

PENELAHAAN KEMBALI STRATEGI PENILAIAN ACUAN NORMA (PAN) DAN PENILAIAN ACUAN PATOKAN (PAP) SEBAGAI PENDEKATAN DALAM PENILAIAN HASIL BELAJAR

Oleh
IKSAN WASESO

Abstrak

Menilai hasil belajar mengajar dalam proses pendidikan formal merupakan salah satu kewajiban bagi guru dan dosen. Namun dalam prakteknya hal itu lebih sering dilakukan secara praktisnya saja, dan dilupakan dasar konsepnya, sehingga kadar manfaatnya menjadi kurang tepat guna. Oleh karena itu kiranya perlu ditinjau kembali konsep dasar apa kiranya, yang secara sadar dan konsekuen, yang sedang dipakai sebagai pendekatan dalam menilai hasil belajar-mengajarnya. Persoalan kuantitas dan kualitas lulusan program pendidikan nampaknya sebagai dua hal yang saling meniadakan. Namun sebenarnya keterpaduan sistem perkuliahan-pengajaran berdasar kompetensi, prinsip belajar tuntas, dan kepastian model PBM-nya akan dapat menyatukan tingkat kualitas dan kuantitas yang cukup tinggi, tanpa melalaikan implikasinya pada pendekatan penilaian prestasi belajarnya, beserta dua prinsip: kontinuitas dan komprehensif. Pokok bahasan ketiga, yang terakhir namun tak kurang pentingnya ialah bahwa penyusunan alat ukur dan alat penjajagannya dalam pendekatan yang merujuk kepada patokan dan kepada norma kelompok, dalam beberapa hal memiliki ciri sifat yang berbeda. Sementara itu lebih sering terjadi prosedur pengujian alat tersebut masih mengikuti pola tradisional prosedur pengujian alat tes psikologika, yang perlu pengamatan khusus kalau hendak dipakai untuk menguji THB. Penelaahan kembali, memahaminya, dan mempraktekkan strategi PAN dan PAP dalam penilaian prestasi belajar dengan segala implikasinya, akan meningkatkan kesahihan dan keterandalan informasi sebagai umpan balik bagi peningkatan kualitas PBM dan keberhasilannya.

PENDAHULUAN

Dalam pembahasan ini telah dipilih jenis permasalahan yang aktual-empirik yang dihadapi sehari-hari dalam proses pelaksanaan program perkuliahan. Justru karena sudah terbiasa melakukannya

itulah maka menjadi tak terasakan lagi sebagai suatu masalah, antara lain hal penilaian prestasi belajar program perkuliahan.

Lebih sering program penilaian ini dilaksanakan secara praktis saja, dan tanpa disadari *konsep dasar* yang melandasinya, sehingga tidak jarang terjadi praktek penilaian meleset dari tujuan pengajaran berdasar kompetensi, dan mempertimbangkan hanya besarnya jumlah lulusannya saja, yang seharusnya juga kualitasnya.

Dalam bagian I makalah ini, akan dikemukakan praktek nyata pendekatan yang diambil para dosen pada umumnya, berdasar hasil pengamatan dan wawancara langsung alamiah dengan mementingkan hal masih ada tidaknya, terasa tidaknya permasalahan pendekatan penilaian hasil belajar. Dalam bagian II, diketengahkan masalah konsep dan praktek keterpaduan pengajaran-perkuliahan berdasar kompetensi, belajar tuntas, dan model proses belajar-mengajar (PBM)nya. Masalah yang nampak penting pada persoalan penentuan tujuan khusus pengajarannya, ialah pada pengajaran sains dan nonsains dalam kaitannya dengan tujuan *pembinaan belajar* pada para mahasiswa.

Mengingat tujuan pengajaran berdasar kompetensi, maka pendekatan penilaian yang tepat ialah penilaian acuan atau rujukan patokan (PAP). Oleh karena itu sesuai dengan ciri keterarahan proses pengajarannya, perlu dikemukakan fungsi penilaian formatif dan sumatif (Bagian III). Dalam praktek, nyatanya ujian tengah semester seringkali dianggap sama dengan penilaian formatif tanpa dilaksanakan tindak-lanjutnya. Nampaknya sumber kesesatan praktek penilaian hasil belajar, terletak pada kurang difahami dan oleh karena itu menjadi kurang disadarinya konsep dasar yang melandasinya, yaitu PAN dan PAP sebagai pendekatan penilaian dan pengukuran prestasi belajar. Bagian IV makalah ini mencoba menganalisis pengertian teoritis-praktis Penilaian Acuan Norma dan Tes Acuan Norma (PAN dan TAN) Penilaian Acuan Patokan dan Tes Acuan Patokan (PAP dan TAP) yang menjadi fungsi pengertian pengukuran dan penilaian, dan bagian V sebagai penutup mencoba memberikan gambaran menyeluruh-merangkul dan memberikan rumusan masalah secara diagnotis untuk kemudian menyarankan suatu upaya *pengatasannya*, meski masih bersifat tentatif. Semoga pembahasan ini memberikan sumbangan perluasan horizon dan kemandirian perspektif dalam melaksanakan tugas mengukur dan menilai prestasi belajar para mahasiswa.

