

SEMINAR LOKAKARYA PEMBUATAN TES & EVALUASI

NO. SURAT	2-11-2001
NO. SURAT	Hodiah
NO. SURAT	KF
NO. SURAT	614/K/2001-81/2
NO. SURAT	371.26 UMI 5:1



Padang, 28 – 29 Juni 1999



Kata Pengantar

Dalam usaha meningkatkan mutu akademik khususnya hasil belajar mahasiswa, Universitas Bung Hatta mengadakan "Semlok Pembuatan Tes & Evaluasi" pada tanggal 28 - 29 Juni 1999 di Aula Gedung B Kampus I Universitas Bung Hatta Ulak Karang. Peserta semlok ini terdiri dari pelaksana akademis: pembantu dekan I, ketua dan sekretaris program studi dalam lingkungan Universitas Bung Hatta.

Pemakalah dan pokok bahasan yang ditampilkan dalam semlok ini adalah:

1. Drs. Kumaidi, M.A., Ph.D.
"Pengujian dan Penilaian Prestasi Belajar"
2. Dr. Imam Sodikun, M.Pd.
"Penulisan Butir Soal, Bentuk Uraian dan Pilihan".
3. Dr. Anas Yasin, M.A.
"Prosedur Penilaian Hasil Tes"


Ketiga pemakalah dalam sajiannya akan memaparkan prosedur dan proses pembuatan tes sampai pada hal-hal yang berkaitan dengan evaluasi. Hal ini dianggap sangat penting dalam usaha meningkatkan prestasi belajar mengajar di lembaga ini. Setiap dosen tentu telah terbiasa membuat tes dan memberi penilaian, namun keragaman/variasi penilaian selalu saja dijumpai yang mungkin pula berdampak negatif yang merugikan mahasiswa karena perbedaan perlakuan. Pertanyaan yang timbul ialah: "sudahkah para dosen memakai teknik pembuatan tes yang dapat mengungkap penilaian hasil belajar yang sesungguhnya? mengingat prestasi belajar itu ditentukan antara lain oleh prosedur dan proses pengukuran, penilaian, dan pengambilan keputusan. Ketiga makalah di atas dimaksudkan untuk memberi bekal ke arah peningkatan keterampilan dalam aspek-aspek tersebut.

Panitia berharap, semoga semlok ini mencapai sasarannya yaitu peningkatan mutu akademis di lembaga pendidikan tercinta ini.

Wassalam
Panitia Semlok

ACARA SEMLOK PEMBUATAN TES & EVALUASI

HARI / TANGGAL	PUKUL	ACARA / PENYAJI / MATERI	MODERATOR / SEKRETARIS
SENIN 28-Jun-99	9.00 - 9.45	- Laporan Panitia - Pembukaan oleh Rektor	Panitia Dr. H. Agustiar Syah Nur, M.A. Yusrita Yanti, S.S., M. Hum.
	9.45 - 10.00	- Istirahat, kudapan	
	10.00 - 12.30	I. Drs. Kumaidi, M.A., Ph.D. Pengujian dan Penilaian Prestasi Belajar	
	12.30 - 13.30	Ishoma	
	13.30 - 16.00	II. Dr. Imam Sodikun, M.Pd. Penulisan Butir Soal, Bentuk Uraian dan Penilaian	
Selasa 29-Jun-99	9.30 - 12.00	III. Dr. Anas Yasin, M.A. Prosedur Penilaian Hasil Tes	Drs. H. Fahmi Isman, M.Si. Drs. Alwin Nazaruddin, M. Hum.
	12.00 - 13.00	Ishoma	
	13.00 - 15.00	Diskusi II, III Dr. Imam Sodikun, M.Pd. Dr. Anas Yasin, M.A.	
	15.00	Penutup	

KETUA,
PANITIA SEMLOK,

SEKRETARIS,
DR. H. AGUSTIAR SYAH NUR, MIAI

DRS. H. SYAMSUDDIN UDIN

PENGUJIAN DAN PENILAIAN PRESTASI BELAJAR

**Oleh
Kumaidi**

Disampaikan Pada Semlok Pembuatan Tes & Evaluasi
Universitas Bung Hatta Padang
28 – 29 Juni 1999

PENGUJIAN DAN PENILAIAN PRESTASI BELAJAR¹

Kumaidi²

1. Pendahuluan

Setiap kali dosen harus melaksanakan penilaian terhadap prestasi belajar mahasiswanya. Penilaian ini dilaksanakan setelah dosen melakukan proses belajar mengajar dan murid melakukan proses pembelajaran. Penilaian ini dilaksanakan karena dosen dan pengelola (dekan, rektor, atau pejabat Depdikbud) akan membuat suatu keputusan. Keputusan dosen biasanya bersifat mikro, sedangkan keputusan pengelola biasanya lebih bersifat makro.

Untuk melakukan penilaian, penilai (dosen, dekan, atau siapa saja) memerlukan informasi tentang berbagai aspek kegiatan belajar siswa selama satu kurun waktu pembelajaran. Kurun waktu tersebut bisa beberapa minggu pembelajaran atau satu semester, satu tahun, bahkan selama beberapa tahun. Informasi tersebut kebanyakan berkaitan dengan kemajuan belajar siswa yang dapat dikumpulkan melalui pengamatan dan/atau pengukuran (pengujian). Oleh sebab itu, pengujian bukan merupakan satu-satunya cara untuk melakukan penilaian.

Penilaian tanpa pengujian mungkin sering dijumpai dalam kehidupan sehari-hari. Misalnya, seorang ayah atau ibu menilai anak-bayinya lebih lambat atau cepat perkembangannya dibandingkan anak tetangga setelah mengamati perilaku mereka. Dalam kegiatan pasar pun sering kali dijumpai proses penilaian tanpa pengujian. Misalnya seorang pedagang perlu menyimpan untuk beberapa waktu dagangannya atau menaikkan harga jual dagangannya, karena isu akan adanya kenaikan gaji pegawai. Hal yang demikian dapat pula terjadi dalam dunia pendidikan, misalnya guru sastra membaca dan menikmati suatu karya sastra; guru lukis mengamati cara muridnya mengekspresikan perasaannya melalui sapuan kuas di atas kanvas.

Dalam berbagai keadaan, penilaian tanpa pengujian memang dapat dilaksanakan. Namun, dalam berbagai keadaan lainnya, penilaian sering memerlukan pengujian. Dan hampir setiap situasi, penilaian dengan pengujian menghasilkan keputusan yang lebih tepat sasaran. Pengambilan keputusan berdasarkan data dan informasi tanpa pengukuran seringkali menghasilkan dampak yang kurang terduga. Di samping itu, mutu keputusan yang dibuat tanpa pengukuran sering kurang stabil, terutama suatu keputusan yang berulang. Oleh sebab itu, ahli penilaian pendidikan menganjurkan keputusan pendidikan yang akan dibuat sebaiknya berdasarkan hasil pengukuran.

Diskusi singkat di atas memperlihatkan tiga peristilahan yang saling terkait yaitu pengukuran, penilaian, dan pengambilan keputusan. Tiga istilah tersebut sering dikacau-artikan oleh sebagian dosen. Pengukuran (pendidikan) diartikan sebagai usaha pengumpulan informasi dalam bentuk angka melalui pengamatan dan pengujian menurut aturan tertentu. Dengan demikian, informasi yang diperoleh dari pengukuran

¹ Makalah disampaikan pada Semlok Pembuatan Tes dan Evaluasi bagi Dosen-dosen Universitas Bung Hatta Padang, 28-29 Juni 1999

² Dosen & Spesialis Pengujian IKIP Padang

(pendidikan) akan berbentuk data kuantitatif atau angka. Oleh sebab itu, apabila dosen memberi ujian kepada mahasiswa dan jawaban mereka diponten dengan skala 0/1 (untuk respons benar/salah dalam tes bentuk pilihan) atau 0-10 (untuk respons dalam tes bentuk uraian), maka dosen tersebut pada dasarnya baru melakukan pengukuran. Dosen belum melakukan penilaian! Informasi dapat saja dikumpulkan tanpa pengukuran, misalnya dengan melakukan pengamatan terhadap aktivitas mahasiswa selama pengajaran berlangsung. Dari informasi yang terkumpul, dosen melakukan penilaian dengan jalan memberi arti dan makna terhadap informasi tersebut. Pemberian arti dan makna ini dapat dicapai melalui perbandingan hasil pengukuran dengan suatu acuan standar. Acuan standar ini dapat berbentuk kriteria ketuntasan belajar (misalnya persen penguasaan atau pencapaian tujuan pengajaran) atau kedudukan relatif prestasi mahasiswa terhadap suatu distribusi prestasi belajar. Dalam dunia pendidikan, penilaian berdasarkan ketuntasan belajar biasanya memberi hasil yang lebih baik. Hasil penilaian biasanya diperlukan dosen untuk pengambilan keputusan, misalnya: Apakah seorang mahasiswa layak diberi predikat "berprestasi sangat baik, baik, cukup, atau kurang"; Apakah dosen perlu mengubah strategi pembelajarannya; Apakah dosen perlu merumuskan ulang sasaran pengajarannya; dan lain sebagainya.

Makalah singkat ini mencoba menjelaskan prosedur dan proses penyusunan tes prestasi belajar untuk materi ajar (suatu bidang studi). Asumsi yang melandasi penyusunan makalah ini, peserta penataran masih memiliki pemahaman yang sangat terbatas tentang prosedur dan proses pengembangan tes dan dosen belum memiliki sistem bank soal yang baik. Oleh sebab itu, secara ringkas, makalah ini akan mendiskusikan penyusunan rancangan tes prestasi belajar dan berbagai pertimbangan pemilihan strategi pengujiannya. Di samping itu, telaah butir soal dan perakitan tes prestasi belajar juga akan dibicarakan dalam makalah ini.

2. Penyusunan Rancangan Tes

Langkah awal dalam penyusunan tes prestasi belajar biasanya dimulai dengan penyusunan rancangan tes. Rancangan tes dapat dipandang sebagai "blueprint" (ingat blueprint dalam teknik sipil) suatu tes. Oleh sebab itu, rancangan tes harus memuat rambu-rambu dan spesifikasi tes yang diinginkan. Spesifikasi tes ini umumnya terdiri dari semua informasi tentang karakteristik tes, yang meliputi: tujuan pengujian, kisi-kisi tes, bentuk tampilan butir soal, panjang tes, standar tes yang diinginkan, dan rencana interpretasi hasil pengujian. Tujuan pengujian perlu dirumuskan dari awal karena tujuan ini akan menentukan strategi pengembangan tes. Tujuan pengujian biasanya dikaitkan dengan fungsi penilaian yang ingin diservis atau dilayani, berkenaan pengumpulan informasi belajar mahasiswa. Dalam lingkup pengajaran, tujuan yang ingin dicapai dosen dapat bervariasi antara lain:

- a. mengetahui kesiapan mahasiswa sebelum pengajaran dilaksanakan. Kesiapan murid ini bertalian dengan penguasaan pengetahuan dasar sebagai prasyarat keberhasilannya mengikuti program pengajaran selama satu semester. Tujuan pengujian ini biasanya dikenal dengan tes penempatan dan pretes. Tes penempatan dan pretes dilaksanakan sebelum pengajaran diberikan.
- b. mengetahui kesulitan belajar siswa dan kelemahan strategi pembelajaran yang dipilih dosen. Yang pertama sering dikenal sebagai tes diagnostik dan dilaksanakan selama pengajaran berlangsung. Yang kedua dikenal sebagai tes formatif, juga dilaksanakan selama pengajaran berlangsung. Perbedaan dari kedua macam tujuan pengujian ini terletak pada target pemberian arti informasi yang masuk. Tes diagnostik lebih menekankan

kepada pemberian arti pada level individu murid, sedangkan tes formatif lebih mengacu kepada kelompok murid (misalnya dalam satu ruang belajar).

- c. memberi grading (predikat keberhasilan belajar) dan sertifikasi murid. Tujuan grading biasanya dilakukan untuk memerikan tingkat pencapaian dan keberhasilan mahasiswa selama mengikuti pengajaran, sedangkan tujuan sertifikasi mahasiswa dilakukan setelah pengacuan terhadap suatu kriteria ketuntasan (mastery) tertentu. Tujuan kelompok ini biasa dikenal sebagai tujuan sumatif yang dilaksanakan di akhir pengajaran. Ini bukan berarti bahwa pengesanan sumatif hanya dilakukan di akhir semester.

Tujuan pengesanan ini sebaiknya dicantumkan dosen dalam rancangan penyusunan tes setelah tujuan pengesanan dirumuskan, dosen perlu menyusun kisi-kisi tes. Penyusunan kisi-kisi tes dimulai dengan analisis sasaran akhir pengajaran. Sasaran akhir pengajaran biasanya ditampilkan dalam sinopsis pengajaran dan dijabarkan lebih lanjut dalam silabus pengajaran yang didokumentasikan dalam Garis-garis Besar Program Pengajaran (GBPP). Dari GBPP ini, dosen biasanya mengembangkan tujuan instruksional khusus (TIK) setiap kali akan menjelaskan bahan ajar baru. TIK ini dipakai sebagai dasar pengembangan tes selanjutnya dan untuk mudahnya bisa dikelompokkan ke dalam beberapa pokok bahasan. TIK yang telah tersusun ini dapat diambil sebagai indikator penguasaan materi ajar oleh mahasiswa. Karena TIK berfungsi sebagai indikator penguasaan materi ajar, maka melalui pengesanan dosen menstimulasi mahasiswa agar mendemonstrasikan kemampuan dan keterampilan yang mereka peroleh selama mengikuti pengajaran. Oleh sebab itu, kisi-kisi tes paling tidak merangkum sasaran akhir pengajaran, uraian materi pengajaran, dan indikator kemampuan mahasiswa.

