayaan dan masyarakat, dan baru secara sekunder mengacu kepada manusia sebagai individu dalam masyarakat. Sebaliknya sikap banyak dipakai dalam dunia psikologi yang terutama menfokuskan pada individu, dan baru secara sekunder mengacu pada kebudayaan dan masyarakat.

Nilai (budaya) diartikan sebagai terdiri dari konsepsi-konsepsi yang hidup dalam alam pikiran sebagian besar warga masyarakat mengenai hal-hal yang harus dianggap amat berharga/bernilai dalam hidup (Koentjaraningrat, 1976:32). Mengkaitkan definisi di atas ke dalam dunia psikologi, maka dari segi si individu dimana suatu nilai hidup dalam dirinya, nilai-nilai tidak lain dari suatu koleksi dari sikap-sikap atau kluster dari sikap-sikap yang semakin inklusif sifatnya. Ditegaskan oleh Newcomb, adalah tujuan-tujuan yang bersifat inklusif di seputar mana pola-pola sikap mungkin diorganisasikan (Martin, 1986:103). Ingin ditekankan di sini adalah bahwa pembicaraan tentang kawasan afektif dapat saja tidak begitu memisahkan antara nilai dan sikap sejauh kita memandang bahwa nilai adalah supra struktur dari sikap-sikap.

Banyak ahli mengemukakan definisi mengenai sikap. Tanpa menyebutkan definisi masing-masing dari ahli-ahli seperti Allport, Krech, Murphy, English & English (Martin, 1986:102) mereka pada akhirnya mempunyai maksud yang hampir bersamaan

mengenai sikap, yaitu kira-kira, kesiapan mental untuk merespon atau beraksi terhadap sesuatu.

Atas dasar kejelasan jalan pikiran mengenai unsur-unsur mental yang terlibat dalam proses afeksi serta mekanismenya, dalam membicarakan sikap kita akan mengambil buah pikiran Fishbein & Ajzen.

Fishbein membedakan empat komponen dalam sistem sikap secara umum, yaitu: keyakinan (belief), sikap (attitude), maksud (intention), dan tingkah laku (behaviour). Keempatnya ini terkait dalam suatu rantai kausalitas, di mana komponen yang mendahului adalah dasar dari komponen berikutnya, dst. Dengan demikian, dasar dari semuanya adalah keyakinan.

Keyakinan mengenai sesuatu objek dapat didefinisikan sebagai probabilitas subjektif seseorang bahwa objek tersebut mempunyai suatu atribut. Istilah objek beserta atributnya dipakai secara generik, dan ia mengacu pada objek fisik, orang, tindakan, peristiwa, ataupun aspek lainnya dari dunia individual. Dalam hal sikap terhadap tingkah laku, yang menjadi pusat perhatian tentunya adalah aspek pelaksanaan dari tingkah laku. Atribut yang dihubungkan dengan 'tingkah laku' ini biasanya adalah konsekuensi atau hasil dari pelaksanaan tingkah laku itu. Ke dalam masing-masing atribut dari objek-objek sikap tersebut terkait respons yang bersifat evaluatif, sehingga disebut respon evaluatif yang bersifat implisit.

Sikap yang dibangun oleh unsur keyakinan mengenai atribut beserta unsur respons evaluatif yang terkait di dalam atribut-atribut dari suatu objek itu, terujud dalam formula berikut:

$$A = f E b e$$

- A = sikap terhadap sesuatu
- b = keyakinan
- e = evaluasi

Dalam suatu situasi tertentu, seseorang itu dianggap mempunyai atau membentuk suatu maksud-berbuat khusus yang mempengaruhi tindakannya selanjutnya. Dalam hal ini "maksud" adalah untuk melakukan tindakan dalam situasi tertentu. Fishbein berikutnya menyatakan adanya dua faktor utama yang menentukan "maksud-berbuat" itu: faktor personal yang berupa "sikap" di atas, dan faktor moral atau normatif. Maka keterkaitan kedua faktor tersebut dalam membentuk "maksud-berbuat", terujud dlam formula ini:

$$B-I = f (w_1A + w_2SN)$$

- B-I = maksud-berbuat
- A = sikap
- SN = norma subjektif
- w = bobot

Sementara itu, norma subjektif adalah berkenaan dengan pengaruh lingkungan terhadap tingkahlaku. Dalam hal ini adalah keyakinan/persepsi si Boy misalnya bahwa kebanyakan orang yang penting bagi si Boy itu berfikir bahwa ia (si Boy) harus atau tidak harus melakukan tindakan tersebut. Ditambah dengan faktor motivasi, maka norma subjektif dibayangkan dalam formula ini:

$$SN = f (E b m)$$

- SN = norma subjektif
- b = keyakinan akan kehendak kelompok referensi
- m = motivasi

Pada akhirnya, "maksud-berbuat" tersebut akan terujud dalam tindakan nyata seseorang. Artinya, jika tidak terjadi perubahan dalam rencana, maka informasi mengenai "maksud-berbuat" seseorang adalah prediktor paling baik bagi tindakan seseorang.

Tanpa peduli amat dengan formula yang dikemukakan, kiranya tidak sukar untuk disimpulkan bagaimana kaitan antara keyakinan- sikap- maksud berbuat- & tindakan/perilaku seseorang. Dalam memandang 'sikap' dalam kaitannya dengan faktor-faktor lainnya tersebut, amat memudahkan kita sampai pada masalah penilaian afeksi yang merupakan maksud pertama dari makalah ini.

PENGUKURAN DAN PENILAIAN SIKAP

1. Permasalahan Taksonomi Tujuan Instruksional Ranah Afektif

Taksonomi tujuan instruksional adalah keseluruhan dari hierarki tujuan-tujuan instruksional dalam menyusun sasaran belajar (TIK). TIK ini adalah unsur yang amat penting dalam pengajaran oleh karena dalam pengembangan sistem instruksional, ini adalah sebagai pedoman untuk menetapkan komponen-komponen lainnya (materi, pengalaman belajar, dan evaluasi).

Taksonomi tujuan instruksional untuk ranah afektif telah coba dikembangkan oleh sejumlah ahli, misalnya Krathwohl dkk, Gephart & Ingle, dan Brandhorst. Taksonomi Krathwohl dkk adalah yang paling terkenal dan luas pemakaiannya, oleh karena yang paling bersifat preskriptif.

Krathwohl dkk setelah melakukan kerja keras menemukan suatu prinsip untuk mengklasifikasikan hierarki mereka, yaitu prinsip internalisasi. Internalisasi adalah proses dengan mana fenomena atau nilai-nilai secara berangsur dan mendalam menjadi bagian dari diri si individu. Hierarki yang ditemukan atas dasar tahap internalisasi nilai tersebut adalah atas dasar tahap internalisasi nilai tersebut adalah: Penerimaan, Merespons, Penilaian, pengorganisasian, dan Karakterisasi. Masing-masingnya ini terbagi lagi ke dalam dua atau tiga sub-sub sehingga terdapat suatu daftar yang luas.

Taksonomi ini ternyata bukan tanpa cacat. Kritik yang pertama-tama dikemukakan adalah bahwa taksonomi ini terlalu umum dan abstrak: taksonomi ini mempersempit dimensi afektif kepada hanya masalah suka dan tidak suka. Selanjutnya dalam studi validasi yang dilakukan para ahli ditemukan misalnya, adalah pola-bercabang terutama pada level Penilaian dan Organisasi, ataupun level Organisasi ternyata tidak dalam urutan seperti yang diharapkan (Martin, 1986:79-83).

MILIK UPT PERPUSTAKAAN
IKIP PADANG

Para ahli mengakui bahwa penggarapan taksonomi afektif belum memberikan hasil se-memuaskan pada taksonomi kognitif. Sampai saat ini belum dijumpai pengembangan alat ukur hasil belajar yang mengacu pada hierarki seperti yang dimaksud dalam taksonomi afektif. Mungkin diantaranya oleh karena taksonomi ranah kognitif. Atas dasar kenyataan tersebut maka pengukuran hasil belajar dan penilaian umumnya untuk ranah ini sifatnya umum: ia tidak sampai membicarakan pengukuran untuk masing-masing tingkat hierarki hasil belajar afektif, seperti yang telah bisa dilakukan untuk ranah kognitif, kecuali kalau rentangan skor yang dijumpai dari pengukurannya dianggap membayangkan tingkat internasionalisasi dari proses pemilikan sikap dan nilai yang dimaksud.

2. Pendekatan Pengukuran

Ada dua pendekatan yang dapat digunakan dalam pengukuran sikap. Pendekatan tersebut adalah:

a. Laporan diri (self report)

Di sini instrumen yang digunakan adalah meminta seseorang melaporkan sendiri sikapnya. Informasi dalam hal ini dapat diperoleh melalui cara-cara lisan (dengan wawancara), ataupun tertulis (kuesioner, skala sikap). Keterpercayaan dari informasi yang diperoleh amat tergantung pada kesediaan yang bersangkutan untuk menjawab secara jujur.

b. Laporan dari orang lain

Disini informasi mengenai sikap seseorang diperoleh melalui laporan orang lain. Pelapor dapat saja berupa seorang yang mempunyai hubungan dekat (misalnya guru, orang tua), atau orang yang tidak mempunyai hubungan dengan orang yang dimaksud. Laporan tersebut dibuat atas dasar pengamatan (observasi) yang dilakukan oleh pelapor terhadap tingkah laku subjek yang diamati.

3. Mengisolasikan Atribut-Atribut yang akan diukur.

Kita berasumsi bahwa perbedaan antara strategi laporan diri dengan observasi dalam pengukuran sikap hanya pada teknik yang digunakan. Mengenai bagaimana kita sampai pada item-item yang terkait dalam self-report dan observasi tidak akan ada perbedaan. Mencari item-item yang hendak diliput, baik untuk strategi self report maupun observasi, adalah apa yang disebut dengan isolasi atribut-atribut untuk diukur.

Secara umum, langkah utama yang ditempuh adalah merumuskan indikator-indikator. Indikator merupakan perincian lebih kecil dari nilai, sehingga nilai tersebut berujud perbuatan-perbuatan yang dapat diukur/diamati melalui pengamatan, pertanyaan, dst.

Atau prosedur yang lebih terarah dikemukakan oleh Popham (1981:333-334) sebagai berikut. Pertama kita coba bayangkan individu bagaimanakah kiranya yang mempunyai atribut-atribut objektif (dalam arti positif) yang dimaksud. Sebagai contoh, bila kita hendak mengukur sikap orang terhadap keikutsertaan dalam atletik, kita bisa membayangkan orang-orang yang secara hipotetis hidup dan diet agar dapat aktif dalam berbagai permainan. Selanjutnya bayangkan pula atribut-atribut orang yang sebaliknya (punya atribut negatif). Lalu langkah berikutnya, pikirkan suatu situasi di mana kedua jenis orang hipotetis kita itu bertingkah laku dalam cara yang amat berbeda. Dengan ini kita dapat menemukan alat ukur yang baik.

Di IKIP, mata-mata kuliah dalam kelompok MEDU seperti Pancasila, PSPB, kita secara kurikuler mau mengembangkan nilai-nilai seperti kerja sama, persatuan, rela berkorban, harga menghargai, cinta bangsa dan tanah air (nilai esensi dalam PSPB) ataupun ujudnya yang lebih terinci dalam bentuk sikap seperti sikap saling menghormati pemeluk antar agama, menghormati kebebasan menjalankan ibadah, saling mencintai, tenggang rasa. dst.

Pada dasarnya nilai dan sikap yang hendak dikembangkan adalah nilai dan sikap-sikap sosial, yaitu yang hendak mengatur hubungan manusia dengan dunia di luar dirinya. Apabila objek sikap apa saja (benda, peristiwa, perilaku) maka kebanyakan aspek yang hendak dikembangkan adalah nilai dan sikap mengenai perilaku (semua yang disebutkan di atas adalah sikap dan nilai mengenai perilaku).

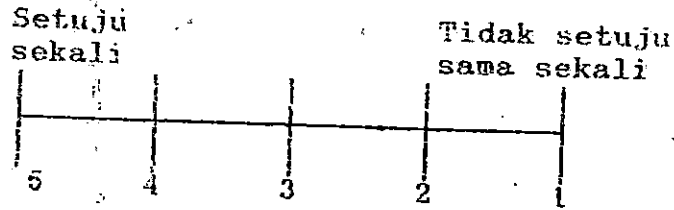
Berpedoman pada cara-cara mengisolasi atribut-atribut objek sikap yang dikemukakan di atas bersifat dari sikap dan nilai yang hendak dikembangkan dan diukur, maka untuk menemukan item-item pengukuran disarankan menggunakan matrik dua faktor. Ambil sebagai contoh "nilai/sikap rela berkorban". Kita melihat adanya dua dimensi utama di dalamnya, (1) apa-apa saja yang dapat dikorbankan, dan (2) situasi apa saja korban itu dilakukan. Atas dasar itu kita menyusun matrik, misalnya sebagai berikut:

Nilai/Sikap: Rela berkorban

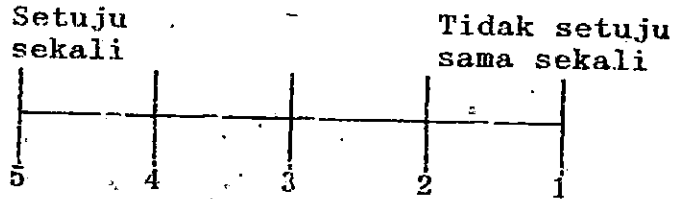
Situasi Jenis	Kemiskinan	Keagamaan	Keterbe- lakangan	dst
1. Material				
2. Tenaga				
3. Ide				
4. Dst				

Lalu dari matrik tersebut dapat dilahirkan item-item pengukuran seperti:

1. Sebagai seorang intelek-tual kita berkewajiban menyumbangkan ide untuk kemajuan rakyat Indonesia



2. Golongan ekonomi lemah membutuhkan bantuan tenaga kita masing-masing



3. Dst

4. Alat Ukur Untuk Strategi Self-Report

Alat ukur untuk strategi self-report yang umum dipakai adalah daftar pertanyaan berbentuk skala (skala sikap). Telah disebutkan di depan bahwa sikap dibangun oleh dua unsur yang menyatu, yaitu keyakinan dan unsur evaluasi, yang keduanya dirangkum dalam formula $A = f (be)$. Maka suatu pertanyaan dalam skala sikap pada dasarnya berupa skala bipolar bertentangan dengan predikat-predikat: baik - buruk, suka - tidak suka, setuju-tidak setuju, dst yang searti.

Dari segi teknis, dunia pengukuran mempunyai setidaknya empat skala sikap standar: Guttman, Sturstone, Likert, dan Semantic Differential. Kelihatannya skala Likert lebih sederhana dan dipakai lebih meluas sehingga disini disarankan untuk memakainya. Dua contoh item yang dikemukakan di atas adalah skala Likert ini.

Skala ini berupa skala lima titik dengan label umumnya: setuju sekali, Setuju, Ragu-ragu, Tidak setuju, Tidak setuju sama sekali. Respon diberi skor dari 5 ke 1. Skor sikap seseorang diperoleh dengan menjumlahkan skor seluruh item.

5. Teknik Observasi

Pada bagian depan telah dikemukakan mekanisme hubungan antara keyakinan - sikap - maksud - tingkah laku. Oleh karena kaitan tersebut, sikap yang sifatnya 'covert' tersebut dapat diramalkan melalui tingkahlaku (overt).

Dalam rangka mengukur dan menafsirkan sikap melalui tingkahlaku, kita mempunyai tiga kriteria tingkahlaku: (1). tindakan-tunggal, yaitu tindakan khusus seseorang pada suatu target, dalam situasi, & waktu tertentu; (2). observasi-berulang, yaitu tindakan khusus yang sama yang diarahkan pada target-target, situasi-situasi, ataupun waktu-waktu berbeda; (3). tindakan jamak, yaitu indeks atas dasar observasi-tunggal atau berulang dari perilaku berbeda. Hal ini dapat digambarkan sebagai berikut ini.

Nilai/Sikap: Rela berkorban

	Observasi				
	1	2	3	4	
1.	B ₁ . 1	B ₁ . 2	B ₁ . 3	B ₁ . 4	R ₁
2.	B ₂ . 1	B ₂ . 2	B ₂ . 3	B ₂ . 4	R ₂
3.	B ₃ . 1	B ₃ . 2	B ₃ . 3	B ₃ . 4	R ₃
4.	B ₄ . 1	B ₄ . 2	B ₄ . 3	B ₄ . 4	R ₄
	M ₁	M ₂	M ₃	M ₄	MR

- B = Observasi tunggal dari perilaku tunggal
 R = Observasi berulang dari perilaku tunggal
 M = Observasi tunggal dari perilaku jamak
 MR = Observasi berulang dari perilaku jamak

Skor untuk observasi berulang maupun tindakan jamak dapat diperoleh melalui penjumlahan maupun rata-rata skor item-itemnya.

kesimpulan dapat ditarik dari tabel ini adalah sebagai berikut. Kriteria tindakan jamak adalah ukuran (measure) untuk sikap terhadap objek, sementara kriteria observasi berulang adalah ukuran (measure) untuk sikap terhadap perbuatan dari perilaku sebagai objek sikap. Artinya, observasi berulang dan observasi jamak adalah indikator dari sikap seseorang terhadap perilaku/objek sebagai objek sikap.

PENUTUP

Pengukuran ranah afektif sampai dewasa ini belumlah setuntas pada ranah kognitif. Dalam makalah ini telah dicoba mengusut permasalahan afektif dari teori sampai ke teknik pengukuran hasil belajarnya (khusus unsur nilai dan sikap). Didalamnya telah terliput suatu ide mengenai teknik pengembangan instrumen serta format penskoran, baik untuk instrumen yang berupa skala sikap maupun observasi. Oleh karena permasalahan ini, yaitu pengukuran masalah afeksi, masih berkembang diharapkan tulisan ini berfungsi menjadi perangsang kearah lahirnya pemikiran-pemikiran, sehingga ide untuk penyusunan buku panduan penilaian berhasil dengan baik.

Rujukan

- Depdikbud, (1984). Kumpulan Hasil Penataran PSPB Kepala-Kepala SGO Seluruh Indonesia. Jakarta: Depdikbud. Pengarang.
- Fishbein, Martin & Ajzen, Icek. (1976). Belief, Attitude, Intention and Behavior: An Introduction to Theory and Reseach. Massachusetts: Addison-wesley publishing Company.
- Koentjaraningrat. (1976). Kebudayaan, Mentalitet, dan Pembangunan. Jakarta: PT Gramedia
- Martin, Barbara L. & Brigs, Leslie J. (1986). The Affective and Cognitive Domains: Integration for Instruction and Research. Englewood Cliffs, New Jersey: Educational Technology Publication.
- Petty, Richard E., Ostrom, Thomas M. & Brock, Tymothy C. ed. (1989). Cognitive Responses in Persuasion. Hillside, New Jersey: Lawrence Erlbaum Associatiates, Publishers.
- Popham, W. James. (1981). Modern Educational Measurement. Englewood Cliffs, New Jersey: Printice-Hall, Inc.

DI BENGKEL KERJA (PSIKOMOTOR)
PENGUKURAN HASIL BELAJAR

IKIP PADANG

DR. AJINTI B. SYARI, M.Sc

PENILAIAN HASIL BELAJAR
DI BENGKEL KERJA (PSIKOMOTOR)

Pendahuluan

Evaluasi hasil belajar di bengkel kerja sama juga halnya dengan evaluasi belajar pada umumnya, yaitu mencakup aspek pengukuran dan pemberian angka terhadap suatu atribut yang sedang di ukur. Dalam paper ini tidak akan dibahas aspek pengukuran dan pemberian angka ("numerical assignment") tersebut tetapi akan menerapkan konsep tersebut pada Evaluasi Hasil Belajar di Bengkel Kerja (EHBBK). Keunikan EHBBK mungkin lebih komprehensif dari evaluasi hasil belajar lainnya (Moore 1980, Grandlund, N. E. 1971, Anastasi A. 1982, Horrocks, J. E. 1964, Bicchi, B. S., Hastings, J. T. dan Madau, G. F. 1971, Abranson, T., Tittle, C. K, dan Cohen 1979). Selain dari itu, banyak para ahli berpendapat bahwa EHBBK harus menerapkan konsep mastery Learning (Smith J, K. 1981, Peng, J. dan Aljufri B. S. 1983, Huynh H. 1976) dan konsep kompetensi. Namun perangkat untuk mengevaluasi kompetensi sampai sekarang masih belum ditemukan perangkat yang dapat diterima oleh para pendidik ketrampilan (Peng, J. dan Aljufri B. S. 1983, Huynh H. 1976, dan Moore 1980). Huynh bertumpu kepada "Binary Classification" teori (teori lulus/tidak) dengan menggunakan dua klasifikasi yang dapat di terapkan untuk jumlah murid lebih dari 15 orang. Sedang Peng J. (1983) dengan menggunakan teori "Multivariate Beta Distribution" menerapkan konsep klasifikasi bertingkat dapat dipakai untuk jumlah murid lebih dari lima orang. Sedang kenyataannya di lapangan dalam kerja bengkel jumlah murid dapat saja kurang dari lima orang. Moore (1980) menggunakan pendekatan faktor analisis dari berbagai aspek

"skill analisis" telah menetapkan berbagai bobot untuk berbagai jenis ketrampilan yang ada (lihat lampiran 1) berbagai literatur yang ditemukan tampaknya ada kesepakatan bahwa penilaian kerja bengkel tidak hanya meliputi satu aspek ranah kognitif saja, tetapi harus meliputi aspek ranah kognitif, afektif, dan psikomotor. Itulah sebabnya, penilaian hasil belajar di bengkel kerja menjadi lebih rumit dan memerlukan pencatatan yang cermat.

Moore (1980) menganjurkan beberapa langkah yang harus dilalui pada waktu menetapkan hasil belajar dalam bengkel kerja: 1. Menetapkan langkah kerja, 2. menetapkan kompetensi yang terkait, 3. menetapkan derajat kesukaran pekerjaan, 4. menetapkan acuan kegiatan, 5. menetapkan kemampuan pelajar untuk menyesuaikan diri dengan lingkungan kerja, dan terakhir 6. menetapkan kebutuhan fisik.

Penetapan langkah kerja diperoleh setelah melakukan "job analisis" yaitu analisis pekerjaan yang diperlukan dengan mengidentifikasi ketrampilan dasar apa yang sebenarnya yang diperlukan, dan kemudian menyusun tata-urutannya. Dalam hal ini tidak dapat dilakukan kalau para guru/dosen tidak melakukan sendiri di labor atau workshop. Urutan inilah nanti yang dikokohkan pada waktu melakukan demonstrasi di depan kelas. Semua uraian inilah nantinya yang digunakan sebagai fondasi dalam menyusun penilaian kerja seseorang:

Dari susunan langkah kerja ini dapat ditarik kesimpulan tentang kompetensi apa yang berhubungan dengan pekerjaan tersebut. Seperti terlihat pada lampiran 1, tampak dari semua pekerjaan batu

yang tertera maka disimpulkan ada tiga macam kompetensi yang perlu dilatih: 1. kemampuan untuk mengidentifikasi alat kerja, 2. kemampuan untuk menggunakan alat kerja secara efektif, dan 3. kemampuan untuk mengenal pentingnya tata-urutan kerja. Tentu saja, tidak sama untuk semua pekerjaan. Proses ini sebenar tidak ada bedanya dengan proses pembuatan "kisi-kisi" seperti yang dianjurkan oleh Gronlund, N. E. (1971), Anastasi, A. (1982), Thorndike, R. L. (1971), Bloom, B. S., Hastings, J. T. dan Madaus, G. F. (1971). Selanjut harus ditetapkan bobot dari profil kemampuan umum yang akan dikembangkan "Qualifications Profile of Critical Aptitudes"

seperti:

G = General Learning Ability = 3

V = Verbal = 3

S = Spatial = 3

P = Form perception = 4

K = Motor Coordination = 3

F = Finger Dexterity = 4

M = Manual Dexterity = 3

Semua ini diperoleh melalui penelitian yang membandingkan berbagai jenis pekerjaan yang sejenis dalam satu unit (dalam hal ini kerja batu). Selanjutnya baru ditetapkan khasanah ilmu penunjang dan ditetapkan pula tingkat/derajat kesukaran untuk melaksanakannya. Dalam hal ini diperlukan kemampuan untuk membaca alat-alat ukur, dan buku petunjuk, kemampuan matematis (untuk melakukan pengukuran yang akurat), alasan mengapa melakukan demikian ("reasoning") subkomponen sangat vital agar setiap mahasiswa yang melakukan pekerjaan mengerti apa dan mengapa ia melakukan suatu pekerjaan yang telah ditetapkan (standar pekerjaan). Dalam teori evaluasi hal yang demikian dikenal dengan

ilmu pendukung (Anastasi, A. 1982, Bloom, B. S., Hasting, J. T. dan Madaus, G. F. 1971, Grondlund, N. E. 1971, dan Thorndike, R. L. 1971).

Selanjutnya perlu di identifikasi Acuan kegiatan pekerjaan "Work Activities Preferences", Untuk ini Moore menetapkan bekerja dengan benda pada level paling rendah (angka 1), sedang keterlibatan dalam proses, cara dan mesin-mesin cukup tinggi (angka 9). Sedang bagaimana menciptakan kepuasan kerja diletakkan pada katagori yang paling tinggi (angka 10). Dapat disimpulkan pemikiran dibelakang kriteria ini bahwa pembinaan kebiasaan kerja adalah unsur pokok yang akan di kembangkan. Hal ini sesuai dengan tuntutan Khazanas (1979). Khazanas mengatakan unsur pokok dalam membina atau melatih suatu ketrampilan adalah membuat segala sesuatu menjadi "kebiasaan" (work habit) dan untuk itu perlu di tanamkan kepuasan akan hasil kerjanya.