BAGIAN I

PAN ATAU PAP : SUATU DESKRIPSI VERIFIKATIF

Dalam menetapkan nilai sebagai keputusan ujian tengah semester dan semester, sebagian terbesar penguji menggunakan "rumus": seberapa banyak yang ditanyakan dapat terjawab sesuai dengan pengertiannya, kehendaknya atau yang diingatkannya sebagai kunci-jawab- betul-nya. Kalau jawabannya tidak "berkenan", maka akan diberikannya "nilai" dibawah 5,6 dan yang diuji harus mengulang. Rumus ini menjadi "rumus umum" hampir untuk segala jenis mata kuliah. Dasar pengambilan keputusan itu, "terasa" lebih pasti dan mantap, kalau bentuk alat ujinya tergolong *tes objektif*. Kalau bentuk "alat" ujinya adalah tes karangan, maka dasar pengambilan keputusan nilainya menjadi terasa kurang mantap, sebab besar kemungkinannya dipengaruhi perasaan subjektif penilaiannya, sehingga sering disebut sebagai *tes subjektif*. Namun rumus umumnya, disadari atau tidak, tetap sama yaitu 5,6 sebagai batas penerimaan, yang telah ditetapkan lebih dulu itu.

Penggunaan bentuk tes objektif, pada umumnya adalah bentuk tes objektif *buatan sendiri* dan *langsung dipakai*, tanpa melalui tahap pengujian taraf kesahihan dan keterandalannya lebih dulu. Pada umumnya, kurang disadarinya *sikap percaya* bahwa telah memenuhi syarat kesahihan dan keterandalan tes objektif buatan sendiri itu, telah mencerminkan sikap anutan pendekatan PAP, yaitu bahwa kesahihan tes objektif buatan sendiri itu telah sah-isih, sah-hukum, dengan *sarat anggapan* bahwa dirinya telah ahli dalam bidang studi yang dikuliahkan dan diujinya itu, serta telah faham prinsip-prinsip penyusunan tes objektif sederhana itu. Namun umumnya, sifat keterandalan tes objektifnya itu belum dikaitkan dengan besarnya kesesatan-baku $E = SB \sqrt{1-rtt}$, karena belum diketahui besarnya taraf keterandalannya rtt. Hal ini berakibat pada angka-biji, dan oleh karena itu juga nilainya menjadi tidak mantap, tidak dapat dipercaya sepenuhnya, namun penguji tetap dengan "berani" menetapkan nilai 5,6 atau 56% dari jumlah soal, menjadi batas penerimaan kelulusannya.

Dari gambaran sikap anutan pendekatan penilaian itu, tidak nampak kecenderungan pada ciri pendekatan PAN yang *berpatokan relatif* terhadap keadaan "apa-adanya" kelompok sebagai acuan atau rujukan norma penilaiannya. Yang nampak ialah kecenderungan-

an kepada pendekatan PAP yang mengacu kepada *kepastian patokan* penilaiannya. Dikatakan "kecenderungan" saja, karena memang belum sepenuhnya memenuhi secara konsekuen pendekatan PAP (uraian bagian III dan IV mendatang).

BAGIAN II

KETERPADUAN SISTEM PERKULIAHAN BERDASAR KOMPETENSI, BELAJAR TUNTAS DAN MODEL PBM: IMPLIKASINYA PADA PENDEKATAN PENILAIAN

Dalam pengembangan sistem pendidikan yang baru (?), dituntut terselenggaranya pendidikan dan pengajaran yang berdasar pada penguasaan kompetensi, yaitu kebiasaan yang "diakui" dalam pengetahuan dan atau keterampilan tertentu. Oleh karena itu, penyediaan dan penyajian mata kuliahnya hendaknya bertujuan memperkuat penguasaan kompetensi yang diperlukan untuk menjalankan tugas-tugasnya nanti. Hal itu berarti bahwa bahan-bahan yang terkandung di dalam setiap mata kuliah yang disajikan harus jelas menunjang bagian kompetensi penting tertentu, dalam rangka tercapainya keseluruhan kompetensi yang diperlukan. Dalam rangka pengembangan sistem pendidikan baru ini, salah satu tujuan penilaiannya ialah untuk menjajagi tingkat pencapaian penguasaan kompetensi yang menjadi tujuan tiap-tiap mata kuliah yang disajikan itu. Karena pendekatan pendidikan dan pengajaran berdasar kompetensi ini berorientasi pada "tingkat penguasaan"-nya, bukan pada tingkat keadaan kemampuan subjek belajarnya, maka kiranya pendekatan penilaian yang tepat ialah pendekatan PAP, dan rumus *tujuan penilaiannya menjadi*: untuk mengetahui-mencek telah dicapai atau belum penguasaan *patokan kompetensi* yang telah ditetapkan untuk tiap-tiap mata kuliah yang disajikan.

Pendirian konsep ini tentu menjadi sulit untuk bisa menerima lagi 56% jawab-betul atau nilai 5,6 sebagai batas minimal kompetensi patokannya, meskipun boleh dinyatakan lulus, yang dengan meragukan itu. Mengenai berapa besarnya patokan minimal kompetensi-patokan, hal itu bergantung kepada besar-kecilnya kepentingan sumbangan kompetensi yang dapat ditarik dari tiap-tiap mata kuliah, terhadap *satuan-terpadu-kompetensi* yang ingin dicapai dan telah ditetapkan sebelumnya.