Setelah kisi-kisi tes ditetapkan, dosen perlu merumuskan bentuk tampilan butir soal dan panjang tes. Bentuk tampilan butir soal meliputi bentuk pilihan (pilihan ganda, benar-salah, atau menjodohkan), isian, dan uraian (esai, hitungan, pembuatan makalah, dan penyusunan neraca keuangan). Setelah bentuk tampilan butir soal ditetapkan, dosen dapat menentukan panjang tes. Panjang tes ditentukan dari perpaduan antara kelayakan kesulitan tes dengan jumlah waktu tersedia. Bentuk tampilan butir soal ikut menentukan panjang tes, sebab sebuah bentuk uraian akan membutuhkan waktu uji yang lebih lama dari sebuah bentuk pilihan atau isian. Tujuan pengesanan juga harus dipertimbangkan dalam memilih bentuk tampilan soal. Tes diagnostik memerlukan rincian langkah yang lebih cermat dari sekedar tes sumatif. Dengan demikian, tes diagnostik memerlukan jumlah soal yang lebih banyak. Bila jumlah soal lebih banyak, butir soal bentuk uraian mungkin kurang layak dipilih.

Standar tes perlu diperhatikan dalam menyusun rancangan tes. Pada waktu yang bersamaan perlu pula dipertimbangkan model interpretasi hasil pengesanan. Standar tes biasanya berkaitan dengan standar butir soal. Dalam teori klasik, standar tes biasanya dikaitkan dengan validitas dan reliabilitas ponten. Model interpretasi apapun yang akan dipakai, validitas ponten perlu mendapat perhatian pertama. Untuk keperluan pengesanan prestasi belajar, validitas yang berkaitan dengan relevansi butir soal dengan indikator atau TIK (item-objective congruence, Hambleton, 1984) merupakan model yang pantas diperhitungkan. Untuk acuan reliabilitas, dosen harus hati-hati sebab pengacuan normatif dan ketuntasan belajar (memakai kriteria keahlian) memerlukan model reliabilitas yang berlainan. Bila interpretasi hasil tes memakai model pengacuan normatif, reliabilitas yang biasa kita kenal (KR-20, KR-21, atau Cronbach Alpha) bisa dipakai. Namun, bila model pengacuan kriteria, reliabilitas keputusan (lulus-tidak lulus,

mastery-nonmastery) harus dipilih (Brennan, 1984). Pengacuan normatif menuntut dosen memilih butir soal yang dapat membedakan prestasi mahasiswa yang berhasil dengan yang tidak, sedangkan pengacuan kriteria, keharusan butir soal membedakan dua jenis prestasi tersebut tidak mutlak.

Seluruh rancangan yang dijelaskan di atas dapat disusun dalam sebuah matriks sebagai berikut:

Model Matriks Rancangan Tes

1. Mata Kuliah :
2. Tujuan :
3. Tempo : menit
4. Semester :

No	Sasaran Akhir	Uraian Materi	Indikator	Banyak Soal	Bentuk Soal
1.	TIU Pokok Bahasan 1	1.1 ... 1.2 ...	1.1.1 ... 1.1.2 ...	5 2	pilihan uraian
2.	TIU Pokok Bahasan 2	dst	dst	dst	dst
3.	TIU Pokok Bahasan 3	dst	dst	dst	dst

Cara pengisian model rancangan tes tersebut dapat dijelaskan secara singkat sebagai berikut:

- a) Mata Kuliah : Tuliskan nama mata kuliah yang akan diujikan;
- b) Tujuan : Tuliskan tujuan utama dari pengujian dan penilaian memakai tes yang dikembangkan, misalnya diagnostik, formatif, atau sumatif (mana yang dipentingkan dari berbagai kemungkinan tersebut);
- c) Tempo : Tuliskan alokasi waktu yang diizinkan bagi siswa untuk menyelesaikan seluruh tes (dalam menit). Penentuan alokasi waktu ini harus disesuaikan dengan banyaknya soal, bentuk butir soal, dan jenis mata ujiannya;
- d) Semester : Tuliskan tes (mata kuliah) tersebut untuk semester ke dan tahun berapa.
- e) Sasaran Akhir : Tuliskan TIU dari setiap pokok bahasan yang dicakup dalam tes; TIU ini biasanya dapat dilihat dalam GBPP suatu mata kuliah;
- f) Uraian Materi: Uraikan jabaran materi menurut pokok bahasan yang diujikan; penguraian ini sebaiknya meliputi

- konsep utama dan bagian-bagian utama suatu pokok bahasan yang akan dijadikan dasar pengujian;
- g) Indikator : Tuliskan ciri-ciri yang dapat diamati dan dijadikan petunjuk penguasaan materi uji dan pencapaian TIU dari seorang siswa; sebaiknya dituliskan dalam ungkapan kata kerja operasional terhadap suatu konsep utama dan mencakup kondisi lingkungan yang diinginkan seorang siswa dapat mendemonstrasikan kemampuan dan keterampilannya;
- h) Banyak Soal : Tuliskan jumlah soal yang akan dibuat untuk setiap indikator atau setiap uraian materi (otomatis setiap pokok bahasan) dan akhirnya jumlah keseluruhan butir soal dalam tes yang dirancang dapat diketahui;
- i) Bentuk Soal : Tuliskan berbagai bentuk butir soal yang akan dipakai (misalnya bentuk pilihan, isian, atau uraian); pemilihan bentuk soal perlu mempertimbangkan uraian materi dan indikator.

3. Pertimbangan Pemilihan Model Pengujian

Dosen seringkali mengalami kesulitan untuk menentukan model pengujian apa yang cocok dengan mata kuliah yang dibinanya. Hal ini mungkin disebabkan oleh bekal pengetahuan teori pengukuran dan pengujian yang dimiliki dosen sangat terbatas. Oleh sebab itu, kita perlu selalu berusaha untuk meningkatkan kemampuan pengujian dosen melalui berbagai interaksi kolegial, misalnya semlok atau workshop seperti saat ini.

Pilihan model pengujian biasanya dikaitkan dengan cakupan TIK yang mungkin ingin dicapai dosen. Dalam sistem persekolahan kita dikenal ada tiga ranah TIK, yaitu kognitif, afektif, dan psikomotorik. Pengujian ranah kognitif paling sering kita jumpai dalam bentuk ujian tertulis di samping kadang dipakai ujian lisan. Bentuk ujian tertulis ini memiliki berbagai model tampilan butir soal yang oleh pemakalah dikelompokkan menjadi tiga bentuk yaitu butir soal bentuk pilihan (termasuk pilihan ganda dan benar-salah), isian, dan uraian (termasuk karangan singkat, perhitungan, dan penyusunan neraca keuangan).

Pengujian ranah afektif dapat mengambil bentuk pengamatan (dengan format pengamatan) atau bentuk angket "self-reported" (swa-lapor diri). Model angket swa-lapor ini bermacam-macam, antara lain bentuk skala Likert, Semantic differential, Thurstone, dan Guttman. Sedangkan pengujian ranah psikomotorik umumnya dilaksanakan dengan simulasi dunia kerja dan percontohan hasil kerja. Pada dasarnya, pengujian ranah psikomotorik dilaksanakan dengan memerintahkan teruji mendemonstrasikan kemampuan dan keterampilannya menyelesaikan suatu pekerjaan tertentu (percontohan hasil kerja) dalam situasi yang mirip dengan tempat kerja sesungguhnya (simulasi). Bentuk tampilan simulasi dan percontohan hasil kerja tersebut dapat bermacam-macam misalnya proyek, kerja praktek, dan tugas-tugas pekerjaan. Bentuk ujian tertulis sebenarnya masih bisa dipakai untuk ujian praktek, misalnya praktek bengkel, labor, menggambar (bentuk keterampilan komunikasi teknik), dan sejenisnya.

Dari pembicaraan tersebut di atas, dosen tentu mempunyai wawasan model pengujian apa yang cocok untuk materi ajar yang disajikannya dalam suatu mata pelajaran. Apakah dia akan memilih ujian tertulis, ujian lisan, atau memerintahkan

siswanya menyelesaikan proyek akhir semester.

Dalam kasus ujian praktek, dosen perlu mempertimbangkan kondisi bengkel kerja dan labor tempat ujian akan dilaksanakan dan persediaan bahan. Apabila dosen memperkirakan bahwa bengkel kerja cukup kecil dan peralatan yang ada terbatas, ada baiknya dosen memilih membuat ujian dalam bentuk proyek. Alasannya, proyek biasanya memerlukan waktu penyelesaian yang lebih panjang dari satu periode tatap muka atau jam praktek, sehingga mahasiswa bebas menyelesaikannya di luar jam praktek dengan pengawasan dosen atau petugas lain. Di samping itu, proyek memang mengizinkan mahasiswa bertanya cara penyelesaiannya kepada orang yang dia anggap dapat membantu memecahkan persoalan yang dia ajukan (hadapi) sebagai topik proyeknya (asal penyelesaian akhir dia lakukan sendiri). Jadi fleksibilitas waktu dan strategi penyelesaiannya lah yang menjadi kekuatan proyek sebagai model pengujian dibanding model lain. Dengan mempertimbangkan berbagai kemungkinan seperti itu, dosen dapat dengan jeli memilih model pengujian apa yang kiranya cocok untuk situasi dan kondisi masing-masing pembelajaran di lapangan.

4. Penulisan Butir Soal Tes Kognitif

Setelah rancangan tes disusun, dosen dapat melanjutkannya dengan menulis butir soal. Seorang penulis soal yang sukses perlu menguasai teknologi penulisan butir soal. Secara garis besar, teknik penulisan butir soal, apapun bentuk tampilannya, sama saja. Namun perlu digarisbawahi bahwa setiap tampilan butir soal memerlukan kelengkapan yang berbeda. Misalnya, untuk butir soal bentuk pilihan, kunci jawaban merupakan gambaran pemahaman yang benar dan pengecoh merupakan semua kemungkinan kesalahpahaman konsep yang diujikan. Oleh sebab itu, alur penulisan dan berfikirnya mungkin lurus dan langsung. Untuk butir uraian, penulis soal harus menampilkan soal dengan jabaran arah jawaban yang lebih jelas (termasuk batasan ruang lingkupnya). Butir soal bentuk pilihan pemontenannya langsung (walau mungkin pakai rumusan tertentu), sedangkan bentuk uraian perlu strategi berbeda (melalui pedoman pemontenan). Tes praktek memerlukan strategi yang lain lagi, termasuk pedoman pemontenan dan pedoman (format) pengamatannya (bila diperlukan).

Dari uraian singkat di atas tampak bahwa seorang penulis soal harus menguasai (dengan baik) materi uji dan juga menguasai teknik penulisan butir soal. Penjelasan selanjutnya dari makalah ini mencoba membicarakan berbagai langkah dan teknik penulisan butir soal serta persiapan lain yang diperlukan dosen untuk penyusunan tes prestasi belajar.

5. Kaidah Penulisan Butir Soal Pilihan

Penjelasan berikut memberi rambu-rambu kepada dosen tentang teori (keterampilan kognitif atau olah pikir) mengembangkan butir soal bentuk pilihan. Mata kuliah yang mungkin diuji dengan butir soal bentuk pilihan ini antara lain, teori ekonomi, akuntansi, matematika ekonomi, ekonometrika, ekonomi makro, dan sejenisnya, (contoh dari program studi ekonomi).

Petunjuk pertama yang dapat diberikan di sini adalah anjuran bagi setiap dosen yang akan menulis soal untuk mencekati rancangannya. Kemudian, dia dianjurkan mengkaji ulang buku pegangan, rumusan uraian materi, dan indikator penguasaan materi uji yang dirumuskan dalam kisi-kisi tes. Dari pencermatan ini diharapkan dosen menemukan rumusan konsep utama dari materi uji secara lebih tajam dan konkrit. Setelah dosen dapat merumuskan konsep utama materi yang akan diujikan, dosen dianjurkan menuangkan gagasan persoalannya dalam suatu pokok soal (stem) memakai bahasa Indonesia yang baik. Usahakan untuk memakai istilah selingkung (jargon)

dalam bahasa Indonesia. Namun, istilah asing yang lazim dikenal mungkin tidak perlu diterjemahkan terutama apabila istilah terjemahan ini justru menyesatkan mahasiswa. Pokok soal (stem) dapat saja berbentuk kalimat tidak selesai atau berbentuk pertanyaan. Alternatif jawaban akan menyesuaikan dengan bentuk pokok soal.

Apabila rumusan pokok soal sudah selesai dituliskan dan (cukup) bagus menurut penulis soal, kunci jawaban harus segera dituliskan. Kunci jawaban untuk pokok soal berbentuk kalimat tidak selesai merupakan *anak kalimat* yang menyelesaikan kalimat pokok soal. Sedangkan pokok soal berbentuk pertanyaan akan memerlukan kunci jawaban berbentuk kalimat utuh. Pada tahap ini, penulis soal tidak perlu memikirkan secara mendalam tentang urutan dan letak kunci jawaban tersebut (misalnya, apakah di a, b, c, dan seterusnya). Tujuan langsung menuliskan kunci jawaban tersebut adalah agar penulis tidak lupa terhadap gagasan awal penulisan soal. Sebab apabila penulis lupa gagasan awal penulisan soal, arah jawaban mungkin akan lenceng. Lebih-lebih apabila pokok soal belum bersih dari kosa kata yang bercabang dan kalimatnya menimbulkan ketaksaan.