Aspek berikut yang tak kalah menariknya adalah menciptakan kemampuan kepada mahasiswa agar mereka mampu menyesuaikan diri dengan situasi kerja yang baru atau lain dari tempat kerja "Work Situations Adaptability". Moore disini meletakkan acuan yang cukup tinggi kepada pengambilan keputusan tentang pencapaian standar dari hasil ukur benda kerjanya (angka 5). Hal lain yang paling tinggi diletakkan adalah bekerja dengan daerah batas toleransi yang diizinkan dengan kepermatan yang di kehendaki (Working within precise limits or standards of accuracy) (angka 10).

Tampaknya kriteria yang digunakan oleh Moore ini tidak berbeda dengan apa yang di anjurkan oleh para ahli Bank Dunia pada waktu

mengembangkan evaluasi ketrampilan di FPTK IKIP PADANG dan Yogyakarta. Pada waktu itu, ditetapkan tiga aspek yang penting untuk di perhatikan: Waktu, Finishing, dan Production. Yang dimaksud dengan waktu di sini adalah waktu yang diperlukan untuk menyelesaikan pekerjaan, makin cepat dapat di selesaikan makin baik dan makin tinggi skor yang diberikan. Sedang Aspek yang ke dua adalah "Finishing", apakah ukuran benda kerja yang diproduksi mahasiswa telah memenuhi standar yang diinginkan atau tidak. Dalam hal ini tentu sama dengan "batas-batas toleransi" yang di perkenalkan oleh Moore. Akhirnya, yang dimaksudkan dengan "Production" adalah jumlah benda yang dapat di hasilkan.

Aspek terakhir yang mendapat perhatian Moore adalah kebutuhan fisik "Physical Demands". Pekerjaan ini dikategorikan pekerjaan ringan (L = Light work) dengan tiga sub-komponen: 1) membungkuk, jongkok, merangkak, dan merunduk (3 = Stooping, kneeling, crawling, and/or crouching), 2) menjangkau, memanipulasi, mengusap, dan meraba (4 = Reaching, handling, fingering, and/or feeling), dan 3) mengamati hasil kerja (6 = Seeing..)

Kalau dibandingkan dua pendekatan ini walaupun mereka mengacu pada hal yang sama tapi penekanan konsep yang melatar belakangi proses penilaian berbeda sama sekali. Konsep yang dipakai oleh FPTK pada waktu mengembangkan sistem pendidikan kejuruan, mengambil alih konsep "kerja" yang ada di industri yaitu dengan memperhatikan benda yang dihasilkan dan mengabaikan faktor manusianya. Mungkin pandangan yang demikian membuahkan "robotisme", "apatisme", yang menimbulkan kurang gairah dan kurangnya "inovasi" karena memang

sejak dari awal mengembangkan konsep kreatifitas.

Melihat dari semua ini, maka penulis mencoba mengembangkan sistem penilaian kerja bengkel, yang meng"akomodasikan" seluruh konsep yang dapat diterangkan ke dalam format 1 berikut ini.

Pada unit tugas isikan berbagai proyek atau tugas kerja yang diberikan misalkan kerja bangku, melipat plat, dan lain sebagainya. Pada kolom kedua isinya berapa lamanya mahasiswa mengerjakan masing-masing tugas tersebut. Kolom berikutnya berikan nilai dari hasil pekerjaan tersebut. Untuk ini Anda dapat menggunakan teknik penilaian seperti yang di anjurkan Moore, atau dengan menggunakan kriteria batas toleransi yang diizinkan. Dalam istilah statistik biasa dinamakan istilah "varians". Apabila mahasiswa dapat mengerjakan tugas sesuai dengan batas toleransi yang diizinkan maka nilainya adalah A, dan bila ada beberapa bagian tidak memenuhi kriteria toleransi yang diizinkan tetapi benda kerja masih dapat dipakai dan dapat berfungsi dengan baik maka nilainya B. Apabila dapat dipakai tapi mungkin umur benda kerja akan turun (karena retak, tidak rata, permukaan kasar dsb) namun masih dapat berfungsi maka nilainya adalah C, demikian seterusnya.

Pada bagian kanan kartu kemajuan digunakan untuk merekam komponen lain seperti aspek afektif, kepribadian, motivasi dan lain sebagainya. Misalnya, yang paling penting dicatat adalah kehadiran mahasiswa sewaktu melaksanakan tugas. Misalnya dalam satu semester dibutuhkan kerja di workshop sebanyak 30 jam, sedang mahasiswa tersebut hadir sebanyak 25 atau mungkin juga mereka menggunakan jam ekstra 40 jam. Dari sini dapat ditera, apakah mahasiswa yang

bersangkutan memiliki kesadaran untuk menyelesaikan tugasnya dengan sempurna sesuai dengan standar industri atau tidak dapat di acuan dalam kehadiran ini.

Komponen berikutnya adalah rangkuman skor yang diberikan dari komponen tugas, ujian, dan kehadiran. Dalam format ini disediakan empat kolom.

Komponen berikut yang tidak kalah pentingnya untuk di rekam adalah sikap mahasiswa terhadap pelaksanaan tugas. Acuan yang dipakai di sini dapat digunakan ranah afektif yang direkomendasikan oleh Krathwohl.

Selanjutnya yang perlu dilihat adalah aspek kerjasama, pertanyaan dapatkan mahasiswa tersebut bekerja sama dengan rekan sekerjanya, mahasiswa, teknisi labor, dan staf pengajar. Hal ini dapat dilihat dari pemakaian alat ataupun pada pemanfaatan bahan praktikum. Kerja sama juga dapat dilihat dari bantuan yang diberikan kepada rekan lainnya.

Aspek terakhir yang perlu juga diperhatikan dapatkan mahasiswa tersebut menyelesaikan tugasnya secara mandiri (ketergantungan)? Banyak para pendidik, tidak memperhatikan faktor ini dalam pemberian nilai mahasiswa. Mahasiswa yang independen tentu akan mendapatkan skor yang lebih tinggi dari mahasiswa yang baru dapat menyelesaikan tugasnya apabila dibantu oleh orang lain.

Pertanyaan berikutnya, bagaimana seharusnya menyimpulkan nilai tersebut? Jadi disini ada tujuh aspek yang diperhitungkan: aktivitas, tugas, ujian, kehadiran, sikap, kerjasama, dan

Kartu Kemajuan

		Nilai Tugas					dit	
		Waktu yg diperlukar Sangat baik Baik Cukup Kurang jelek						
Unit Tugas		A	B	C	D	E		
							Nama	
							BP.	
							Hadir dalam jam	
							Diperlukan	
							hadir	
							Persent	
							Nilai Akhir	
							Aktivitas	
							Tugas	
							Ujian	
							Hadir	
							Sikap	
							Kerjasama	
							Ketertarikan	

Gambar 1. Kartu Evaluasi Kemajuan Belajar

(Aljufri B. Sverif 1989)

ketergantungan. Misalnya saya akan memberikan bobot sebagai berikut:

Nilai akhir = $0,6 \times \text{aktivitas} + 0,9 \times \text{tugas} + 0,6 \times \text{ujian} + 0,2 \times \text{kehadiran} + 0,8 \times \text{sikap} + 0,4 \times \text{kerjasama} - 0,4 \times \text{ketergantungan}$.
Semua patokan ini tentu sangat tergantung kepada materi yang diajarkan dan kepada komposisi tugas yang diberikan.

Lampiran I.

Kriteria Evaluasi Hasil Belajar Kerja Batu (Menurut Moore, A. B. 1980)

Work Performance Benchmarks

Drew a line 162.5 cm long using chalk.

Handled trowel to scoop mortar up off the mortar board and throw it on the board until it made a rounded mound.

Placed a full concrete block at one end of the chalkline to serve as a guide.

Manipulated trowel to lay a mortar bed in a straight line from the middle of the block to end of chalkline.

Manipulated trowel to make a groove in the mortar in order to spread it out.

Placed a full block at each end of the mortar bed.

Wrapped the mason's line around the corner blocks to establish a sight line and hooked the corner block to the concrete block near the top.

Hung line level on the sight line to level the line.

Placed block with mortared ends next to an end block; the width of the line away from the sight line.

Created joints 1 cm wide by tapping block with trowel handle.

Repeated buttering and placed another block in place.

Handled carpenter's level to crate level blocks at top, ends, and sides.

Constructed second course blocks following same procedure as for first course, but placing a half block at each end of the wall.

Handled carpenter's level to level the top, ends, and sides of blocks.

Constructed the third course of block, placing full block at each end.

Handled carpenter's level to level the top, ends, and sides of blocks.

Constructed the third course of block, placing full block at each end.

Handled carpenter's level to level the top, end, and sides of
Manipulated the trowel to clean the joints.

Cleaned mortar board and tools.

Related Competencies

Ability to identify small tools.

Ability to use small tools.

Ability to follow sequential direction to complete a task.

Qualifications Profile

Critical Aptitudes

G = General Learning Ability = 3

V = Verbal = 3

S = Spatial = 3

P = Form perception = 4

K = Motor Coordination = 3

F = Finger Dexterity = 4

M = Manual Dexterity = 3

Worker Function Level = .384

G.E.D. = 3

Mathematical = 3

Reasoning = 3

Language = 2

Work Activities Preferences

1 = Activities dealing with things

9 = Activities involving processes, methods, or machines

10 = Activities involving satisfaction from work on or producing things

Work Situations Adaptability

8 = Making decisions using standards that can be measured or checked

10 = Working within precise limits or standards of accuracy

Physical Demands

1 = Light work

3 = Stooping, kneeling, crawling, and/or crouching

4 = Reaching, handling, fingering, and/or feeling

6 = Seeing

SANDARAN PUSTAKA

- Abramson, T., Tittle, C. K., dan Cohen, L. (1979). Handbook of Vocational Education Evaluation. California: Beverly Hill, Sage.
- Anastasi, A. (1982). Psychological Testing (Edisi keempat). New York: MacMillan.
- Birnbaum, A. (1968). Statistical Theories of Mental Test Scores. California: Menlo Park, Addison Wesley.
- Bloom, B. S., Hasting, J. T. dan Madaus, G. F. (1971). Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill.
- Cenci, L. dan Weaver, G. G. (1968). Teaching Occupational Skills.
- Grondlund, N. E. (1971). Measurement and Evaluation in Teaching (Edisi kedua). New York: Macmillan.
- Hollnocks, J. E. (1964). Assessment of Behavior: The Methodology and Content of Psychological Measurement. Ohio: Columbus, Charles E. Merrill.
- Huynh, H. (1976). On the reliability of decisions in domain-referenced testing. Journal of Educational Measurement, 19, 253-264.
- Huynh, H. (1978a). Psychometrika, 43, 317-325.
- Huynh, H. (1978b). Computation and inference for two reliability indices in mastery testing based on the beta-binomial model. Research Memorandum 78-1, Publication Series in Mastery Testing. University of South Carolina College of Education.
- Kryspin, W. J. dan Feldhusen, J. F. (1974). Developing Classroom Tests: A Guide for Writing and Evaluation Test Items. Minnesota: Minneapolis, Burgess.
- Lord, F. M. (1980). Application of Item Response Theory to Practical Testing Problems. New Jersey: Hillsdale, Lawrence Erlbaum.
- Madaus, G. F., Scriven, M. S. dan Stufflebean, D. L. (1983). Evaluation Models: Viewpoints on Education and Human Services Evaluation. Boston: Kluwer-Nijhoff.
- Micheels, W. J. dan Karnes, M. R. (1950). Measuring Educational Achievement. New York: McGraw-Hill.

- Moore, A. B. (1980). Project discovery: Vocational Evaluation Manual. Iowa: Red Oak. Experience Education.
- Peng, C. J. dan Aljufri B. S. (1983). On reliability estimation of multiple classification in criterion-referenced testing. Proposal Submitted for Spencer Foundation Grant, Indiana: Bloomington, Indiana University.
- Rossi, P. H., Freeman, H. E. dan Wright, S. R. (1979). Evaluation: A systematic approach. California: Beverly Hills, Sage.
- Thorndike, R. L. (1971). Educational Measurement (Edisi kedua). Washington D. C.: American Council on Education.
- Thorndike, R. L. dan Hagen, E. (1969). Measurement and Evaluation in Psychology and Education (Edisi ketiga). New York: John Wiley.
- Udinsky, B. F., Osterlind, S. J. dan Lynch, S. W. (1981). Evaluation Resource Handbook: Gathering, Analyzing, Reporting Data. California: San Diego, Edits.
- Worthen, B. R. dan Sanders, J. R. (1973). Education Evaluation: Theory and Practice. CA: Belmont. Wadsworth.

PERTIMBANGAN PENENTUAN SKOR DAN NILAI AKHIR SEMESTER

Oleh: Kumaidi
FPTK IKIP Padang

Disampaikan pada:

**Seminar Penyusunan Buku Penilaian Hasil Belajar
IKIP Padang tanggal 22 dan 23 Januari 1990**

**PERTIMBANGAN PENENTUAN SKOR DAN NILAI
AKHIR SEMESTER *)**

Oleh: Kumaidi
FPTK IKIP Padang

Pendahuluan

Dalam berbagai kesempatan kita, guru atau dosen, selalu dihadapkan pada situasi untuk menentukan skor atau nilai akhir yang merupakan gabungan atau komposit dari beberapa sumber informasi atau tes. Keadaan ini baru saja kita lalui beberapa waktu yang lalu ketika kita harus melaporkan hasil belajar mahasiswa selama semester Juli-Desember 1989.

Penggabungan skor dari beberapa sumber informasi atau tes ini hampir tidak dapat dihindarkan, karena guru atau dosen hampir tidak mungkin hanya mengandalkan satu kali pengukuran untuk mendapatkan informasi kemajuan belajar mahasiswanya.

Sumber informasi ini bisa bermacam-macam, misalnya keaktifan mahasiswa dalam diskusi kelas, penyelesaian tugas terstruktur, penulisan makalah, penyelesaian proyek akhir semester, ujian tengah semester, dan ujian akhir semester. Ditinjau dari segi jenis informasinya pun bisa beragam. Ada informasi tentang hasil belajar yang berbentuk olah pikir, ketrampilan tangan (manual), dan mungkin juga perubahan sikap

*) Makalah ini disampaikan pada Seminar Penyusunan Buku Panduan Penilaian Hasil Belajar IKIP Padang tanggal 22 dan 23 Januari 1990

mahasiswa sebagai "*dampak*" dari suatu mata kuliah (pendidikan dan pengajaran selama satu semester). Variasi jenis dan bentuk informasi ini akan menambah sulit penentuan *model* penggabungan skornya.

Makalah singkat ini mencoba untuk memberi masukan dalam pencarian model penggabungan skor untuk menentukan nilai akhir seorang mahasiswa mengikuti suatu mata kuliah. Masukan ini terbatas pada dasar pertimbangan yang mungkin dipakai secara operasional untuk keperluan guru atau dosen sehari-hari, terutama di akhir semester apabila mereka dihadapkan pada keadaan penentuan prestasi mahasiswa dari berbagai sumber informasi atau tes. Model penggabungan skor yang akan dibicarakan adalah model regresi yang disajikan secara sederhana agar mudah dipahami dan dilaksanakan.

Penggabungan Skor Dari Beberapa Tes

Prosedur penggabungan beberapa skor biasa dilaksanakan dengan mengasumsikan bahwa skor masing-masing tes merupakan suatu komponen dari skor gabungan. Hal ini sering dilakukan untuk mencari total skor dari suatu tes yang terdiri dari beberapa item atau butir soal. Guru atau dosen biasanya menganggap bahwa butir soal merupakan *komponen* dari suatu tes. Dengan demikian total skor tes seorang murid atau mahasiswa ditentukan oleh jumlah atau fungsi dari skor butir soal tersebut.

Dosen biasanya memakai penjumlahan, karena model ini merupakan model paling sederhana. Penjumlahan merupakan model regresi termudah, karena secara *apriori* dosen menganggap bahwa koefisien regresi dari masing-masing komponen ditetapkan samadengan satu. Apabila kita tidak dapat menerima bahwa setiap butir soal atau skor tes memiliki bobot sumbangan yang sama, maka biasanya dosen menetapkan (masih secara *apriori*) dengan persentase tertentu, misalnya skor butir soal satu samadengan 10, butir soal dua samadengan 30, dan sebagainya. Ini sering kita lakukan dengan dasar pertimbangan profesionalitas kita sebagai "pakar" dari suatu mata kuliah dan penilaian hasil belajarnya. Hal yang sama dapat pula kita berlakukan untuk menetapkan skor dan nilai akhir mahasiswa dalam suatu mata kuliah yang diukur dan dievaluasi dengan berbagai cara dan saat.

Prosedur yang diceritakan di atas dapat pula diaplikasikan secara apa adanya untuk kasus nilai akhir semester. Akan tetapi, dosen sering kurang yakin bahwa cara-cara pengujian atau pemantauan perubahan penguasaan materi dan kompetensi mahasiswa memiliki prediktibilitas yang sama, maka dia menggabungkan sumber informasi tidak dengan penjumlahan langsung tetapi dengan menetapkan besar sumbangan atau bobot masing-masing tes. Penentuan secara *apriori* terhadap bobot suatu tes dalam menetapkan skor atau nilai gabungan (komposit) dapat diterima karena memang dosenlah yang paling mengetahui bagaimana cara menilai perkembangan

prestasi dan kompetensi mahasiswanya dalam mata kuliah yang dibinanya. Oleh sebab itu, seandainya masing-masing dosen memiliki model pembobotan atau koefisien regresi tertentu yang selama ini dipakai, model tersebut dapat dipertahankan sepanjang hasil akhir yang diperoleh dapat menggambarkan prestasi mahasiswanya secara tepat. Permasalahan yang sering dijumpai adalah banyak dosen yang kurang memperhatikan aspek evaluasi hasil belajar ini secara integral, akibatnya, dari tahun ke tahun, banyak mahasiswanya yang gagal. Sumber kegagalan selalu saja secara mudah dikatakan mahasiswa yang malas.

Situasi demikian tentu tidak dapat dipertahankan terus menerus. Apabila memang banyak mahasiswa gagal, dosen harus mencermati instrumennya dan mencoba mencari faktor penyebab kegagalan mahasiswa. Apabila faktor penyebabnya adalah tes yang terlalu sulit bagi mahasiswa, maka model penggabungan skor dapat dipakai sebagai jembatan penghubung dalam menetapkan prestasi mahasiswa. Lebih-lebih apabila dosen memang menyimpan dokumentasi jawaban mahasiswa terhadap tiap tes yang dibuatnya dari tahun ke tahun. Dengan demikian, populasi mahasiswa dan universe (populasi) butir soal untuk suatu mata kuliah dapat dikembangkan. Dengan demikian "contour" penyebaran prestasi dalam acuan butir soal dan populasi mahasiswa dapat dibuat, yang akhirnya model ini dapat dipakai untuk mengontrol sistem pemberian nilai akhir (sebagai indikator prestasi mahasiswa). Model ini sebaiknya dikembangkan oleh setiap dosen di IKIP Padang ini agar

"makna" setiap "nilai" yang diberikan kepada mahasiswa dapat dikontrol relatif stabil dari waktu ke waktu. Sistem pengembangan ini dapat dibantu oleh Pusat Testing dan Evaluasi Pendidikan IKIP Padang, untuk setiap dosen yang memerlukannya.

Yang penulis maksudkan dalam paragraf di atas adalah tingkat kesulitan tes seharusnya terkontrol dari tahun ke tahun dan dari semester ke semester. Apabila setiap kali mata kuliah tersebut ditawarkan, diikuti oleh sejumlah mahasiswa, maka pengambil mata kuliah itu, setiap semesternya, dapat dipandang sebagai sampel mahasiswa (yang mengambil mata kuliah dimaksud). Tes yang dibuat juga hanya merupakan sekumpulan sampel item (butir soal) dari suatu universe (populasi) item dalam mata kuliah itu. Dengan demikian, berdasarkan pemikiran ini, maka terdapat dua sampel, yaitu sampel mahasiswa dan sampel item. Kedua macam sampel ini akan menentukan model skoring dan penilaiannya. Tingkat "kecakapan" mahasiswa yang diukur diusahakan terkontrol dalam lingkup populasi dan tingkat kesulitan tes juga dikontrol dalam level universe itemnya. Oleh sebab itu, pemberian nilai akhir mahasiswa seharusnya ditentukan oleh tingkat kecakapan atau prestasi mahasiswa dan tingkat kesukaran item (atau tes) dalam level populasi dan universe ini. Implementasi praktis dari konsep ini, tentu mengakibatkan atau sering timbul kemelesetan estimasi "nilai" atau "skor" sebagai gambaran kecakapan mahasiswa sebagai akibat kekurang-akuratan tes menggambarkan universe item.

Oleh sebab itu, daerah batas beda indeks prestasi (misalnya antara nilai A dan B, B dan C, dan seterusnya) merupakan daerah kritis yang seharusnya mendapat perhatian dan "concern" dosen, karena daerah kritis ini, seharusnya memiliki standar "error" pengukuran yang terkecil.

Konsep pemikiran di atas menuntut "awareness" dosen akan rumit dan kompleksitasnya permasalahan skor atau nilai akhir mahasiswa dalam suatu mata kuliah. Dalam level praktis, kita sering dihadapkan pada pemecahan pragmatis, yaitu dengan penetapan skor komposit secara sederhana, misalnya:

$$Y = X_1 + X_2 + \dots + X_n$$

dimana: Y = skor komposit;

X = skor komponen (masing-masing tes dari yang pertama sampai yang ke n).

Model penggabungan di atas dapat saja diubah menjadi model rata-rata dan bukan model penjumlahan, yang pada dasarnya sama saja, tergantung dari rentang skor akhir yang diinginkan. Model di atas dapat pula disesuaikan dengan keinginan dosen untuk memberi pembobotan yang berbeda-beda atas komponen yang satu terhadap komponen lainnya, misalnya:

$$Y = a_1X_1 + a_2X_2 + \dots + a_nX_n$$

dimana: a_1 s/d a_n = bobot yang diinginkan untuk masing-masing sumber informasi atau tes (yang merupakan komponen dari komposit Y).

Model yang terakhir ini biasa dipakai orang dan sering disebut model regresi.

Permasalahan yang dihadapi dalam pemakaian kedua model di atas akan tetap sama, yaitu: Apakah model penggabungan atau penyusunan skor terbaik terhadap prestasi sebenarnya dari seorang mahasiswa dalam suatu mata kuliah? Apakah Y akan memberi tingkatan derajat kompetensi mahasiswa dalam suatu mata kuliah? Dua pertanyaan tersebut tidak akan langsung terjawab oleh model Y mana pun! Oleh sebab itu, dosen, berdasarkan pengalamannya bertahun-tahun membina suatu mata kuliah, harus yakin bahwa model penggabungan yang dipakai merupakan prediktor terbaik terhadap prestasi mahasiswa dalam suatu mata kuliah secara keseluruhan. Untuk mendapatkan keyakinan dimaksud, dosen perlu mencermati universe tes dan populasi mahasiswanya. Populasi mahasiswa ini akan semakin baik setiap kali sang dosen mengajar mahasiswa "baru" dalam mata kuliah binaannya. Apapun model penggabungan komponen skor menjadi skor komposit yang dipilih, dosen perlu mengetahui karakteristik komposit yang diinginkannya.

Problema penentuan skor komposit di atas tidak hanya ditemui dalam penggabungan skor komponen dari tes yang mengukur satu ranah (misalnya hanya ranah olah pikir atau kognitif saja), akan tetapi juga akan ditemui apabila dosen ingin menggabungkan informasi dari ranah yang berbeda. Problema penggabungan yang pertama mungkin sifatnya lebih "ringan" dibandingkan dengan problema penggabungan yang kedua, yang terdiri dari berbagai ranah, misalnya X_1

merupakan informasi ranah afektif, X_2 ranah kognitif dan X_3 ranah motorik. Model ini akan menjadi lebih rumit dikarenakan dosen perlu mengkaji pola hubungan antar ranah menurut teori psikologi (pendidikan), yang biasanya kurang menjadi perhatian dosen bidang studi. Lain halnya apabila komponen skor tersebut masih dalam satu ranah (misalnya kognitif saja), sebab permasalahannya hanya terbatas mencari pola hubungan dalam materi ajar, yaitu antar sub-pokok bahasan.