Belajar tuntas merupakan konsep strategis PBM yang berkeyakinan bahwa pemberian kesempatan belajar yang sesuai dengan kemampuan pelajarnya, dan keterlibatan pelajar dalam proses belajar-mengajar yang berkualitaslah, yang akan menentukan keberhasilan belajar. Dengan strategi belajar tuntas ini, diharapkan 75%-95% dari pelajar peserta dapat mencapai tingkat penguasaan bahan pelajaran perkuliahannya sebesar katakanlah 80%-90% (Modul 26 DIK Akta V B, Penggalan 2) Ini berarti bahwa konsep belajar tuntas telah menetapkan *rentangan batas minimal* patokan kompetensi, untuk dapat dikatakan telah menguasai bahan pelajaran-perkuliahan. Ada beberapa ciri strategi belajar tuntas, sama dengan strategi pengajaran berdasar kompetensi yaitu bahwa: (i) tujuan pengajaran ditetapkan lebih dulu, (ii) memperhatikan-memperhitungkan perbedaan individu, dan (iii) pendekatan penilaiannya adalah pendekatan PAP yang dilakukan *secara kontinyu*. Perbedaannya ialah pada penetapan besarnya patokan penguasaan bahan, dan yang penting pelaksanaan pendekatan penilaiannya, secara kontinu (dibahas dalam bagian III).

Strategi PBM pada konsep belajar tuntas dan pengajaran berdasar kompetensi, sama-sama diarahkan kepada pencapaian tujuan pengajaran-perkuliahan. Oleh karena itu, kegiatan PBM dan instrumen evaluasinya harus juga berorientasi kepada tujuan pengajaran-perkuliahan yang telah ditetapkan lebih dulu itu. Untuk jenis mata kuliah sosial dan humaniora tidak nampak timbul persoalan. Pada mata kuliah pendidikan sains yang menggunakan metoda "menemukan-sendiri", akan sulit diterima penetapan lebih dulu dari tujuan pengajarannya itu karena hal itu akan melumpuhkan kreativitas, yang justru kemampuan yang kreatif itu harus dikembangkan melalui proses belajar. Lebih-lebih lagi para mahasiswa telah dianggap sebagai manusia penalar yang bertanggung jawab untuk mengembangkan kemampuan penalarannya sendiri. Oleh karena itu, untuk menjajagi keberhasilan proses belajar kreatif, tidaklah cukup kiranya hanya dengan bentuk pengukuran tes objektif sederhana, terutama dalam ilmu-ilmu sosial dan humaniora.

Dapatlah disimpulkan sementara bahwa praktek dan konsep perkuliahan berdasar kompetensi, belajar tuntas, dan model PBM hendaknya merupakan paduan yang selaras dengan pilihan pendekatan penilaiannya.

BAGIAN III

FUNGSI PENILAIAN FORMATIF DAN SUMATIF DALAM PERKULIAHAN BERDASAR KOMPETENSI

Kegiatan penilaian hasil belajar yang sampai kini berjalan, sebagian terbesar masih menuruti sistem terbuka tertutup. *Terbuka*, dengan mengumumkan dan menempelkan nilainya di papan pengumuman; *tertutup*, dengan mengabaikan fungsi umpan balik nilai ujiannya, baik pada tengah semester, apalagi pada ujian semesternya. Sikap tertutup itu benar-benar menutup kemungkinan perbaikan cara belajar dan atau cara penyajian secara objektif praktis-empiris, barangkali juga pilihan bahannya kurang sesuai. Hampir tidak pernah secara sadar dilakukan ujian-penjajagan setiap satuan mata kuliah "karena beberapa hal", antara lain sibuk mengasuh banyak mata kuliah yang tidak semuanya menjadi mimbar keahliannya, atau belum membudaya dalam penyiapan SAP (Satuan Acara Perkuliahan) sehingga keterarahan proses perkuliahannya sulit dibayangkan dan dirancangan satuan-satuannya. Kalau demikian maka penilaian secara kontinu dalam rangka pembinaan pencapaian tingkat penguasaan patokan kompetensi, tidak dapat dilakukan. Disinilah letak fungsi *penilaian formatif* sebagai tes diagnostik kemajuan belajar, yang mampu mengidentifikasi kesalahan-kesalahan belajar secara khusus setiap pesertanya. Dari informasi hasil tes formatif ini sebagai umpan balik, dapat dirancangan upaya perbaikan dan pembinaan belajar-mengajarnya. Oleh karena itu tes formatif ini sering disebut sebagai tes pembinaan, yang lazim dipakai dalam rangka pendekatan PAP dalam proses belajar tuntas. Di samping itu, tes pembinaan juga mempunyai fungsi mengukuhkan cara belajar yang sudah betul, yang cocok, dan fungsi korektif terhadap butir-butir tes formatif itu sendiri; ketiga fungsi tersebut dapat segera nampak dengan menyusun tabel analisis butirnya.