Setelah kunci jawaban dirumuskan, penulis soal baru menuliskan pengecoh. Penulisan pengecoh tersebut harus mempertimbangkan kemungkinan kesesatan pemahaman siswa terhadap makna dan arti konsep yang diujikan; ketidaktepatan aplikasi konsep dalam pemecahan persoalan di lapangan; ketidakcermatan mahasiswa dalam menganalisis dan menghitung. Apabila keseluruhan butir soal selesai dituliskan dosen dianjurkan untuk segera mengedit dan membersihkan rumusan (akhir) butir soalnya. Rumusan akhir ini kemudian didokumentasikan dalam sebuah *kartu soal* (lihat model kartu soal di bawah). Untuk soal yang akan dimasukkan dalam kartu soal, dosen perlu memutuskan kunci jawabannya akan ditaruh dalam urutan ke berapa (pastikan kunci akan diletakkan di a, b, c, dan sebagainya).

Beberapa petunjuk berikut mungkin bermanfaat bagi dosen untuk membantu menuliskan butir soal bentuk pilihan (ganda):

- a) Rumuskan pokok soal yang hanya memuat informasi (secara jelas, komplit, dan disepakati oleh para ahli) yang diperlukan siswa untuk menjawab soal. Rumusan ini mestinya langsung pada sasaran (to the point);
- b) Pilih kosa kata yang tepat sehingga tidak menimbulkan keraguan atau makna ganda dan dapat menimbulkan ketaksaan kalimat;
- c) Susunlah (anak) kalimat dalam pokok soal memakai tata bahasa Indonesia yang baku dan benar;
- d) Usahakan informasi, pernyataan, atau perintah dalam pokok soal tidak menempatkan teruji dalam posisi '*menerjemahkan*' atau menera (melakukan judgment);
- e) Usahakan pokok soal dan alternatif jawaban (kunci dan pengecoh) memiliki hubungan yang makul (logis) dan benar menurut tata bahasa;
- f) Usahakan kunci jawaban merupakan pilihan terbaik dan pengecoh harus makul, namun *jelas bukan* merupakan pilihan yang benar;
- g) Usahakan alternatif jawaban merupakan pilihan yang homogen baik dilihat dari sudut pandang tata bahasa, isi, penekanan, maupun panjang (anak) kalimat;
- h) Usahakan pokok soal tidak memiliki penunjuk arah (clue atau hint) jawaban yang benar atau salah;
- i) Usahakan antar butir soal tidak merupakan butir soal yang saling bertautan, terutama jawabannya butir yang dahulu merupakan prasyarat untuk jawaban butir soal selanjutnya. Dengan kata lain, butir soal harus merupakan butir yang berdiri sendiri;

- j) Dosen tidak perlu takut untuk memberi alternatif jawaban yang berbeda antara butir soal yang satu dengan lainnya. Artinya, jumlah alternatif jawaban tidak harus sama (misalnya seluruh butir soal pilihan harus memiliki alternatif jawaban lima).

Di samping itu, dosen dianjurkan untuk merumuskan konsep utama (untuk menuliskan butir soal dalam bentuk apa pun) yang diujikan dalam bentuk pernyataan positif atau negatif yang dapat dinyatakan benar atau salah. Pernyataan ini biasa dikenal sebagai proposisi. Untuk memberi gambaran yang lebih konkrit dari pemakaian model proposisi di atas, di bawah ini diberikan contoh penulisan butir soal berkaitan dengan "Reliabilitas".

Proposisi: Reliabilitas ponten adalah keakuratan skala pengukuran dalam menggambarkan besaran yang diukur. Konsep ini dijelaskan dalam bentuk stabilitas pengukuran, kesetaraan pengukuran, dan konsistensi internal hasil pengukuran.

Contoh butir soal:

- 1) Bentuk Benar-Salah:

B-S Estimasi reliabilitas tes memakai koefisien korelasi antara hasil pengujian dua tes yang setara (ekuivalen) juga dikenal sebagai koefisien stabilitas.

- 2) Bentuk pilihan ganda:

Salah satu cara untuk meningkatkan koefisien reliabilitas suatu tes adalah

- a. menambah butir soal dengan karakteristik yang mirip dengan butir yang sudah ada
- b. merakit butir-butir soal dengan daya beda yang bervariasi dari rendah ke tinggi
- c. memperpanjang waktu pengujian, sehingga teruji mempunyai waktu cukup menjawab semua soal
- d. menambah sembarang butir soal pada tes yang telah ada

Contoh berikut menunjukkan pemakaian kartu soal sebagaimana dijelaskan di muka dengan memakai butir pilihan ganda di atas.

Mata Kuliah : Dasar-dasar Evaluasi Pendidikan Kode MK : DAK 214/IKIP Padang Penulis Soal : Kumaidi Uraian Materi : Konsep Keliabilitas Ponten		
Nomor soal	Indikator	Butir Soal
5	Siswa dapat menghitung koefisien reliabilitas belah dua, apabila diketahui koefisien korelasi antar belah.	Apabila koefisien korelasi belah dua (split-halves) suatu tes sebesar 0,75, maka estimasi reliabilitas seluruh panjang tes adalah a. 0,56 b. 0,75 c. 0,90 d. 0,86

Model Kartu Soal

6. Kaidah Penulisan Butir Soal Uraian

Kaidah pertama sama saja dengan butir soal bentuk pilihan, yaitu pemahaman dan pengkajian rancangan tes serta penguasaan materi uji secara baik oleh dosen. Dosen masih perlu merumuskan konsep utama materi uji dalam bentuk proposisi. Selanjutnya petunjuk di bawah dapat dipedomani agar dihasilkan rumusan butir soal bentuk uraian yang baik.

- a) Rumuskan gagasan persoalan dalam bentuk kalimat pertanyaan atau perintah secara jelas dengan informasi yang cukup (tidak kurang dan berlebihan);
- b) Berilah batasan ruang lingkup dan kedalaman jawaban secara jelas dan lugas. Batasan ini, di samping membantu siswa memfokuskan pikirannya dan merumuskan jawabannya, juga membantu dosen menetapkan model dan standar pemontenannya;
- c) Tuliskan rumusan gagasan dengan kalimat dan kosa kata bahasa Indonesia baku dan benar. Hindari pemakaian istilah asing, kecuali istilah yang padanan bahasa Indonesianya justru membingungkan mahasiswa;
- d) Gunakan kata tanya atau perintah, misalnya mengapa, apa yang dimaksud dengan, jelaskan, uraikan, dan sebagainya. Hindarkan rumusan soal yang dapat dijawab YA atau TIDAK;
- e) Hindarkan pemakaian kalimat yang dapat menyinggung perasaan kelompok mahasiswa;
- f) Hindarkan pemakaian kosa kata dan kalimat yang meragukan artinya dan menimbulkan ketaksamaan bahasa;
- g) Lengkapi butir soal dengan peta, gambar, sketsa, grafik, atau sejenisnya (bila diperlukan) secara jelas dan terkait langsung dengan soal;
- h) Lengkapi setiap butir soal dengan rencana pemontenannya dan jelaskan ponten maksimum yang dapat diperoleh setiap mahasiswa, apabila setiap

soal tidak memiliki ponten (skor) yang sama.

Contoh berikut, dalam bidang pengukuran dan penilaian pendidikan dengan memakai proposisi (konsep reliabilitas) di muka, akan dapat memperjelas maksud pedoman di atas.

3) Bentuk uraian:

Kenapa harga koefisien Alpha dari Cronbach lebih stabil dibandingkan dengan koefisien reliabilitas belah-dua (split-halves)? Jelaskan secara ringkas tidak lebih dari lima kalimat. (maksimum 4 angka)

Contoh butir soal bentuk uraian lain yang dapat ditunjukkan disini berasal dari 'teori-kejuruan teknik pengerjaan logam'. Butir soal ini ditulis untuk teori keterampilan kerja dengan mesin bor untuk barang cetakan. TiUnya mungkin berbunyi: "Murid terampil bekerja dengan mesin pengerjaan logam". Uraian materinya mencakup pemakaian mesin bor di bengkel (workshop) mesin. Indikator pengembangan (butir soal) tes mungkin berbunyi "Murid dapat:

- a) menjelaskan cara penyenteran (penempatan senter bor) pada lubang hasil cetakan;
- b) memilih jenis mata bor yang dipakai untuk mengeboran awal barang hasil cetakan;
- c) memilih ukuran mata bor yang tepat untuk penyelesaian akhir (sebelum dihaluskan dengan reamer) lubang 70 mm dengan toleransi H8;
- d) membaca dan mengartikan toleransi (H8);
- e) menjelaskan cara pengeboran agar dihasilkan lubang sesuai dengan gambar kerja."

Dengan indikator itu, contoh butir soal permasalahan ini dapat dilihat dalam Lampiran 1 makalah ini.

Butir-butir soal dalam bentuk uraian inipun sebaiknya juga dimasukkan ke dalam model kartu soal. Tujuannya agar mudah disimpan (dibankkan) dan diarsipkan untuk pemakaian kemudian hari, terutama bila butir soal tersebut (setelah dianalisis) merupakan butir soal yang layak dipakai lagi.

7. Penulisan Tes Praktek

Tes praktek biasanya dilakukan bagi mata kuliah praktek dan tes bentuk ini mungkin ditemui di jurusan Pendidikan Teknik Mesin (PTM) IKIP. Untuk PTM IKIP, tes pratek biasanya mengambil tempat di bengkel kerja. Mahasiswa yang sudah diberi latihan selama satu semester atau beberapa minggu (kurang dari satu semester) dapat saja diberi ujian (tes) praktek.

Tes praktek hendaknya mencoba menguji tingkat "mastery" atau ketuntasan belajar dengan standar minimal tertentu. Fokus pengujian hendaknya diarahkan kepada penguasaan keterampilan (psikomotorik) murid terhadap jenis-jenis keterampilan yang diajarkan. Namun, pengujian terhadap aspek pendidikan lainnya bukan hal yang mustahil, misalnya aspek sikap murid terhadap pekerjaan dan pemakaian alat (afektif), serta aspek kerja sama dengan teman lainnya.

Dalam pemilihan strategi pengujian praktek, dosen perlu memperhatikan tujuan pengujiannya. Misalnya dosen ingin mengadakan penilaian sumatif, maka dosen harus menyusun suatu sampel pekerjaan yang dapat 'memancing' atau 'memaksa' dan menstimuli murid mendemonstrasikan atau menampilkan semua keterampilan yang telah dipelajarinya selama selang waktu tertentu (misalnya satu semester). Hasil akhir dari suatu pekerjaan dapat merupakan indikator penguasaan keterampilan yang

diujikan.

Di samping itu, faktor keefektifan pemakaian waktu dan efisiensi pemakaian alat dan bahan juga merupakan unsur penting untuk diperhatikan. Kecepatan murid menyelesaikan tugas (pekerjaan) yang diberikan juga memberi gambaran tentang penguasaan keterampilan yang diujikan. Efisiensi pemakaian bahan dan alat yang tepat juga menunjukkan keterampilan (serta penguasaan informasi tentang pekerjaan) murid.

Faktor keselamatan kerja, baik ditinjau dari aspek lingkungan, orang, maupun benda kerja merupakan unsur penting dari indikator penguasaan keterampilan yang diujikan dan sikap kerja di bengkel. Oleh sebab itu, semua faktor di atas perlu dipertimbangkan sebagai sumber informasi pemberian 'indeks prestasi' (ponton, biji) keterampilan murid di bengkel kerja (laboratorium, studio). Kemudian dosen harus mempertimbangkan strategi pengamatan dan pengumpulan informasi penguasaan murid.

Dosen mungkin saja mengembangkan format pengamatan yang akan dipakai sebagai pedoman pengumpulan informasi. Strategi penilaian dengan jalan format pengamatan memang baik, namun bukan tanpa kendala. Misalnya, dalam suatu pengujian, jumlah murid 20 orang dan diawasi oleh dua orang dosen, bagaimana dosen pengawas tersebut dapat mengamati dengan cermat seluruh murid yang sedang menyelesaikan tugas ujiannya? Permasalahan ini tentu sulit dijawab. Oleh sebab itu, dosen harus memilih, dari semua faktor penting di atas, faktor mana yang paling kritis.

Penatar mencoba berargumentasi bahwa faktor-faktor yang perlu mendapat perhatian dalam tes praktek adalah unsur *hasil akhir, keefektifan waktu, dan keselamatan kerja* merupakan unsur utama. Hasil akhir dapat diamati setelah semua murid peserta ujian menyelesaikan tugasnya. Hasil akhir yang baik ditentukan oleh *ketelitian penyelesaian tugas* (sesuai dengan syarat-syarat pengerjaan dan toleransi). Keefektifan pemakaian waktu bisa dicatat dengan *jumlah waktu* yang dipakai murid untuk menyelesaikan tugasnya. Faktor keselamatan kerja dapat dilihat *ada tidaknya kerusakan alat dan benda kerja, serta kecelakaan kerja* selama ujian praktek dilaksanakan.

Dari tiga unsur tersebut, dosen perlu menetapkan kriteria atau standar pencapaian murid. Caranya, dosen harus mencoba membuat model terhadap tugas pekerjaan yang diberikan. Catat berapa lama dosen memerlukan waktu untuk menyelesaikannya. Kemudian murid diberi waktu sampai dengan, misalnya, satu-setengah kali waktu yang diperlukan dosen. Hasil akhir dosen juga bisa menjadi standar tampilan benda kerja. Ketelitian pengerjaan dapat memakai ukuran tertentu, misalnya standar kesalahan atau toleransi yang diizinkan. Murid yang prestasinya hebat, misalnya mereka yang dapat menyelesaikan tugas praktek dalam batas toleransi dan waktu kerjanya 15% lebih cepat dari waktu nominal, serta tidak terjadi kecelakaan kerja; demikian seterusnya.

Pada dasarnya, pengembangan instrumen pengujian praktek dilakukan dengan pengkajian terhadap kodrat dan hakekat atau 'nature' keterampilan yang ditilik. Oleh sebab itu, model pengembangan instrumen dapat dilaksanakan dengan:

- a) pengkajian keterampilan yang ingin diujikan;
- b) perumusan contoh tingkah laku melalui 'job' atau pekerjaan yang diperlukan untuk menampilkan keterampilan dalam butir a);
- c) penentuan situasi lingkungan, strategi pengamatan, dan standar, serta norma kelulusan;
- d) perumusan petunjuk pelaksanaan pengujian.

