

**PEMBAHARUAN SISTEM EVALUASI  
DALAM SKALA MIKRO DAN SKALA MAKRO  
PADA PENGAJARAN BIOLOGI DI SMU**

**Oleh : Bambang Subali**

**FMIPA Universitas Negeri Yogyakarta**

Diterima : 8 April 2001 / disetujui : 29 Mei 2001

**Abstract**

The article discussed about the status and function of evaluation which is proportionally using both macro and micro scales on Biology, specially in Senior High School. The discussion was applied by combining theories and the facts exist. The result concluded that macro scale evaluation on Biology in Senior High School by using the national assessment (Ebtanas), which is only measuring the cognitive ability should be used as a mean to monitor and describe the quality of education in improving the policy and not to determine the graduation or entrance selection. The result of analysis on Ebtanas outcomes should be distributed to the teachers in order to make any improvement on teaching quality based on cognitive aspect. The graduation or student achievement on Biology related to the notion of IPA which contains the product and process skill was better to be assessed by using authentic assessment so that the students' work could be the orientation of teaching-learning achievement.

**Key words: macro scale evaluation, micro scale evaluation, authentic assessment.**

**Pendahuluan**

Mempermasalahkan pembaharuan pendidikan di SMU pada khususnya ataupun pembaharuan pendidikan pada umumnya tidak akan terlepas dari permasalahan pembaharuan kurikulum, pembaharuan strategi (termasuk metode dan media pembelajaran), dan pembaharuan sistem evaluasi. Selain itu terkait pula permasalahan pembaharuan/pengadaan sarana-prasarana dan pembaharuan sistem penggajian.

Dari sisi kurikulum misalnya, pembaharuan pendidikan di SMU sudah dilakukan sejak digulirkannya Kurikulum 1975 sebagai penggantian yang cukup revolusioner terhadap Kurikulum 1968. kemudian Kurikulum 1984 sebagai penyempurnaan Kurikulum 1975, dan akhirnya berlakunya Kurikulum 1994 beserta suplemennya masih mengandung banyak perdebatan.

Pembaharuan strategi pembelajaran secara teoretik sudah dilakukan,

mengingat sejalan dengan pembaharuan kurikulum juga disertai anjuran penerapan strategi yang sesuai. Demikian pula dalam hal sistem evaluasinya, setiap penggantian kurikulum disertai anjuran penerapan teknik evaluasi sesuai dengan karakteristik keilmuannya. Evaluasi dalam skala makro pun seperti Ebtanas dimunculkan sebagai salah satu solusi untuk mermantau dan memetakan mutu pendidikan nasional..

Dari segi perbaikan/pengadaan sarana-prasarana banyak permasalahan yang mengemuka. Demikian pula jika dilihat dari sistem penggajian. Rendahnya penghargaan terhadap profesi guru, pemberian tunjangan fungsional yang hampir sama antar guru pada jabatan fungsional yang berbeda dicurigai pula sebagai kendala terhadap upaya perbaikan pendidikan.

Artikel ini hanya akan memfokuskan pada pembahasan yang berkait dengan pembaharuan sistem evaluasi

dalam skala makro dan skala mikro-khususnya dalam mata pelajaran Biologi di SMU mengingat upaya perbaikan pembelajaran tidak dapat lepas dari sistem evaluasi yang diterapkannya.

### **Pengertian Evaluasi Skala Makro dan Skala Mikro**

Evaluasi lebih luas dari penilaian (*assessment*) karena evaluasi merupakan subsistem dalam sistem pendidikan. Melalui evaluasi dapat diketahui kemajuan dan perkembangan penyelenggaraan pendidikan dari waktu ke waktu. Melalui evaluasi pula dapat diketahui sejauh mana tindakan yang telah dikerjakan sebelumnya benar-benar berharga. Hasil-hasil evaluasi merupakan informasi yang sangat penting bagi pengambil kebijakan untuk melangkah lebih jauh dalam mengimplementasikan program pendidikan yang diselenggarakan. Dalam skala mikro, setiap pengajar sebagai komponen penyelenggaraan program pendidikan pun memiliki tanggung jawab untuk mengevaluasi setiap program pembelajaran yang diselenggarakan. Bagi guru hasilnya dipakai untuk memperbaiki program yang disusun, sedangkan bagi siswa dipakai untuk memperbaiki cara belajarnya agar dapat lebih berhasil (Depdikbud 1997: 2-3).