Kalau pendirian perkuliahan berdasar kompetensi itu betul-betul dilaksanakan, maka pendekatan penilaian yang paling tepat ialah pendekatan PAP, dengan melakukan *ujian-ujian pembinaan* secara kontinu sebagai alat diteksi kemajuan belajar. Kenyataan yang terjadi sekarang ialah dalam penilaian menggunakan prinsip pendekatan PAP, namun sepanjang proses pelaksanaan program perkuliahan tidak atau belum secara sadar dilakukan ujian pembi-

naan (tes formatif) dengan segala fungsinya, bahkan justru sikap pendekatannya ialah PAN, yaitu dengan membiarkan mereka belajar sendiri-sendiri, tidak terstruktur, tidak mandiri, melainkan semuanya sendiri, hanya kadangkala diselingi beberapa nasehat cara belajar. Kalau peristiwa yang bertentangan ini telah melanda secara luas, dapatlah dipradugakan bahwa hal itu merupakan salah satu sebab menurunnya kualitas lulusan perguruan tinggi. Seharusnya justru dengan pendekatan PAP dalam sistem perkuliahan berdasar kompetensi itulah bertumpu harapan akan meningkatkan sekaligus *kuantitas dan kualitas* lulusan perguruan tinggi.

Tes pembinaan dilakukan sepanjang waktu satu program secara kontinu, pada setiap atau beberapa satuan perkuliahan. Sedang pada akhir satu program perkuliahan semester dilakukan tes keseluruhan (tes sumatif), yang butir-butirnya merupakan sampel isi keseluruhan program, dan oleh karena itulah bobot penilaiannya menjadi 2 berbanding 1 terhadap ujian tengah semester. Apa yang bisa terjadi dalam menetapkan *nilai akhir semester*? Ada beberapa praktek pendirian: (i) nilai ujian semester itulah nilai akhir semesternya. Atau nilai ujian ulangannya, (ii) rerata dari nilai ujian tengah semester ditambah nilai ujian semester dengan bobot nilai 1 berbanding 2. Pendirian pertama antara lain mereratakan ujian utama dengan ujian ulangan atau berapa pun nilai yang diperolehnya pada ujian ulangan yang di atas 5,6 maka nilai akhirnya adalah C, atau nilai ujian ulangan itulah nilai akhirnya.

Para penilai yang berpendirian pertama, mempunyai dasar pikiran bahwa ujian tengah semester merupakan ujian penjajagan dan pembinaan. Hasilnya dipakai sebagai umpan balik metoda perkuliahannya atau kualitas bahannya, dan kalau nilai itu kemudian dianalisis bersama yang diuji, akan merupakan umpan balik terhadap perbaikan atau peningkatan cara belajarnya; juga bentuk instrumen penilaiannya akan mempengaruhi cara belajarnya! Kalau hasil ujian tengah semester dianggap sebagai kredit ujian setengah bahan perkuliahan, berarti penilai berpendirian yang kedua, dengan anggapan "semuanya telah beres", dan kalau hasilnya tidak memuaskan, semata-mata karena yang diujilah yang kurang siap atau tidak menguasai bahan ujiannya. Ungkapan "semuanya telah beres" ialah dalam arti keterandalan dan kesahihan alat ujinya telah memenuhi syarat sebagai alat uji yang "baik", termasuk juga daya-

pembeda dan taraf kesukaran tes itu. Persoalannya sekarang ialah pendirian yang mana dan kapan pendirian itu lebih terandalkan?

Mempertimbangkan kenyataan praktek secara praktis yang menggunakan penilaian pendekatan PAP, dan alat uji buatan sendiri itu, maka pada keadaan tes yang *belum dibakukan*, kiranya akan lebih terandalkan kalau dilakukan tes-pembinaan (tes formatif) sebelum ujian tengah semester. Kalau hasilnya belum mencerminkan "keberesan cara penyampaian dan cara belajar", padahal sudah dilakukan perbaikan, maka lebih baik hasil ujian tengah semester dianggap sebagai tes pembinaan, dan kalau perlu dilakukan beberapa kali lagi, sebelum ujian semester. Memang 'berat' tugas guru, namun demi kecintaan kepada yang dididiknya, apalah artinya berat itu. Inilah pendirian pertama: nilai ujian semester sebagai nilai akhir semesternya. Kalau instrumen ujian tengah semester (dan semesternya) *telah dibakukan*, maka dapat dipakai pendirian kedua: nilai akhir adalah rerata nilai tengah semester dengan nilai semester dengan bobot 1:2. Pendirian ini juga sama tuntutan 'berat'-nya, ialah dalam proses analisis butir, taraf kesukaran, dan daya beda butir-butirnya, dan harus selalu melakukan perbaikan dalam selang sebanyak-banyaknya 4 tahun (kurun S1). Baik untuk tes bentuk objektif, maupun tes karangan, jarang dipikirkan cara pembakuannya.

Kalau hal-hal tersebut dilakukan secara komprehensif dan kontinu, maka penilai benar-benar telah melaksanakan pendekatan PAP secara konsekuen, dan dapatlah diharapkan akan diperoleh kuantitas dan kualitas kelulusan yang tinggi. Sekaligus dengan demikian *persyaratan* penilaian dalam pendekatan perkuliahan berdasar kompetensi telah terpenuhi.