Pengkajian keterampilan dimaksudkan untuk menentukan bagian atau elemen

keterampilan apa saja yang seharusnya diliput oleh pengujian praktek. Pada tahap ini, penguji harus dapat menetapkan jenis keterampilan esensial yang seharusnya dimiliki oleh seorang lulusan. Keterampilan esensial yaitu keterampilan wajib yang harus dimiliki, sampai tingkatan minimal tertentu, lulusan sesuai dengan kesepakatan pakar profesi. Pengkajian ini dilakukan penguji (dosen) dengan memedomani GBPP dan jabatan karakteristik lulusan yang ingin diuji atau sasaran pengajarannya sesuai dengan norma profesi dan standar industri (dunia kerja).

Langkah ini pada dasarnya merupakan modifikasi analisis pekerjaan atau jabatan atau keterampilan apa saja yang harus dimiliki oleh seorang lulusan untuk suatu jabatan tertentu. Di samping keterampilan praktek, keterampilan tambahan apa saja yang harus dimiliki oleh si (calon) pekerja. Semua keterampilan esensial ini harus dapat tercakup dalam satu atau dua jenis pekerjaan yang dapat ditampilkan dalam satu atau dua periode ujian.

Permasalahan yang juga sering dihadapi dosen di ujian praktek adalah kurangnya melibatkan semua keterampilan yang telah diajarkan. Untuk ini disarankan, dosen membuat tugas ujian yang meliputi sebagian besar keterampilan yang telah diajarkan dan terangkum menjadi satu tugas (job). Kemudian dosen melakukan analisis pengerjaan (semacam analisis pekerjaan) terhadap tugas (job) yang diberikan.

Salah satu pendekatan analisis pekerjaan dapat dijelaskan sebagai berikut:

- a) Perikan jenis pekerjaan yang harus dilakukan murid atau teruji;
- b) Rincikan komponen atau bagian dari jenis keterampilan yang perlu dimiliki untuk kesuksesan penyelesaian tugas jenis pekerjaan;
- c) Perikan jenis alat, bahan, dan lain-lain hal yang diperlukan teruji untuk menjamin keterlaksanaan tugas pekerjaan saat pengujian berlangsung;
- d) Perikan prosedur pengerjaan tugas pekerjaan yang paling efisien dan efektif dan berbagai alternatifnya, terutama seandainya perlengkapan atau mesin yang diperlukan dipakai teruji lainnya (prosedur ini penting untuk penyusunan format pengamatan);
- e) Jelaskan berbagai tindakan pengamanan yang mungkin diperlukan pekerja (teruji) agar dia dapat bekerja secara aman, terutama untuk bagian keterampilan yang kritis dan perlu peringatan khusus;
- f) Perikan kondisi lingkungan tempat kerja atau praktek yang dapat menjamin kerja optimal dari para teruji;
- g) Jelaskan hal-hal penting dari setiap komponen atau bagian pekerjaan yang akan mendapat pemeriksaan dan standar pemeriksaan (penilaian) yang akan dianut. Untuk butir terakhir ini sebaiknya penguji telah mencoba membuat master benda kerja dan mencatat waktu yang diperlukan sebagai acuan;
- h) Jelaskan jumlah waktu yang diizinkan bagi setiap teruji untuk menyelesaikan tugas pekerjaannya.

Setelah analisis jabatan tersebut dilakukan, penguji harus mampu menyusun suatu perangkat uji. Perangkat uji ini harus mencakup contoh keterampilan yang harus didemonstrasikan teruji dalam selang waktu yang ditentukan. Perangkat uji harus dapat diselesaikan dalam waktu tersedia oleh penguji dan kemudian diperhitungkan (misalnya satu-setengah kali waktu yang diperlukan penguji) toleransi waktunya agar teruji yang tingkat prestasinya rata-rata dapat menyelesaikan tugas ujian dengan baik. Usahakan perangkat uji tersebut cukup sederhana, tetapi menampilkan banyak aspek keterampilan wajib dengan tingkat sesukaran yang layak bagi calon pekerja (pemula) dalam bidang uji.

Langkah selanjutnya, penguji harus mampu memerikan kondisi dan situasi

lingkungan (bengkel kerja atau laboratorium) yang diharapkan teruji mampu menampilkan kemampuan dan keterampilannya secara maksimal. Bahan dan alat apa saja yang harus disediakan penguji, catatan atau perangkat apa saja yang tidak boleh dibawa teruji, persyaratan keamanan kerja apa saja yang harus diperhatikan teruji selama pelaksanaan ujian, dan lain sebagainya yang dapat menjelaskan kondisi lingkungan tempat pelaksanaan ujian. Kemudian standar akademik atau keterampilan yang bagaimana seseorang teruji dapat dinyatakan memenuhi syarat kelulusan minimal. Bagaimana cara pengamatan dan pengumpulan data dilaksanakan? Apakah terbatas pada hasil akhir teruji dalam bentuk benda kerja jadi? Bagaimana penguji dan pengawas ujian harus mengamati kegiatan dan berbagai aspek lain yang sangat penting dari teruji yang biasanya diperlukan di dunia kerja (misalnya kerjasama calon pekerja)? Format pengamatan apa saja yang diperlukan pengawas atau pelaksana ujian? Semua pertanyaan ini harus mampu diungkapkan dan dirumuskan oleh penguji agar pelaksanaan uji profesi dapat berjalan secara lancar dan sukses.

Akhirnya, setelah contoh pekerjaan dan spesifikasi persyaratan lingkungan tempat ujian dapat dirumuskan, penguji harus mampu merumuskan petunjuk pelaksanaan ujian. Petunjuk pelaksanaan ini harus mengikat pengawas ujian (terhadap apa yang harus dilakukan) dan teruji (tingkah laku apa yang mungkin mengurangi dan menambah prestasinya) dalam tingkatan yang berbeda. Petunjuk pelaksanaan tersebut perlu memperhatikan dan mempertimbangkan kondisi rata-rata tempat kerja, sehingga teruji dihadapkan pada situasi yang didesain seolah-olah tempat kerja mereka di masa yang akan datang. Hal demikian akan menguntungkan teruji karena memberi pengalaman nyata.

8. Penyusunan Strategi Pemontenan

Bagi dosen yang menggunakan tes tertulis bentuk butir pilihan, strategi pemontenan bukan masalah. Hal ini disebabkan model pemontenannya biasanya langsung baik memakai prosedur pemberian hukuman atau bukan. Pemberian hukuman biasanya diyakini sebagai prosedur yang "fair" karena adanya kecurigaan terhadap faktor keberuntungan. Murid yang tidak memiliki ilmu berkemungkinan menjawab benar karena 'kebetulan' memilih jawaban benar. Namun, pendapat di atas tidak seluruhnya tepat sebab, yang dinamakan kebetulan tersebut menganggap semua orang memiliki peluang yang sama.

Pemontenan tes berbutir soal bentuk uraian dan tes praktek betul-betul memerlukan strategi tertentu, agar dosen dapat membatasi kemungkinan berperannya unsur subjektivitas mereka. Strategi pemontenan ini termasuk bagian yang dipersiapkan dosen begitu mereka selesai menulis butir soal bentuk terakhir ini (uraian dan tes praktek). Penulisan strategi pemontenan ini, termasuk menuliskan model jawaban yang diharapkan akan diberikan oleh mahasiswa, adalag agar dosen tidak lupa jawaban apa sebenarnya yang diharapkan akan diberikan oleh mahasiswa ketika membaca butir soal.

Penggabungan skor atau ponten untuk menentukan skor akhir atau "nilai" semester selalu diperlukan dosen. Alasannya, dosen biasanya melakukan lebih dari satu kali pengujian atau pengesanan. Di samping itu, ponten dapat juga diperoleh dari keaktifan mahasiswa dalam kelas, penyelesaian tugas-tugas terstruktur, penulisan makalah, penyelesaian tugas akhir semester, ujian tengah semester, ujian akhir semester, catatan kemajuan belajar dalam labor, dan skala perubahan sikap mahasiswa. Variasi sumber dan bentuk "ponten" atau informasi prestasi belajar mahasiswa ini perlu dicermati oleh dosen; kemudian digabungkan untuk mendapatkan gambaran utuh dan ponten "kumulatif" prestasi belajar mahasiswa yang akhirnya akan ditrasfer menjadi "nilai" akhir semester. Nilai akhir semester dalam makalah ini perlu dibedakan dengan

skor atau ponten akhir semester. Istilah nilai akhir semester mencakup kandungan arti dan makna suatu skor (angka) prestasi belajar. Sistem penilaian di perguruan tinggi biasanya mengenal sistem nilai huruf A (sangat baik) s/d E (gagal atau sangat jelek).

Untuk tujuan penggabungan, dosen seringkali perlu mengetahui kontribusi masing-masing sumber informasi terhadap variasi nilai akhir semester yang ingin dibentuknya. Dalam kondisi ini, dosen dapat mencermati sasaran akhir pengajaran dan perkuliahan yang diamanatkan oleh kurikulum sebagai pedoman. Maksudnya, sasaran akhir pengajaran biasanya diurai dan dikembangkan menjadi beberapa komponen dalam bentuk pokok dan sub-pokok bahasan. Oleh sebab itu, skor gabungan dapat dipandang sebagai totalitas informasi prestasi belajar mahasiswa dalam satu semester. Karena skor gabungan merupakan totalitas, maka model penggabungan yang paling sederhana mungkin berbentuk model penjumlahan seluruh komponen.

Penulis makalah ini menganjurkan, sebelum dosen memutuskan untuk memakai model penjumlahan dalam penggabungan skor, dosen perlu mengkaji kodrat dan hakekat mata kuliah yang dibinanya. Tujuan pengkajian ini untuk menemukan interaksi antar sub-pokok bahasan yang disajikan dalam struktur perkuliahan mingguan dan pemberian tugas mahasiswa. Kodrat dan hakekat mata kuliah ini terutama diperlukan untuk menentukan bobot kontribusi setiap komponen dalam model penggabungan skor.

Dengan demikian, setiap komponen akan mendapat perhatian yang layak sesuai dengan hasil analisis TIK dan sasaran akhir yang ingin dituju. Model pembobotan ini hendaknya dikomunikasikan kepada mahasiswa agar mereka dapat mengantisipasi semua kegiatan perkuliahan dan menyelesaikan tugas-tugas yang diberikan secara tepat.

Untuk memberi gambaran tentang model penggabungan skor dimaksud (model penjumlahan) dapat dirumuskan secara matematis sebagai berikut:

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

dimana:

a_1, a_2, \dots, a_n = bobot yang diinginkan untuk masing-masing sumber informasi yang merupakan komponen dari skor atau ponten akhir;

Y = skor atau ponten akhir semester;

X_1, X_2, \dots, X_n = skor atau ponten masing-masing sumber informasi.

Dari hasil penggabungan tersebut, Y kemudian ditransfer ke dalam sistem huruf yang berlaku.

Dalam pelaksanaannya model penggabungan di atas dapat pula dikombinasikan dengan pendekatan skor-z (z-scores), yaitu dengan jalan mengubah setiap skor komponen menjadi skor-z masing-masing komponen. Maksud pengubahan skor mentah menjadi skor-z ini tidak lain agar kontribusi masing-masing komponen sebagaimana dirumuskan dalam model pembobotan di atas dapat dicapai, karena variasi skor ternyata mempunyai pengaruh yang sangat kuat terhadap variasi skor gabungan. Pengubahan skor mentah menjadi skor-z berarti pula penyeragaman skala skor sehingga semua skor mentah memiliki varians yang sama (yaitu satu).

Model apapun yang akan dipilih dosen, penulis makalah menganjurkan agar dosen mempertimbangkan pilihannya dengan matang. Dengan demikian dosen tidak terkecoh dengan berbagai prosedur pengolahan yang rumit, namun diharapkan dosen tidak terlalu menyederhanakan persoalan ini hanya karena malas.

Dalam mentransfer skor gabungan ke dalam "nilai" sistem huruf, dosen selalu menganggap bahwa batasan A (misalnya 80-100%), B (65-79%) dan sebagainya

merupakan persentase penguasaan mahasiswa terhadap perkuliahan. Menurut penulis makalah, anggapan tersebut kurang tepat. Penguasaan materi perkuliahan hanya sekedar yang terukur oleh tes saya rasa anggapan yang terlalu disederhanakan (simplified). Penulis makalah lebih cenderung mengatakan bahwa penransferan skor tadi berkaitan dengan skor ideal (baik teoritis maupun praktis).

Bila dosen memilih pendekatan skor ideal (teoritis) maka dosen tersebut harus sudah siap dengan kemungkinan mahasiswanya tidak ada yang memperoleh "nilai" A, dan banyak yang hanya mendapat "nilai" C. Oleh sebab itu, penulis makalah lebih cenderung menganjurkan untuk memilih pendekatan skor ideal (praktis) dengan memakai skor maksimum yang dicapai mahasiswa sebagai skor ideal. Kata-kata skor maksimum ini dikaitkan dengan skor maksimum dalam "populasi" mahasiswa. Dengan demikian, kriteria dan standar penilaian yang dipakai dosen diharapkan akan dapat lebih stabil atau mantap dari waktu ke waktu dan dari tahun ke tahun. Model skor mahasiswa dalam populasi, ini pun dianjurkan untuk dipakai pada pendekatan norma melalui model PBN.

Bila pendekatan "populasi" mahasiswa akan dikembangkan, dosen dianjurkan untuk mengembangkan sistem informasi prestasi mahasiswa dalam setiap mata kuliah yang dibinanya. Sistem informasi ini membutuhkan sistem pendataan yang baik dari tahun ke tahun. Statistik prestasi mahasiswa dapat disusun dan kecenderungan sentral setiap terjadi perubahan statistik dengan mudah dapat dihitung secara manual (tanpa komputer) oleh dosen. Untuk itu dianjurkan dosen berkonsultasi dengan buku-buku statistik pendidikan.