Hasil evaluasi skala mikro hendaknya dapat dimanfaatkan oleh pihak yang berkepentingan, yaitu baik guru sendiri sebagai perancang program pembelajaran, dan subyek didik sebagai pihak yang berkedudukan sebagai pelaku dalam proses belajar.

Hasil evaluasi skala makro diharapkan bermanfaat lebih luas, karena banyak pihak yang terkait. Hasil evaluasi diharapkan dapat dimanfaatkan oleh orang tua yang mempercayakan anaknya kepada sekolah. Sekolah dan institusi di atasnya juga harus dapat memanfaatkan hasil evaluasi untuk

memper-baiki dan meningkatkan mutu penyelenggaraan pendidikan. Dengan demikian kebijakan baru yang akan diterapkan benar-benar mampu meningkatkan harkat dan martabat bangsa.

Permasalahannya adalah sistem evaluasi yang bagaimana yang dapat dipertanggungjawabkan yang sesuai dengan hakikat belajar Biologi pada jenjang pendidikan di SMU.

### **Pengajaran Biologi di SMU**

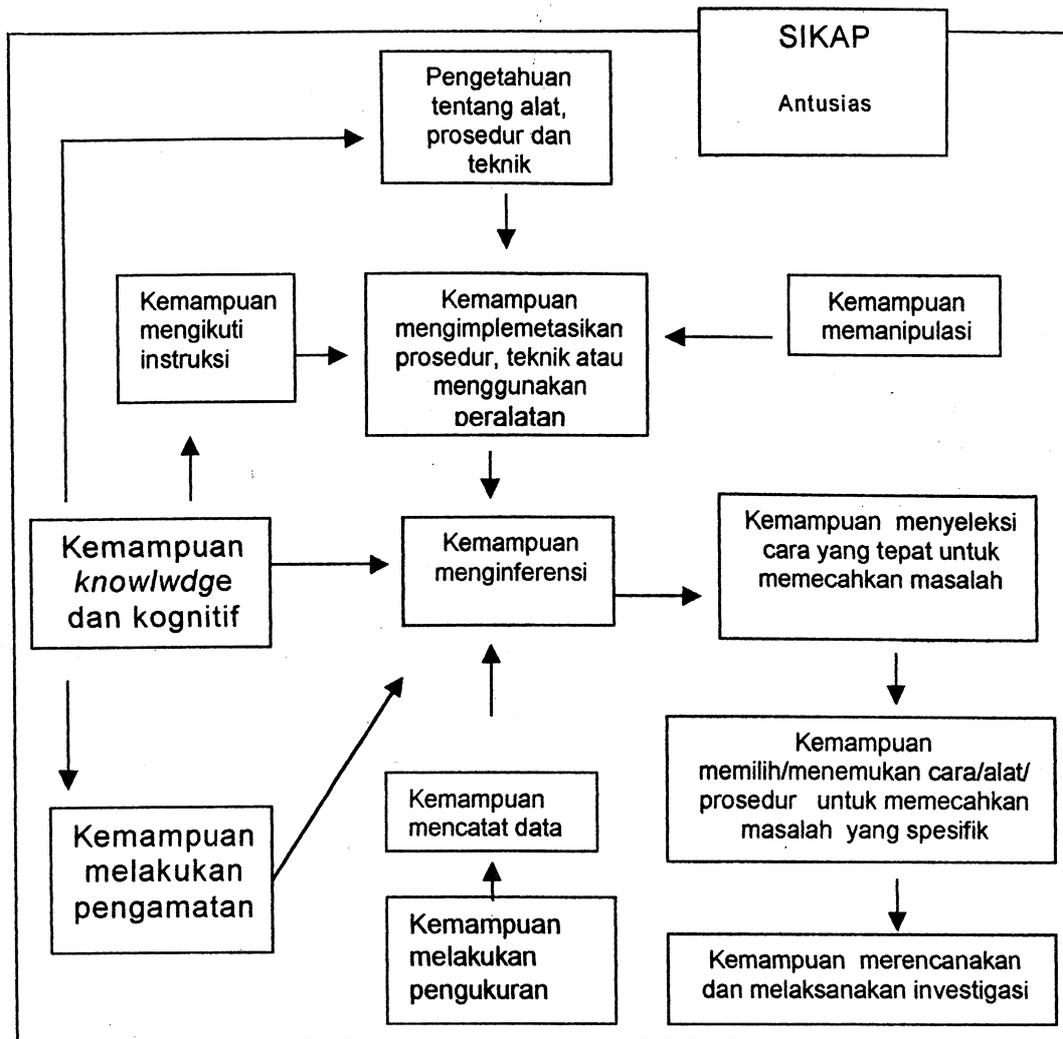
Dalam buku Petunjuk Teknis Mata pelajaran Biologi SMU (Depdikbud, 1995) dikemukakan secara detail tentang pengertian GBPP dan komponen-komponennya. Dalam buku tersebut juga dimuat prinsip-prinsip pembelajaran baik menyangkut pendekatan, metode, pengelolaan kelas dan laboratorium, dan penilaian (dalam arti *assessment*), beserta contoh-contohnya. Demikian pula tentang perencanaan pembelajaran yang memberikan arahan tentang cara membuat persiapan mengajar lengkap dengan contoh-contohnya serta model-model pelaksanaan pembelajaran yang memberikan wawasan kepada guru tentang pelaksanaan pembelajaran yang baik.