Sebagai *catatan penting* untuk pendirian kedua, yang menggunakan bobot nilai, bahwa kalau nilai semester lebih besar daripada nilai tengah semesternya, maka hasil nilai akhir akan lebih besar daripada jika nilai semester lebih kecil daripada nilai tengah semesternya (coba buktikan). Dari kenyataan itu timbullah *pendirian ketiga* yaitu dengan tes terbakukan, tes objektif *dan* bentuk karangan, untuk ujian tengah semester dan semesternya, bahannya separoh-separoh dari keseluruhan bahan program perkuliahannya. Nilai akhir diperoleh dari rerata nilai tengah semester ditambah nilai semesternya, dengan bobot 1:1.

Dari ketiga pendirian dasar itu, dapat dikembangkan ragam pendekatan dengan rentangan paling-tidak-terandalkan sampai dengan paling-terandalkan sesuai dengan maksud penggunaan hasil penilaiannya, untuk pembinaan atau untuk pengambilan suatu keputusan, dan bukannya menggabungkan nilai tes pembinaan dengan hasil tes akhir. Ada petunjuk yang disarankan untuk memperoleh nilai akhir, yaitu: rerata nilai tes formatif ditambah nilai sumatif, dibagi dua. Kalau benar petunjuknya demikian, berarti menyalahi pengertian tes pembinaan dan mengikuti akibat fatal catatan pendirian kedua tadi. Demikianlah lebih banyak terungkap kelemahan penilaian dalam praktek nyatanya, karena menyalahfungsikan tes formatif, tengah semester, dan semesternya, serta cara menentukan rerata nilai, akibat dari kurang akuratnya dalam memahami fungsi tes pembinaan dan tes sumatif dalam kaitannya dengan pendekatan perkuliahan berdasar kompetensi dan prinsip belajar tuntas.

BAGIAN IV

PAN DAN PAP SEBAGAI PENDEKATAN PENILAIAN DAN PENGUKURAN PRESTASI BELAJAR

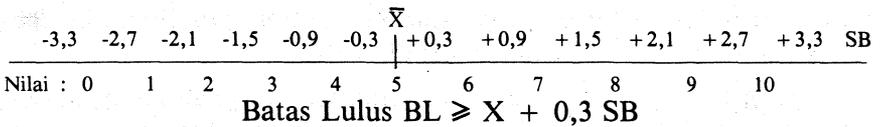
Sejak awal pembicaraan, istilah dengan singkatan PAN dan PAP belum disatubahasakan, sehingga kita masing-masing hidup dalam persepsinya sendiri dan terhayati masalahnya secara pribadi.

PAN dan PAP berasal dari terjemahan NRE (*norm-referenced evaluation*), dan CRE (*criterion-referenced evaluation*), yang huruf P-nya dapat dibunyikan sebagai penilaian atau pengukuran (*measurement*). Sedang kata "reference" akhir-akhir ini diterjemahkan sebagai "rujukan", yang dulu sering diterjemahkan dengan "acuan", dan menjadi lebih mudah pengucapannya, daripada PRN dan PRP. Untuk selanjutnya akan dipakai PAN dan PAP. Norma diartikan sebagai konsensus, persetujuan bersama, yang sifatnya relatif terhadap keadaan "apa adanya" yaitu keadaan nyata prestasi kelompok yang dinilai. Oleh karena itu disebut juga dengan singkatan PAR (R = relatif). Criterion diterjemahkan dengan patokan (PAP), bukan bakuan (standar). Istilah yang dijumpai adalah NRE dan CRE dan bukan NRM atau CRM. Oleh karena itu, selanjutnya PAN dan PAP kepanjangan dari penilaian dan bukan pengukuran.

Kalau perlu, pengukuran acuan norma dan patokan akan diistilahkan-singkatkan sebagai TAN dan TAP (T = tes).

Pengertian penilaian dalam pendidikan seringkali kalut, karena keadaannya. Ada yang menyamakan pengukuran dengan penilaian, misalnya dalam menjajagi prestasi belajar dengan THB — tes hasil belajar — akan diartikan mengukur hasil belajar, atau juga menilai hasil belajar, tanpa memandang arti khusus kedua istilah itu. Namun dalam *menilai*, atas dasar data lengkap pribadi siswa, dipandang sebagai suatu *deskripsi kualitatif* perilaku, mencakup semua cara penajagan yang tidak bergantung kepada hasil pengukuran saja. Pengukuran sendiri diartikan sebagai suatu *deskripsi kuantitatif*, yang berupa skor, angka, nilai, tentang perilaku dalam proses belajarnya. Jadi menurut batasan itu, pengertian penilaian lebih luas, lebih mencakup, daripada pengertian pengukuran. Tanpa atau dengan pengukuran maka penilaian selalu mengenakan pernilaian-nya (value judgement). Kiranya akan memadai kalau dalam penilaian pendidikan, terutama dalam pendidikan formal, penilaian diartikan sebagai suatu upaya yang bertataaturan dalam menentukan taraf pencapaian sasaran pendidikan-pengajaran-perkuliahan. Karena PBM berorientasi kepada tujuan perkuliahan, maka penilaiannya selalu berorientasi kepada TKP-nya (tujuan khusus perkuliahannya). Memanglah THB dikembangkan dari TKP ini. Berlandaskan kepada pengertian penilaian itu, maka program penilaian prestasi belajar haruslah dikembangkan dalam bentuk pengukuran — tes objektif — dan nonpengukuran: tes karangan sebagai suatu upaya penajagan (assessment).