Untuk memberi gambaran aplikasi pendekatan di atas, berikut diberikan ilustrasinya. Contoh penggabungan ponten yang ditunjukkan ini dimaksudkan untuk memperjelas berbagai konsep yang diterangkan di muka. Model penggabungan yang dipakai adalah model regresi sebagaimana dijelaskan di depan.

Misalkan seorang mahasiswa mendapat ponten rata-rata dari berbagai tugas dan ujian sebagai berikut: Tugas Rumah (83); Ujian Tengah Semester (60); Proyek Akhir Semester (70); dan Ujian Akhir Semester (90). Dalam penjelasan di awal kuliah, dosen akan memberi bobot masing-masing sumber informasi sebagai berikut: Tugas Rumah (1); Ujian Tengah Semester (2); Proyek Akhir Semester (3); dan Ujian Akhir Semester (4). Ponten akhir setelah dihitung dengan model penggabungan skor (halaman 6, di muka) diperoleh:

$$Y = [83 + (2)(60) + (3)(70) + (4)(90)]/[1 + 2 + 3 + 4]$$

$$= 77,3.$$

Seandainya dari keseluruhan mahasiswa yang mengikuti kuliah dengan menyelesaikan tugas dan ujian yang diberikan dosen mengharapkan total skor ideal teoritis sebesar 100, maka mahasiswa (yang berponten gabungan 77,3) tersebut akan mendapat persentase skor sebesar 77,3%. Harga persentase ini dapat dipakai sebagai dasar dan indikator penguasaan murid terhadap kesanggupannya menyelesaikan tugas belajar. Apabila dosen memakai standar grading IKIP Padang, maka mahasiswa tersebut akan memperoleh "grade" (indeks prestasi kuliah) C (lihat Buku Pedoman IKIP Padang 1991-1993, halaman 48).

Namun, seandainya dosen tidak memakai acuan standar skor ideal empiris, hasil kesimpulan "grade" mungkin lain. Misalkan, untuk kasus contoh di atas, dosen menemukan skor maksimum empiris yang diperoleh mahasiswa (dalam populasi mereka) sebesar 81, maka mahasiswa yang mendapat skor 77,3 akan memperoleh persentase sebesar 95,4%. Harga persentase ini apabila ditransfer ke dalam "grade" huruf, mengikuti Buku Pedoman IKIP Padang 1991-1993 (halaman 48), maka

mahasiswa yang sama akan mendapat indeks prestasi perkuliahan sebesar A.

Perbedaan kesimpulan "nilai" atau "grade" atau "indeks prestasi perkuliahan" tersebut, pada dasarnya mirip kalau tidak sama (persis) dengan kasus mengukur suhu badan manusia. Yang satu memakai "skala Celsius" sedang ukuran lainnya memakai "skala Fahrenheit". "Nilai huruf" atau "grade" yang berbeda bukan berarti maknanya berbeda, tetapi lebih disebabkan oleh "skala" yang dipakai berbeda.

9. Telaah Butir Soal

Setelah butir soal ditulis lengkap, sesuai dengan jumlah yang diinginkan, dosen perlu melakukan telaah ulang terhadap soal-soal yang ditulisnya. Hal ini dilakukan agar kesalahan teknis yang tidak perlu dapat dihindarkan. Telaah ulang ini sebenarnya bisa saja ditujukan kepada konsep yang diujikan, misalnya untuk menjawab pertanyaan apakah konsep utama yang mendasari suatu soal merupakan konsep penting dan benar. Di samping itu, pemaparan gagasan (dalam bentuk soal) sebagai wahana untuk merangsang murid berfikir perlu pula dicermati. Pertanyaan yang perlu dijawab dalam pencermatan ini adalah, menurut bidang keahlian yang diajarkan, tampilan gagasan dalam bentuk butir soal yang disajikan itu apakah sudah merupakan cara terbaik. Cara pemaparan gagasan sudah tepat atau belum, artinya apakah bahasa yang dipakai sudah efisien. Berbagai pertanyaan dan persoalan di atas hanya dapat dijawab dengan jalan melakukan telaah ulang terhadap butir soal.

Telaah ulang oleh penulis soal (dosen sendiri) mempunyai kelemahan sebab dosen bisa saja bias terhadap apa saja yang ditulisnya. Dosen bisa menganggap soalnya sudah baik! Walaupun, apabila soal tersebut disajikan kepada sejawat lain, mereka mungkin mengalami kesulitan memahami maksud penulis soal. Dampak anggapan soal sudah baik tersebut antara lain, apabila murid tidak menjawab benar, si dosen lantas menyalahkan muridnya yang "BODOH", "MALAS BELAJAR", dan sejenisnya. Anggapan ini sudah umum. Padahal bisa saja setelah ditelaah ulang dan silang oleh sejawat lainnya, pengungkapan pertanyaan kurang baik. Banyak kata yang tidak dimengerti oleh muridnya, menimbulkan pemahaman yang berbeda, dan lain sebagainya. Seandainya keadaan terakhir ini yang terjadi, maka para mahasiswa sebenarnya tidak seratus persen salah. Mungkin letak persoalannya adalah komunikasi dosen dan mahasiswa (melalui butir soal) yang tidak lancar. Oleh sebab itu, dianjurkan dosen melakukan telaah silang dengan sejawatnya.

Apabila dari telaah silang diketahui bahwa soal sudah tampil dalam bentuk terbaiknya, maka kesalahan jawaban mahasiswa bisa berarti memang mahasiswa tidak mencapai prestasi yang diinginkan. Baru kemudian analisis soal dan tes, setelah terkumpul data empiris, akan membuktikan tingkat kesulitan soal. Persoalannya kemudian berkembang, apakah butir soal yang dibuat dosen sesuai dengan "kemampuan" mahasiswanya. Apabila ternyata tes terlalu sulit, maka dianjurkan dosen membuat skala pemontenan yang disesuaikan. Problema skala pemontenan merupakan persoalan lain yang akan dibahas pada kesempatan lain.

Secara ringkas dapat disimpulkan bahwa telaah butir soal dapat dibagi menjadi telaah segi materi uji, teknis penulisan, dan penuangan gagasan menjadi tampilan soal. Dari segi materi, perlu dipikirkan ketepatan rumusan konsep yang diujikan dan ketepatan soal dengan indikator. Segi teknis mencoba melihat apakah soal mengandung kemelahan teknis (bahasa, kosa kata, hubungan pokok soal dan alternatif jawaban). Segi gagasan mencoba menganalisis kejelasan kasus persoalan yang ditampilkan dengan konsep yang diujikan.

10. Perakitan Tes

Bagian terakhir ini akan membahas perakitan butir soal menjadi suatu tes yang utuh. Tes harus disadari hanya merupakan contoh (sampel) soal dalam suatu keahlian dari keseluruhan universe (populasi) butir soal. Secara teoritis, suatu materi ajar dapat ditulis jumlah soal yang tak terhingga. Oleh sebab itu, apapun yang ditampilkan dalam suatu tes, dia hanya merupakan wakil butir soal dalam suatu universe. Dasar pemikiran ini yang dipakai sebagai landasan pengembangan rancangan tes dengan sebaran butir soal yang merata terhadap pokok bahasan yang dicakup suatu materi ajar.

Bila dosen ingin merakit tes, hanya butir soal yang "baik" saja yang dipakai. Butir soal yang baik artinya yang secara teoritis tidak mengandung kesalahan konsep, secara teknis tidak mengandung kesalahan bahasa, dan merupakan tampilan yang dapat dicerna secara mudah. Hanya butir soal yang lolos dari telaah ulang dan silang yang pantas dirakit menjadi tes.

Dalam kesempatan ini, dosen perlu pula melengkapi tes yang dibuat dengan petunjuk umum dan khusus penyelesaian tes. Misalnya, bagaimana mahasiswa menjawab, dimana dia harus menjawab, batasan yang bagaimana jawaban yang syah, dan sejenisnya. Kemudian tes siap dicetak untuk diujikan.

Daftar Pustaka

- Anastasi, A. (1976). *Psychological Testing* (edisi keempat). New York: Macmillan.
- Brennan, R. L. (1984). Estimating the dependability of the scores. Dalam Ronald A. Berk (Ed.), *A guide to criterion-referenced test construction* (hal. 292-334). Baltimore, MD: The Johns Hopkins University Press.
- Hambleton, R. K. (1984). Validating the test scores. Dalam Ronald A. Berk (Ed.), *A guide to criterion-referenced test construction* (hal. 199-230). Baltimore, MD: The Johns Hopkins University Press.
- Kumaidi (1988). *Petunjuk mempersiapkan tes hasil belajar untuk materi keteknikan*. Makalah disampaikan pada seminar staf pengajar jurusan pendidikan teknik mesin FPTK IKIP Padang 12 November 1988.
- Kumaidi (1989). *Pembuatan tes hasil belajar*. Padang: Kanwil Depdikbud.
- Kumaidi (1989). *Pengembangan tes perbuatan*. Padang: Kanwil Depdikbud Sumbar.
- Kumaidi (1990). *Uji profesi: Suatu pemikiran awal*. Makalah disampaikan pada Penataran/Lokakarya Pemantapan Proses Belajar Mengajar Pendidikan Menengah Kejuruan 18-23 November di Padang: Kanwil Depdikbud Sumbar.
- Linn, R. L. (1989). *Educational Measurement* (edisi ketiga). New York: American Council on Education.
- Mehrens, W. A., & Lehmann, I. J. (1984). *Measurement and evaluation in education and psychology* (edisi ketiga). New York: Holt, Rinehart, and Winston.
- Millman, J., & Greene, J. (1989). The spesification and development of test of achievement and ability. Dalam Robert L. Linn, *Educational Measurement* (edisi ketiga, hal. 335-366). New York: American Council on Education.
- Siegel, A. I. (1986). Performance tests. Dalam Ronald A. Berk (Ed.), *Performance assessment: Methods and applications* (hal. 121-142). Baltimore, MD: Johns Hopkins University Press.
- Sukesti, L. S. H. (1990). *Penulisan soal tes prestasi belajar bentuk soal uraian*. Jakarta: Pussisjian Balitbang Dikbud.
- Ward, C. (1980). *Designing a scheme of assessment*. Cheltenham, UK: Stanley Thornes.
- Ward, C. (1981). *Preparing and using constructed-answer questions*. Cheltenham, UK: Stanley Thornes.

**PENULISAN BUTIR SOAL
BENTUK URAIAN DAN PILIHAN**

**Oleh
Imam Sodikun**

Disampaikan Pada Semlok Pembuatan Tes & Evaluasi
Universitas Bung Hatta Padang
28 – 29 Juni 1999

Penulisan Butir Soal Bentuk Uraian dan Pilihan

Imam Sodikun

A. Pendahuluan

Penilaian merupakan bagian yang selalu dilakukan pada setiap kegiatan, termasuk dalam proses pembelajaran. Penilaian ini dilakukan setelah dosen (guru) melaksanakan proses pengajaran dan siswa melaksanakan proses pembelajaran. Hal ini perlu dilakukan karena dosen ataupun lembaga akan membuat keputusan. Keputusan ini bisa berskala makro (lembaga/sekolah) dan mikro (kelas). Keputusan makro biasanya berorientasi kepada kepentingan program pengajaran (sekolah) dan keputusan mikro berorientasi kepada kepentingan studi. Jadi penilaian bukan hanya untuk kepentingan mendapatkan angka atau skor saja, tetapi sampai pada menentukan keputusan dan tindak lanjut (Ngalim Purwanto, 984).

Selanjutnya, dalam melaksanakan penilaian ini dosen (penilai) membutuhkan data atau informasi tentang berbagai hal yang berkaitan dengan kegiatan belajar siswa untuk satu kurun waktu pembelajaran tertentu. Kurun waktu ini bisa mingguan, bulanan, catur wulan, semester dan tahunan, tergantung kepada kebutuhannya. Informasi yang diperlukan ini biasanya berkaitan dengan kemajuan (perkembangan) belajar mahasiswa yang dapat dijaring melalui pengamatan dan atau pengujian (testing).

Untuk melaksanakan kegiatan penilaian yang baik, sering diperlukan pengujian terlebih dulu. Memang kadang-kadang orang bisa melakukan penilaian tanpa didahului dengan pengujian. Misalnya guru menilai bahwa siswanya Aman malas karena sering terlambat datang tanpa menguji/menanyakan alasan keterlambatannya. Hal tersebut tentunya tidak tepat, karena penilaian kasus ini akan berbeda jika dilakukan setelah menanyakan alasan keterlambatan secara benar. Oleh karena itu keputusan yang akan diambil pun akan lebih tepat jika didahului dengan adanya pengujian untuk mendapatkan data (informasi) sebagai bahan memberikan penilaian.

Dengan demikian, hampir dalam setiap situasi, penilaian yang disertai dengan pengujian atau didasari dengan data/informasi akan menghasilkan keputusan yang lebih tepat pada sarannya. Sebaliknya pengambilan penilaian yang tanpa didasari pengujian (data) hasilnya sering tidak tepat. Di samping itu, kadar keputusan yang diambil tanpa pengukuran lebih dulu sering tidak stabil, terutama pada keputusan yang sifatnya berulang atau berlanjut. Hal ini terjadi karena pada dasarnya hasil pengujian ini

merupakan data/fakta/informasi yang dapat dipakai sebagai standar (patokan) untuk memberikan keputusan penilaian. Apabila standar ini sendiri tidak pasti, maka penilaian yang diambil juga akan labil (tidak mantap). Oleh karena itu, para pakar penilaian pendidikan menganjurkan agar keputusan pendidikan yang akan diambil (ditetapkan) seyogyanya didasarkan pada hasil pengujian/pengukuran (Ebel, 1979).