Dalam menentukan skor komposit ini perlu disadari oleh setiap dosen bahwa karakteristik skor komposit tetap akan dipengaruhi oleh karakteristik skor komponennya. Oleh sebab itu, seandainya seorang dosen menyadari bahwa ujian akhir semester merupakan ujian komprehensif dari bahan-bahan kuliah yang diajarkan selama satu semester, atau komprehensif terhadap kemampuan-kemampuan yang ingin didemonstrasikan mahasiswa setelah mendapat bimbingan dari dosen selama satu semester, maka wajar apabila dosen tersebut memberi bobot terbesar kepada skor ujian akhir semester daripada skor ujian lainnya dalam menentukan model skor kompositnya. Sebaliknya, apabila dosen menganggap bahwa ujian akhir semester tidak berbeda dari ujian lainnya dalam semester itu, maka bobot ujian akhir semester juga tidak berbeda dari bobot ujian lainnya dalam menentukan skor komposit. Demikian pula hendaknya bagi dosen yang meletakkan fokus perkuliahannya dengan memberi tugas-tugas (penulisan makalah, proyek observasi, studi kasus, dan sebagainya) sebagai bagian dari

proses belajar yang diberikan kepada mahasiswanya, maka setiap tugas akan mendapat bagian yang layak dalam menentukan skor komposit tadi. Dengan demikian, dosen dapat mengolah dan menentukan model penetapan skor komposit atau skor gabungan berpedoman kepada tujuan pengujian dan makna yang diinginkan terhadap skor atau nilai akhir yang diberikan sebagai "label" keberhasilan atau prestasi mahasiswa dalam mengikuti perkuliahan. Dalam aspek inilah dosen merupakan orang yang paling tahu cara dan tepat menentukan bobot atau model skor komposit atau skor gabungan dari berbagai model pengukuran yang dilakukannya selama satu semester.

Penetapan Indeks Prestasi Akhir

Apabila saya berbicara tentang Indeks Prestasi (IP) Akhir dalam makalah ini mohon dipahami sebagai "nilai" akhir seorang mahasiswa dalam suatu mata kuliah, sebagai "label" prestasinya mengikuti kuliah selama satu semester. Jadi IP di sini tidak sama dengan IP yang selama ini dikenal di IKIP Padang (atau IP yang dijelaskan dalam Buku Pedoman IKIP Padang). IP dalam makalah ini sering pula dinyatakan, di IKIP Padang, dalam "nilai huruf" A, B, C, D, dan E dalam model standar lima yang kita pakai.

Untuk menetapkan IP mahasiswa ini paling tidak ada dua pendekatan, yaitu pendekatan pengacuan kepada patokan tertentu dan pengacuan kepada norma tertentu. Kedua pendekatan tersebut selanjutnya dikenal sebagai pendekatan

penilaian acuan patokan (PAP) dan penilaian acuan norma (PAN). Apapun model yang akan dianut dalam menentukan IP mahasiswa, sebenarnya kedua-duanya tetap beracuan kriteria. Bedanya, PAP mengacu kepada patokan tertentu sebagai kriteriannya, sedangkan PAN mengacu kepada kelompok tertentu sebagai kriterianya. Acuan patokan dalam PAP biasanya diterjemahkan dan bersumber kepada *tujuan pengajaran khusus* (TIK), sedangkan acuan norma biasanya memakai *kelompok pengambil tes* sebagai dasar pembentuk norma dimaksud. Dosen sebaiknya menetapkan model acuan mana yang akan dipakai.

Menurut penulis makalah, PAP sebaiknya dipakai untuk penilaian kecakapan bentuk keterampilan motorik, sedangkan PAN sebaiknya dipakai untuk penilaian kecakapan kognitif dan afektif (misalnya untuk sikap dan minat). Dalam memakai PAP untuk menetapkan IP mahasiswa, dosen perlu melakukan "penelitian" atau "pencermatan" terhadap tingkatan-tingkatan ketrampilan yang diinginkan dimiliki oleh mahasiswanya. Misalnya, untuk ketrampilan mengetik ditetapkan bahwa mahasiswa yang memperoleh "nilai A" adalah mahasiswa yang memiliki kecakapan sebagai seorang pengetik kelas profesional (atau apapun yang oleh dosen mengetik dipandang layak sesuai dengan tingkatan training yang diberikan) dengan rata-rata kesalahan ketikannya sebanyak tidak lebih dari 5 karakter (huruf) per seratus hentakan dan kecepatan mengetiknya tidak kurang dari 45 kata per menit; mahasiswa yang memperoleh "nilai B" adalah mahasiswa yang memiliki kecakapan mengetik sebagai seorang pengetik semi-profesional dengan kesalahan

pengetikan 6-15 karakter (huruf) per seratus hentakan dan kecepatan ketikannya antara 40-45 kata per menit. Dengan definisi yang jelas seperti itu maka interpretasi dan pemberian "nilai akhir" mahasiswa akan baku dan jelas.

Demikian pula hendaknya untuk pemakaian PAN dalam menetapkan IP mahasiswa dalam keterampilan kognitif dan afektif. Norma apa yang akan dipakailah yang perlu dibakukan, disamping pembakuan tes yang akan dipakai. Pembakuan tes dapat dilaksanakan apabila dosen memiliki sistem dokumentasi tes dan jawaban tiap butir soal dari setiap mahasiswa yang pernah mengambil suatu mata kuliah. Pembakuan ini sebenarnya dapat diserahkan kepada tenaga ahli atau sistem yang ada di IKIP Padang. Dalam hal ini Pusat Testing dan Evaluasi Pendidikan IKIP Padang mungkin dapat membantu.

Problema pembakuan norma untuk penginterpretasian dan penetapan "nilai akhir" mahasiswa tersebut secara ringkas dapat dijelaskan sebagai berikut. Populasi pengambil tes dapat dihimpun melalui pendokumentasian mahasiswa pengambil tes dari semester ke semester dan dari tahun ke tahun (siapa pun dosen yang mengajarnya, asal kurikulum dan silabusnya sama). Universe butir soalnya juga perlu dihimpun dari tahun ke tahun, agar tes yang dipakai setiap saat dapat di"equate"kan (atau diskalakan pada metrik yang sama).

Setiap tahun atau semester kelompok mahasiswa yang mengambil suatu mata kuliah dipandang sebagai sampel mahasiswa, demikian pula hendaknya tes yang dipakai merupakan

sampel dari universe butir soal. Dengan demikian, populasi mahasiswa pengambil mata kuliah akan semakin besar dan norma akan semakin matap dan stabil. Penetapan "nilai akhir" mahasiswa ditetapkan atas dasar standing mahasiswa dalam populasi dan bukan dalam sampel. Oleh sebab itu, semakin lama semakin mantap dan stabil acuan penetapan "nilai" tersebut. Pada suatu masa kuliah mungkin saja terjadi bahwa mahasiswa dalam kelompok tersebut memiliki jumlah mahasiswa yang ber"nilai" A lebih sedikit dari tahun sebelumnya karena kelompok mahasiswa tersebut memang lebih "bodoh" dari tahun sebelumnya, atau sebaliknya pada masa itu "nilai" A lebih banyak diberikan dari tahun sebelumnya karena memang kelompok mahasiswa pada tahun berjalan lebih "pandai" dari kelompok sebelumnya. Ini tidak jadi soal karena memang standing mereka relatif terhadap populasi mahasiswa.

Alternatif lain adalah penggabungan model PAP dan PAN. Dalam model ini, PAP dipakai untuk menetapkan batas mana yang "wajar" dinyatakan lulus dan mana yang tidak. Kemudian, PAN dipakai untuk menetapkan mahasiswa mana yang wajar menerima "nilai" A, B, dan C (dengan anggapan bahwa mahasiswa yang lulus terendah diberi nilai C). Dengan demikian yang lulus pasti bagi mereka yang memiliki kecakapan minimal yang telah ditetapkan (apapun batas minimal tersebut), sedangkan yang diberi "nilai" tertinggi tentu memiliki prestasi yang terbaik (pada saat pengukuran diadakan). Sekali lagi, untuk maksud ini Pusat Testing dan Evaluasi Pendidikan IKIP Padang dapat ditugaskan untuk membantu dosen mengembangkan populasi

mahasiswa dan universe butir soalnya. Dengan demikian dosen yang kurang menguasai masalah-masalah teknis pengukuran tidak perlu merasa takut dan sulit.

Penutup

Untuk mengakhiri diskusi kita dalam makalah singkat ini, penulis mencoba merangkum isi makalah sebagai berikut:

- 1) Setiap dosen hampir pasti setiap akhir semester selalu dihadapkan pada penentuan skor komposit dan "nilai akhir" mahasiswanya. Oleh sebab itu perlu disusun strategi penetapan skor komposit dan nilai akhir.
- 2) Model penggabungan skor komponen menjadi skor komposit yang paling sederhana adalah model regresi (baik dengan atau tanpa bobot). Model ini mungkin sudah dipraktikkan oleh setiap dosen di IKIP Padang.
- 3) Bagi yang memakai model pembobotan, bobot sebaiknya ditetapkan oleh dosen atau tim dosen dengan memperhatikan "tujuan" dan "makna" skor atau nilai komposit yang diinginkan.
- 4) Dosen perlu menyadari bahwa skor komposit selalu dipengaruhi oleh karakteristik skor komponennya. Oleh sebab itu, dengan menyadari adanya pengaruh skor komponen dalam skor komposit ini dosen perlu merumuskan tujuan pengukuran dan makna nilai akhir secara cermat.

- 5) Penetapan "nilai" akhir sebagai "label" prestasi seorang mahasiswa perlu memperhatikan makna nilai yang diinginkan dan acuan yang dipakai.
- 6) Ada dua alternatif pokok dan satu alternatif gabungan yang dapat dipakai untuk menetapkan "nilai akhir" tersebut, yaitu PAP dan PAN. PAP perlu memikirkan kompetensi lulusan sesuai dengan tingkatan "nilai" yang ingin diberikan, sedangkan PAN perlu disusun norma yang mengacu kepada populasi mahasiswa yang mengambil mata kuliah tersebut dengan mengontrol tes yang dipakai melalui pembentukan universe butir soal.
- 7) Pusat Testing dan Evaluasi Pendidikan IKIP Padang, sebagai lembaga yang dibentuk oleh IKIP Padang untuk maksud pengembangan sistem testing dan evaluasi pendidikan, siap membantu masalah-masalah teknis yang diakibatkan oleh dilaksanakannya butir 6) di atas.

Akhirnya harapan penulis, semoga makalah ini dapat memberi sumbangan pemikiran sebagaimana yang diharapkan. Di samping itu tanggapan dan diskusi kita kali ini akan sangat berguna dalam pengembangan kajian-kajian bidang testing dan evaluasi pendidikan ini.

Kepustakaan

- American Educational Research Association, American Psychological Association, & the National Council on Measurement in Education. (1985). Standards for educational and psychological testing. Washington, Dc: American Psychological Association.
- Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. New York: Holt, Rinehart, and Winston.
- Ebel, R. L., & Frisbie, D. A. (1986). Essentials of educational measurement. (edisi keempat). Englewood Cliffs, NJ: Prentice-Hall.
- Chiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). Measurement theory for the behavioral sciences. San Francisco, CA: Freeman and Company.
- Gulliksen, H. (1950). Theory of mental tests. New York: Wiley & Sons.

PEMILAHAN BELAJAR DI LABORATORIUM

1990

1990

Disampaikan pada:

Seminar Pengajaran 22 dan 23 Januari 1990

di Jurusan Pendidikan dan Pengembangan Sumber

Manusia dan Pembelajaran Universitas Pendidikan Indonesia

1990

PENILAIAN BELAJAR DI LABORATORIUM *)

RUSYDI RUSYID

Pendahuluan

Sebagai pengajar di IKIP Padang kita selalu dihadapkan kepada masalah bagaimana menetapkan nilai mahasiswa sebagai hasil dari usaha belajar mereka. "Buku Pedoman" IKIP Padang telah memuat sistem penilaian yang pada pokoknya didasarkan kepada: (1) kehadiran dalam kuliah, (2) keaktifan mahasiswa dalam kelas, (3) tugas terstruktur, dan (4) ujian. Dasar penilaian ini tentu saja tidak tepat diaplikasikan terhadap kegiatan belajar di laboratorium. Pola belajar di laboratorium ditekankan kepada pendekatan proses disamping pendekatan konsep dan pendekatan lingkungan, sedangkan pola belajar di kelas mengacu kepada pendekatan konsep. Yang tak kalah menariknya, bermacam-macam format penilaian kegiatan belajar di laboratorium yang terdapat dalam literatur atau yang diperoleh dosen lewat penataran tetap kurang familiar bagi sebagian besar staf pengajar.

Oleh sebab itu paper ini mencoba membahas masalah tentang komponen-komponen apa saja yang harus dinilai dalam kegiatan belajar di laboratorium dan bagaimana cara menilai dan pengembangan tes/instrumennya "Penilaian Belajar di Laboratorium" Bahasan ini diharapkan dapat dijadikan suatu model untuk menetapkan nilai mahasiswa dalam kegiatan belajar di laboratorium.

Penilaian Belajar di Laboratorium dimaksudkan sebagai suatu proses mengambil keputusan yang mencakup kegiatan

*) Paper ini disiapkan untuk Seminar dalam rangka penyusunan Buku Panduan Penilaian Hasil Belajar Mengajar IKIP Padang, 22 dan 23 Januari 1990.

pengukuran (measurement) dan penilaian (evaluation) terhadap kegiatan belajar di laboratorium. Mengukur adalah membandingkan sesuatu dengan suatu ukuran (bersifat kuantitatif), sedangkan menilai adalah mengambil keputusan "baik-buruk" atau "lulus tidak lulus" terhadap apa yang diukur (bersifat kualitatif).

Tinjauan Ringkas Tentang Penilaian

Penilaian sebagai kegiatan mengukur dan menilai memiliki tiga makna, yaitu: (1) makna bagi mahasiswa, (2) makna bagi dosen, dan (3) makna bagi institut. Penilaian bermakna bagi mahasiswa karena dengan penilaian mahasiswa dapat mengetahui sejauh mana dia telah berhasil mengikuti/menerima pelajaran yang diberikan dosennya. Penilaian bermakna bagi dosen karena dengan penilaian dosen dapat mengetahui: mahasiswa yang telah menguasai pelajaran atau belum, materi pelajaran yang diberikan tepat atau tidak, dan metoda mengajar yang digunakan sesuai atau kurang. Penilaian bermakna bagi institut karena dengan penilaian institut mengetahui apakah kondisi belajar-mengajar yang diciptakan sesuai dengan harapan atau belum, apakah kurikulum tepat atau kurang tepat, dan apakah institut sudah memenuhi standar atau belum.

Secara garis besar alat penilaian dapat dibedakan ke dalam dua golongan, yaitu (1) tes, dan (2) bukan tes (non-tes). Kedua alat ini selanjutnya dapat disebut teknik penilaian-teknik tes dan teknik non-tes.

Yang tergolong teknik tes adalah sebagai berikut:

- (1) Tes diagnostik
- (2) Tes format
- (3) Tes sumatif

Sedangkan yang tergolong teknik non-tes adalah sebagai berikut:

- (1) Skala bertingkat
- (2) Kuesioner
- (3) Daftar cocok
- (4) Wawancara
- (5) Pengamatan
- (6) Riwayat hidup

Tes merupakan serangkaian pertanyaan, latihan, atau alat lain yang dapat digunakan untuk mengetahui kemampuan, ketrampilan, intelegensi, bakat yang dimiliki individu ataupun kelompok. Dengan demikian tes bersifat komprehensif.

Penilaian yang bagus menurut tes yang baik.

Tes yang baik haruslah memenuhi persyaratan:

- (1) Valiliditas
- (2) Reabilitas
- (3) Objektivitas
- (4) Praktibilitas

Pola Pendidikan IPA

Aktifitas mendasar dari IPA tentu saja melakukan observasi yang hati-hati baik bersifat kualitatif yang sering melibatkan eksperimen. Ahli IPA mencoba menemukan hubungan yang umum yang mempersatukan observasi tersebut. Suatu pernyataan verbal yang singkat atau persamaan matematika yang menyimpulkan bermacam-macam observasi dan pengalaman dikenal dengan hukum ilmiah atau hukum suatu penjelasan tentatif terhadap hukum disebut hipotesa. Hipotesa yang berguna adalah hipotesa yang mengandung ramalan yang kebenarannya dapat diuji (ditolak atau diterima) melalui eksperimen. Suatu hipotesa yang telah teruji berkali-kali kebenarannya disebut teori. Suatu teori dapat membantu mempersatukan area yang

luas dan dapat memberikan dasar penjelasan hukum-hukum.

Sebagai contoh ialah teori atom Dalton yang dapat menjelaskan hukum Froust, hukum Lavoiser, dan hukum Dalton.

IPA mengantarkan anak didik memiliki kecerdasan, mempunyai sifat ingin tahu, berfikir secara logis, mengembangkan sikap ilmiah, memberi bekal yang dapat diaplikasikan dalam kehidupan di tengah-tengah masyarakat, serta melahirkan rasa percaya dan kagum kepada sang Pencipta. Tidakkah salah mengatakan bahwa alih teknologi didasari atas IPA (Biologi, Kimia, dan Fisika).

Peranan IPA seperti disebut di atas melahirkan pola-pola pendidikan IPA, yaitu pola:

- (1) Pendekatan Proses
- (2) Pendekatan Konsep
- (3) Pendekatan Lingkungan/Aplikasi

Dengan pendekatan proses anak didik dapat memahami proses-proses alamiah, dan sekaligus dapat melakukan kegiatan-kegiatan yang melahirkan penemuan.

Dengan pendekatan konsep anak didik akan kaya dengan pengetahuan, sedangkan dengan pendekatan lingkungan anak didik akan dapat menggunakan pengetahuan dan ketrampilannya demi kehidupan di tengah-tengah masyarakat.

Kegiatan Belajar di Laboratorium

Ada tiga tingkat kegiatan yang menonjol di laboratorium: (1) membuat persiapan, (2) melaksanakan eksperimen, atau (3) menulis laporan. Ketiga tingkat kegiatan ini saling terkait satu sama lain dan oleh karena itu ketiganya harus dilaksanakan.

Persiapan sangat diperlukan karena tanpa persiapan pelaksanaan eksperimen tidak akan berjalan mulus bahkan mungkin tak bisa dilaksanakan sama sekali atau kalau dilaksanakan juga akan dapat mengundang bahaya. Pengalaman

menunjukkan bahwa mahasiswa yang memasuki laboratorium untuk melaksanakan eksperimen tanpa persiapan akhirnya setelah jam habis dia pulang tanpa membawa hasil apa-apa. Sang mahasiswa tidak tahu apa yang harus dikerjakannya, tidak tahu apa bahan dan alat yang diperlukannya. Disamping itu dengan persiapan dapat diketahui apakah ada bahan yang harus disiapkan atau alat yang harus diset sebelum jam pelaksanaan eksperimen. Misalnya pembuatan reagen.

Persiapan yang baik tidak hanya sekedar menjadikan lancarnya pelaksanaan eksperimen, tetapi jauh dari itu harus dapat memperhitungkan bahaya yang mungkin timbul akibat pemakaian zat kimia, penggunaan listrik ataupun penggunaan gas dan sistim air serta bagaimana alternatif penanggulangan bahaya tersebut. Oleh karena itu pengenalan sifat dan karakteristik bahan dan alat mutlak diperlukan.

Pemahaman terhadap bahaya dan karakteristik zat terutama zat kimia membawa konsekuensi bagaimana dan dimana eksperimen harus dilaksanakan. Apakah dilaksanakan dalam almari asam atau ruang terbuka dan bagaimana cara menyimpan bahan-bahan kimia tersebut. Tanpa pemahaman ini bisa terjadi penggunaan eter diruang terbuka dan menyimpan hidrogen peroksida dalam botol tak berwarna. Tampaknya lucu, tetapi mengerikan.

Sebelum melaksanakan aksperimen beberapa hal yang harus dipertimbangkan ialah sebagai berikut:

- (1) Kebutuhan bahan dan alat
- (2) Efek kesehatan dari bahan-bahan kimia
(carcinogenicity, mutagenicity, tetratogenicity, toxicity)
- (3) Karakteristik dari bahan yang digunakan
karakteristik zat kimia antara lain:
flammability, eksplosivity, corrosiveness,
pyrophoricity, melting points, boiling point, flash points, solubility

- (4) Kemungkinan bahaya dari zat kimia yang tertumpah
- (5) Peralatan (sistem air, sistem gas, sistem listrik, sistem pemanas)
- (6) monitoring selama proses
- (7) Penyimpanan bahan dan alat
- (8) Pencucian

Pada pelaksanaan eksperimen dapat dilihat skil apa yang dituntut dan bakal diperoleh mahasiswa. Disini pendekatan proses berperanan besar. Pada dasarnya ada enam macam skil yang diharapkan diperoleh mahasiswa setelah melaksanakan kegiatan laboratorium (melaksanakan eksperimen), yaitu:

- (1) Observational skills, seperti kemampuan mengidentifikasi, mengklasifikasi, mendeteksi perubahan dan perbedaan
- (2) Recording skills, seperti kemampuan menyampaikan dengan menggunakan bahasa lisan/tulisan, diagram, label, dan grafik
- (3) Measurement skills, seperti kemampuan membaca skala, mengukur unit, mengestimasi harga/nilai
- (4) Manipulative skills, seperti kemampuan menguasai peralatan, mengatur gerak, menyesuaikan dan mengontrol
- (5) Procedural skills, seperti kemampuan memilih dan menggunakan bagian tertentu dari alat, mengimplementasikan suatu prosedur atau teknik
- (6) Following instructions, seperti kemampuan mendemonstrasikan instruktur lisan/tulisan atau gambar

Penulisan laporan sangat berharga dalam kegiatan belajar di laboratorium, karena melaporkan hasil penemuan menuntut kemampuan yang sangat kompleks, sehingga dengan laporan dapat melihat kecerdasan dan kemampuan seseorang. Tidak berarti apa-apa bila penemuan yang tidak komunikatif.

Pada dasarnya laporan suatu eksperimen memperlihatkan kemampuan seseorang dalam: (1) merumuskan masalah, (2) berhipotesa, (3) mengungkapkan prosedur eksperimen, (4) menyusun dan menganalisa data, (5) menyimpulkan, dan (6) melihat masalah baru dari hasil penemuan (future work). Oleh sebab itu laporan eksperimen umumnya dilaporkan dalam bentuk format tertentu, misalnya:

- Perumusan Masalah
- Hipotesa
- Rancangan dan Prosedur Eksperimen
- Data dan Analisis Data
- Diskusi
- Kesimpulan

Akhirnya kegiatan laboratorium hendaklah ditutup dengan ujian (ujian semester) untuk mengetahui sejauh mana mahasiswa telah menguasai skil dan konsep sesuai dengan tujuan eksperimen.

Cara Menilai Kegiatan Laboratorium

Walaupun ada tiga tingkat kegiatan yang menonjol dalam belajar di laboratorium, tetapi untuk menentukan/menetapkan nilai mahasiswa data cukup diperoleh dari: (1) persiapan, (2) laporan, dan (3) ujian semester. Mengapa kegiatan pelaksanaan eksperimen tidak menjadi pertimbangan dalam menentukan nilai mahasiswa? Jawabnya, karena kegiatan ini baru dapat diukur setelah kegiatan satu semester berakhir. Oleh sebab itu untuk mengetahui penguasaan kemampuan mahasiswa dari apa yang

diharapkan pada pelaksanaan eksperimen dilakukan ujian tertulis. Sedangkan pada kegiatan pelaksanaan eksperimen tes yang digunakan adalah tes diagnostik.

Ketiga komponen ini diboboti 20%, 40%, dan 40% masing-masing untuk persiapan, laporan, dan ujian semester. Sedangkan masing-masing komponen diberi skor dengan rentangan 10 - 100. Dengan demikian untuk memperoleh skor seorang dapat digunakan rumus:

$$N = \frac{20.P + 40.Q + 40.R}{100}$$

N = skor mentah
P = skor persiapan
Q = skor laporan
R = skor ujian semester

Selanjutnya skor mentah ini dinyatakan sebagai "derajat penguasaan" dan untuk menetapkan nilai dipakai kriteria yang telah ditetapkan dalam Buku Pedoman IKIP Padang sebagai berikut:

Derajat penguasaan	Nilai
90% - 100%	A
80% - 80%	B
65% - 79%	C
55% - 64%	D
< 55%	E

Pengembangan Tes/Instrumen

(1) Persiapan

Persiapan dapat berupa ujian lisan atau berupa rencana tertulis. Baiknya ujian lisan ataupun ujian tertulis maksudnya sama, yaitu melihat kesiapan mahasiswa terhadap objek eksperimen. Apakah mahasiswa sudah tahu objek eksperimennya? Apakah mahasiswa sudah tahu bahan dan alat yang akan digunakan? Apakah mahasiswa sudah tahu sifat dan karakteristik bahan dan alat yang digunakan? Apakah

mahasiswa sudah memiliki alternatif penanggulangan masalah seandainya timbul bahaya yang disebabkan oleh bahan dan alat yang digunakan?

Ujian lisan sebelum memulai eksperimen sudah dilaksanakan di jurusan Pendidikan Kimia sejak mulai berdirinya tahun 1965, tetapi rencana tertulis boleh dikatakan belum. Persiapan dalam bentuk rencana tertulis ini dapat dilaksanakan dengan mempedomani format yang ada pada lampiran paper ini.

(2) Laporan

Sehubungan dengan laporan format seperti telah dikemukakan dapat digunakan. Bentuknya ialah:

- Perumusan masalah
- Hipotesa
- Rancangan dan Prosedur Eksperimen
- Data dan Analisa Data
- Diskusi
- Kesimpulan

(3) Ujian Semester

Ujian semester sebaiknya berupa ujian praktek yang dilengkapi dengan ujian tertulis. Ujian praktek diarahkan untuk melihat keenam skil dasar yang telah dikemukakan di atas. Sedangkan ujian tertulis diarahkan untuk melihat penguasaan konsep.

Kesimpulan

1. Skor untuk menetapkan nilai mahasiswa dapat diperoleh dari: persiapan, laporan, dan ujian smester dengan bobot masing-masing 20%, 40%, dan 40% yang selanjutnya disebut derajat penguasaan.
2. Penetapan nilai sesuai dengan kriteria penilaian yang telah ditetapkan dalam Buku Pedoman IKIP Padang.
3. Pengembangan tes instrumen hendaklah ditujukan kepada lancarnya jalan eksperimen mampunya mahasiswa melaporkan penemuan ilmiah dan dikuasainya konsep-konsep yang penting.