PAN adalah pendekatan penilaian yang merujuk kepada keadaan prestasi kelompok apa adanya, dengan menetapkan batas kelulusan sebagai batas penerimaannya, berdasarkan prestasi kelompok itu, yaitu *di atas rerata kelompok di tambah 0,3 SB*. Sifat norma kelompoknya yaitu pada pengambilan SB (Simpangan Baku sekor seluruh kelompok) dan 0,6 SB sebagai satuan nilainya. Ini berarti bahwa nilai anggota kelompok sangat bergantung atau relatif terhadap keadaan prestasi kelompok itu. Pembentukan skala satuan 0,6 SB, tentu saja berdasar asumsi normalitas prestasi kelompok itu. Dengan *nilai bakuan sebelas* akan diperoleh nilai dari 0 s/d 10 yaitu dari jarak skala -3,3 SB s/d +3,3 SB.



Dgr.4.1.: SKALA BAKUAN SEBELAS

Rerata ditambah 0,3 SB itu merupakan patokan, namun *besar-nya* bergantung kepada kualitas prestasi kelompok, dan oleh karena itu tak dapat ditentukan lebih dulu, kecuali rumusnya (BL dan satuan skala nilainya).

Dapatlah dimengerti bahwa untuk nilai yang sama diperlukan besarnya skor yang berbeda-beda, yaitu bergantung kepada kualitas anggota kelompoknya, yang terlihat dari besarnya rerata skor dan simpangan bakunya. Oleh karena itu, pendekatan PAN akan tidak menguntungkan kalau kualitas kelompok itu cukup tinggi. Perlu diingat bahwa yang dimaksud dengan kelompok, keluasannya bisa berkisar dari kelompok kelas sekolah formal sampai dengan kelompok nasional. Misalnya kemampuan berbahasa Inggris untuk memahami teks ilmiah dapat dibakukan atas kelompok perguruan tertentu, yang dasar patokannya diambil dari keadaan tuntutan wajar apa adanya agar mampu mengikuti perkuliahan dalam bahasa Inggris sebagai bahasa pengantarnya. Dalam hal ini meskipun atas dasar kemampuan kelompok, namun pendekatannya sudah menggunakan PAP, karena telah ditetapkan batas kelulusannya lebih dulu, misalnya 500 untuk TOEFL, dan bukan berdasar $\bar{X} + 0,3 SB$ dari kelompok di tempat ujian itu.

Dalam pendekatan PAP maka *besar-nya skor* sebagai batas penerimaan kelulusannya ditetapkan lebih dulu, misalnya 70% untuk P4 dinyatakan lulus dengan nilai Cukup atau C. Nilai C pada pendekatan PAP tidak lagi 56% dari skor total, atau nilai 5,6, melainkan bergantung kepada pentingnya hal yang diujikan itu yang secara minimal harus dikuasai. Pada beberapa lembaga pendidikan, dituntut 75% s/d 95%; misalnya untuk profesi guru atau penerbang pesawat udara; bahkan 100% untuk penerbang pesawat ruang angkasa luar. Jadi, besarnya skor sebagai pertanda kompetensi penampilan yang dapat diterima, telah ditetapkan lebih dulu. Bagi jenis kompetensi tertentu, batas minimal itu merupakan suatu *kemutlakan* yang harus dicapai, namun besarnya tuntutan itu, *relatif* terhadap

jenis tugas nyatanya yang nanti akan dihadapi. Pedoman menentukan besarnya tuntutan pada setiap jenis kompetensi yang diharapkan itu tidak ada, dan oleh karena itu disebut *arbitrer* yaitu menurut kuasanya yang ahli di bidang kompetensi itu. Pembicaraan ini semua atas dasar asumsi bahwa semua persyaratan alat pengukuran dan penjangannya telah dipenuhi(!), dan tidak perlu lagi asumsi normalitas distribusinya.

Kalau pada PAN, hubungan *angka skor* dan *angka nilai* tidak tetap karena bergantung kepada kualitas kelompoknya, maka pada PAP hubungan itu tetap, yang biasa diubah dulu ke dalam persenan (%) kalau nilai tertinggi 100; atau diubah ke dalam persepuluhan, kalau nilai tertingginya 10. Dengan kata lain, setiap skor mempunyai harga nilai berapa persen atau berapa persepuluh. Pada PAN agar skor dapat dinilai, maka diubah dulu ke dalam skor baku, yaitu dalam satuan SB (0,6SB). Hubungan yang tidak tetap pada PAP ialah antara angka nilai dengan *huruf nilai*. Kalau telah disepakati huruf C sebagai nilai batas kelulusan (Cukup), pada PAN $C = 5,6$ sedangkan pada PAP, C bergerak dari 6,5 s/d 7,5. Bahwa pada PAP konversi skor ke dalam persenan menjadi nilainya, menjadi alasan pendirian yang menyamakan pengukuran dan penilaian, sedang pada PAN, lain pengukuran, lain penilaiannya.

Persoalannya sekarang, bagaimana dengan pengukuran yang merujuk kepada norma dan kepada patokan, dalam kaitannya dengan tes formatif dan tes sumatif?.