Dari uraian tersebut, maka selanjutnya dirasa perlu untuk membicarakan/mendiskusikan bagaimana mengkonstruksi suatu bahan pengujian yang baik. Bahan pengujian yang dimaksud adalah bentuk soal ujian yang bersifat uraian dan pilihan.

B. Sistematika Pembahasan

1. Pengertian
2. Perencanaan tes
3. Konstruksi butir tes
4. Pengadministrasian tes

1. Pengertian

- Test is a measure containing a set of questions, each of which can be said have a correct answer (Ebel & Frisbie, 1986)
- Test is any planned, intrusive procedure or series of tasks used to obtain observation (Sax, 1980).
- Measurement is limited to quantitative descriptions of pupil behavior (Norman E. Gronlund, 1971).
- Evaluation is the systematic assessment of the worth or merit of some objects (Daniel L. Stufflebeam dan Anthony J. Shinfield, 1985).

Ketiga istilah (unsur) tersebut saling berkaitan yang saling mendukung. Oleh karena itu penilaian hasil belajar baru dapat dilakukan dengan baik dan benar bila telah menggunakan informasi yang diperoleh melalui pengukuran hasil belajar yang menggunakan tes sebagai alat ukurnya. Hasil pengukuran ini dipakai sebagai dasar atau bahan untuk menentukan penilaian atau evaluasi.

2. Perencanaan Tes

Untuk membuat tes yang baik perlu suatu langkah yang ditempuh melalui prosedur yang sistematis. Prosedur yang baik ini dimulai dari langkah perencanaan pengujian sesuai dengan tujuan. Hal ini menjadi penting, sebab tes yang dilakukan tanpa rencana yang baik dapat menjadi sia-sia, bahkan dapat mengganggu proses pencapaian tujuan. Sebagai pedoman membuat perencanaan tes tersebut dapat ditempuh melalui pertimbangan sebagai berikut:

1. Menentukan tujuan
2. Mengambil sampel dan pemilihan butir tes
3. Tipe tes yang akan digunakan
4. Format butir tes
5. Jumlah butir tes, dan
6. Distribusi tingkat kesukaran butir tes.
7. Membuat kisi-kisi tes
8. Menulis dan merakit butir tes.

3. Konstruksi Butir Tes

Setelah ditempuh langkah pembuatan perencanaan tes selanjutnya ditempuh pembuatan (konstruksi) butir tes. Untuk mengkonstruksi butir tes, perlu didasari pada hal-hal sebagai berikut:

- 1). Tes hasil belajar harus dapat mengukur materi yang dipelajari dalam pembelajaran sesuai dengan tujuan instruksional yang tercantum di dalam kurikulum yang berlaku.
- 2). Tes hasil belajar disusun sedemikian rupa, sehingga benar-benar mewakili bahan yang telah dipelajari.
- 3). Pertanyaan (tes) ini hendaknya disesuaikan dengan aspek-aspek tingkat belajar yang diharapkan.
- 4). Tes ini hendaknya disusun sesuai dengan tujuan penggunaan tes itu sendiri (untuk keperluan apa) !
- 5). Tes ini juga hendaknya dapat digunakan untuk memperbaiki proses pembelajaran.
- 6). Tes juga disesuaikan dengan pendekatan pengukuran yang dianut (PAN / PAP).

Setelah mempertimbangkan hal-hal tersebut selanjutnya memilih/menentukan bentuk butir tes yang akan digunakan. Secara garis besar bentuk butir tes dibagi menjadi dua yaitu: butir tes bentuk uraian (esei) dan pilihan (objektif). Kedua bentuk tes tersebut mempunyai karakteristik yang berbeda, di samping juga ada persamaannya yaitu sama-sama mengukur kemampuan penguasaan materi belajar. Selanjutnya marilah kita bicarakan masing-masing bentuk tes tersebut.

a. Butir tes uraian

1). Pengertian

Tes uraian adalah butir soal yang mengandung pertanyaan atau tugas yang jawaban atau penyelesaian soal tersebut harus dilakukan dengan cara mengekspresikan pikiran peserta tes.

2) Keunggulan

Soal tipe uraian ini mempunyai beberapa kelebihan jika dibandingkan dengan tipe butir soal lain. Kelebihan yang dimaksud adalah:

- Tes uraian dapat digunakan dengan baik untuk mengukur hasil belajar yang kompleks.
- Menekankan kepada pengukuran kemampuan dan keterampilan mengintegrasikan berbagai buah pikiran.
- Lebih dapat meningkatkan motivasi bagi peserta tes.
- Lebih mudah menyusun butir soalnya.
- Sangat menekankan kemampuan menulis.

3). Kelemahan

- Kesulitan dalam memeriksa/mengoreksi dengan baik.
- Waktu mengoreksi relatif lama.
- Reliabilitas rendah.
- Jawaban kadang-kadang disertai dengan bualan.
- Tidak bisa mengungkap materi pelajaran secara luas (menyeluruh).

4). Klasifikasi tes

Tes uraian terdiri atas:

- tes uraian terbuka/bebas
- tes uraian terstruktur
- tes uraian terbatas.

5). Pedoman penyekoran

- perhatikan pembobotan/kerumitan jawaban
- ketepatan/tingkat jawaban soal
- tentukan rentangan skor.
- buat rambu-rambu jawaban.

b. Butir Tes Pilihan

1). Pengertian

Tes pilihan adalah butir tes yang telah mengandung alternatif jawaban yang harus dipilih atau dikerjakan oleh peserta tes.

2). Keunggulan

- mudah melakukan skoring
- sifatnya lebih objektif
- dapat mengukur kemampuan lebih luas (keseluruhan)
- Sudah tersedia jawaban yang diinginkan.

3). Kelemahan

- lebih sulit membuat soalnya
- lebih mudah menebak/menerka jawaban
- sulit menanyakan hal-hal yang bersifat kompleks.
- kurang memacu motivasi untuk belajar.

Lebih menekankan kepada ingatan.

4). Klasifikasi

Klasifikasi tes pilihan

i) benar- salah

ii) menjodohkan

iii) pilihan ganda

- pilihan ganda biasa

- pilihan ganda analisis hubungan antar hal

- pilihan ganda analisis kasus

- pilihan ganda kompleks dan

- pilihan ganda yang menggunakan diagram, grafik, tabel atau gambar

5) Pedoman Penyekoran

Butir soal bentuk pilihan cara memeriksanya mudah, cepat dan objektif. Namun demikian untuk lebih meningkatkan objektifitas perlu diperhatikan hal-hal sebagai berikut:

- apakah kunci jawaban sudah merupakan jawaban yang paling tepat ?
- apakah jawaban pengecoh sudah berfungsi ?
- untuk pengecekan lakukan dua atau tiga orang korektor.
- Gunakan transparan yang berisi lubang-lubang kunci jawaban.

Untuk mengkonstruksi butir tes diperlukan suatu bentuk matriks yang disebut kisi-kisi soal (tes). Matriks ini mengandung unsur-unsur sasaran belajar, materi, indikator, banyak soal dan bentuk soal. Matriks inilah dipakai sebagai dasar untuk menulis butir-butir tes. Selanjutnya penulisan matriks yang dimaksud dapat ditampilkan sebagai berikut:

Matriks Rancangan Tes

1. Matakuliah :
2. Tujuan :
3. Waktu :
4. Semester :

No.	Sasaran akhir (TIU)	Uraian Materi	Indikator (TIK)	Banyak Soal	Bentuk Soal
1.	TIU pokok Bahasan 1	1.1. ...	1.1.1....	5	Pilihan
		1.2. ...	1.1.2....	2	Uraian
		1.3. ...	1.2.1....	1	Uraian
2.	TIU Pokok Bahasan 2	2.1....	2.1.1...	1	Uraian
		2.2....	2.1.2....	4	Pilihan
		2.3....	2.2.1...	3	Pilihan
3.	TIU Pokok Bahasan 3	dst.	dst.	dst	dst

Cara pengisian matriks rancangan tes tersebut dapat dijelaskan sebagai berikut:

1. No. : cukup jelas
2. Matakuliah : cukup jelas
3. Tujuan : diagnostik, formatif, sumatif atau yang lain
4. Waktu : waktu untuk mengerjakan soal
5. Semester : ganjil/genap/tahun
6. Sasaran akhir : Tuliskan TTU setiap pokok bahasan (GBPP)
7. Uraian materi : jabaran materi menurut pokok bahasan yang diujikan
8. Indikator : ciri-ciri yang dapat diamati (TIK)
9. Banyak soal : jumlah soal untuk setiap indikator (pokok bahasan)
10. Bentuk soal : cukup jelas

Penulisan Butir Tes

Setelah matriks rancangan tes dibuat, dapat dilanjutkan dengan menulis butir tes. Dalam membuat butir tes perlu dipedomani hal-hal sebagai berikut:

- 1). pemakaian bahasa Indonesia yang baku dan benar.
- 2). penggunaan kosa kata yang tidak rancu.
- 3). menggunakan kata tanya atau perintah
- 4). Bila menggunakan bentuk pilihan, setiap alternatif jawaban memiliki akar dan dasar aspek yang sama (homogen), tetapi kunci jawaban merupakan alternatif terbaik dan benar, sedangkan pengecoh (distraktor) menunjukkan kesalahpahaman tertentu.
- 5). Setiap butir mengacu kepada TK / TTU.

Menulis Butir Soal Uraian

Tipe tes esei secara umum dapat dibagi menjadi dua jenis, yaitu tes uraian bebas (extended response) dan tes uraian terbatas (restricted response). Dasar pembedaannya adalah atas dasar besarnya kebebasan peserta tes untuk mengorganisasikan, menulis dan menyatakan pikiran dan gagasannya.

a. Tes uraian bebas.

Contoh:

- 1) Uraikan peranan pemuda dalam perjuangan mencapai kemerdekaan Indonesia sejak tahun 1908 sampai dengan tahun 1928. Dalam uraian hendaknya terdapat contoh-contoh organisasi pemuda yang ada pada jaman itu beserta nama para pemimpinnya. Uraian tidak boleh lebih dari 2 halaman.
- 2) Jelaskan jenis dan kehidupan ikan di laut!

3) Anda telah menjadi mahasiswa Universitas Bung Hatta. Coba jelaskan bagaimana jalannya perkuliahan yang pernah anda ikuti !

371.26

Umi

S:1

b. Tes uraian terbatas

Contoh:

- 1) Dalam GBHN kita mengenal ada 8 jalur pemerataan pembangunan. Sebutkan urutan jalur yang dimaksud dan pilih salah satu jalur, kemudian uraikan secara jelas beserta contoh-contohnya !
- 2) Bandingkan taktik dan strategi perjuangan mencapai kemerdekaan antara Ir. Soekarno dan Drs. Moh. Hatta !
- 3) .Sebut dan jelaskan tiga cara mencegah erosi di lahan kritis !

Menulis Butir Soal Pilihan

Secara umum tipe soal bentuk pilihan dapat dibagi menjadi 3 bentuk, yaitu:

1. Benar – Salah (true – false)
2. Menjodohkan (matching)
3. Pilihan Ganda (multiple choice)

Dari bentuk pilihan ganda dapat dikembangkan menjadi:

- a. Pilihan Ganda biasa
- b. Pilihan Ganda analisis antar hal
- c. Pilihan Ganda analisis kasus
- d. Pilihan Ganda kompleks
- e. Pilihan Ganda yang menggunakan diagram, grafik, tabel atau gambar.

Contoh Benar (B) dan Salah (S):

- B – S Bung Hatta dilahirkan di Bukittinggi : lemah.
- B – S Pemikiran Bung Hatta tentang hak asasi manusia telah diabadikan dalam pasal-pasal UUD 1945 : lebih baik.
- B – S Bila dibulatkan kedua angka di belakang koma, maka Pi bernilai sama dengan 3.41.
- B – S Bila air dipanaskan dari 15 c menjadi 50 c maka volumenya akan tetap.

2. Contoh soal tipe menjodohkan:

Cocokkan pernyataan yang ada pada kolom pertama dan kedua

Lemah :

- | Kolom pertama | Kolom kedua |
|---------------------|----------------------------------|
| 1. Ir. Soekarno | A. Pemerataan hasil pembangunan |
| 2. Pertanian | B. Mahluk laut yang menghasilkan |
| 3. Persaingan bebas | C. Proklamator kemerdekaan |
| 4. Ikan paus | D. Demokrasi liberal |

Lebih baik:

Kolom pertama

1. Soekarno
2. Suharto
3. Moh. Hatta
4. Sri Sultan Hamengkubuwono

Kolom kedua

- A. Bapak Koperasi Indonesia
- B. Bapak Palang Merah
- C. Bapak Pramuka
- D. Bapak Revolusi

3a. Contoh soal pilihan ganda biasa

Lemah:

Kebanyakan hewan hidupnya di dalam air, bernafas dengan:

- A. insang
- B. kulit
- C. paru
- D. insang dan paru.

Lebih baik:

Ikan paus bernafas dengan

- A. insang
- B. kulit
- C. paru
- D. insang dan paru.

3b. Contoh soal pilihan ganda analisis hubungan antarhal:

Untuk soal-soal berikut ini pilihlah:

- A. jika pernyataan benar, alasan benar dan keduanya menunjukkan hubungan sebab akibat.
- B. Jika pernyataan benar, alasan benar tetapi tidak menunjukkan hubungan sebab akibat.
- C. Jika salah satu dari pernyataan tersebut salah.
- D. Jika kedua pernyataan tersebut salah.

Soal (pernyataan)

Frekuensi detak nadi seseorang yang baru lari cepat akan naik

Sebab

Pada waktu lari cepat denyut nadi jantung bertambah cepat.