Lampiran

EXPERIMENTAL PLAN SHEET

I. Chemist Responsible for Experiment _____

Date _____

II. Objective of Experiment or Synthesis (include a short description of the overall scientific problem.

III. Description of Experiment (include procedure to be followed, diagram of apparatus to be, references for procedure). Add additional pages as needed.

IV. Chemicals to be used:

TABLE 1

Chemical	B.P	M.P	F.P	TLV-TWA	TLV-STEL	A.L
----------	-----	-----	-----	---------	----------	-----

(B.P. = boiling point. M.P. = melting point. F.P. = flash point. A.L. = level which would result if all of material you intend to use were vaporized in the room at one time as discussed in section 5. Show calculations in your notebook for A.L.)

Table 2

Chemical	Ref	Cor	Pyu	Car	Mut	Ter	Flam.	Exp.	Stab
----------	-----	-----	-----	-----	-----	-----	-------	------	------

(Cor = Corrosive, Pyr = Pyrophoric; Car = Carcinogenic; Mut = Mutagenic, Ter = Teratogenic; Flam = Flammable; Exp = Explosive; Stab = Stable Chemical. Answer yes, no or ? in the above table. The number (s) of the bibliographic source you used for the information in Tables 1 and 2 should be noted in the Ref column.)

V. Are any chemicals which you intend to used incompatible?
Give reference for answer.

VI. Can the hood you intend to used keep the level of the
most toxic chemical or carcinogen in the table above,
below its TLV-TWA. See section 2.c. xxiii and Hood Test
Notice.

Record Hood Test Notice Data Below.

VII. Give spill clean up procedures.

VIII. Give emergency procedures for equipment, water or electrical failure. Give fire fighting methods.

Kepustakaan

- Bryce, T.G.K.; et al., 1983, Techniques for Assesment of Practical Skills in Foundation Science, Heineman Educational Books, Glasgow.
- Depdikbud Ditjen Dikti, 1982/1983, Pendekatan Khusus Pengajaran IPA Kecenderungan Bidang Studi IPA.
- Ebel, R.E.; Frisbie, D.A. , 1986, Essentials of Educational Measurement, Prentice Hall, Inc., Englewood Cliffs, New Jersey.
- IKIP Padang, Buku Pedoman 1989-1990, Percetakan IKIP Padang, Padang.
- Miller, E.J., 1986, Advanced Laboratory Manual.
- Rangel, L. Tobing, 1979, Cara Menilai Kegiatan Laboratorium, P3G Depdikbud, Jakarta
- Suhaimi Arikunto, 1984, Dasar-dasar Evaluasi Pendidikan, PT: Bina Aksara.

THE UNIVERSITY OF CHICAGO
DEPARTMENT OF CHEMISTRY
57 SOUTH EAST ASIAN AVENUE
CHICAGO, ILLINOIS 60607
TEL: 773-936-3700

Office : Dr. Zamil MA

PENNINGTON BRYANT BARRAS

PENILAIAN BELAJAR BAHASA

Penilaian belajar bahasa terhadap mahasiswa dari program studi bahasa ataupun nonbahasa di IKIP Padang perlu ditingkatkan. Ini tidak berarti bahwa penilaian hasil belajar bahasa pada masa-masa yang lampau di IKIP Padang tidak baik. Peningkatan ini dilakukan dalam rangka meningkatkan mutu lulusan IKIP Padang yang akan datang.

Sampai saat ini sistem penilaian hasil belajar bahasa oleh dosen-dosen yang berlainan pada jurusan-jurusan yang berbeda belum mempunyai suatu standar yang dapat dipakai sebagai pedoman dalam penilaian.

Dengan mengemukakan istilah "standard" di sini, penulis tidak bermaksud menyeragamkan cara penilaian belajar bahasa di IKIP Padang, tapi memenuhi permintaan panitia seminar bahasa yang akan diuraikan dalam makalah ini adalah:

- 1) perencanaan pengembangan tes/instrumen bahasa sampai dengan penskoran (scoring);
- 2) format-format yang diperlukan untuk pengadministrasian skor;
- 3) usul pertimbangan untuk menetapkan skor akhir; dan
- 4) lain-lain yang dirasa perlu.

Dalam pengembangan instrumen untuk penilaian belajar bahasa di IKIP Padang, kenyataan dalam belajar bahasa yang ada di institut ini sampai sekarang perlu dipertimbangkan. IKIP Padang mempunyai enam fakultas dengan 27 jurusan dan satu balai bahasa. Semua jurusan nonbahasa mempunyai mata kuliah bahasa Inggris (MKBS) dan bahasa Indonesia (MKDU). Hanya dua

jurusan saja yang merupakan jurusan pendidikan bahasa yaitu jurusan pendidikan Bahasa Indonesia dan jurusan Pendidikan Bahasa Inggris dengan program-program studinya masing-masing. Sedangkan balai bahasa hanya mempunyai program untuk penataran bahasa Inggris.

Jurusan-jurusan non-Bahasa Inggris mempunyai mata kuliah Bahasa Inggris. Mata kuliah ini diberikan agar mahasiswa mempunyai kemampuan membaca karangan ilmiah dalam menambah pengetahuan di bidangnya, serta untuk memperoleh informasi tentang ilmu dan teknologi. Kenyataan menunjukkan bahwa informasi tentang ilmu pengetahuan dan teknologi banyak yang ditulis dalam bahasa Inggris. Orang-orang yang mampu membaca dalam bahasa Inggris dan mempunyai kemauan untuk itu akan lebih beruntung dibandingkan dengan orang yang tidak mempunyai kemampuan dan kemauan untuk itu.

Pengajaran bahasa Inggris di jurusan-jurusan non-Bahasa Inggris hendaknya mempunyai tujuan seperti di atas sehingga bahan dan metoda mengajar bahasa Inggris dapat disesuaikan dengan maksud seperti itu. Dengan demikian penilaian belajar bahasa Inggris dapat didasarkan kepada bahan yang diberikan dan tujuan yang akan dicapai.

Dari uraian di atas dapat disimpulkan bahwa salah satu instrumen yang dapat dikembangkan untuk penilaian hasil belajar bahasa Inggris di jurusan-jurusan non-Bahasa Inggris adalah tes kemampuan membaca.

Penilaian belajar bahasa pada Jurusan Bahasa Inggris berbeda dari yang dijelaskan sebelumnya. Penilaian hasil belajar untuk tiap-tiap mata kuliah kebahasaan tergantung

kepada relevansi mata kuliah yang bersangkutan dengan "skill" bahasa, yaitu Listening, Speaking, Reading, dan Writing. Penilaian hasil belajar mahasiswa pada mata kuliah yang menyangkut Listening diutamakan pada kemampuan mendengarnya; yang berhubungan dengan Speaking ditekankan pada kemampuan pengungkapannya; yang bertalian dengan Reading difokuskan pada kemampuan untuk mengerti, menghayati, dan mengaplikasikan isi bacaan; sedangkan yang berkaitan dengan Writing lebih diperhatikan mengenai kemampuannya dalam mengungkapkan sesuatu melalui tulisan.

Penilaian belajar bahasa akan dapat dilakukan antara lain dengan: 1) Achievement Test; 2) Proficiency Test; 3) Diagnostic Language Test; dan 4) Aptitude Test (Finocchiaro and Sako, 1983). Valette (1977) berpendapat sejalan dengan Finocchiaro and Sako yaitu: 1) the aptitude test; 2) the progress test; 3) the achievement test; dan 4) the proficiency test.

The language aptitude test digunakan untuk mengukur kesiapan mahasiswa sebelum mulai belajar bahasa.

The progress/diagnostic language test dilaksanakan untuk mengetahui penguasaan mahasiswa bagi bagian-bagian tertentu dalam belajar.

The language achievement test diberikan untuk melihat penguasaan mahasiswa dari keseluruhan bahan yang diberikan.

Sedangkan the language proficiency test digunakan untuk mengukur kelancaran dan ketepatan berbahasa mahasiswa.

Uraian tentang penilaian belajar bahasa Inggris dalam makalah ini akan lebih diperinci mengenai language

achievement test untuk jurusan-jurusan non-Bahasa Inggris dan untuk Jurusan Bahasa Inggris ditambah lagi dengan proficiency test.

Dalam mengembangkan instrumen penilaian belajar bahasa pada Jurusan Pendidikan Bahasa Inggris dan Jurusan Pendidikan Bahasa Indonesia ada beberapa komponen yang harus diperhatikan: 1) komponen ketrampilan; 2) komponen informasi, dan 3) komponen teori. Machmoed (1989:4) menjelaskan ketiga komponen ini sebagai berikut:

Komponen ketrampilan meliputi (i) penguasaan bahasa yang akan diajarkan (oleh mahasiswa sebagai calon guru bahasa), (ii) teknik pengajaran dan aktifitas kelas, termasuk pengelolaan belajar, evaluasi, dan lain sebagainya. Komponen informasi antara lain berisi informasi tentang pendidikan, kurikulum yang harus didikuti oleh guru, bahan pengajaran, serta tentang hakekat bahasa. Sedangkan komponen teori terdiri dari filsafat pendidikan [bahasa], psikologi [bahasa], linguistik, linguistik terapan dan lain sebagainya.

Dengan mempertimbangkan ke tiga komponen ini maka pengembangan instrumen-instrumen penilaian belajar menjadi lebih kompleks. Namun demikian, inti sari dari penilaian belajar bahasa ini (termasuk bahasa Indonesia dan bahasa Inggris) adalah manfa'at yang dapat diambil untuk memberi balikan yang sangat berharga terhadap proses belajar mengajar serta untuk memberi informasi kepada semua pihak yang memerlukan hasil penilaian itu, antara lain pembelajar, pengajar, dan pihak lain yang memerlukannya.

PENGEMBANGAN INSTRUMEN PENILAIAN

Penilaian belajar dapat dilakukan dengan menggunakan formative dan summative evaluation. Formative evaluation dilaksanakan untuk setiap unit bahan pelajaran untuk memberikan balikan baik terhadap pembelajar ataupun pengajar. Sedangkan summative evaluation dilakukan untuk melihat hasil belajar dari satu mata kuliah tertentu.

Pada sebahagian besar mata kuliah bahasa pengertian-pengertian dan kemampuan-kemampuan yang pokok harus dikuasai sebelum dilanjutkan kepada unit-unit berikutnya. Untuk ini mastery test untuk mengetahui penguasaan materi dapat digunakan. Instrumen ini memerlukan standar minimum yang harus dicapai oleh setiap mahasiswa, biasanya 75% sampai dengan 80%, dalam penguasaan materi. Jadi evaluasi seperti ini didasarkan kepada penguasaan individu yang dibandingkan dengan kriteria minimum yang harus dicapai oleh semua pembelajar. Inilah yang dimaksudkan dengan criterion-referenced evaluation.

Sebagai tambahan informasi yang diperoleh melalui criterion-referenced evaluation, maka norm-referenced evaluation perlu pula diberikan untuk melihat kemajuan dan penguasaan satu kelompok/kelas dibandingkan dengan yang lainnya.

Untuk mengembangkan instrumen/test yang akan dipakai dalam penilaian belajar bahasa seperti tersebut di atas, Harris (1969) mengemukakan ciri-ciri tes yang baik, yaitu semua tes yang baik hendaklah mempunyai tiga macam kualitas:

reliabilitas (reliability), validitas (validity), dan praktis (practicality). Sedangkan Finocchiaro (1983) menambahkan satu aspek lagi yaitu comprehensiveness.

Reabilitas dari suatu tes menunjukkan bahwa tes dapat mengukur kemampuan pembelajar secara konsisten untuk waktu yang berbeda, tanpa proses belajar. Seorang mahasiswa akan memperoleh nilai yang sama kalau dia dites dengan menggunakan tes yang sama pada waktu yang lain. Validitas dari suatu tes menyatakan bahwa tes mengukur dengan tepat apa yang seharusnya diukur. Seorang mahasiswa yang diuji kemampuan membacanya, maka dalam tes yang diberikan tidak akan ada celah untuk menilai jawaban mahasiswa mengenai ejaan (spelling). Dalam mengembangkan instrumen hendaknya diperhatikan (content validity), validitas empiris (empirical validity), dan validitas format atau bentuk soal (face validity).

Praktis dari suatu tes dapat dilihat dari segi ekonomis (economy), kemudahan dalam melaksanakan dan menilai (ease of administration and scoring), dan kemudahan dalam memberikan interpretasi (ease of interpretation).

Tes dapat dikatakan comprehensive apabila tes itu dapat mencakup (mewakili) semua bagian penting dari bahan yang diajarkan.

Tes/Instrumen Penilaian Belajar Bahasa pada Jurusan-Jurusan non-Bahasa

Telah diuraikan sebelumnya bahwa fokus pengajaran bahasa Inggris pada jurusan-jurusan itu adalah untuk meningkatkan kemampuan membaca sesuatu yang ditulis dalam bahasa Inggris.

Oleh karena itu, pengembangan tes/instrumen penilaian hasil belajar mahasiswa juga dititikberatkan kepada kemampuan membaca itu. Jadi instrumen yang digunakan haruslah merupakan bacaan (passage) beserta pertanyaan-pertanyaan yang baik mengenai isi bacaan itu. Pertanyaan yang baik adalah pertanyaan yang hanya bisa dijawab oleh mahasiswa setelah dia betul-betul mengerti bacaan itu (Person and Johnson, 1978; Lamberg and Lamb, 1980; and Joni, 1983)

Beberapa ahli teori dan peneliti mengemukakan pendapat mengenai tingkat-tingkat pertanyaan (the levels of reading comprehension questions) yang dapat dipakai sebagai bahan pertimbangan untuk membuat pertanyaan yang baik bagi suatu bacaan, antara lain Bloom, Anderson, Herber, dan Ruddel.

Bloom (1956) membagi tingkat pertanyaan (the levels of reading comprehension questions): knowledge, comprehension, applied, synthesis, analysis, dan evaluation questions.

Anderson dan kawan-kawan (1969) membaginya menjadi: literal, interpretive, reference, dan generalization.

Herber (1970) mengelompokkan pertanyaan-pertanyaan itu menjadi: literal, interpretive, dan applied reading comprehension questions.

Sedangkan Ruddel (1974) mengategorikannya dengan: literal, applicative, dan applied reading comprehension questions.

Dari teori-teori yang disebutkan di atas, tingkat pertanyaan yang dikemukakan oleh Herber, yaitu literal, interpretive, dan applied reading comprehension questions,

dapat dipedomani dengan mudah dalam membuat pertanyaan untuk wacana yang dipakai sebagai instrumen penilaian belajar bahasa.

Pertanyaan-pertanyaan yang diajukan sebaiknya lebih banyak diberikan pada tingkat yang ke dua dan tiga. Tingkat yang pertama, literal adalah tingkatan yang paling rendah sehingga mahasiswa dapat menjawab pertanyaan itu walaupun dia tidak mengerti isi wacana secara keseluruhan. Herber (1970) dan Lamberg dan Lamb (1980) mendukung pendapat tersebut.

Penemuan penelitian (Guszake, 1967) menunjukkan bahwa sebahagian besar dari guru membaca (reading teachers) memberikan pertanyaan yang tingkatnya paling rendah itu sehingga instrumen yang seperti itu tidak dapat mengukur kemampuan membaca yang lebih tinggi dari mahasiswa.

Sejalan dengan yang diuraikan di atas, Zainil (1989:ii) telah menyampaikan penemuan penelitiannya sebagai berikut:

Penelitian ini telah mengungkapkan bahwa mahasiswa Program S1 Jurusan Pendidikan Bahasa Inggris IKIP Padang lebih banyak dilatih dengan dengan literal reading comprehension questions sehingga mereka lebih mampu menjawab (pertanyaan) sejenis itu dari pada jenis lain. Mereka yang mampu menjawab lebih banyak pertanyaan interpretive dan applied reading comprehension questions ternyata memperoleh nilai yang lebih tinggi pula.

Berdasarkan penemuan di atas penulis menyarankan agar dalam mata kuliah Reading Comprehension pada Program S1 Jurusan Pendidikan Bahasa Inggris IKIP Padang dipakai buku teks bahasa Inggris yang memuat pertanyaan literal, interpretive, dan applied reading comprehension questions dengan penekanan pada interpretive dan applied reading comprehension questions.

Selanjutnya dapat ditambahkan lagi faktor-faktor yang membantu dalam pengembangan instrumen/tes untuk digunakan dalam penilaian belajar bahasa Inggris pada jurusan-jurusan selain Jurusan Pendidikan Bahasa Inggris, seperti jawaban dari pertanyaan, beberapa buah pertanyaan untuk satu wacana dan berapa panjangnya wacana itu.

Untuk menjawab pertanyaan itu dapat dipedomani rekomendasi yang diberikan oleh Harris (1969:60) bahwa panjang wacana hendaknya berkisar antara 100 sampai dengan 250 kata dan jumlah pertanyaan yang diberikan hendaklah tujuh sampai dengan sembilan buah. Pertanyaan-pertanyaan itu hendaknya diberikan sesuai dengan urutan yang mulai dengan pertanyaan yang menghendaki jawaban yang umum lalu pertanyaan yang menghendaki jawaban yang paling khusus (Best, 1971:166).

Untuk membuat pertanyaan yang baik, maka perlu penjelasan berikut:

- 1) Literal reading comprehension questions adalah pertanyaan yang dapat dijawab oleh mahasiswa sekalipun dia tidak mengerti isi wacana dengan baik. Jawaban pertanyaan yang seperti ini tertulis dalam kalimat-kalimat yang ada dalam wacana, sehingga mahasiswa tersebut menjawab pertanyaan itu hanya dengan menyalin bagian-bagian atau kalimat-kalimat yang ada dalam wacana saja. Pertanyaan yang seperti ini sebaiknya diberikan porsi yang paling rendah.
- 2) Interpretive reading comprehension questions adalah pertanyaan-pertanyaan yang hanya bisa dijawab oleh mahasiswa setelah mereka membaca tidak hanya yang tertulis tapi juga yang tersirat (to read between the lines) untuk memperoleh

arti dalam konteks (implicit meaning), ide (implied meaning), dan lain-lain (Lamberg dan Lamb, 1980).

- 3) Applied reading comprehension questions adalah pertanyaan-pertanyaan yang hanya bisa dijawab dengan benar oleh mahasiswa setelah mereka mengerti lebih jauh dari arti kata-kata yang tertulis (beyond the words), bahkan sering mereka mengambil pengertian lebih jauh dari isi teks (often beyond the entire text) untuk sampai kepada implikasi dan pengambilan kesimpulan (implications and judgments) seperti yang dikemukakan oleh Lamberg dan Lamb (1980).

Pengembangan Tes/Instrumen untuk Penilaian Belajar Bahasa pada Jurusan Bahasa Inggris

Tes/instrumen yang dikembangkan berdasarkan uraian yang sudah dikemukakan dapat digunakan pada mata kuliah yang relevan, seperti Reading I, II, III, dan IV di Jurusan Bahasa Inggris, yaitu mata kuliah yang dapat digunakan untuk mengembangkan salah satu kemampuan (skill) dalam belajar bahasa. Kemampuan dalam belajar bahasa adalah : Mendengar (Listening), Berbicara (Speaking), Membaca (Reading), dan Menulis (Writing). Untuk mengembangkan tes/instrumen penilaian bagi setiap kemampuan itu, Vallette (1977) memberikan banyak cara menilainya, diantaranya adalah sebagai berikut:

a. Speaking

- 1) The Foreign Service Institute (FSI) Rating Procedure

Tes FSI ini digunakan untuk menilai kemampuan berbicara. Instrumen yang digunakan berisi lima komponen yaitu: Accent, Grammar, Vocabulary, fluency, dan

comprehension (Valette, 1977: 157 - 160).

2) The Clark Four Scale System (CFSS)

Clark (1972) yang diuraikan kembali oleh Valette (1977:161) mengemukakan empat komponen dalam menilai kemampuan berbicara yaitu Pronunciation, Vocabulary, Structure, dan Fluency dengan tingkatan 1 sampai dengan 4.

3) The Schulz Communicative Competence Scale (SCSS)

Tes ini hanya mempunyai empat komponen, yaitu Fluency, Comprehensibility, Amount of Communication, dan Quality of communication (Valette, 1977: 161 - 162).

4) The Language Assessment Scales (LAS)

De Avila dan Duncan (1983) mengembangkan LAS, suatu tes standar untuk menilai kelancaran berbicara bahasa asing. Tes ini terdiri dari satu set lengkap dengan format-format untuk penskoran dan analisis serta interpretasinya dan terdiri dari komponen Pairs, Lexical, Phonemes, Comprehension, dan Production.

Semua tes/instrumen yang digunakan untuk menilai kemampuan berbicara bahasa Inggris yang telah diuraikan di atas adalah tes-tes standar international yang dapat dipakai sebagai pedoman untuk pengembangan tes/instrumen penilaian belajar bahasa untuk mata kuliah Speaking dan English Proficiency pada jurusan bahasa Inggris. Oleh karena itu dosen-dosen mata kuliah tersebut dapat mempedomani salah satu dari tes itu dalam mengembangkan instrumen penilaian karena tes-tes itu dilengkapi dengan format pengadministrasian skor serta cara penilaian dan interpretasi hasil penilaian itu.

b. Listening

Dalam semua instrumen yang diuraikan pada (a) di atas terdapat bagian yang digunakan untuk mengukur kemampuan mendengar. Bagian-bagian ini dapat dipedomani untuk pengembangan tes/instrumen yang akan digunakan untuk menilai kemampuan mahasiswa dalam mendengar (listening).

c. Reading

Tes/instrumen yang akan digunakan untuk melihat kemampuan membaca mahasiswa Jurusan Pendidikan Bahasa Inggris dapat dikembangkan dengan cara pengembangan tes/instrumen penilaian kemampuan membaca pada jurusan-jurusan non-Bahasa.

Pertanyaan-pertanyaan yang diberikan untuk sebuah wacana hendaklah meliputi literal, interpretive, dan applied reading comprehension questions dengan memfokuskan kepada kedua tipe pertanyaan yang terakhir. Pertanyaan-pertanyaan itu hendaknya disusun dengan mempedomani pola-pola pemikiran yang ada dalam wacana itu, baik pola pemikiran untuk pengembangan alenia (paragraph) atau pola pemikiran untuk pengembangan esai, antara lain: 1) Illustration-Example; 2) Definition; 3) Comparison; 4) Contrast; 5) Comparison & Contrast; 6) Cause-Effect; 7) classification; dan 8) Chronological Order. (McWhorter, 1986, hal. 89 - 106).

d. Writing

Green (1963) menyarankan beberapa hal untuk membuat tes "esai" (writing) antara lain:

- 1) Pertanyaan (perintah) haruslah ditulis dengan jelas, ringkas, dan tepat (unambiguous).
- 2) Kesahihan tes itu hendaklah dilihat dari isi dan tujuannya.
- 3) Keterandalannya, banyak tergantung kepada jumlah dan tipe pertanyaan (perintah) yang diberikan.

Dia juga mengemukakan dua cara untuk menilai writing: 1) the point score method; dan 2) the sorting method. Yang pertama dapat dilakukan dengan beberapa langkah yaitu: 1) menyediakan kunci (jawaban) yang berisi aspek-aspek penting untuk setiap bahagian (pertanyaan). Berikan nilai untuk bahagian itu bersama-sama dengan nilai bagian setiap aspek yang sudah ditentukan untuk bagian itu; 2) Baca nomor atau jawaban dari soal yang sama untuk setiap esai dan berilah nilai dengan membandingkan semua jawaban itu. Lakukan hal yang seperti ini untuk nomor-nomor atau bagian-bagian lainnya; dan 3) Jumlahkan semua nilai untuk satu esai dan lakukan hal ini untuk seluruhnya. Dari nilai yang sudah ada kemudian dibandingkan dengan nilai yang sudah ditentukan sebagai kriteria untuk lulus. Dari sinilah kita menentukan nilai-nilai akhir untuk semua yang mengambil ujian esai itu.

Yang kedua dapat dilakukan dengan langkah-langkah berikut: 1) Baca keseluruhan esai dengan cepat; dan 2) Kelompokkan sesuai dengan kenyataan yang ada; dan 3) Berikan nilai untuk satu kelompok dan kemudian ambil kelompok itu untuk menjadi standar bagi kelompok-kelompok lainnya.

Dalam melaksanakan kedua cara itu, penulis esai sebaiknya tidak diketahui identitasnya, karena hal ini akan

berpengaruh secara subyektif kepada pemberian nilai. Juga dianjurkan untuk memeriksa semua esai itu pada satu waktu yang sudah ditetapkan. Kalau ini tidak dilakukan maka penilai akan dapat dipengaruhi oleh faktor waktu, kondisi yang berbeda, dan barangkali kegairahan untuk memeriksa dengan sungguh-sungguh juga berbeda.

Dalam menilai yang seperti ini ada beberapa faktor yang sering mempengaruhi penilai antara lain: 1) kualitas dari tulisan (Quality of handwriting) yang sering sekali membuat penilaian menyimpang (bias); 2) tata bahasa (grammar); 3) ejaan (spelling); 4) kosa kata (vocabulary); 5) organisasi ide-ide (organization of ideas); 6) kenecisan (neatness); dan 7) kemampuan mengarang (composition ability). Oleh karena itu akan lebih baik jika setiap faktor ini diberi bobot dan dinilai untuk semua esai. Kemudian nilai total untuk tiap-tiap esai dibandingkan dengan kriteria untuk lulus yang sudah ditetapkan. Dari sinilah nilai-nilai yang lain dibandingkan dan ditentukan.