Kalau niatnya "mengukur" *pendidikan* untuk memaknainya terhadap suatu pembandingan, maka diperlukan *kelompok normatif* (dalam teori tes, pembandingan ini disebut sebagai kriterium), karena guru-pendidik tidak bisa secara etis manusiawi, menuntut semauanya bahwa subjek didiknya harus begini-begitu, dengan dirinya atau menurut pikirannya sebagai ukurannya (kriterium). Guru-pendidik harus memahami apa yang sepatutnya menjadi tujuan yang harus dicapai, dari segi perkembangan fisik-sosial-kejiwaan subjek didiknya. Untuk itulah maka perlu dipikirkan bahan, cara penyajian, media, dan menetapkan yang sepatutnya itu sebagai normanya (TAN) bukan kriteriumnya.

Lain halnya kalau niatnya mengukur hasil *latihan kerja* untuk tugas tertentu, maka diperlukan ukuran patokan (kriterium !) minimal agar dapat melaksanakan kerja dengan efektif dan selamat.

Dengan kata lain, diperlukan pengukuran-rujukan-patokan (TAP).

Dalam kaitan dengan tes formatif sebagai tes pembinaan dalam proses perkuliahan yang memiliki TKP (tujuan khusus pengajaran), maka tepat dipakai TAP (tes acuan patokan) namun dengan pendekatan PAN dalam pertimbangan psikologis-diagnostis. Dalam proses belajar yang tidak memiliki TKP yang ditentukan lebih dulu, maka akan tepat dipakai TAN (tes acuan norma) sebagaimana dilakukan dalam PBM di pendidikan nonformal. Adapun untuk ujian tengah semester, tepatnya dipakai TAP dan PAP, seperti juga pada ujian semester, sebagai tes sumatifnya. Pembicaraan ini berdasar asumsi telah dipenuhinya sarat tes yang "baik". Kalau tes buatan sendiri dan belum dibakukan, maka sebaiknya dipakai PAN saja, baik dalam tes formatif, tes tengah semester dan tes sumatifnya, sebab pembakuan tes dalam ilmu sosial dan pendidikan, pada hakekatnya adalah PAN dalam memperkirakan koefisien keterandalan, analisis butir dan daya-bedanya, juga taraf kesukaran butirnya.

Persoalannya sekarang ialah bagaimana syarat-syarat tes yang berorientasi kepada patokan (TAP), prosedur dan pengertiannya sama dengan yang TAN? Beberapa istilah pokok, sama namun isinya berbeda, misalnya istilah kesahihan, keterandalan dan taraf kesukaran tes. Jenis keterandalan yang baik untuk TAN adalah koefisien konsistensi internal, sedang untuk TAP adalah koefisien stabilitas atau ekuivalen, sebab yang penting adalah keseluruhan tesnya.

Bagi TAP, yang penting ialah dapat difahaminya maksud butir tesnya. Kalau isi kemampuannya memang ada, maka jawabannya tentu akan sama. Jadi butir yang pesannya dapat ditangkap secara sama oleh semua pembacanya itulah butir yang terandakan. Mengenai *kesahihan TAP*, karena masalah pokoknya ialah *pola gambaran yang baku*, maka yang tepat ialah kesahihan isi, yang gambarannya sesuai dengan patokannya. Sedangkan untuk TAN, tentu kesahihan prediktif.

Perbedaan yang menarik ialah pengertian taraf kesukaran butir dan indeks diskriminasinya. Karena tujuan TAP untuk menjajagi kemampuan telah memenuhi patokannya atau belum, maka taraf kesukaran butirnya harus diukur dari sukar-mudahnya hal yang harus dipenuhi, keadaan tugasnya nanti, dan atau tuntutan kebiasaan yang

cocok, perlu, dan cukup. Dasar rujukannya ialah hal-hal yang telah ditetapkan harus dikuasai. Pada TAN selalu diorientasikan kepada kemampuan kelompoknya. Mengenai indeks diskriminasi butir, pada TAP tidak menjadi persoalan. Boleh saja butir yang baik itu tak seorang pun dapat menyelesaikannya. Yang penting ialah apakah butir TAP itu mampu mengukur efektivitas hal yang dididikkan-dilatihkan. Untuk itu, cukup dengan *indeks kepekaan*: selisih jawaban betul pada pascates-prates, dibagi jumlah-tetap peserta pada pra- dan pasca-tesnya.

Dapat digambarkan keadaan-keadaan :

PRA-TES :	salah semua	betul semua	sebagian betul
PASCA-TES :	1. salah semua	2. betul semua	3. lebih banyak
	4. betul semua	5. salah semua	6. lebih sedikit

Dengan indeks itu telah cukup memberikan gambaran, dan tidak perlu melakukan uji-t atau anava. Indeks kepekaan yang ideal tentu peristiwa no. 4 dan yang umum: no. 3; yang fatal: no. 5 yang memalukan: no. 6 dan yang membingungkan: no. 1, masing-masing perlu diteliti sebab musababnya.

BAGIAN V RANGKUMAN, KESIMPULAN DAN SARAN-SARAN

Dalam menelaah kembali kebijakan menilai prestasi belajar dengan pendekatan PAN dan atau PAP dapatlah dirangkumkan hal-hal sebagai berikut.