3c. Contoh soal pilihan ganda analisis kasus:

Kasus:

“ Kadit Lantas Polda Jatim Letkol Pol. Dra. Watie Sumarno menjelaskan jumlah kecelakaan lalu lintas di Jatim bulan Januari – Nopember 1996 sebanyak 70% kasus atau meningkat 4,87% dibandingkan tahun 1995 periode yang sama. Meningkatnya kecelakaan lalu lintas itu antara lain karena terhentinya Operasi Zebra menjadi operasi rutin lalu lintas. Di samping itu para pengguna jalan raya hanya berdisiplin jika ada petugas “

Pertanyaan

Meningkatnya kecelakaan lalulintas di Jatim bukan hanya disebabkan oleh terhentinya Operasi Zebra tetapi juga disebabkan

- A. Pengawas lalu lintas yang kurang jumlahnya
- B. Volume kendaraan di jalan makin meningkat
- C. Pengawas yang terlibat dikurangi jumlahnya
- D. Potensi polisi lalu lintas belum dikerahkan maksimal.

3d. Contoh soal ganda kompleks

Untuk soal berikut ini pilihlah

- A. jika (1) dan (2) benar
- B. jika (1) dan (3) benar
- C. jika (2) dan (3) benar
- D. jika semuanya (1,2,3) benar.

Soal

Salah satu vitamin yang larut dalam lemak adalah Vitamin A yang terdapat dalam

- (1) minyak, ikan dan telur
- (2) bayam, ikan dan telur
- (3) air susu dan wortel

3e. Contoh soal pilihan ganda menggunakan diagram, gambar, grafik atau tabel

Tabel 3. Rerata suhu dan curah hujan di kota Padang

Udara (oC	Jan	Feb	Mart	Apr	Mei	Juni	Juli	Agust	Sept	Okt	Nop	Des
Suhu udara	28,1	29,2	29,3	29,3	29,8	29,0	30,1	31,2	30,4	30,1	29,5	30,1
Curahhujan (mm)	1	4	23	86	25	27	0	1,2	2,5	4,2	3,1	3,0

Pertanyaan: Manakah yang benar untuk kota Padang?

- A. Bulan terpanas suhu udaranya adalah bulan yang sedikit curah hujannya.
- B. Setiap bulan selalu turun hujan di kota Padang.
- C. Terjadi dua kali musim hujan dalam setahun di kota Padang.
- D. Waktu yang paling baik untuk menanam padi di kota Padang adalah pada bulan Juni.

4. Pengadministrasian tes

Setelah menulis butir tes, sebaiknya dilakukan telaah ulang dan silang untuk melihat barangkali masih ada kesalahan/kekeliruan baik dari segi konsep, isi, bahasa dan sebagainya. Setelah telaah ulang dan silang dilakukan, barulah soal-soal tersebut dirakit menjadi satu tes yang siap diujikan. Di bagian awal dari tes ini harus ada petunjuk yang jelas mengenai cara mengerjakan, dan pembobotan nilai (skor) yang digunakan.

Ada dua hal yang perlu dipertimbangkan dalam pengadministrasian tes, yaitu:

1. Penyuntingan naskah tes
2. Penggandaan naskah tes

ad.1. Penyuntingan naskah tes

Suatu naskah tes terdiri atas beberapa butir soal. Penyusunan butir tes tersebut menjadi satu perangkat tes haruslah mempertimbangkan beberapa hal yang memungkinkan peserta tes dapat mengerahkan kemampuan terbaiknya dalam mengerjakan tes tersebut. Hal ini berarti bahwa pertimbangan utama dalam penyuntingan tes adalah untuk kepentingan peserta tes. Pertimbangan tersebut antara lain perangkat tes harus ada petunjuk mengerjakan soal dan cara skoringnya. Tingkat kesukaran disusun mulai dari soal yang paling mudah terus sampai yang paling sulit.

ad.2 Penggandaan naskah tes

Setelah naskah tes selesai disunting, langkah berikutnya adalah penggandaan naskah tes. Prosedur penggandaan ini harus dapat menjamin kerahasiaan naskah tes, di samping adanya jaminan hasil penggandaan tersebut tidak akan mengganggu konsentrasi peserta tes dalam melaksanakan tes.

Penggandaan tes sebaiknya terpisah antara lembaran tes dari lembaran jawaban. Pemilahan ini akan lebih memudahkan bagi peserta tes dalam menentukan tempat menjawab tes yang diberikan. Bagi dosen hal ini sangat memudahkan dalam penyekoran, karena tidak lagi digalaukan oleh lembaran tes yang dapat mengganggu penyekoran.

Penutup.

Untuk menulis butir soal perlu dipahami dulu konsep dasar tes, materi/bahan ujian, kelemahan, kelebihan tes uraian tes pilihan, dan tujuan tes itu sendiri. Setelah memilih bentuk tes, kemudian buatlah kisi-kisi soal beserta tentukan kebutuhan banyaknya butir soal, termasuk cara skoringnya. Kemudian buatlah butir-butir soal sesuai dengan kemampuan peserta tes. Selanjutnya buatlah perangkat tes yang sudah siap digunakan, lengkap dengan petunjuk mengerjakan dan penyekorannya.

Daftar Rujukan

- Ebel, R.L. & Frisbie, D.A. 1986. *Essentials of Educational Measurement*, Englewood Cliffs, N.J.: Prentice Hall.
- Gronlund, M.E. 1971. *Measurement and Evaluation in Teaching*. New York: The Macmillan
- Kumaidi. 1992. *Penyusunan Tes Prestasi Belajar*. Padang: Pusat Testing IKIP Padang.
- Linn, R.L. 1989. *Educational Measurement* (Edisi ketiga). New York: American Council on Education.
- Purwanto, Ngalim M. 1988. *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran* (edisi keempat). Bandung: C.V. Remaja Karya.
- Sax. G. 1980. *Principles of Educational and Psychological Measurement and Evaluation*. Belwont, CA: Wadworth.
- Zaidul, Asmawi dan Nasoetion, Mochl. 1996. *Penilaian Hasil Belajar*. Jakarta: Dikti, Depdikbud.

Prosedur Penilaian Hasil Tes

Anas Yasin

**Disampaikan pada Seminar dan Lokakarya Sistem Penilaian Hasil Tes
Universitas Bung Hatta
28 – 29 Juni 1999**

PROSEDUR PENILAIAN HASIL TES

Anas Yasin

Pendahuluan

Dalam dunia pendidikan, penilaian hasil belajar merupakan sesuatu yang mendapat perhatian oleh para ahli pendidikan. Penilaian adalah alat ukur untuk menentukan apakah pembelajar telah berhasil dengan proses belajarnya. Di lapangan terlihat kenyataan bahwa sulit untuk mendefinisikan keberhasilan pembelajar. Tidak ada kriteria yang jelas dan tepat secara matematis dalam menilai keberhasilan pembelajar. Terdapat cara yang beragam dalam memberikan nilai tergantung pada bentuk tes yang diberikan. Ada bermacam bentuk tes yang diberikan dosen dalam ujian, diantaranya adalah pilihan ganda dan esai. Cara menilai hasil tes pilihan ganda yang pernah teramati adalah:

- (1) dosen menghitung jumlah jawaban yang benar dari sejumlah soal dan kemudian mengalihkan langsung dari nilai yang ada menjadi nilai huruf (A, B, C, D, dan E) dengan membuat patokan sendiri. Misalnya, dari tes dengan jumlah soal 100, dosen menetapkan 75 ke atas mendapat nilai A, dari 60 sampai dengan 74 ~~74~~, B, dari 50 sampai 59, ~~59~~, C, dan seterusnya.
- (2) Jika jumlah soal bukan 100 (misalnya 40), dosen menghitung yang benar dan mentransferkannya ke dalam bentuk persentase. Jika dari 40 soal, misalnya yang betul 25, nilai ini ditransfer menjadi $25/40 \times 100\% = 62.5$. Sesudah itu, dosen membuat kriteria sendiri dalam mengalihkan nilai angka menjadi nilai huruf, seperti pada butir 1 di atas.
- (3) Dari kedua bentuk di atas, dosen mungkin juga mengacu kepada patokan yang sudah ditetapkan oleh universitas dalam mentransfer nilai dari angka menjadi A.

Cara penilaian yang dilakukan dosen yang terpantau di lapangan untuk tes esai adalah:

- (1) Setelah membaca secara keseluruhan hasil tes, dosen serta merta menetapkan nilai huruf, misalnya B, untuk jawaban yang dianggap baik dan nilai C atau D untuk jawaban yang menurut dosen kurang bagus.
- (2) Dosen memberikan angka pada setiap jawaban esai dengan bobot tertentu secara merata dan kemudian menjumlahkannya. Dari hasil yang diperoleh dosen, ditetapkan nilai hurufnya.
- (3) Dosen memberikan bobot yang berbeda pada masing-masing soal, kemudian menjumlahkan nilainya dan memberikan nilai huruf menurut patokan yang ditetapkan sendiri.

Dari beberapa contoh di atas permasalahan yang kelihatan adalah dosen tidak mempunyai teknik tertentu dalam memberi nilai. Makalah yang sederhana ini menyuguhkan sebuah teknik menilai hasil tes yang dikemukakan oleh Cronbach (1984).

Beberapa Kendala dalam Menilai dan Pemecahan Masalahnya

Seperti yang telah dikemukakan di atas, beberapa masalah dalam menilai ditemukan di lapangan. Pada tes esai, misalnya, penilaian sangat subjektif. Untuk tes yang sama, dosen yang berbeda mungkin memberikan nilai yang berbeda pula. Seorang dosen memberikan nilai tinggi sedangkan yang lain memberikan nilai rendah. Juga terdapat perbedaan

terdapat yang demikian. Kita dapat mengatakan bahwa tinggi X dua kali tinggi Y. Tetapi kita tidak dapat menentukan dengan pasti bahwa penguasaan X tentang suatu materi kuliah dua kali penguasaan Y yang bernilai separo dari nilainya.

Argumentasi yang sama berlaku untuk tes penalaran. nilai 80 berarti dua kali nilai 40. Tetapi tes tidak mencakup semua permasalahan yang dapat dipecahkan oleh setiap orang. Andaikata semua pemecahan masalah yang ada diujikan, rentangannya mungkin menjadi 140 - 180 atau 1040 - 1080. Dalam bernalar, anak kecilpun sudah memperlihatkan kemampuan bernalarnya. Angka nol untuk suatu kemampuan berarti tidak ada kemampuan sama sekali.

Perbedaan skor mentah biasanya tidak memperlihatkan jarak yang sebenarnya antara individu. Misalnya, pada nilai sebuah matakuliah, Adam mendapat nilai 53, Tantri 56, dan Hasan 59, selisih nilai sama, yaitu 3. Apakah memang benar pada realitanya perbedaan kemampuan antara Adam dengan Tantri sama dengan antara Tantri dengan Hasan? Kita tidak yakin betul dengan hasil ini, karena perbedaan angka tergantung kepada item yang digunakan. "Perbedaan yang sama" akan bermakna hanya apabila kriteria telah ditetapkan. Interval yang sama pada suatu skala tidak akan sama pada skala lain.

Dapat disimpulkan bahwa nilai yang "adi!" yang diperoleh mahasiswa tergantung pada apakah tes yang diberikan sudah mendekati baku, walaupun tidak mencapai tingkat baku yang diharapkan. Jika cara penulisan tes sudah mengikuti cara yang telah ditetapkan, untuk menghindari salah-interpretasi nilai mentah berikut ini dikemukakan jenis-jenis acuan dalam penilaian.

Acuan Norma, Acuan Domain, dan Acuan Patokan/Kriteria

Acuan norma adalah acuan yang menjelaskan posisi seseorang di antara yang lain yang mengambil tes; *acuan domain* adalah acuan yang menjelaskan tingkat kesulitan yang dapat dicapai dalam sejenis tugas yang diberikan padanya; sedangkan *acuan kriteria* adalah angka sebagai penanda apakah seseorang dapat atau tidak dapat mencapai patokan tertentu yang telah ditetapkan.

Contoh:

Acuan norma:

Bahar lebih baik daripada 80% teman-temannya yang ikut mengambil matakuliah Pemasaran pada tingkat tertentu.

Acuan domain:

Nilai Bahar menunjukkan bahwa dia dapat mengerjakan dua per tiga dari semua soal yang diberikan pada tingkat kompleksitas yang ditentukan.

Acuan Patokan/Kriteria:

Mahasiswa yang telah mencapai tingkat tertentu dapat dikatakan telah mencapai kriteria yang ditetapkan untuk dapat terus ke tingkat berikutnya (Nitko, 1980).

Hampir semua test dapat diinterpretasikan dengan menggunakan tiga jenis penilaian di atas. Pemilihan yang akan dipakai dalam penilaian tergantung pada teknik yang mana yang relevan dengan yang diterapkan oleh penguji.

Aspek Statistik dalam Penilaian

Ada dua jenis statistik yang dapat dipakai dalam penilaian: (1) persentil dan (2) angka baku (standard score).