Selanjutnya Finocchiaro (1983) menjelaskan bahwa writing terdiri dari beberapa komponen, yaitu 1) Mechanics (Spelling, pronunciation, paragraphing); 2) Grammar (morphology dan syntax); 3) Appropriateness of vocabulary; 4) Information and logical idea produced; and 5) organization of ideas.

Dengan mempedomani hal-hal yang disebutkan di atas maka tes/instrumen penilaian untuk writing akan dapat dikembangkan dengan baik oleh dosen yang bersangkutan. Komponen-komponen yang dapat dinilai untuk writing diberikan dalam tabel (Finocchiaro, 1983: 144 - 145, 150 - 151, dan 156).

PENETAPAN SKOR AKHIR

Penetapan skor akhir dapat dilihat baik untuk tiap-tiap mata kuliah ataupun untuk indeks prestasi semester bagi setiap mahasiswa untuk setiap semester dan indeks prestasi kumulatif bagi setiap mahasiswa untuk selama berada dalam satu program studi pada jurusan tertentu.

Pertama, penetapan skor akhir untuk mata kuliah tertentu dapat dilakukan dengan menggunakan instrumen serta penskoran yang sesuai dengan mata kuliah masing-masing seperti yang sudah diuraikan di atas, misalnya menggunakan LAS yang sudah dimodifikasi untuk mata kuliah English Proficiency.

IKIP Padang mempunyai pedoman untuk menetapkan skor akhir bagi penguasaan mahasiswa untuk satu mata kuliah tertentu sebagai berikut: Buku Pedoman IKIP Padang (1988/1989:39)

Derajat penguasaan	Nilai
90% - 100%	A
80% - 89%	B
65% - 79%	C
55% - 64%	D
< 55%	E

Nilai A diberi bobot 4, B 2, C 3, D 1, dan E 0 (Buku pedoman IKIP Padang, 1989:40)².

Kedua, IP (Indeks Prestasi) semester dan IP kumulatif bagi setiap mahasiswa di IKIP Padang diperoleh dengan menggunakan rumus: (Buku Pedoman IKIP Padang, 1989: 40 - 41).

$$IP = \frac{\sum (B_n \times K)}{\sum K}$$

IP = Indeks Prestasi Semester atau Kumulatif

B_n = Angka nilai yang diperoleh untuk setiap mata kuliah dalam periode tertentu (semester atau seluruh program yang ditempuh).

K = Harga SKS (sistem kredit semester) masing-masing mata kuliah.

Dari rumus ini kelihatan bahwa setiap nilai dari mata kuliah yang berbeda-beda (setelah diberi bobot) dikalikan dengan harga SKS dan dijumlahkan serta kemudian dibagi lagi dengan jumlah semua harga SKS yang diambil. Barangkali penjumlahan yang seperti ini kurang tepat karena yang dijumlahkan adalah hal yang berbeda sama saja dengan menjumlahkan duku dan durian yang secara kebetulan keduanya adalah buah-buahan.

IP semester yang diperoleh dengan cara beginipun digunakan pula untuk menentukan nasib mahasiswa dalam mengambil jumlah jam kredit (mata kuliah) pada semester berikutnya. (Buku Pedoman IKIP Padang, 1989: 26)

Pedoman Penetapan Beban Studi

SO	Indeks Prestasi Semester Yang Lalu		Maksimal semester berikutnya dalam SKS
	S1	S1-Thesis	
0,00 - 1,00	0,00 - 1,00	0,00 - 1,00	15
1,00 - 1,75	1,00 - 2,00	1,00 - 2,00	19
> 1,75	> 2,00	2,00 - 2,50	22
---	---	>2,50	23

Kalau dikaji secara ilmiah apakah cara ini sudah tepat? Banyak sekali kemungkinan-kemungkinan yang bisa berlainan untuk semester yang akan datang kalau dibandingkan dengan semester sebelumnya; misalnya tingkat kesukaran mata kuliah; bobot yang berbeda yang diberikan dosen untuk SKS yang sama; situasi, kondisi, dan fasilitas yang berlainan pula; kekurangan mampunya pada semester yang lalu belum tentu terjadi persis sama untuk semester berikutnya, karena memang ada mata kuliah yang dirasakan lebih sulit dari yang lainnya; mungkin saja seseorang mahasiswa lemah dalam satu mata kuliah, tapi mungkin sangat baik untuk mata kuliah lainnya; barangkali seorang mahasiswa kurang sehat untuk beberapa waktu pada semester sebelumnya sehingga dia tidak punya waktu banyak untuk belajar, sedangkan dia punya kemampuan dan semangat yang tinggi pada semester berikutnya.

Oleh karena cara penjumlahan dengan memakai rumus di atas, menjumlahkan sesuatu yang berbeda, tidaklah benar kalau dilihat dari segi statistik, maka perlu dicari jalan untuk memperbaikinya. Caranya adalah sebagai berikut:

1. Cari Z-score untuk setiap mata kuliah (Hatch & Farhadi, 1982:66)

$$Z = \frac{X - \overline{X}}{S}$$

X = Score

\overline{X} = mean

S = Standard deviation

Jadi mean dan standard deviation untuk setiap mata kuliah harus dicari lebih dahulu sebelum menggunakan rumus ini.

2. Jumlahkan semua Z-score untuk semester atau seluruh program. (Secara statistik sudah benar karena sudah menjumlahkan yang sejenis).
3. Tentukan IP semester atau kumulatif.

Dengan cara begini dapatlah seorang mahasiswa dibandingkan dengan mahasiswa lainnya untuk menentukan ranking.

Namun demikian IP semester ini sebaiknya tidak digunakan untuk membedakan mahasiswa dalam mengambil jumlah jam kredit untuk semester berikutnya. Sebaiknya IKIP menentukan jumlah jam kredit maksimum dan minimum yang dapat diambil oleh setiap mahasiswa dengan memperhitungkan lama satu program dan jumlah jam kredit untuk program itu sehingga dapatlah rata-rata jam kredit untuk satu semester. Kemungkinan lebih atau kurang sedikit hendaknya ada jarak (range) yang juga ditentukan oleh IKIP. Untuk bergerak dalam jarak itu, berikanlah kebebasan kepada PA yang bersangkutan. Carroll (1963) mendukung ide itu dan dia menyatakan bahwa kalau

MILIK UPT PERPUSTAKAAN
IKIP PADANG

mahasiswa: diberikan waktu yang dibutuhkannya untuk belajar, sebahagian besar (95%) dari mereka akan menguasai materi yang diberikan. Dia mempunyai satu "model of learning" yang belum dibantah orang sampai sekarang dan model ini banyak sekali dijadikan referensi oleh profesor-profesor yang mengajarkan bahasa asing.

Model itu adalah:

$$\text{Degree of learning} = f \left\{ \frac{\text{Time actually spent}}{\text{Time needed}} \right\}$$

"Time actually spent" equals time allowed and perseverance.

"Time needed" equals aptitude, quality of instruction, and ability to understand instruction.

Pengembangan tes/instrumen sampai dengan penskoran (scoring) serta format-format yang diperlukan untuk pengadministrasian skor seperti yang sudah diuraikan sebelumnya sudah ada dalam buku-buku teori tentang penilaian hasil belajar bahasa (Language Testing). Teori-teori itu jarang sekali dilaksanakan oleh dosen-dosen bahasa. Menurut pengalaman penulis serta pemantauan terhadap teman-teman sejawat selama 30 tahun mengajar Bahasa Inggris, maka sebahagian besar dari dosen-dosen bahasa belum melakukan sesuai dengan apa yang seharusnya dilaksanakan dalam penilaian belajar bahasa. Yang sering dilakukan adalah menilai secara global dimana subyektifitas banyak berpengaruh atau menjumlahkan skor dari setiap soal yang diberikan dan kemudian mentransfer nilai angka ke dalam nilai huruf: A, B, C, D, E, dan BL.

KESIMPULAN

Dari uraian di atas dapat disimpulkan bahwa kenyataan di lapangan menunjukkan sebahagian besar dosen melaksanakan penilaian belajar bahasa dengan mengambil pedoman yang sangat minimum dari teori-teori yang ada. Oleh karena itu, penilaian belajar bahasa hendaknya dilaksanakan sesuai dengan proses belajar mengajar. Untuk ini, formative dan summative evaluation dapat dilaksanakan. Dengan kata lain, suatu kriteria (criterion referenced evaluation) perlu dibuat untuk setiap unit bahan kuliah atau untuk semua bahan kuliah dalam satu semester.

Sebagai bahan pertimbangan, faktor-faktor berikut dapat diperhatikan atau dinilai: 1) kehadiran; 2) pekerjaan terstruktur dan mandiri (homework); 3) partisipasi; dan 4) proyek atau paper dan/atau ujian. Untuk penilaian hasil belajar baik bagi setiap unit, bahan kuliah atau bahan kuliah untuk selama satu semester, yang perlu diperhatikan adalah komponen-komponen dengan bobot penilaian untuk setiap komponen-komponen yang dimaksud.

Kesimpulan ini didukung oleh undang-undang No. 2 Tahun 1989, Bab XII, Pasal 43 bahwa penilaian belajar merupakan bagian yang integral dari proses belajar mengajar. Lebih khusus lagi, IKIP Padang telah menggariskan bahwa penilaian hasil belajar haruslah didasarkan kepada kehadiran, keaktifan di dalam kelas, tugas terstruktur dan ujian (Buku Pedoman IKIP Padang, 1989, hal 38 - 39).

DAFTAR BACAAN

- Anderson, Jonathan H. dan Milicent E. Poble. (1969). "Reading in Specific Content Areas", dalam Efficient Reading: A Practical Guide. Sydney: McGraw-Hill Book Company, hal.118 - 135
- Best, John W. (1981). Educational Research. New Jersey: Prentice Hall, Inc.
- Bloom, Benyamin S., dan kawan-kawan. (1965), Taxonomy of Educational Objectives. New York: Longman, Green.
- Carroll, J.B. (1963). A Model of School Learning, Teachers College Record, 64, hal. 723 - 733.
- Cohen, Andrew D. (1980). Testing Language Ability in the Classroom. Rowley, Massachusetts: Newbury House Publishers, Inc.
- Finocchiaro, mary dan Sydney, Sako. (1983) Foreign Language Testing: A Practical Approach. New York: Regents Publishing Company, Inc.
- Green, John A. (1963). Teacher Made Tests (2nd ed.). New York: Harper & Row, Publishers
- Gusrak, Frank J. (1967). "Perspective in Reading", dalam The Evaluation of Children Achievement, Nemark: International Reading Association, Hal.97 - 109.
- Harris, David P. (1969). Testing English as a Second Language. New Delhi: McGraw Hill, Inc.
- Hatch, Evelyn dan Hosein Farhady. (1982). Research Design and Statistics for Applied Linguistics. Rowley, Massachusetts: Newbury House Publishers, Inc.
- Herber, Harold L. (1970). Teaching Reading in Content Areas. New Jersey: Prentice Hall, Inc.
- IKIP Padang. (1989 - 1990). Buku Pedoman. Air Tawar, Padang: Percetakan IKIP
- Lado, Robert. (1961). Language Testing. Hongkong: Wing Tai Cheung Printing Co, Ltd.
- Lamberg, Walter J. dan Charles E. Lamb. (1980). Reading Instruction in the Content Areas. Boston: Hangton Miffin Company.
- Machmoed, Zaini. (1989). Sebuah Model Teori Proses Belajar-Mengajar Bahasa. Malang: IKIP Malang (makalah yang disampaikan pada kongres ISP, 1989, di Jakarta).

- McWhorter, Katherine T. (1986). Guide to College Reading. Boston, Little Brown and Company
- Micheels, William J. dan M. Ray Karnes. (1950). Measuring Educational Achievement. New York: McGraw - Hill Book Company, Inc.
- Valette, Rebecca M. (1967). Modern Language Testing. New York: Harcourt Brace Javanovich, Inc.
- Zainil. (1989). Kelancaran Berbicara Bahasa Inggris Mahasiswa Jurusan Pendidikan Bahasa Inggris FPBS- IKIP Padang. Padang: Pusat Penelitian IKIP Padang
- Zainil, Yetti. (1989). A Study of Types of Reading Comprehension Questions. Padang: Jurusan Pendidikan Bahasa Inggris FPBS- IKIP Padang (Tesis).

PANITIA SEMINAR
PENYUSUNAN BUKU PANDUAN PENILAIAN HASIL BELAJAR
MAHASISWA IKIP PADANG

PENILAIAN BELAJAR SENI, MUSIK DAN TARI

Oleh : Dr. Ibenzani Usman

PADANG, 22 - 23 JANUARI 1990

PENILAIAN BELAJAR SENI, MUSIK, DAN TARI

Oleh : Dr. Ibenzani Usman

Disampaikan pada forum seminar
Evaluasi Pendidikan IKIP Padang
di Aula FPTK IKIP Padang tanggal
22 dan 23 Januari 1990

Pendahuluan

Pertama-tama kami sampaikan ucapan terima kasih yang sebesar-besarnya kepada Rektor IKIP Padang, melalui Pusat Testing dan Evaluasi Pendidikan IKIP Padang selaku penyelenggara seminar, yang telah memberi kehormatan dan kepercayaan kepada kami menyampaikan makalah dalam kesempatan ini. Selain ucapan terima kasih kami juga ingin memohonkan maaf kepada segenap sidang hadirin atas makalah ini yang sesungguhnya berada di luar disiplin keilmuan kami. Namun bertolak dari modal dasar pengalaman mengajar di Jurusan Seni Rupa dan Keterampilan Kerajinan FPBS IKIP Padang selama lebih dari 26 tahun dan pernah pula membantu-bantu di ASKI Padang Panjang dan Jurusan Sendratasik FPBS IKIP Padang, kami coba menyampaikan beberapa bahan masukan melalui judul "Penilaian Belajar Seni, Musik, dan Tari" ke dalam forum ini sejauh kemampuan yang ada pada kami. Mudah-mudahan ada yang dapat terjaring kedalam sasaran seminar ini.

Jadi ruang lingkupnya tidak dapat keluar dari penilaian belajar di bidang studi seni rupa dan sendratasik.

Kalau kami tidak salah tafsir, permasalahan terungkap pada Terms of Reference yang disampaikan kepada kami adalah sebagai berikut : Bagaimana mempersempit variabelitas pengertian dan

hakekat antero bidang studi sehingga secara makro dapat diwujudkan suatu sistem dan mekanisme kerja penilaian yang menghasilkan "nilai" dalam kesatuan makna.

Secara mikro tentu hal ini akan langsung berhadapan dengan dilema yang tidak mudah mencari jalan keluarnya. Hal ini dapat dipahami mengingat perbedaan hakekat bidang studi di seantero jurusan dan fakultas yang terdapat di IKIP Padang, seperti telah disinyalir juga di dalam TOR, sehingga model-model Penilaian dengan Acuan Norma (PAN) atau Penilaian dengan Acuan Patokan (PAP) memberi pengertian yang berbeda-beda pula bagi setiap dosen. Namun demikian kembali dilihat secara makro, tentu ada celah-celah yang mungkin dapat diisi sebagai upaya kearah pencapaian tujuan yang diharapkan. Se jauh apa yang pernah kami amat, celah-celah itu ada pada variabelitas itu sendiri.

Variabelitas Bidang Studi Seni

Seni adalah wujud budaya fisik masyarakat suatu bangsa. Sebagai wujud budaya fisik seni mampu memroyesikan wujud budaya tangkai laku berpola yang diekspresikan dalam hal-hal yang bersifat seremonial. Sebagai wujud budaya fisik seni juga mampu membangun citra (image) yang terkandung dalam wujud budaya ideal yang menguak tabir nilai, norma, aturan khusus dan perangkat keyakinan suatu masyarakat dan bangsa¹⁾. Sekalipun seni itu merupakan hasil ciptaan orang perorangan, kenyataannya pun akan demikian juga.

Dari ungkapan di atas ada benarnya bila suatu teori mengungkapkan bahwa suatu karya seni selalu didukung oleh tiga macam nilai, yaitu nilai filsafat, nilai kejiwaan dan nilai fisik itu sendiri²⁾. Nilai-nilai itu bersumber pada unsur-unsur

bidang filsafat, bidang ilmu jiwa dan fisik seni seperti garis, warna, nada, irama, gerak dan sebagainya. Oleh karena itu studi seni memerlukan bantuan ilmu pengetahuan dasar yang dipungut dari disiplin ilmu pengetahuan lain seperti, ilmu sejarah, Sosiologi, anthropologi budaya, dan filsafat. Ilmu-ilmu pungut itu terhimpun dalam beberapa mata kuliah seperti :

- di Jurusan Seni Rupa : Sejarah Seni Rupa, Filsafat Seni / Estetika, Apresiasi/Tinjauan seni.
- di Jurusan Sendratasik: Sejarah Musik/Tari, Estetika Musik / Tari, Apresiasi Musik/Tari dan Tinjauan Musik dan Tari.

Semua matakuliah di atas termasuk kepada kelompok mata kuliah pengetahuan dasar seni.

Seni sebagai wujud budaya fisik mempunyai wajah seribu satu muka. Hal ini terjadi karena seni memang merupakan hasil dari dunia kreativitas. Tidak pernah hasil suatu karya seni merupakan pengulangan dari karya seni lainnya. Itu pulalah yang menyebabkan variabelitas bidang seni itu amat luas sekali. Walau demikian keadaannya variabelitas dalam seni itu masih dapat dikelompokkan kedalam tiga bagian besar. Kelompok yang menjurus kepada aspek penglihatan (Seni Rupa), pendengaran (musik) dan kedua aspek dan gerakan (tari)³⁾. Masing-masing kelompok inipun masih ber-variabilitas. Akan tetapi bila hal ini dikaitkan dengan studi seni itu sendiri, hal ini akan bermuara di alam kreativitas. Kreativitas adalah suatu proses yang di dalamnya terdapat tiga tahapan kegiatan. Tahap permulaan, disebut juga masa inkubasi atau incept, adalah tahap munculnya inspirasi, perenungan dan pematangan konsep gagasan. Tahap tindak nyata, dimana seluruh potensi pelaku proses kreatif (misalnya seniman) dikerahkan secara aktif dan intensif. Dalam tahap ini pelaku

mengerahkan semacam instrumental input berupa keterampilan, sarana, dan lingkungan; serta fundamental input berupa originalitas, apresiasi dan identitas. Kemudian tahap akhir, yaitu tahap penyelesaian dari suatu karya (Karya seni misalnya).⁴⁾

Betapapun bentuk dan sifatnya proses kreatif bidang seni, dalam studi seni sebagai proses belajar mengajar, hal ini akan tertampung dalam satu pola saja, yaitu proses belajar mengajar yang dituangkan kedalam jenis mata kuliah praktikum. Dalam hal ini mahasiswa diberikan teori atau pengetahuan terapan kemudian mereka melakukan kegiatan mandiri mengeksperimentkan atau mempraktekkan teori terapan tersebut. Mata kuliah demikian dapat dicontohkan seperti:

-di Jurusan Seni Rupa : Menggambar Bentuk, Menggambar Ilustrasi, Seni Lukis, Seni Patung, Seni Grafis, Seni Keramik, dsb.

-di Jurusan Sendratasik : Program Studi Tari : Gerak Dasar tari, Praktek Tari, Komposisi Tari, Koreografi dsb.

Program Studi Musik: Teori Musik, Direksi, Praktek vokalia, Praktek instrumentalia, Harmoni dan aransemen dan sebagainya.

Semua mata kuliah tertayang di atas termasuk kedalam kelompok matakuliah praktikum. Kami yakin bahwa istilah praktikum ini akan membawa asosiasi kepada mata-mata-kuliah yang terdapat pada jurusan-jurusan MIPA. Kami tidak tahu persis dimana letak kesamaannya. Kalau perbedaannya tentu banyak sekali. Namun matakuliah praktikum di Jurusan Seni Rupa dan Sendratasik, bentuk kegiatannya seperti sudah dijelaskan di atas yaitu, teori terapan ditambah eksperimentasi.

Memang ada mata kuliah yang aktivitasnya bersifat asah terampil. Dalam hal ini mahasiswa lebih banyak diarahkan kepada kemahiran mewujudkan sesuatu. Misalnya seperti terdapat di:

-Jurusan Seni Rupa : Menggambar Anatomi, Menggambar Proyeksi, dan seni tulis.

-Jurusan Sendratasik: Gerak Dasar Tari, Notasi Tari, Solfegio, dan Dikte Musik.

Mata-mata kuliah di atas adalah mata-mata kuliah yang mendukung mata-mata kuliah praktikum. Tugas-tugas yang diberikan dalam mata kuliah ini lebih banyak bersifat latihan-latihan. Maksudnya adalah dengan melakukan banyak latihan dalam menggambar anatomi, maka mahasiswa akan mudah menggambar manusia atau hewan secara wajar. Begitu juga dengan latihan gerak dasar tari, mahasiswa akan mudah menarikan sebuah karya tari dengan wajar. Dengan banyaknya berlatih dalam solfegio dan dikte musik, mahasiswa akan mudah menggelarkan karya-karya komponis terkemuka secara wajar. Mata kuliah semacam ini sebetulnya tidak banyak, akan tetapi tidak ada salahnya bila dikelompokkan juga kedalam kelompok Matakuliah Bantu Praktikum.

Jadi Bidang Studi Seni memang banyak variabilitasnya. Namun variabilitas itu hanyalah dalam bentuk media, sarana dan wujudnya. Dalam variabilitas itu juga dapat ditemui celah-celah kesamaan hakekat dan pengertian seperti terlihat dalam pengelompokan, proses kreatif, proses belajar mengajar dan tiga wujud mata kuliahnya : Mata Kuliah Ilmu Pengetahuan Dasar Yang Teoritis, Mata Kuliah Praktikum yang teoritis terapan dan eksperimentasi, dan Matakuliah Bantu Praktikum yang penuh tugas latihan.

Penilaian Belajar Seni

Pada Variabilitas Bidang Studi Seni di atas sudah terlihat tiga kelompok matakuliah yang masing-masingnya tentu memerlukan model penilaian yang mengacu kepada norma/kriteria dalam bidang

seni itu sendiri. Norma dan kriteria ini sebetulnya sudah ada di dalam masing-masing kelompok mata kuliah. Barangkali yang diperlukan itu adalah semacam matrik dengan lajur-lajur penampung aspek-aspek ukur yang sesuai dengan keperluan masing-masing kelompok mata kuliah.

Sebelum sampai kepada pemilihan dan penetapan matriks yang akan dipakai baiklah ditinjau mengenai apa yang akan dinilai itu. Yang akan dinilai itu tidak lain dari pada perkembangan individu dalam studi seni. Dengan perkembangan berarti kita dihadapkan kepada pertanyaan apa yang telah diperbuat individu itu kemaren, apa pula yang sedang diperbuatnya sekarang dan bagaimana prospek masa depannya. Para ahli di bidang pendidikan seni telah menemukan beberapa kriteria sebagai pegangan bagi dosen ybs untuk mengamati perkembangan individu si mahasiswanya. Conant dan Randall⁵⁾ menawarkan empat kelompok kriteria untuk menilai perkembangan individu melalui belajar seni, yaitu Perkembangan kepribadian secara menyeluruh, kompetensi sosial, Kreativitas, dan Paham estetika serta cakup berseni.

Dalam hal perkembangan individu secara total, ada delapan kemampuan yang harus dimiliki, di antaranya mampu berinisiatif, mampu berkonsentrasi dalam menghadapi ekspresi karya seni, dan sebagainya. Dalam kompetensi sosial, mampu menghayati keunikan keunikan, mampu kerja sama dan mampu melihat perkembangan orang lain. Kreativitas menghendaki lima macam kemampuan diantaranya gairah bereksperimen dengan ide materi dan alat. Paham estetika dan cakup berseni juga ada empat mampu dan dua kemauan.

Keempat kriteria tersebut (perinciannya lihat lampiran) memerlukan pengamatan terus menerus. Sekurang-kurangnya sebanyak 16 kali tatap muka di dalam kelas atau studio. Dalam kaitannya

dengan sistem penilaian IKIP Padang (Buku Pedoman 1989-1990; 38-39), kehadiran dalam kuliah, Keaktifan mahasiswa di dalam kelas atau studio, dan tugas terstruktur sangat menentukan sekali, di samping hasil-hasil ujian yang dicapainya.

Perihal kehadiran dan keaktifan mahasiswa di kelas dan studio, idealnya diperlukan semacam checklist yang kriterianya tersebar perkembangan pribadi dan kompetensi sosial tawaran Conant diatas. Alangkah baiknya bila para dosen (seni) dapat mempersiapkan checklist menurut matakualiah yang di binanya. Kemudian dalam hal kreativitas dan pemahaman estetika serta kemampuan berseni, ini memang memerlukan matrik tersendiri. Matrik yang cocok ini dapat dipedomani Matrik unsur dan prinsip disain dari W.J.Gephart's (Lihat lampiran 2 hal 290).^{5A)}

Prinsip Disain ada 8 yaitu Unity, Conflict, Dominance, Repetition, Alternation, Balance, Harmony dan Gradation. Unsur disain ada 7 yaitu Line, Value, Color, Texture, Shape, Size, dan Direktion. Kedelapan Prinsip dihadapkan kepada ketujuh unsur disain tersebut. Hasilnya suatu matrik dengan 56 kotak. Lalu kita lihat hasil tugas terstruktur mahasiswa, misalnya karya seramik atau tembikar. Lalu kita lihat hubungan prinsip dan unsurnya. Line dengan Unity baik, beri tanda V, line dengan Conflict kurang, beri tanda (-), Line dengan Dominance, baik beri lagi tanda (V) dan seterusnya, Begitu juga hubungan value dengan conflict, dan seterusnya, dan seterusnya sehingga ke 56 kotak itu terisi semuanya. Kalau semua kotak berisi 56 (V), ini berarti bahwa mahasiswa tersebut telah mencapai tingkat keberhasilan 100%. Dia dapat diberi A. Bila 45 (V) berarti menguasai 70%. Dia dapat diberi B. Bila 35 (V) berarti 60% atau C, 15 (V) = 25% = E.