1. Dalam praktek penilaian hasil belajar dalam program perkuliahan, pada umumnya penilai masih merujuk kepada suatu "patokan semu" 5,6 atau 56% jawab betul atau C sebagai batas lulusannya. Penentuan nilai akhir dengan menghitung rerata dari nilai tengah semester dan semesternya dengan bobot 1:2, atau nilai semester itulah nilai akhirnya; dan nilai tengah semester sebagai tes pembinaan saja, sebagai umpan balik perbaikan belajar dan mengajarnya. Jarang yang melakukan tes

pembinaan secara dasar dan konsekuen sampai dengan tindak lanjutnya. Alat THB belum semua merupakan bentuk gabungan tes objektif dan tes karangan, dan walaupun disusun tes objektif buatan sendiri, langsung dipakai dan dijadikan dasar keputusan, padahal belum diketahui kesesatan bakunya. Prinsip yang dipakai adalah pendekatan PAP-semu, dengan dukungan keahlian bidang studinya.

2. Praktek perkuliahan berdasar kompetensi perlu ditegaskan dan dikendalikan, dengan memperjelas gambaran kompetensi yang akan dicapai oleh setiap mata kuliah yang disajikan, manamana yang menjadi prasyarat untuk yang mana. Ketegasan model PBM-nya perlu disadari dan dilakukan secara konsekuen sehingga proses belajar tuntas dan asas kompetensinya dapat dipadukan, dan akan memberi jalan secara pasti dalam menetapkan besarnya patokan sebagai batas penerimaannya.
3. Kalau ingin konsekuen menganut pendekatan PAP, hendaknya dilaksanakan ujian-ujian pembinaan secara kontinu dan dijadikan sumber informasi umpan balik perbaikan penyajian dan penyusunan program remedialnya. Nilai akhir mengambil keputusan, tidak mencampurkan dengan nilai ujian pembinaannya. Tes formatif menggunakan TAP dengan penilaiannya menggunakan pendekatan PAN.
4. Kalau berniat akan menggunakan PAP dan cara belajar tuntas, kiranya harus berorientasi pada kompetensi dan disiapkan SAP (Satuan Acara Perkuliahan) yang merujuk kepada silabus programnya. Dalam menyusun instrumen penilaiannya, jangan terperosok ke dalam prosedur klasik untuk tes psikologis, kecuali dalam hal-hal yang khusus.
5. Kesimpulan umum dan saran khususnya ialah bahwa perlu bagi setiap pengajar-pendidik dan calon guru untuk memperdalam secara konseptual dan praktis tentang seluk beluk pengukuran dan penilaian prestasi belajar, dengan mempertimbangkan pendekatan pengajarannya, agar dapat disusun rencana pengajarannya sesuai dengan; (i) hakekat PBM-nya, (ii) pendidikan berdasar kompetensi, (iii) ketepatan pilihan PAN-TAN atau PAP-TAP-nya, secara terpadu. Perumusan TKP hendaknya memberikan peluang berkembangnya kreativitas pelajarnya.

KEPUSTAKAAN

1. AMIEN, DR., *Proses Belajar Mengajar*, NKK IKIP Yogyakarta: Oktober 1982.
2. ANONYMOUS, *Belajar Tuntas*, Modul 26, Jakarta: PPIPT. Depdikbud., 1982-3.
3. ANONYMOUS, *Evaluasi Hasil Belajar*, Modul 16, Jakarta: PPIPT., Depdikbud, 1982-3.
4. ANONYMOUS, *Konsep Lanjut Evaluasi Hasil Belajar*, Modul 38, DIK, Jakarta: PPIPT. Depdikbud, 1983.
5. ANONYMOUS, *Penilaian dalam Pendidikan*, Buku III D, Jakarta: PPIPT Depdikbud, 1981.
6. ANONYMOUS, *Penyelenggaraan Pendidikan dan Penilaian dalam Sistem Kredit Semester*, Jakarta: Depdikbud, 1980, buku IV.
7. ASLANIAN, C.B., *Improving Educational Evaluation Methods*, London: Sage Publications, 1981.
8. FERNANDES, H.J.X., *Comparisons of CRM and NRM Tests*, Unesco: April 1983.
9. POPHAM, W.J., *Criterion-Referenced Measurement*, New Jersey: Educ'l Technology Publ., 1971.
10. POPHAM, *Educational Evaluation*, New Jersey: Prentice Hall, Inc., 1975.
11. WASESO, I., *Analisis Item: Implikasinya pada Penyusunan Bank Data Item dan THB.*, Solo: UNS, Mei, 1979.
12. WASESO, I., *Beberapa Catatan Pokok Materi Evaluasi Belajar*, Yogyakarta; UII, September, 1982.
13. WASESO, I., *Cara Menyusun Tes Hasil Belajar*, Klaten: PPSP, Penlok Guru-Guru, Maret, 1973.
14. WASESO, I., *Eksplorasi Sistem Konversi Sekor kedalam Nilai Tes Hasil Belajar*, Yogyakarta: DepMeth. Penel. FIP IKIP YOGYAKARTA, Juli 1974.
15. WASESO, I., *Kapita Selekta Hal dan Teori Belajar*, Yogyakarta: Otopress, 1983, Oktober.
16. WASESO, I., *Penggunaan Beberapa Cara dan Teknik Penyekoran dan Penilaian Prestasi Belajar*, Yogyakarta: Penlok BPA*UGM, 1972.
17. WASESO, I., *Pengukuran dan Pernilaian Prestasi Belajar*, Yogyakarta: Otopress, 1983, Juni.