Persentil

Ranking persentil menggambarkan proporsi kelompok yang berada di bawah seseorang. Apabila 37 orang dari 40 orang berada di bawah seseorang, ini berarti 2,5 berada di atasnya dan 37,5 berada di bawahnya. Karena 37,5 adalah 94% dari 40, maka angka persentilnya adalah 94. Dengan demikian orang yang berada di tengah kelompok adalah pada persentil yang ke 50 (yang disebut median). Langkah-langkah yang diikuti dalam menentukan persentil adalah sebagai berikut:

1. Mulai dengan nilai mentah; tentukan nilai tertinggi dan terendah; hitung rentangannya:

Contoh:

54	42	35	51	50	66	36	47	21	40	58
32	48	66	35	25	45	41	49	46	40	51
48	57	31	53	50	38	53	52	58	28	45
51	49	42	41	45	60	24	46	50	56	38
48	38	41	44	57	41	37	17	49	64	41
47	28	49	22	55	52	43	59	43	25	25
33	51	54	61	26	43	56	43	40		

Nilai tertinggi = 66; Nilai terendah = 17; Rentangan = 49

2. Apabila rentangannya lebar, pakai "class interval". Jumlah class interval yang disarankan adalah 15 ke atas.
3. Teli masing-masing nilai.
4. Tulis jumlah teli di dalam kolom yang disediakan.
5. Untuk menentukan frekuensi kumulatif, jumlahkan dari bawah. Tulis jumlah N.
6. Untuk menentukan persentase kumulatif, bagi frekuensi kumulatif dengan N, kalikan dengan 100.
7. Buatlah diagram persentase kumulatif terhadap nilai.
8. Buatlah garis kurva sesuai dengan angkanya.
9. Tentukan persamaan persentil sebuah nilai dengan membaca kurva

Contoh:

Nilai	Tali	Frekuensi	Frekuensi Kumulatif	Persentase Kumulatif
65 - 70	//	2	75	100
60 - 64	///	3	73	97
55 - 59	### ///	7	70	93
50 - 54	### /// ///	13	62	83
45 - 49	### /// ///	14	49	65
40 - 44	### /// ///	16	35	47
35 - 39	/	6	19	25
30 - 34	/// /	3	13	17
25 - 29	///	6	10*	13
20 - 24	### /	3	4	5*)
15 - 19	///	1	1	1
	/	-----	0	0
		75 N		

* 10 kasus berada di bawah 29.5

*) 5% kasus berada di bawah 24.5

Norma merupakan kerangka acuan. Apabila median berada jauh di bawah norma umum, pengulangan ujian mungkin dapat dilakukan atau perencanaan pengajaran harus diubah. Norma merupakan dasar perbandingan seluruh tes.

Angka Baku

Ada dua hal yang terlibat dalam penentuan angka baku dalam penilaian hasil tes: (1) rata-rata dan (2) simpangan baku.

Rata-rata (\bar{X}) dan Simpangan Baku (SD)

Cara kedua untuk menyimpulkan performansi kelompok adalah dengan menggunakan rata-rata (\bar{x}) dan simpangan baku (standard deviation = SD). Rata-rata adalah hasil pembagian jumlah keseluruhan angka dengan banyaknya angka. Rata-rata disebut varian distribusi. SD menggambarkan penyebaran angka. SD adalah rata-rata penyimpangan dari rata-rata kelompok. Dengan menggunakan SD, kita dapat menentukan berapa jauh angka seseorang menyimpang dari rata-rata kelompok.

Rumus:

$$\bar{x} = \frac{\sum X}{N}$$

$$SD = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N - 1}}$$

Contoh:

Misalkan ada 10 pengikut tes dengan nilai berikut ini:

46 47 47 48 48
48 50 50 55 60

$$\Sigma X = 499$$

$$N = 10$$

$$\bar{X} = 499 : 10 = 49.9$$

X	X ²
46	2116
47	2209
47	2209
48	2304
48	2304
48	2304
50	2500
50	2500
55	3025
60	3600
$\Sigma X = 499$ $(\Sigma X)^2 = 249001$	$\Sigma X^2 = 25121$

$$\begin{aligned} SD &= \sqrt{\frac{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}{N - 1}} \\ &= \sqrt{\frac{25121 - \frac{249001}{10}}{10 - 1}} \\ &= \sqrt{\frac{220.9}{9}} \\ &= \sqrt{24.54} \\ &= 4.954 \end{aligned}$$

SD dipakai ketimbang rentangan untuk menentukan penyebaran nilai karena SD lebih stabil dari satu sampel ke sampel yang lain.

Skala Konversi

Skala nilai baku yang berdasarkan rata-rata dan simpangan baku mempunyai tujuan yang sama dengan skala persentil. Nilai baku memperlihatkan berapa SD seseorang

berada di atas atau di bawah rata-rata. Pada contoh di atas didapatkan angka rata-rata (χ) 49.9 dan angka SD, 4.954.

Berikut ini adalah langkah-langkah mengkonversikan nilai mentah menjadi nilai baku dengan rata-rata 0 dengan masing-masing SD di atas rata-rata dihitung sebagai satu unit atau disebut dengan skala "nol + satu". Konversi z, walaupun penting dalam perhitungan statistik, tetapi jarang dipakai dalam penilaian hasil tes. Nilai tes biasanya ditempatkan pada skala "50 + 10" -- di mana rata-ratanya 50 dan SD sama dengan 10. cara ini memungkinkan untuk menyatakan setiap nilai dalam angka positif. Langkah-langkah adalah seperti berikut:

1. Mulailah dengan nilai mentah yang akan dikonversikan.

Misalkan $\chi = 49.9$ dan $SD = 4.94$.

2. Untuk mendapatkan nilai z, dicari penyimpangan masing-masing nilai mentah (t) dari χ dibagi dengan SD. Misalkan $t = 55$

$$z = \frac{t - \chi}{SD}$$

$$= \frac{55 - 49.9}{4.954}$$

$$= \frac{55 - 49.9}{4.954}$$

$$= \frac{5.1}{4.954}$$

$$= \frac{5.1}{4.954}$$

$$= 1.03$$

3. Untuk mendapatkan angka pada sakala 50 + 10, nilai z dikalikan dengan 10 dan ditambah 50.

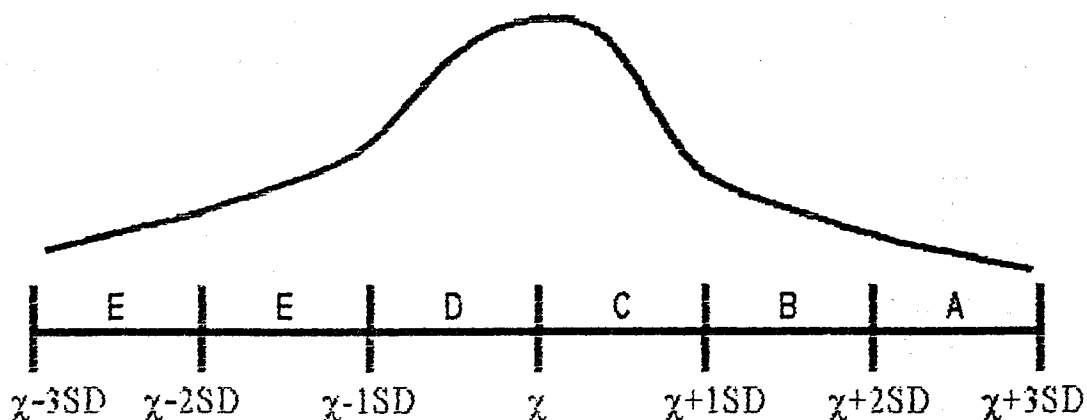
$$\begin{aligned} & 50 + 10(z) \\ &= 50 + 10(1.03) \\ &= 50 + 10.3 \\ &= 60.3 \end{aligned}$$

4. Untuk mengalihkan nilai angka menjadi nilai huruf, dapat dipedomani patokan yang telah diberikan oleh masing-masing institusi yang mungkin mempunyai acuan yang berbeda. Contohnya, batas lulus yang diberikan oleh UNAND mungkin berbeda dengan yang diberikan oleh UBH atau IKIP Padang. IKIP Padang, misalnya, menggunakan skala berikut:

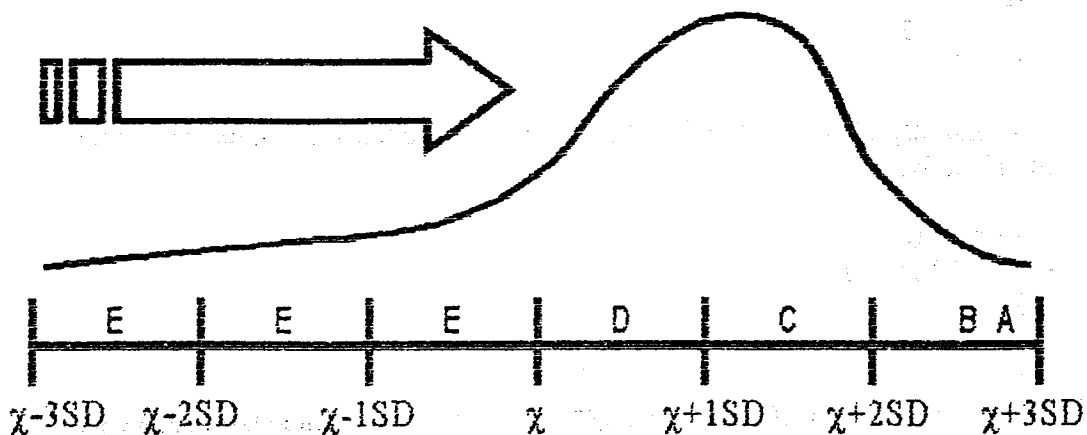
90 - 100 %	A	55 - 64 %	D
80 - 89 %	B	< 55 %	E
65 - 79 %	C		

Pengalihan Nilai Mentah menjadi Nilai Huruf

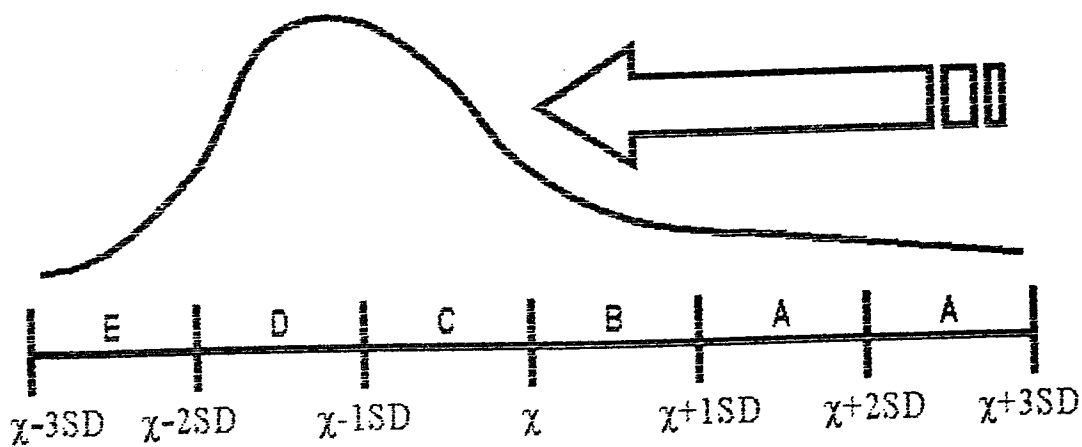
Nilai mentah secara langsung dapat dialihkan menjadi nilai huruf tanpa memakai skala konversi 50 ± 10 dengan menetapkan posisi angka dari χ ke atas dan dari χ ke bawah. Penetapan angka dimulai dari angka yang terletak antara χ dan 1SD, ($\chi + 1SD$) dan ($\chi + 2SD$), dan ($\chi + 2SD$) dan ($\chi + 3SD$) kemudian angka yang terletak antara χ dan -1SD, ($\chi - 1SD$) dan ($\chi - 2SD$), dan ($\chi - 2SD$) dan ($\chi - 3SD$), seperti yang terlihat pada diagram di bawah ini:



Apabila kurva menunjukkan posisi normal maka kita dapat meletakkan nilai A pada posisi antara ($\chi+2SD$) dan ($\chi+3SD$) dan seterusnya ke bawah sampai dengan E.



Tetapi apabila kurva menunjukkan bahwa angka menumpuk ke atas, ini mengindikasikan bahwa tes yang kita berikan terlalu mudah dan terlalu sederhana. Tes yang begini tidak menggambarkan tes yang representatif. Dalam keadaan yang seperti ini, kita dapat menggeser nilai B, C, D, dan E ke posisi di atasnya, seperti pada gambar di atas.



Apabila kurva memperlihatkan angka menumpuk ke bawah, ini menunjukkan bahwa ada kesalahan yang terjadi, seperti tes yang kita berikan terlalu sulit atau pembelajar tidak secara normal menguasai bahan pelajaran karena kesalahan strategi mengajar yang pada gilirannya harus dibenahi oleh dosen. Dalam keadaan yang begini kita dapat pula menggeser nilai satu langkah ke bawah.

Penggeseran nilai seperti di atas dilakukan apabila dosen tidak mungkin membuat sebuah tes yang baku (standar). Pembuatan tes yang baku membutuhkan pengujian dan percobaan berkali-kali. Setelah dicobakan berkali-kali dengan persyaratan pembuatan tes yang benar dan hasilnya sudah reliabel, penilaiannya pun dapat kita percaya dan kita dapat menggunakan teknik penilaian yang pertama di atas.

Penutup

Masih terdapat sistem penilaian yang belum baku yang dipakai oleh dosen dalam menilai hasil ujian mahasiswa, terutama ujian tengah dan akhir semester. Untuk mendapatkan hasil penilaian yang absah, sistem penilaian yang baku harus diterapkan. Ini tentu saja harus didahului dengan pembuatan tes yang baku pula.

Referensi

- Cronbach, Lee J. (1984). *Essentials of Psychological testing*. New York: Harper and Row, Publishers.
- Downing, Douglas. (1983). *Statistics: the Easy Way*. Woodbury: Barrons.
- Elzey, Freeman F. (1971). *A Programmed instruction to Statistics*. Belmont: Brooks/Cole Publishing Company.
- Gay, L.R. (1987). *Educational Research*. Columbus: Merrill Publishing Company.
- Hatch, Evelyn. (1982). *Research Design and Statistics for Applied Linguistics*. London: Newbury House Publishers, Inc.

Nunnally, Jun C. (1978). *Psychometric Theory*. New York: McGraw-Hill Book Company.

Runyon, Richard P. (1976). *Fundamentals of Behavioral Statistics*. London: Addison-Wesley Publishing Company.