Matrik di atas hanya untuk menilai tugas-tugas terstruktur pada mata kuliah Praktikum Jurusan Seni Rupa. Jurusan Sendratasik juga bisa memanfaatkan dengan catatan harus bisa ditransformasikan prinsip dan unsur yang terdapat dalam mata kuliah musik dan tari. Misalnya dalam mata kuliah program studi Musik, Prinsip disainnya adalah partitur yang berisikan Melody, Harmony, Rhythm, Dynamic, ekspresi, Interpretation, dan teknik. Sedangkan unsur-unsur disainnya (The Physical Characteristics) adalah, Tone, Pitch, Loudness, Timbre dan Duration.⁶⁾ Ketujuh prinsip ini dihadapkan kepada lima unsur musik ini diperoleh tiga puluh lima Kotak yang akan menentukan taraf keberhasilan mahasiswa untuk mata kuliah praktikum tersebut. Begitu juga dalam tari, Prinsip dan unsur disainnya dapat dicari kedalam azas-azas koreografi. Kesemuanya ini juga bisa dihadapkan kepada ujian-ujian dalam mata kuliah praktikum.

Pada mata kuliah pengetahuan dasar seni dan bantu praktikum kiranya yang diperlukan cuma sebuah tabulasi score. Sebab yang ingin kita ketahui itu adalah seberapa jauh kemampuan mahasiswa menyerap pengetahuan dan teori-teori seni selama aktif dalam proses belajar mengajar atau sejauh mana penguasaan keterampilan yang dimiliki mahasiswa selama menjalani latihan demi latihan pada mata kuliah bantu praktikum tersebut.

Secara garisbesarnya penilaian belajar seni ini dapat digambarkan sebagai berikut. (1) Mata kuliah Prog.Stu.Sen ada tiga kelompok; pengetahuan dasar, bantu praktikum dan Praktikum seni. (2) Untuk m.k. dasar dan bantu cukuplah mengandalkan model tabulasi score biasa, Untuk m.k. praktikum diperlukan suatu matrik tersendiri. (3) Menyesuaikan diri dengan sistem yang sudah ada (Buku Pedoman, IKIP Padang '90), unsur wajib-

hadir, aktivitas dikelas dan studio perlu diwujudkan dalam suatu checklist. (4) Tugas-tugas terstruktur, dan ujian-ujianya, disesuaikan dengan bentuk mata kuliahnya, tabulasi score atau matrik. (5) Total score keseluruhannya untuk satu mata kuliah ditawarkan konsep tabel di bawah ini.

DAFTAR NILAI.....
 JURUSAN SENI RUPA
 MATA KULIAH: MENGGAMBAR ILLUSTRASI
 dst.

No	NAMA	No BP	Hadir	Tugas	Mid	Sem	akhir	
			Aktif	Tstr			ANK	HRP
3	Badu	9023123	bobot 0,1	0,2	0,2	0,5	7,1	B
			7	8	5,5	7,5		
			0,7	1,6	1,1	3,7		

Keterangan : Si Badu BP 9023123

Hadir dan aktif	70%	bobot 0,1 =	0,7
Tugas ter str	80%	bobot 0,2 =	1,6
Mit Tes Sem	55%	bobot 0,2 =	1,1
Ujian akhir Sem	75%	bobot 0,5 =	3,7

Total angka	7,1
huruf	B

Pengakhiran (tapi belum berakhir)

Sebagai gambaran tentang 'penilaian belajar seni, musik dan tari' sudah kami sajikan secara ideal seperti tertera di atas. Kami katakan ideal karena kami sudah antarkan masalah penilaian ini sampai ujung jalan gang-gang terkecil, Namun sampai sejauh manakah keikhlasan hati para dosen rela mengantarkan mahasiswa sampai ke ujung-ujung jalan ini.

Kami tidak melihat perbedaan hakekat dan pengertian dalam variabilitas bidang studi seni. Wajahnya memang seribu satu muka, akan tetapi proses kreatif, proses belajar dan teknik evaluasinya berada atau diperadakan dalam satu pola.

Pola seperti tergambar di atas

Demikianlah sumbangan pikiran sebagai bahan masukan dari kami sebagai wakil dari bidang studi seni. Bidang seni memang unik, tapi jelas tidak seunik matematik. Mudah-mudahan ada manfaatnya. Sebagai akhir kata dari makalah ini ingin kami sampai sebuah kalimat yang unik di bawah ini.

"Sabukik depek dibarikan, nan sapingik tahan dahulu".

Terus terang, saya pribadi sungguh tidak mengerti hakekat kalimat ajaib ini. Mudah-mudahan para hadirin juga tidak mengerti. Karena kita sebagai guru yang baik tidak pernah tahu apa itu "nan sapingik", apa lagi memiliki yang "sapingik" itu. Apa lagi yang akan kita tahan-tahan.

Sekianlah. Terima kasih.

Padang, 22 januari 1990

6) Howard D. McKinney and W.R. Anderson, Discovering Music, American Book Company, New York, 1952, h. 43-57,

Catatan Tambahan:

- 1) Hampir dalam beberapa bukunya Koentjaraningrat menyampaikan konsep wujud budaya ini yang dapat kami interpretasikan seperti kalimat tersaji di atas; di antaranya adalah, Prof. Dr. Koentjaraningrat, Kebudayaan Mentalitet dan Pembangunan, Penerbit P.T. Gramedia, Jakarta 1974, h. 15 - 16.
- 2) Herbert Read, The Meaning of Art, Pelican Books, London, 1954, h. 35.
- 3) Thomas Munro, The Arts and Their Interrelations, The Press of Case Western Reserve University, Cleveland and London, 1969, h. 193.
- 4) Prof. Ahmad Sadali, dalam Agus Sachari, SD Int., (Ed), Seni, Desain & Teknologi, Penerbit Pustaka, Bandung, 1986, h. 1 - 17.
- 5) Prof. Dr. Howard Conant and Dr. Arne Randall, Art in Education, Chas. A. Bennett Peoria, Illinois, 1963, h. 205, 206.
- 5a) William J. Gephart, dalam Nick L. Smith (ed), New Techniques for Evaluation, Sage Publications, Beverly Hill, London, h. 286

The key to successful evaluation is a concept of growth measurement, as contrasted with the determination of status. The touchstone of growth is the individual himself—*what he was, what he is, and what he gives promise of becoming.*

Criteria Suggested for Evaluating Individual Growth Through Art

EVIDENCE OF INDIVIDUAL GROWTH IN:

Total personality development

Individual uniqueness

- Ability to think for oneself, to use initiative
- Ability to identify or put something of oneself into one's work
- Ability to concentrate upon creative art expression to a point where the individual is not easily distracted.
- Ability to express moods and sincere feelings through art expression
- Ability to work to the capacity of one's ability
- Ability to evaluate oneself
- Ability to receive and profit by deserved praise and constructive criticism
- Ability to derive personal satisfaction and justifiable pride from accomplishments in art

Social competence

- Ability to enjoy others' uniqueness
- Ability to work cooperatively with others
- Ability to evaluate growth in others

Creativity

- Willingness to experiment with ideas, materials, and tools
- Interest in art activities which involve personal value judgments and unique working methods
- Enjoyment of the learning process as well as the end product
- Ability to express ideas in two- or three-dimensional form as well as in words
- Ability to express oneself through abstract as well as realistic symbols
- Ability to work creatively, with steadily decreasing interest in the use of non-creative devices

Aesthetic understanding and art ability

- Ability to produce art works which possess increasingly significant form. This comes about intuitively, in children, as a result of deepened interest and sincere expression in art activities. On this level, the quality of *form* in the art work is merely an indication of deepened interest and sincerity, and is not something to be evaluated in itself. On the adolescent and adult

levels, however, the ability to produce art works which possess increasingly significant form becomes more important to both student and teacher.

Ability to observe design in everyday environment

Willingness to attempt and ability to solve increasingly more complex problems

Ability to make use of increasingly more complex processes and tools

Ability to choose and correctly use best-suited tools and equipment

Use of art knowledge in selection of personal clothing and grooming styles, living quarters, manufactured items, and art objects

5) Prof. Dr. Howard Conant and Dr. Arne Randall, Art in Education, Chas. A. Bennett Peoria, Illinois, 1963

LAMPIRAN II

ALTERNATIVE CRITICAL PERSPECTIVES 299

Terry Denny, thank you for the opportunity to read a fascinating story. And thank you for the opportunity to springboard from that story to make these five points:

- (1) The descriptive research process and the evaluation process in application are markedly different.
- (2) the criteria we use for assessing the quality of them are also different.
- (3) Composing an evaluation is an important ingredient in our work. Evaluations do not just happen, we create them and we need heuristics to help us.
- (4) Evaluations need to compel the mind. We need to deal with the major and minor decisions in a way in which our attention handles all of them, but at the level of intensity they demand.
- (5) By helping our clients complete the value picture, we help them become more informed about the options. Our basic assumption is that a more informed decision is a better decision.

William Gephart, who is currently an independent consultant in Bloomington, Indiana, has invested twenty-seven years in the study of painting. For the past twenty years he has worked almost exclusively with watercolors. His paintings are in public and private collections in seven countries. He has had several one-man shows and has participated in numerous group shows; his paintings have been displayed in galleries in Florida, South Carolina, Kentucky, Indiana, and Michigan. Dr. Gephart has been making contributions to the understanding of the evaluation process for the past fifteen years through: co-authorship, with Daniel Stufflebeam and five others, of Educational Evaluation and Decision Making (F.E. Peacock, 1971); as editor of Phi Delta Kappa's CEDR Quarterly; as founder and Co-Director of the National Symposia for Professionals in Evaluation and Research (NSPER); and as founder and past President of the Evaluation Network.

THE WORK OF Terry Denny (1978a) is some of the most delightful writing about education I have encountered. There is a quality in his writing that is reminiscent of a long-forgotten television series entitled "You Are There." In those programs, through the magic of television, we were transported along with a journalism guide to experience some of the major events in history. Terry Denny is capable of giving us this same treatment. Through the use of words, and very simple words at that, Denny makes you know the education program in a community. That's not quite right. Through the magic of his words, Terry Denny puts you in a specific education setting. You feel the people, you know them. Their triumphs, their fears, their successes, their failures, encompass you. You are there.

His work at first appears so simple. He quotes people. He interacts with the people who make up an education setting or an education program and records their statements. How could anything be more simple? In the midst of this simplicity, though, is a thorny problem. How do you decide on the order or sequencing of the hundreds of individual statements that have such remarkable characteristics, so fresh, or so full of sadness, so excited, optimistic, pes-

simistic, biased, naive? One thing is clear; the statements Denny chose to record in his spiral notebook or via a small recorder could not have come in the sequence in which they are presented. Someone insightfully pieced them together in a way that magically makes you feel that you are a part of that scene. That someone was Terry Denny. I covet his skill.

My assignment is to present a meta-evaluation of Terry Denny's work. In particular, we are focused on a report entitled "Some Still Do: River Acres, Texas," a report prepared in 1977. A meta-evaluation is, in short, an evaluation of an evaluation. A meta-evaluation should have an evaluation to which it is attached. Given that fact, I have a dilemma. "Some Still Do: River Acres, Texas" is not an evaluation. Let me use Denny's words. "My task as I saw it was to describe what people said and did about the teaching and learning of science, and mathematics from kindergarten through twelfth grade in the River Acres Independent School District, a suburban/rural setting in the Houston area. Not to evaluate it" (Denny, 1978a, p.1).

One response to that dilemma would be to thank Denny for some delightful reading matter, and to stop there. By the time I finish, you may wish I'd taken that option. I want to act for the next few minutes as if Denny's task *was* to evaluate, not to describe. I will focus on three items, three aspects of the evaluation process. I label these: (1) composing the evaluation; (2) compelling the mind; and (3) completing the value picture. I picked these three topics to structure my "meta-evaluation" of Denny's work because of my recent involvement in a project at the Northwest Regional Educational Laboratory. For quite a number of years I have invested my energies in two areas; the process of evaluation; and painting watercolor pictures. Because of accomplishments in those two areas, I was asked to contribute to a new book on alternative evaluation methods (Gephart, 1981). My first reaction to that invitation was excitement. What could be better than a chance to mix my vocation and avocation? My first efforts were extremely frustrating, so frustrating that I stopped trying. It didn't seem that I could do it. What I had accumulated was not making any progress. The materials I examined were not helping me get a better understanding of the evaluation process at all. I considered contacting the people in charge of the project with an admission that I couldn't do the job.

Before taking that step, I had a conversation with myself. In it, two questions were asked that made me turn the corner: "What are you looking at, and how does that relate to what you have done for the

past 25 years?" The answers: I was looking at the literature on aesthetic criticism, and that was *not* what I had looked at or studied for the past 25 years. I had studied painting. And more specifically, I had studied painting representational watercolor pictures. That was what I had invested my leisure time in. That was what got me the invitation to contribute to the book. Over that span of time I had had an immense amount of trial and error activity with watercolor paints. And, I had carefully examined the works of many very good watercolor painters.

Before giving up, I decided to write a description of how I paint a watercolor painting, a description of a process that starts with a decision to paint a particular subject and ends with the framing of that piece. I recognize that this is *my* way of doing a representational watercolor, not necessarily *the* way. I do not believe there is a single way. Nor do I mean that I *always* follow that particular sequence of activities in that particular order. It is, however, the set of procedures I go through most of the time.

When I completed the description of that process, I stepped back and asked myself which of these things, which of these activities, have a parallel in the evaluation process and what is the nature of that parallel? There are over twenty points in which I feel there is a parallel between the process of painting a representational watercolor and the process of doing an educational evaluation. Three of these points are important here: composing an evaluation, compelling the mind, and completing the value picture.

Composing the Evaluation

When I encounter a subject that I feel I need to paint, there are some things I must do before I put my brush to the paper. The first, and most important, is the construction of a composition. Picking a subject to paint does not automatically establish the composition. Picking a subject I want to paint is the recognition that there is some feeling about that object or scene that stirs within me, something I want to try to capture on paper. For example, a favorite subject for me is the Gulf Coast. The sand dunes, the sea oats, the sea gulls are things I enjoy very much. Have you ever had the experience of walking across a warm stretch of sugar-white sand and suddenly getting a glimpse of the waves peacefully washing the shore? There is a combination of excitement, of awe, and of peacefulness that comes together for me when I make that trek. When I become conscious of that kind of feeling, I very often have picked something I must paint. Before I can paint it, I have to design a composition that has a chance

of conveying that feeling to a viewer. The specific stretch of beach may not do it. There may be too much clutter at that particular point. The waves may be too high. There may be too many people occupying the space. I need to explore the use of the ingredients that make up that setting in such a way that it communicates that excitement, awe, and peacefulness all in one.

The work of Edgar Whitney (1974), in a book entitled *The Complete Guide to Watercolor Painting*, has been a tremendous help in structuring this necessary exploration. One of the chapters in his book deals with what Whitney describes as the eight elements of a composition: unity, dominance, conflict, repetition, alternation, balance, harmony, and gradation. The second chapter deals with seven techniques used to accomplish the eight elements. Those techniques are line, color, value, shape, texture, direction, and size. For a number of years I used those two chapters as the basis for ideas when I felt dissatisfied with composition.

On rereading Whitney's work several years ago, it seemed clear to me that an 8 x 7 matrix was being discussed. I began to recognize that the eight elements of design are variables. Any given composition has some degree of unity about it. In some compositions all of the components, all of the pieces, all of the aspects fit together and belong. In other compositions there are items that do not. And in still others everything seems to be by itself. Which of these is right? Well, all of them; it depends upon the artist's purpose. If I want to communicate a feeling of chaos, I will deliberately include things in the composition that do not appear to belong. Unity is a *dimension* of composition, a variable, and it can be used purposively by the artist. Conflict, too, is a dimension of a composition. In some pictures there is an extensive amount of conflict, while in others peacefulness is clear. For example, if I place many vertical lines in the picture space, I have taken a step toward reducing the amount of conflict in the composition. If I have equally strong lines at right angles, the potential for communicating conflict is heightened.

By constructing a matrix of the items in these two chapters, I have created the heuristic tool displayed in Figure 5.1. As I engage in the process of designing a composition, I can turn to that matrix for guidance in asking and answering questions related to the composition work. The 56 cells in that matrix represent directions of thinking that I need to consider. I do not mean to suggest that any composition has to have something that fills all 56 cells. Rather, I have a heuristic device that pushes me to think about possible solutions to the problem of composing a picture that will communicate my message to a viewer

		DESIGN PRINCIPLES							
		Unity	Conflict	Dominance	Repetition	Alternation	Balance	Harmony	Gradation
DESIGN ELEMENTS	Line								
	Value								
	Color								
	Texture								
	Shape								
	Size								
	Direction								

Figure 5.1 Design Principles and Elements

SOURCE: W. J. Gephart's adaptation of E. Whitney's chapters on Design Principles and Elements, 1977.

In the composition of an evaluation the same situation exists. There are a number of elements in the evaluation process, and there are a number of techniques that we can use to try to accomplish each of those elements. The combination of elements and techniques can be overwhelming! We need a heuristic tool that helps us keep the different points of concern in mind while we deal specifically with any single aspect.

A three-dimensional figure with multiple layers, tiers, and columns serves as such a heuristic device for me in the composition of an evaluation. Think for a moment about such a figure. Figure 5.2 pictures a cube, the three dimensions of which represent three facets of the evaluation process. When we are invited to design an evaluation, we are being asked to help an individual (or groups of individuals) to apply a set of value depicting variables to some group of options that are going to be considered. The three dimensions of the cube are, therefore, labeled with those three facets. The vertical dimension I label as options, the horizontal dimension as value depicting variables, and the third as decision maker levels.

Let me give some illustrations of these three dimensions. At the end of a test run of an educational program we very often face three

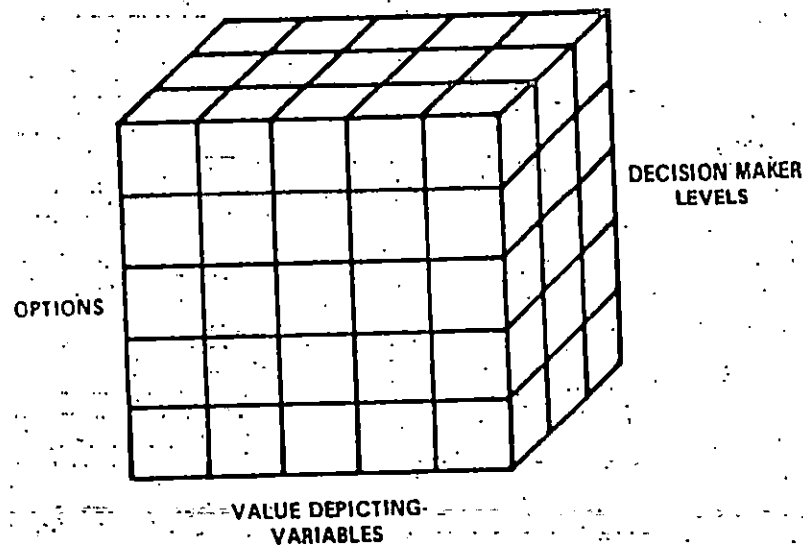


Figure 5.2 Heuristic Device for the Composition of an Evaluation

options. In some cases, the information we have gained may indicate that the program we have tested is so good that we need to build it into the ongoing program. In other instances, we may find that the program has a great deal of promise *but* is weak in this or that respect. That information suggests that we try the program again with some modifications built in. On still other occasions, we may find that the test program is a disaster. These three items are options that are faced when we engage in a trial run of an educational program. Another illustration of the options dimension might be seen in an evaluation being conducted to help in decisions about adopting a specific textbook series. The options in this case are each of the different textbook series that might be used. There is still another illustration. When we evaluate a student's work we have in mind a number of qualitative descriptors that might be used to communicate quickly with others about the quality of that work, or about the capability of that student. The options in this evaluative area are the different qualitative descriptors that might be used.

The second facet of the cube is that of the variables we find useful in depicting the value of the different options. This might include things like dollar costs, political acceptability, capability of the staff,

amount of student growth, and the like. One illustration of the decision-maker levels facet, the third dimension of our cube, can be seen in the terms "participant," "institution," and "society." These three are levels of decision making. And, when we think about decisions made at different levels, we can see slight changes in the options and subtle changes in the value depicting variables.

When I start thinking about an evaluation effort, I visualize that cube. I do not know at the outset how many options are listed on that vertical dimension or how many value depicting variables are listed on the horizontal dimension, or for that matter, how many levels of decision making may fit. My first task is to interact with the people who want the evaluation conducted in such a way that we make explicit the options to be considered, the list of variables to be used, and the decision level, or levels, toward which the effort needs to be focused. Through interactions with the person requesting the evaluation, those items are finalized; the three, five, or seven, or how many options are defined. The variables to be used in describing the value of those options are then selected, and the decision-making level specified. My next task is to structure the activities that will provide the data for each cell in the cube. The specification of those activities accomplishes the evaluation design.

Composing the Evaluation Applied to "Some Still Do: River Acres, Texas"

The first time I read through Terry Denny's paper with the idea of doing a meta-evaluation of it, I felt troubled. I could not see a heuristic tool that helped set out the options in a decision to be made by some person or group. I read it and reread it and became more troubled. The options did not show up. And then the breakthrough came. Denny, by his own words, did not do an evaluation. Rather, he described what people say and do about teaching and learning in River Acres. And it is a beautiful description. Let me share with you just a bit of it.

It pleases me to write this story without the additional burden of formally judging the merit and the teachers practice I observed. The fact that *I* was there and not *you* is of huge importance, of course.

The Study

Seeing may be believing, but I need more. I never see the picture worth a thousand words. It occurs to me that a very few words can represent a thousand pictures; can represent unobservable feelings;

can reveal tomorrow's hopes and yesterday's fears which shape today's actions. My story is largely teachers' words. Students, parents, administrators and others with something to say about River Acres contribute to its telling. But it's mostly a story of and by teachers. It wasn't supposed to turn out this way. The deeper I went, the more I needed a place to park my mind to keep it out of trouble. I found it in the teacher's words.

I began my study by looking at science instruction and then interviewing teachers and students about what they did, why they did it. Hour after hour I saw teachers working, doing what teachers always have done. Students fell into their rightful places, too. Frequently after an observation and interview, a teacher would say something in a few sentences that summed up several hours of observing for me.

Teachers would say their piece. I'd tape it or write it down, and then I'd read it back to them. They'd say, "Hey, that's a good idea. But the parents would never go for it." Or, "That sounds too radical." When I'd say, "I am going to mention it in my report," I heard, "Okay, but don't mention me." The one I liked the best was, "Okay, you be the second Joan of Arc and I'll come behind and beat the drums" [Denny, 1978a, p. 1-2].

Denny's story about River Acres, Texas schools is a very artistic piece. Its substance clearly lays out the cognitive items necessary to clearly picture mathematics, science, and social studies instruction as it goes on in River Acres schools. But it also gives you the emotional, the affective attributes of that operation as well. It is a beautifully artistic description.

But it is not an evaluation. At the end of his section on the elementary schools he says, "The elementary schools of River Acres are pleased with what they are achieving with their children" (Denny, 1978a, p.39). Denny's writing to this point does not make clear to me what specific options might be being considered nor does it give me information that weights those options differentially. As such, I do not classify it as an evaluation.

From this point, the story of River Acres moves to the two junior high schools. At the close of the section on Westland, Denny (1978a, p. 55) says, "I cannot pull all this together for the reader or myself. Westland science, mathematics, social studies are not together in the minds of the students, the faculty, or the district curriculum writers. They are separate aspects of school instruction. Best leave them that way. On to Eastland!" He closes the segment on Eastland with the statement, "There are times to observe and not to observe. It is like

learning never to buy a pair of shoes in the morning. Well, I bought out of social studies at Eastland one morning in October. A rather weak bridge to transport us to the high school? The distance is short: let's go anyway" (Denny, 1978a, p. 74). His treatment of the high school is summarized with the following, "I left River Acres hoping it would be there when I returned, knowing that it would not, and feeling that it would be better than it was. Like ice fishing, trying to explain to somebody else is ridiculous" (Denny, 1978a, p. 125).

My comments here are not critical of Terry Denny or of his report. The truth of the matter is that Terry Denny did not do an evaluation, and attempting to apply an evaluative heuristic just does not work. It is unproductive to search for options that are being dealt with when, in fact, the intent was not to identify and clarify options and to determine their relative worth. Rather, the intent was to describe.

Denny described well. I recognize that I have said that numerous times and at the risk of boring you even further, I'll repeat it again. It is a beautiful description of an educational program. I have a reason for continuing this "meta-evaluation" in the face of information that the item being evaluated is not in itself an evaluation. This set of circumstances gives me the opportunity to show that descriptive research and evaluation are different processes. As such, criteria or procedures that we use to evaluate them must differ. I cannot use the evaluative heuristic described in my earlier statements to determine the value of Denny's descriptive study. I cannot ask, "Did he compose an evaluation well?" He composed a description.

Compelling the Mind

The accomplished painter recognizes that his completed picture will have a major interest point, and numerous minor interest points. His task in composing and rendering the picture is to arrange those minor interest points and the major center of interest in such a way that the viewer's eye is compelled to move from one to the other, staying within the boundaries of the picture space. In doing so the artist is compelling the eye. There is a distinct parallel between this concept in painting and in evaluation. An evaluation typically focuses on a major decision and just as typically touches a number of minor decisions. The evaluator has the task of arranging the work on these major and minor decisions in such a way that it compels the mind to attend to the minor decisions *and* the major decision. In doing this, the treatment of those minor decisions must direct the mind towards the major decision. If the work concentrates so heavily on the minor

decisions that it reduces the concern for the major decision, the evaluative report is of lesser value. If the work is done in such a way that we cannot distinguish between the minor decisions and the major decision, the evaluation is of lesser quality.

The concern for "compelling the mind" applies in two places in the evaluation process. It must be considered as a part of the thinking when the evaluation activity is composed. It also is extremely important as a structure for designing an evaluation report. To make this point clearer, let me go back to a pictorial illustration and a painter's thinking. Here I would like to use a piece of work done by Ted Kautsky, one of the outstanding watercolorists in the United States during the last century. Kautsky did a painting entitled "Motif Number One, The Wharf and Fishing Shanty at Brockport, Massachusetts." His major point of interest was a graceful sea gull silhouetted against the dark, dark shadows of that wharf. That center of interest, however, created a major design problem. The lines of the wharf in perspective were very strong, and they moved the eye out of the picture frame very, very rapidly. When a viewer's eye leaves the picture frame the sensory input from the painting ceases. Kautsky did not want the viewer's eye to leave his picture that rapidly. By adding more detail and attention to the fishing shed at the end of the wharf, he slowed that movement down somewhat. But the diagonal line at the roof of the building paralleled the diagonal lines of the wharf. Again, the eye was propelled out of the picture frame. He tried to place a number of small boats out in the water. Their vertical masts served as an interception of that diagonal movement. But the intensity of those masts, necessary to convince the viewer that they are in the distance, was not strong enough to stop the directing of the eye out of the picture frame. He tried strengthening the foreground as a way of stopping the rapid exit from the picture frame. That, too, added more minor interest points, but did not set an arrangement in which the eye moves from minor interest point to minor interest point, to the major, to minor interest point, minor interest point, to the major. His design was complete when he added in the foreground two very strong, dark posts that might have been holding up a wharf or pier that no longer exists. The addition of these two darks in the foreground served to intercept the eye as it is being propelled out of the picture by the diagonals of the wharf. They served to bring the eye back to the foreground, to the rocks in the lower right foreground, and then back to the major point of interest, that graceful sea gull silhouetted in the darkness of the shadow under the wharf. In this picture your eye can move from item to item for an extensive amount of time and not be

propelled out of the picture frame. It is composition through which Kautsky compels the eye to stay with the scene.

Evaluation has a similar problem. We start with concern for a major decision. We try to identify the options inherent in it, the value descriptive variables that will help us determine its worth, and the individual or group that will make the final choice. Permit me to use the textbook adoption situation as an illustration. In states that have a state-approved textbook adoption process, a committee at the state level decides that textbook series A, B, C, D, and E are acceptable for use in the schools and that the individual school or school system can make the choice as to which of these to include in their work. At first, the options seem to be the different textbooks. In such an evaluation task the major decision is "Which textbook series shall we use?" There are some minor decisions, one of which shows up very quickly. That minor decision is, "Shall we use one textbook series universally in our school or shall we use some combination?" Another minor decision is "What combination?" To respond to the evaluation request, a request that a school be helped in deciding which textbooks to purchase, we must attend to those minor decisions as well as the major one. It is extremely easy to get caught up in the response to one of the minor decisions and not serve the major. It is also extremely easy to try to attend only to the major decision and not help the school with those minor decisions. Either case creates an evaluation that is less valuable than it might be.

I can apply the concept of compelling the mind to Terry Denny's study. His report compelled my mind. I was fascinated. I stayed with the reading to the point that I felt I knew that school system. I felt I knew the people. I could close my eyes and see individuals with whom I worked express the same kinds of interest, frustration, satisfaction, and so on. Terry Denny did compel my mind through expressive skills that are, in my perception, quite rare. Most of us are incapable of writing a description that has the captivating quality of this man's work.

When I was studying Denny's work with the idea that I was doing an evaluation of his evaluation, the concept of compelling the mind was problematic. I was searching for elements in his presentation that would direct me from minor decision to minor decision to major decision. I did not find them. My relief, when I finally decided that we could not call Denny's report an evaluation, was large. At the same time, that recognition made me think again; the evaluation and research processes are markedly different in their application, and the criteria that we use for appraising one cannot be used identically in the

appraisal of another. The concept of compelling the mind, I believe, goes across processes, but its application is considerably different when I use it in relation to a research study than in relation to an evaluation.

Completing the Value Picture

Most artists do not strive to duplicate the subject they are painting. Photographic realism is an approach, but not one used by most artists. Usually, the camera used artfully on a subject achieves photographic realism better than does a painting medium. Most artists try only to suggest the basic elements of the item being painted. In so doing, we call upon the viewer to complete the details of that item.

This technique is used extensively by watercolorists. It is extremely difficult to control watercolors. When an oil painter touches a loaded paint brush to a canvas, the paint stays where that touch occurred. Watercolors, however, work with the surface. Capillary action and gravity cause water to move from one point to another. Because of these factors, it is much more difficult for a watercolorist to get exacting detail.

Many watercolorists have learned to use this difficulty to advantage. That advantage accrues from perceptual differences. It is well known that if each of us looks at the same tree, the perceptual image we form will have distinct differences. If, as a watercolor painter, I portray the basic ingredients of a particular kind of tree, an elm tree for example, or a pine, those essential characteristics will be used by a viewer to complete that picture in his or her mind. Again, a piece of work by Ted Kautsky is helpful. In this picture Kautsky presents some trees, several rather tall ones and a shorter one. A background of a stand of trees off in the distance provides a dark area. When a viewer examines that scene, he or she completes a mental picture of those trees. For example, that smaller foreground tree is defined with considerable detail in each viewer's mind. You know that it is a small pine tree. Its general shape, its bark, its needles are all there. If we enlarge that picture and focus just on that small pine tree, it is rather startling to notice that the trunk of the tree is simply white paper surrounded by a dark background, resulting in the outline of a tree trunk. That is not the image that most of us will have formed when we looked at it as a part of the overall composition.

I said that many watercolorists have learned to use this difficulty to advantage. If I had painted the absolute details of that small pine tree and presented it to a number of viewers, because of perceptual

differences it would be an inaccurate presentation for many of them. By carrying out the painting in a way that gives the viewer the opportunity to complete the picture, it becomes an accurate representation for each viewer.

Completing the picture occurs in evaluation as well, but instead of perception, we are dealing with valuing. The selection of an option from some set of options is basically a valuing activity. Each of us has a value base which we use in deciding to pick option A or options B, C, or D. Just as there are perceptual differences when people look at an object, there are value differences when people consider choosing one item over others. An evaluator's task is not to present *the* singular absolute description of the value of those options, but rather to identify the ingredients used by the decision makers in their valuing of those options and to create the circumstances in which the decision makers will use those ingredients systematically in their valuing.

The application of this point to Terry Denny's paper has the same message as does the application of the two earlier ones. I can apply the idea of completing the value picture to Denny's work. Through the statements of teachers, he has given me the basis for completing the picture. But Denny's task was descriptive. It was not his task to push me to differentially value a set of items. Rather, his task was to describe how the people there value those items. Again, I think he did that and did it masterfully.

Completing the value picture, I believe, asks us to change markedly our approach to evaluation. Systematic evaluation has evolved out of an environment in which the research process is revered. One of the basic assumptions of the research process is the idea that truth is external to the individual and that if we can but perfect our measuring tools, we eventually will know that truth. When we try to determine the relative value of a number of options, we are in a different ball game. The assumption that there is a single way to value the different options *does not hold us*. The worth of items A, B, and C will not always be the same from person to person. If we, as evaluators, employ a process that assumes that there is a singular set of value descriptors for the options on which we focus, we will not serve our evaluation clients well. We need to employ a process that assumes that value perspectives are multiple. In doing so, our work, particularly our reporting, needs to push the decision makers to as thorough and broad a valuing activity as possible. It is extremely rewarding to me to see that some of the newer models of the evaluation process are moving in this direction.

**BUKUPANDUAN PENILAIAN HASIL BELAJAR
DI IKIP PADANG**

Oleh:

**DR. Phil. Yanuar Kiram
FPOK IKIP Padang**

**Makalah disajikan pada Seminar Penyusunan
Buku Panduan Penilaian Hasil Belajar IKIP Padang
tanggal 22 dan 23 Januari 1990**

Prakata

Dengan segala keterbatasan yang dimiliki, baik yang berhubungan dengan ilmu pengetahuan maupun yang berkaitan dengan pengalaman, penulis telah berusaha untuk menuangkan buah pikiran kedalam makalah ini dengan harapan dapat memberikan sumbangan pikiran pada Panitia Ad-hoc penyusunan buku panduan penilaian hasil belajar mahasiswa IKIP Padang.

Sadar akan segala keterbatasan dan kekurangan-kekurangan yang penulis miliki, maka penulis yakin sepenuhnya, bahwa makalah ini masih jauh dari kesempurnaan yang diinginkan. Oleh karenanya, dengan segala kerendahan hati penulis mengharapkan tegur sapa, kritik dan saran-saran yang membangun dari para peserta seminar yang terhormat, sehingga memungkinkan penulis dapat belajar lebih banyak dan menimba pengalaman yang bermanfaat, baik untuk perbaikan makalah ini, maupun dalam rangka perbaikan penulisan pada masa-masa yang akan datang.

Atas segala sumbangan pikiran dari para peserta seminar, terlebih dahulu penulis menghaturkan terima kasih yang setulus-tulusnya. Semoga pokok-pokok pikiran yang tertuang dalam makalah ini dapat memberikan sedikit sumbangan terhadap pengelolaan proses belajar mengajar di IKIP Padang.

Terima kasih yang setulus-tulusnya juga penulis sampaikan kepada penyumbang pikiran, sehingga memungkinkan terwujudnya makalah yang sederhana ini.

Akhirnya, penulis mengucapkan terima kasih kepada Panitia seminar yang terhormat yang telah memberikan kesempatan pada penulis untuk dapat menimba ilmu dan pengalaman dalam seminar ini.

Padang, 18 Januari 1990

PENULIS

BEBERAPA SARAN TENTANG PENILAIAN HASIL BELAJAR PERKULIAHAN PRAKTEK OLAAHRAGA DI FPOK IKIP PADANG

Oleh. Dr. Phil. Yanuar Kiram
FPOK IKIP Padang

LATAR BELAKANG MASALAH:

Perencanaan, Pelaksanaan dan Penilaian terhadap hasil belajar adalah merupakan tiga komponen utama yang tidak hanya harus tergabung dalam suatu sistem pendidikan secara keseluruhan, melainkan juga harus dapat saling mendukung satu dengan yang lain dalam suatu proses secara berkelanjutan, dalam rangka mencapai tujuan pendidikan atau dalam rangka mencapai hasil belajar yang optimal. Pencapaian tujuan pendidikan atau pencapaian hasil belajar secara optimal hanya akan dapat dimungkinkan bila ketiga komponen utama tersebut dikelola dengan teliti, matang dan profesional. Ungkapan ini menunjukkan bahwa tidak satupun diantara ketiga komponen tersebut yang dapat diabaikan di dalam mengelola system pendidikan secara keseluruhan, atau dengan kata lain bahwa kelemahan pada salah satu komponen tersebut akan dapat mengakibatkan kegagalan dari usaha-usaha pendidikan yang dilakukan.

Penilaian hasil belajar memiliki kedudukan yang amat penting dalam system pendidikan. Bahkan penilaian hasil belajar mempunyai peranan yang penting dan menentukan baik dalam perbaikan, peningkatan maupun dalam pengembangan system pendidikan secara keseluruhan, karena penilaian hasil belajar akan memberikan seperangkat informasi atau data-data dari usaha-usaha yang dilakukan di dalam proses belajar mengajar yang dapat diolah, dianalisa dan diinterpretasikan kedalam berbagai aspek yang terlibat dalam pengelolaan proses belajar mengajar. Dari hasil pengolahan, analisa dan interpretasi informasi atau data-data yang diperoleh dari penilaian hasil belajar, akan didapatkan gambaran atau kenyataan tentang proses belajar mengajar yang diselenggarakan. Kemudian dari

kenyataan-kenyataan yang ditemukan tersebutlah seharusnya seorang tenaga pengajar merencanakan selanjutnya baik dalam perbaikan, peningkatan maupun pengembangan system pendidikan atau proses belajar mengajar.

Melihat pentingnya kedudukan dan peranan penilaian hasil belajar terhadap system pendidikan secara keseluruhan, maka penilaian tersebut haruslah dilakukan secara teliti dan profesional dengan memperhatikan prinsip-prinsip ilmiah. Oleh karenanya melakukan penilaian terhadap hasil belajar bukanlah suatu hal yang mudah, tetapi adalah merupakan suatu hal yang kompleks sulit dan rumit. Kesulitan dan kerumitan tersebut juga akan ditemukan di dalam melakukan penilaian hasil belajar keterampilan olahraga seperti yang di selenggarakan oleh FPOK IKIP Padang, sebagai salah satu lembaga kependidikan yang mengemban tugas untuk mempersiapkan tenaga-tenaga yang terampil dan terdidik sebagai tenaga pengelola dan pelaksana pendidikan olahraga di sekolah.

Belajar keterampilan olahraga memiliki karakter-karakter, khusus, bila dibandingkan dengan bidang-bidang ilmu pengetahuan yang lain seperti matematika, kesenian, bahasa dsb. Pada umumnya pada bidang-bidang ilmu yang lain aspek yang dominan dalam kegiatan belajar adalah kognitif dan Affektif, sedangkan di dalam belajar ketrampilan olahraga di samping kedua aspek tersebut, maka kemampuan dan ketrampilan motorik adalah merupakan aspek yang dilatih, dipelajari dan dinilai. Oleh karena FPOK- IKIP Padang mempersiapkan tenaga-tenaga terdidik dan terlatih untuk melaksanakan pendidikan olahraga di sekolah (guru olahraga), maka penilaian terhadap hasil belajar ketrampilan olahraga tidak hanya dilakukan pada aspek kemampuan dan ketrampilan motorik saja, tetapi juga dilakukan terhadap aspek kognitif dan efektif.

Penilaian terhadap aspek motorik bertujuan antaran lain untuk mengetahui hal-hal sebagai berikut:

- Sejauh mana individu yang belajar telah menguasai kemampuan-kemampuan dan ketrampilan-ketrampilan motorik yang sudah diajarkan dan sekali gus untuk mengetahui sejauh

mana ketrampilan dan kemampuan-kemampuan motorik yang masih belum tercapai.

- Penilaian terhadap aspek motorik juga bertujuan untuk mengumpulkan data dalam rangka memberikan nilai hasil belajar (Rothig, 1983, hal 232).

Landasan berpikir digunakan di dalam penilaian kemampuan motorik adalah: bahwa seorang guru olahraga harus dapat memberikan contoh-contoh atau mendemonstrasikan gerakan-gerakan yang akan dipelajari oleh para individu yang belajar. Pendemonstrasian gerakan oleh guru, adalah merupakan salah satu cara pemberian informasi tentang gerakan-gerakan yang akan dipelajari (Yanuar Kiran, 1989, 127). Semakin baik dan semakin lengkap serta semakin jelas seorang guru olahraga dapat mendemonstrasikan gerakan, maka semakin baik dan semakin lengkap pula informasi-informasi tentang gerakan yang akan dipelajari dapat diserap oleh individu-individu yang belajar. Semakin benar gerakan-gerakan yang di demonstrasikan akan semakin memungkinkan para individu yang belajar dapat menyimpan bentuk-bentuk gerakan yang benar kedalam ingatannya. Bentuk-bentuk gerakan yang benar yang tersimpan didalam ingatannya adalah merupakan salah satu modal untuk dapat mempelajari ketrampilan olahraga dengan baik dan sekaligus merupakan bantuan yang sangat positif sekali bagi yang belajar untuk mempelajari gerak-gerak tersebut. Oleh karenanya guru-guru olahraga dituntut untuk dapat mampu menguasai kemampuan dan ketrampilan-ketrampilan motorik dengan baik. Untuk dapat mengarahkan dan mengontrol proses belajar-mengajar tersebut, maka FPOK- IKIP Padang telah menuangkan tingkat kemampuan ketrampilan olahraga yang harus dicapai oleh mahasiswa FPOK- IKIP Padang kedalam silabi (lihat silabi FPOK- IKIP Padang, 1987).

Dari sisi lain dapat diartikan bahwa seorang guru olahraga adalah merupakan salah satu pusat informasi bagi anak didik, terutama tentang berbagai aspek ilmu pengetahuan olahraga yang akan diajarkan. Dalam hal ini, pertanyaan yang harus dijawab antara lain adalah: Kemampuan apakah yang harus

dimiliki oleh seorang guru olahraga sebagai nara sumber dari anak didiknya di dalam proses belajar mengajar. Untuk menjawab pertanyaan ini di atas, ada dua bentuk kemampuan yang dapat dikemukakan. Pertama adalah kemampuan di dalam penguasaan ilmu pengetahuan yang akan diajarkan dan kedua adalah kemampuan di dalam menginformasikan atau menyajikan ilmu pengetahuan tersebut kepada anak didik. Kemampuan penguasaan ilmu pengetahuan yang di maksudkan di sini adalah meliputi penguasaan ilmu pengetahuan olahraga yang akan diajarkan dan ilmu-ilmu penunjang lainnya. Sedangkan kemampuan menginformasikan atau menyajikan ilmu pengetahuan, meliputi aspek-aspek yang berhubungan dengan didaktik dan metodik. Oleh karenanya penyelenggaraan proses belajar mengajar di FPOK- IKIP Padang tidak hanya diberikan dalam bentuk praktek tetapi juga diberikan dalam bentuk teori. Selanjutnya, untuk dapat mengarahkan dan memperbaiki proses belajar mengajar, khususnya di bidang teori, maka diperlukan pengontrolan terhadap hasil belajar. Pengontrolan tersebut hanya mungkin dapat dilaksanakan bila terdapat informasi-informasi ataupun data-data tentang proses belajar mengajar tersebut untuk mendapat data-data tersebut, maka diperlukanlah penilaian ataupun tes.

Penilaian terhadap hasil belajar teori ini adalah bertujuan, antara lain: untuk mengetahui sejauh mana teori-teori baik yang menyangkut aspek-aspek ilmu pengetahuan olahraga maupun yang berhubungan dengan didaktik dan metodik, yang dikuasai oleh mahasiswa. Dari hasil tes dan penilaian itulah nanti akan dapat dijadikan pedoman untuk dapat menetapkan hasil belajar mahasiswa. Mampu bekerjasama dengan individu atau dengan kelompok lain, mampu menempatkan diri sesuai dengan apa yang dimilikinya dalam kelompok lain, mampu berkomunikasi, berinteraksi dan berintegrasi dengan individu ataupun dengan kelompok lain, adalah merupakan aspek-aspek yang tergolong dalam komponen affektif di dalam proses belajar mengajar keterampilan olahraga. Komponen afektif ini adalah juga merupakan faktor penting dalam rangka membentuk

kepribadian seorang calon guru olahraga. Oleh karenanya aspek ini juga harus mendapat perhatian di dalam penetapan hasil belajar mahasiswa.

Dari uraian-uraian di atas, maka telah tergambar secara umum bentuk kemampuan yang harus dimiliki oleh seorang mahasiswa FPOK IKIP Padang yang dilihat dari tiga aspek yang terlibat di dalam proses belajar mengajar keterampilan olahraga yaitu aspek motorik, kognitif dan afektif. Lalu sekarang timbul pertanyaan, bagaimanakah menilai bentuk kemampuan yang dilihat dari ketiga aspek tersebut. Penilaian hasil belajar mahasiswa FPOK IKIP Padang telah dilaksanakan dengan melibatkan ketiga aspek tersebut di atas. Namun demikian masih terdapat perbedaan dari segi pembobotan penilaian terhadap aspek motorik, kognitif dan afektif. Sebahagian dosen ada yang memberikan bobot penilaian lebih besar pada aspek motorik dari pada kognitif, sementara staf pengajar yang lain menuntut kemampuan kognitif lebih besar dari kemampuan motorik. Contoh yang lebih konkrit misalnya; salah satu perkuliahan praktek olahraga memberikan pembobotan 60% untuk kognitif dan 40% untuk motorik, sementara perkuliahan praktek yang lain memberikan bobot penilaian 40% untuk motorik dan 60% untuk kognitif. Bahkan ada yang memberikan perbedaan pembobotan yang lebih besar yaitu 75% motorik dan 25% kognitif. Dari ketiga aspek tersebut, maka aspek afektif mendapatkan pembobotan yang terkecil. Dari hasil observasi dan wawancara yang penulis lakukan terhadap beberapa orang staf pengajar FPOK IKIP Padang dalam rangka penulisan makalah ini, maka dapat penulis simpulkan beberapa hal sbb:

- Belum ada patokan yang permanen di dalam menetapkan pembobotan penilaian terhadap ketiga aspek tersebut.
- Masih terdapat beberapa perbedaan alasan-alasan yang dikemukakan di dalam memberikan bobot penilaian.

Beberapa staf pengajar yang memberikan pembobotan yang lebih besar terhadap aspek kognitif mengemukakan beberapa alasan sbb:

- FPOK IKIP Padang mendidik calon guru dan bukan atlit. Oleh karenanya kemampuan kognitif lebih dituntut dari kemampuan motorik. Alasan ini mempunyai latar belakang pemikiran, bahwa seorang guru olahraga harus mampu meberikan bimbingan kepada peserta didik, harus mampu menyajikan materi pelajaran dengan baik, harus mampu menganalisa dan mengoreksi kesalahan-kesalahan yang terjadi di dalam pelaksanaan gerakan dan harus mampu memberikan alasan-alasan ilmiah tentang aspek-aspek yang menyangkut dengan keolahragaan itu sendiri. Untuk itu diperlukan penguasaan teori-teori olahraga baik yang berhubungan dengan olahraga secara umum, didaktik dan metodik, maupun yang menyangkut dengan ilmu olahraga itu sendiri.
- Alasan lain yang dikemukakan oleh beberapa staf pengajar FPOK IKIP Padang yang memberikan pembobotan nilai lebih besar pada aspek kognitif adalah: bahwa di dalam memberikan perkuliahan teori satu jam di kelas, mahasiswa akan dapat memperoleh ilmu pengetahuan lebih banyak bila dibandingkan dengan satu jam kuliah praktek olahraga yang lebih dititikberatkan pada pengulangan-pengulangan latihan gerakan. Dengan demikian mahasiswa lebih banyak menerima ilmu pengetahuan olahraga melalui perkuliahan teori.

Dari beberapa orang staf pengajar FPOK IKIP Padang yang memberikan pembobotan nilai lebih besar pada aspek motorik mengemukakan alasan antara lain:

- Bahwa inti aktivitas olahraga adalah gerak atau kemampuan gerak. Seorang guru olahraga yang memilih kemampuan motorik yang baik akan dapat mendemonstrasikan gerakan dengan baik. Pendemontrasian gerakan yang baik akan lebih membantu peserta didik dalam memahami bentuk-bentuk keterampilan olahraga yang akan dipelajari.
- Alasan lain yang dikemukakan adalah: bahwa proses belajar keterampilan olahraga mahasiswa FPOK secara systematis dan

intensif hanya berlangsung selama masa perkuliahan di FPOK. Setelah mereka jadi guru olahraga di sekolah mereka akan disibukkan oleh tugas-tugas mengajar. Oleh karenanya selama masa perkuliahan praktek olahraga di FPOK, mereka harus digembleng sedemikian rupa, sehingga memiliki kemampuan dan keterampilan olahraga yang baik.

- Alasan berikutnya yang dikemukakan adalah: bahwa kemampuan fisik untuk dapat menguasai kemampuan dan keterampilan olahraga sangat terbatas. Keterbatasan tersebut dipengaruhi antara lain oleh faktor usia. Sedangkan kemampuan penguasaan ilmu pengetahuan secara teoritis (kognitif) masih dapat diteruskan pada usia yang lebih lanjut.

Dari kedua kelompok staf pengajar FPOK IKIP Padang yang berbeda di dalam pembobotan penilaian hasil belajar mahasiswa dengan alasan-alasan seperti yang dikemukakan pada bagian terdahulu, maka sangat sulit untuk menetapkan mana yang harus diterapkan didalam memberikan penilaian hasil belajar mahasiswa. Hal ini disebabkan karena alasan-alasan yang dikemukakan yang dapat diterima. Keadaan tersebut masih berlangsung sampai sekarang pada saat makalah ini diturunkan. Oleh karenanya seminar ini akan memberikan manfaat yang sangat besar sekali untuk pemecahan masalah yang sedang dihadapi, melalui pemberian masukan-masukan oleh peserta seminar yang terhormat.

Beberapa Permasalahan di dalam Penilaian Hasil Belajar mahasiswa di FPOK IKIP Padang

Untuk melakukan penilaian terhadap hasil belajar mahasiswa IKIP Padang telah menetapkan dengan melakukan dua model pendekatan, yaitu penilaian acuan norma (PAN) dan penilaian acuan patokan (PAP) (buku pedoman IKIP Padang 1989-1990, hal 39). Dalam melakukan penilaian terhadap hasil belajar mahasiswa, FPOK IKIP Padang juga menggunakan dua model pendekatan penilaian tersebut. Untuk perkuliahan-perkuliahan yang hanya bersifat teori (tidak ada praktek

lapangan, dan praktek labor), penggunaan dua model pendekatan penilaian tersebut tidak menemukan kesulitan yang berarti. Kesulitan di dalam memberikan penilaian baru ditemui dalam memberikan penilaian hasil belajar mahasiswa pada perkuliahan-perkuliahan yang bersifat praktek dan teori atau istilah khusus dalam perkuliahan olahraga adalah kognitif dan motorik. Misalnya perkuliahan bulu tangkis, bola basket, sepak bola, atletik dan sebagainya. Perkuliahan-perkuliahan ini menuntut beberapa bentuk penguasaan yaitu:

- penguasaan kemampuan dan keterampilan motorik, yaitu kemampuan untuk melakukan/melaksanakan teknik-teknik gerakan yang dituntut pada setiap cabang olahraga yang dipelajari.
- penguasaan kemampuan bermain
- penguasaan kemampuan pemecahan tugas gerakan yang dituntut. (Ketiga bentuk kemampuan diatas adalah berbentuk kemampuan atau keterampilan motorik).
- penguasaan ilmu pengetahuan setiap cabang olahraga yang dipelajari misalnya, sejarah, organisasi dan peraturan pertandingan.
- penguasaan teori setiap cabang olahraga, misalnya pengetahuan tentang prinsip-prinsip pelaksanaan gerakan, menemukan kesalahan-kesalahan gerakan, mendiagnosa kesalahan dan melakukan koreksi-koreksi terhadap kesalahan yang terjadi pada unjuk kerja yang diperlihatkan.
- penguasaan ilmu pengetahuan didaktik dan metodik (bentuk-bentuk penguasaan yang terakhir ini adalah bersifat kognitif).

Bila pengertian hasil belajar dipandang sebagai tingkat atau derajat penguasaan bahan kuliah oleh mahasiswa, maka bahan kuliah yang dimaksud pada perkuliahan-perkuliahan yang bersifat praktek olahraga adalah tingkat penguasaan motorik dan kognitif. Pertanyaannya sekarang adalah bagaimana menggabungkan nilai kedua aspek tersebut didalam penilaian hasil belajar mahasiswa. Jelasnya pertanyaan ini adalah: berapa prosentasekah sebaiknya aspek motorik kognitif dan

afektif yang dinilai. Pertanyaan ini juga dapat diformulasikan kebentuk pertanyaan yang lain yaitu: manakah yang lebih penting antara kemampuan teori/kognitif dengan kemampuan keterampilan olahraga/motorik. Pertanyaan ini akan dapat lebih diperjelas lagi dengan mengajukan contoh yang sederhana sbb:

Manakah yang akan kita pilih dalam menghadapi dua alternatif dibawah ini.

- seorang guru olahraga A memiliki ketrampilan olahraga yang baik, tetapi kurang bisa menyajikan pelajaran dengan baik.
- seorang guru olahraga B, kurang memiliki keterampilan olahraga yang baik, tetapi mampu menyajikan pelajaran dengan baik.

Permasalahan lain yang penulis temukan dari hasil observasi dan wawancara dengan staf pengajar di FPOK IKIP Padang adalah penerapan Penilaian Acuan Patokan (PAP). Penerapan penilaian dengan melakukan penilaian dengan melakukan pendekatan melalui model PAP berarti seorang staf pengajar telah menentukan terlebih dahulu patokan-patokan yang menjadi pedoman pemberian nilai. IKIP Padang misalnya menetapkan patokan tingkat penguasaan materi perkuliahan yang telah disajikan sbb:

Derajat Penguasaan	Nilai
90% - 100%	A
80% - 89%	B
65% - 79%	C
55% - 64%	D
< 55%	E

(Sumber : Buku Pedoman IKIP Padang 1989-1990, hal 39).

Salah satu landasan berfikir di dalam menggunakan model PAP adalah: "untuk mengerjakan sesuatu" dibutuhkan kemampuan yang relevan dengan kualitas tertentu" (Dept-Dikbud, Universitas Terbuka, 1984, 1985, Buku II : Modul Evaluasi Hasil Belajar,

hal 27). Bila ungkapan ini dianalisa, maka dapat ditarik suatu kesimpulan yang sederhana yaitu: "untuk dapat menjadi guru olahraga, dibutuhkan kemampuan yang berkualitas sesuai dengan tugasnya". Menuntut kemampuan tertentu dengan memberikan patokan tertentu di dalam proses belajar mengajar adalah sangat baik. Tetapi perlu dipertimbangkan apakah sarana dan fasilitas serta pelayanan mengajar yang dibutuhkan untuk mendapatkan kualitas kemampuan tertentu telah tersedia atau terlaksana sebagaimana mestinya. Dalam hal ini dapat dikemukakan mata kuliah atletik dasar I di FPOK IKIP Padang sebagai contoh. Perkuliahan atletik dasar I "meliputi pengetahuan dan pemahaman mengenai sejarah, peraturan perlombaan, administrasi, organisasi, gerakan tehnik dasar lari 100 m, 400 m, Lompat Jauh, Lempar Cakram, serta dapat mengajarkannya" (FPOK IKIP Padang 1987, Silabi Mata Kuliah Jurusan Pendidikan Olahraga SI hal, 102). Mata kuliah ini memiliki materi yang sangat padat dan luas. Dalam perkuliahan ini, mahasiswa dituntut tidak hanya menguasai teori-teori yang luas, tetapi juga dituntut untuk menguasai kemampuan-kemampuan dan keterampilan motorik yang kompleks dari 4 nomor atletik yaitu lari motorik yang kompleks dari 4 nomor atletik yaitu lari 100 m, 400 m, lompat jauh dan lempar cakram. Sedangkan jumlah pertemuan secara teoritis hanya 18 kali. Pertanyaannya sekarang adalah sbb:

- apakah jumlah 18 kali pertemuan tersebut telah memungkinkan mahasiswa dapat menguasai tehnik-tehnik gerakan yang baik dari 4 cabang olahraga tersebut ? apa lagi ditambah dengan teori.
- apakah peralatan yang dibutuhkan sudah menunjang proses belajar mengajar atau proses penguasaan keterampilan olahraga ?

Pertanyaan-pertanyaan di atas dapat diformulasikan sebagai berikut: apakah dengan keadaan yang digambarkan di atas, mahasiswa mampu mencapai target atau patokan yang telah ditentukan ?

Dari informasi-informasi yang diperoleh sementara ini tentang penerapan PAP pada beberapa mata kuliah praktek, terdapat beberapa keluhan dari mahasiswa misalnya "tidak lulus" atau "nilai rendah. Akibatnya, timbul beberapa asumsi sbb:

- Penetapan PAP untuk beberapa mata kuliah cenderung hanya menuntut kualitas tertentu dari hasil belajar, tetapi mengabaikan proses penyampaian kemampuan tertentu, yang dipengaruhi oleh banyak faktor.
- Asumsi kedua adalah : materi perkuliahan terlalu luas dan padat, sehingga penerapan PAP untuk menilai hasil belajar mahasiswa menemui masalah.

Sebahagian besar perkuliahan praktek olahraga di FPOK IKIP Padang dibagi atas 3 tingkatan sbb:

- Tingkat Dasar, yaitu penguasaan tehnik-tehnik dasar dari cabang olahraga yang diajarkan. Kemampuan yang dituntut adalah kemampuan koordinasi kasar.
- Tingkat pendalaman. Kemampuan yang dituntut adalah penguasaan keterampilan motorik pada cabang olahraga yang diajarkan secara halus.
- Tingkat spesialisasi. Kemampuan yang dituntut adalah pemantapan dan penstabilan kemampuan dan keterampilan motorik dari beberapa cabang olahraga yang diambil oleh tingkat spesialisasi ini mahasiswa tidak hanya dituntut untuk mampu mengajar dengan baik, tetapi juga dituntut untuk mampu melatih dengan baik.

Permasalahan dalam penilaian untuk setiap tingkatan di atas adalah sama seperti pada masalah terdahulu, yaitu kesulitan di dalam pembobotan antara penilaian teori/kognitif dan praktek/motorik. Perbedaannya adalah, pada bahagian terdahulu bersifat pembobotan secara umum, sedangkan sekarang pembobotan untuk setiap tingkat belajar perkuliahan praktek. Jelasnya adalah: bagaimanakah pembobotan penilaian yang layak untuk setiap tingkatan ?

Beberapa Saran untuk Pemecahan Masalah:

Seperti yang telah diuraikan pada bahagian terdahulu, bahwa terdapat ketidakseragaman para staf pengajar di FPOK IKIP Padang dalam pemberian bobot penilaian hasil belajar praktek olahraga mahasiswa FPOK IKIP Padang. Ketidakseragaman tersebut bersumber dari pandangan-pandangan dan alasan-alasan yang berbeda terhadap bentuk dan tingkat kemampuan yang harus dimiliki oleh tamatan FPOK IKIP Padang yang dihubungkan dengan tugas-tugas yang dihadapinya sebagai guru olahraga di sekolah.

Selubungan dengan adanya ketidakseragaman tersebut penulis akan coba mengemukakan beberapa saran-saran untuk penilaian hasil belajar keterampilan olahraga mahasiswa FPOK IKIP Padang yang kiranya saran-saran ini dapat memperkecil perbedaan-perbedaan pandangan tersebut. Jalan tengah yang penulis ajukan disini adalah sbb:

1. Bagi mata kuliah praktek olahraga yang hanya ditawarkan 1 semester (tidak ada pendalaman dan spesialisasi) maka pembobotan penilaian hasil belajar adalah sbb: kemampuan motorik diberi bobot penilaian 40% dan kemampuan kognitif 60%. Sedang penilaian hasil belajar dilakukan melalui pendekatan model Penilaian acuan norma (PAN).-Landasan berpikirnya adalah sbb: untuk dapat menguasai bentuk-bentuk keterampilan motorik olahraga dengan baik, diperlukan pengulangan-pengulangan latihan yang systematis, terarah dan kontiniu. Pencapaian hasil belajar keterampilan motorik yang baik sangat ditentukan oleh waktu yang tersedia. Pada mata kuliah praktek olahraga yang hanya ditawarkan untuk 1 semester dengan jumlah pertemuan sebanyak 18 kali termasuk di dalamnya beberapa kali pertemuan teori, maka tidak memungkinkan bagi mahasiswa untuk dapat menguasai keterampilan-keterampilan motorik yang baik. Sedangkan dengan bekal perkuliahan 1 semester tersebut, mahasiswa dituntut untuk mampu mengajarkan cabang olahraga yang dipelajarinya. Oleh

karena dalam hal ini, kemampuan menginformasikan atau penguasaan kemampuan mengajar haruslah merupakan titik berat kegiatan belajar mengajar. Untuk itu perkuliahan sebaiknya diarahkan lebih banyak pada penguasaan kemampuan teoritis baik yang berhubungan dengan pengetahuan tentang tehnik-tehnik gerakan yang diajarkan, maupun yang berkaitan dengan didaktik dan metodik. Karena dengan jumlah pertemuan yang tidak memadai dalam satu semester untuk dapat menguasai ketrampilan motorik dengan baik, maka kekurangan izin harus diimbangi dengan kemampuan teoritis yang memadai. Untuk itulah pembobotan penilaian aspek kognitif lebih besar (60%) dari pada aspek motorik (40%). Untuk dapat mengarah pada pencapaian tujuan tersebut, maka tentulah penguasaan pengetahuan didaktik dan metodik, penguasaan dalam hal analisa, diagnosa dan koleksi-koleksi gerakan serta pemahaman terhadap tehnik-tehnik gerakan serta pemahaman terhadap tehnik-tehnik gerakan yang diajarkan lebih ditekankan di dalam kegiatan belajar mengajar.

Alasan memakai pendekatan PAN dalam penilaian juga didasarkan pada pemikiran yang sama pada uraian terdahulu, karena adalah merupakan suatu kesulitan bagi mahasiswa untuk dapat mencapai patokan-patokan penilaian tertentu dengan jumlah pertemuan yang terbatas dan materi yang terbagi dua yaitu teori dan praktek. Penerapan PAP pada situasi dan kondisi belajar seperti yang diuraikan di atas akan cenderung menghasilkan nilai-nilai yang rendah. Nilai yang rendah tersebut belum tentu diakibatkan oleh ketidakmampuan mahasiswa dalam belajar, tetapi juga bisa disebabkan karena terbatasnya waktu yang tersedia, atau sarana dan prasarana yang tidak memadai.

2. Bagi mata kuliah yang ditawarkan dalam bentuk dasar maka pembobotan penilaian juga disarankan seperti di atas, yaitu 60% untuk kognitif/teori dan 40% motorik dengan penerapan PAN sebagai model penilaian. Landasan pemikiran dalam hal ini adalah pase-pase perkembangan belajar

keterampilan motorik. MEINEL (1977, 282) mengemukakan ciri-ciri kemampuan motorik seseorang yang berada pada tingkat belajar pertama/"dasar" antara lain adalah sbb:

- belajar pada tingkat dasar di tandai dengan penguasaan kemampuan koordinasi gerakan secara kasar.
- Gerakan-gerakan yang dituntut hanya dapat dilakukan pada situasi dan kondisi yang membantu
- Pelaksanaan gerakan baru hanya dapat memenuhi struktur-struktur dasar gerakan
- sangat membutuhkan bantuan dari luar.

Berdasarkan pada keterbatasan-keterbatasan kemampuan motorik yang dimiliki oleh individu yang berada pada tingkat belajar dasar ini, maka penulis menyarankan agar pembobotan penilaian hasil belajar mahasiswa dalam aspek motorik lebih kecil dari aspek kognitif (40% motorik dan 60% kognitif). Materi perkuliahan teori lebih dititik beratkan pada penguasaan:

- didaktik-metodik dan
- pemahaman teknik-teknik gerakan.

Penggunaan model PAN dalam penilaian juga dilandaskan pada ciri-ciri kemampuan motorik dengan segala keterbatasan yang dimiliki oleh mahasiswa yang berada pada belajar tingkat dasar.

3. Pada perkuliahan tingkat pendalaman, penulis menyarankan agar memberikan bobot penilaian yang sama yaitu 50% kognitif dan 50% motorik dengan menggunakan PAP untuk penilaian hasil belajar. Landasannya adalah:

- bahwa mahasiswa yang mengambil matakuliah pendalaman ini dituntut untuk dapat melaksanakan gerakan yang diajarkan dengan kemampuan koordinasi halus. Ini berarti bahwa tuntutan dalam kemampuan motorik lebih berat bila dibandingkan dengan tuntutan tingkat dasar. Oleh karenanya pembobotan aspek motorik pada tingkat ini juga lebih besar yaitu 50% bila dibandingkan pada tingkat dasar yaitu 40%

- Pembobotan 50% untuk aspek kognitif/teori (10% lebih banyak dari tingkat dasar), bukanlah berarti pengurangan terhadap materi teori, melainkan perimbangan dalam bentuk lanjutan dari materi yang telah diperoleh pada belajar tingkat dasar. Pada belajar tingkat dasar pembobotan teori lebih besar dari tingkat dasar pendalaman (60%). Ini berarti bahwa materi teori telah membahas secara luas tentang pengetahuan-pengetahuan yang dibutuhkan untuk cabang olahraga yang diajarkan, terutama masalah didaktik dan metodik. Dengan demikian penyajian teori pada tingkat belajar pendalaman adalah merupakan pendalaman yang lebih dititik beratkan pada analisa gerak, analisa dan mendiagnosa kesalahan serta koreksi terhadap kesalahan-kesalahan yang terjadi. Artinya pembahasan materi teori tidak lagi meluas, tetapi dipersempit dan memiliki titik berat tertentu.
- Perimbangan pembobotan penilaian hasil belajar teori dan praktek sebesar 50% dan 50% dapat diartikan sebagai tuntutan yang berimbang terhadap kemampuan yang harus dimiliki oleh mahasiswa. Perimbangan tuntutan kemampuan tersebut berarti bahwa mahasiswa yang telah mengambil mata kuliah pada tingkat pendalaman, dituntut tidak hanya mampu melakukan gerakan-gerakan dengan baik, tetapi juga mampu menyajikan atau menginformasikan bahan pengajaran dengan baik.
- Usulan untuk menerapkan model PAP dalam penilaian mempunyai landasan berpikir, bahwa mahasiswa yang berada pada tingkat pendalaman ini, telah memiliki kemampuan-kemampuan dasar motorik yang didapatnya pada belajar tingkat dasar. Dengan kata lain, dia telah memiliki pengalaman dasar-dasar kemampuan motorik yang dapat digunakannya untuk berlatih sendiri (mandiri) ditambah dengan latihan-latihan yang diperolehnya pada perkuliahan pada tingkat pendalaman. Kondisi yang demikian telah memungkinkan mahasiswa untuk dapat

mencapai patokan-patokan penilaian yang telah ditetapkan.

3. Bagi mata kuliah yang disajikan tiga semester dalam bentuk proses yang berkelanjutan yaitu dasar-pendahuluan-dan spesialisasi, maka penilaian hasil belajar mahasiswa pada tingkat spesialisasi adalah sebagai berikut yaitu: 60% Motorik dan 40% kognitif dengan merapkan PAP sebagai model pendekatan dalam penilaian hasil belajar. Sebagai bahan informasi, dapat disampaikan hasil di sini, bahwa mahasiswa yang berada atau yang mengambil mata kuliah praktek olahraga pada tingkat spesialisasi diarahkan pada penguasaan keterampilan-keterampilan motorik yang tinggi, sesuai dengan cabang olahraga yang dipilihnya. Kreteria-kreteria kemampuan motorik olahraga pada tingkat spesialisasi ini didasarkan pada kreteria-kreteria yang harus dimiliki oleh peserta didik yang berada pada tingkat belajar motorik ketiga. MEINEL (1977, 282-284) mengemukakan ciri-ciri kemampuan motorik seseorang yang berada pada fase belajar tingkat ketiga antara lain adalah sebagai berikut:

- seseorang yang berada pada fase belajar tingkat ketiga telah memiliki kemampuan koordinasi halus.
- kemampuan koordinasi halus yang telah dikuasainya adalah konstan atau stabil. Mampu melaksanakan gerakan yang dituntut secara bervariasi dalam situasi dan kondisi yang berbeda-beda atau situasi dan kondisi yang dipersulit.
- Memiliki kemampuan penyesuaian yang tinggi di dalam perubahan situasi yang terjadi secara cepat.
- Memiliki kepekaan yang tinggi terhadap jalannya gerakan yang sedang dilaksanakannya. Dengan kata lain, mampu memonitor sendiri tentang jalannya gerakan yang sedang dilaksanakannya. Kepekaan yang tinggi ini didapatkan dari hasil melakukan gerakan-gerakan secara berulang-ulang.

Berdasarkan pada kemampuan-kemampuan dan keterampilan-keterampilan motorik yang harus dicapai oleh peserta didik yang berada pada tahap belajar tingkat ke III, maka perlu diberikan kesempatan yang luas pada mahasiswa untuk dapat melakukan latihan-latihan yang terarah di bawah bimbingan dosen. Kemampuan-kemampuan dan keterampilan-keterampilan motorik seperti yang digambarkan di atas adalah merupakan kemampuan dan keterampilan yang berkualitas tinggi dan kompleks. Sejalan dengan tingginya kemampuan motorik yang dituntut tersebut, maka disarankan agar kemampuan keterampilan motorik diberikan bobot penilaian yang lebih besar yaitu sebanyak 60%

Materi perkuliahan teori pada tingkat spesialisasi ini seharusnya lebih dititik beratkan pada pendalaman terhadap ilmu melatih dan teori gerak. Pilihan ini sejalan dengan tujuan yang ingin dicapai pada jurusan Kepeleatihan FPOK IKIP Padang, yaitu mempersiapkan tenaga-tenaga yang mampu mengajar dan melatih keterampilan olahraga. Pengetahuan teori yang bersifat didaktik dan metodik telah diperdapatnya pada perkuliahan tingkat dasar dan pendalaman *).

Pada tingkat spesialisasi ini, penerapan PAP dalam penilaian dapat dianjurkan. Anjuran ini dilandaskan pada pemikiran, bahwa mahasiswa yang mengambil mata kuliah spesialisasi telah memiliki kemampuan-kemampuan dan keterampilan-keterampilan motorik yang baik, sehingga memungkinkan bagi mereka untuk mencapai patokan-patokan yang telah ditetapkan. Selain dari itu penerapan PAP dalam hal ini juga wajar, untuk dapat mencapai perbedaan kemampuan dan keterampilan motorik yang nyata, bila dibandingkan pada tingkat belajar dasar dan pendalaman.

*) Oleh karenanya materi teori pada tingkat spesialisasi lebih dititikberatkan pada ilmu melatih, teori gerak dan sport medicine.

Berikut ini disampaikan ringkasan memberikan pembobotan penilaian perkuliahan praktek olahraga yang disarankan dalam makalah ini. Ringkasan tersebut dikemukakan dalam bentuk tabel di bawah ini.

Sifat mata kuliah Praktek	Bobot Penilaian		Pendekatan Penilaian
	Kognitif	Motorik	
- Dasar, atau yg disajikan hanya 1 semester	60 %	40 %	PAN
- Pendalaman	50 %	50 %	PAP
- Spesialisas	40 %	60 %	PAP

Beberapa Informasi Tentang Penilaian Hasil Belajar Keterampilan Olahraga di FPOK IKIP Padang

Dalam melakukan penilaian hasil belajar keterampilan olahraga, staf pengajar FPOK IKIP Padang menggunakan tes-tes keterampilan motorik khusus untuk setiap cabang olahraga. Tes-tes keterampilan motorik tersebut adalah tes-tes yang sudah baku. Dengan kata lain bahwa tes-tes tersebut telah memenuhi standar atau kriteria sebagai alat pengukur. Bentuk-bentuk tes yang dipergunakan tersebut ada yang berasal dari Amerika dan ada yang sudah dimodifikasi untuk ukuran orang Indonesia.

KESIMPULAN:

1. Masih terdapat ketidakseragaman tentang pembobotan penilaian hasil belajar praktek olahraga di FPOK IKIP Padang
2. Kesulitan di dalam penyeragaman kebijaksanaan di dalam menilai hasil belajar tersebut, juga disebabkan karena setiap cabang olahraga memiliki dan menuntut keterampilan-keterampilan metodik yang spesifik.

SARAN-SARAN

1. Se jauh literatur yang pernah ditelusuri, penulis masih belum berhasil menemukan patokan-patokan tentang penilaian aspek kognitif dan motorik. Oleh karenanya akan merupakan hal yang sangat positif sekali, bila FPOK- IKIP Padang melakukan diskusi-diskusi secara intensif atau melaksanakan seminar untuk masalah penilaian yang dimaksud.
2. Sebaiknya beberapa penelitian yang ada di FPOK IKIP Padang diarahkan pada pemecahan masalah penilaian hasil belajar.

Daftar Bacaan

- Ausschuss Deutscher Reibseraicher. Sport Planen Durchfiikren- Auswerfer. Schorndorf bei Sttutgart 1987.
- Baumann.H. Bewegungslehre. Frankfurt am Main, 1984
- Dept-Dik-Bud. Universitas Terbuka. Buku II Modul Evaluasi Hasil Belajar, 1984/1985
- Dept-Dik-Bud. Universitas Terbuka. Buku II Modul Pembuatan Alat Evaluasi Keterampilan Olahraga, 1984/1985
- Digel.H. Lekren im Sport, Reinbek bei Hamburg, 1983.
- Fakultas Pendidikan Olahraga dan Kesehatan. Silabi Mata Kuliah Jurusan Pendidikan Olahraga, 1987.
- Fakultas Pendidikan Olahraga dan Kesehatan, Silabi Jurusan Pendidikan Kepeatihan, 1987.
- Fetz. F. Allgemeine Methodik der Leibesübungen, Wien, 1988.
- Grossing. S. Einführung in die Spordidaktik, Wiesbaden, 1988.
- Institut Keguruan dan Ilmu Pendidikan (IKIP) Padang Buku Pedoman, 1989 - 1990.
- Institut Keguruan dan Ilmu Pendidikan (IKIP) Padang Kumpulan Makalah Penataran Penilaian Hasil Belajar Angkatan II, 1989
- Kiram; Yanuar, Prinzipien des Lekrens von Sportlichen Bewegungen Unter besonderer Berücksichtigung indonesicher Belange, Darmstadt, 1989.
- Lutan, Rusli, Belajar Keterampilan Motorik. Pengantar Teori dan Metode, Jakarta, 1988
- Meinel, K. Bewegungslehre, Berlin, 1977.
- Rothig, Sportwissenschaftliches Lexikon, Schorndorf, 1983
- Singer, R.N. Motorisches Lernen und Menschliche Leistung, Bad Homburg 1985
- Nurkancana, W. Evaluasi Pendidikan, Surabaya, 1982.
- Weineck, J. Optimales Bewegungslernen, Erlangen, 1983
- Wirjasantosa, R. Supervisi Pendidikan Olahraga, 1984.