Berkait dengan pendekatan konsep dan pendekatan keterampilan proses di dalam buku tersebut diuraikan secara detail pengertian konsep, cara mengembangkan konsep dan sub-konsep serta cara menghubungkan antar konsep sebagai bahan ajar yang lebih bermakna. Dikemukakan pula Biologi sebagai bagian dari IPA terdiri dari produk yang terdiri atas fakta, konsep, prinsip, teori dan hukum; serta proses IPA yang meliputi berbagai keterampilan. Misal: (a) keterampilan mengamati menggunakan sebanyak mungkin panca indera, mengumpulkan fakta yang relevan, mencari kesamaan dan perbedaan, mengklasifikasikan; (b) keterampilan dalam menafsirkan hasil

pengamatan seperti mencatat secara terpisah setiap jenis pengamatan, dalam menghubungkan-hubungkan hasil pengamatan, (c) keterampilan menemukan suatu pola dalam seri pengamatan, dan dalam mencari kesimpulan hasil pengamatan, (d) keterampilan dalam meramalkan apa yang akan terjadi berdasarkan hasil-hasil pengamatan, (e) keterampilan menggunakan alat/bahan dan mengapa alat/bahan itu digunakan, (f) keterampilan dalam menerapkan konsep, baik penerapan konsep dalam situasi baru, menggunakan konsep dalam pengalaman baru untuk menjelaskan apa yang sedang terjadi, maupaun dalam menyusun hipotesis. Keterampilan lainnya misal: (a) keterampilan dalam merencanakan kegiatan seperti menentukan alat bahan yang akan digunakan, menentukan variabel, menentukan variabel tetap/bebas dan variabel berubah/tergayut, menentukan apa yang diukur dan diamati, menentukan cara dan langkah kerja, (b) keterampilan cara mengorganisasi baik dalam bentuk grafik, tabel atau yang lainnya, (c) keterampilan cara mengolah hasil-hasil pengamatan. Keterampilan IPA juga menyangkut keterampilan dalam berkomunikasi seperti (a) keterampilan menyusun laporan secara sistematis, (b) menjelaskan hasil percobaan atau pengamatan, (c) cara mendiskusikan hasil percobaan, (d) cara membaca grafik atau tabel, dan (e) keterampilan mengajukan pertanyaan baik bertanya apa, mengapa dan bagaimana, maupun bertanya untuk meminta penjelasan serta keterampilan mengajukan pertanyaan yang berlatar belakang hipotesis.

Menurut (Bryce dkk. 1990: 2). Pembelajaran IPA, mengembangkan kemampuan dasar (*basic skill*) sebagai kemampuan yang terendah, kemudian meningkat ke kemampuan proses (*process skill*) dan menuju ke kemampuan investigasi (*investigation*

*skill*) sebagai kemampuan yang tertinggi. Kemampuan dasar mencakup: (a) kemampuan melakukan pengamatan (*observational skill*), (b) kemampuan mencatat data (*recording skill*), (c) kemampuan melakukan pengukuran (*measur-ment skill*), (d) kemampuan meng-impelemntasikan prosedur (*procedural skill*), dan (e) kemampuan mengikuti instruksi (*following instructions*). Kemampuan proses meliputi: (a) kemampuan menginferensi (*skill of inference*) dan (b) kemampuan untuk menyeleksi berbagai cara/prosedur (*selection of procedures*). Kemampuan investigasi berupa kemampuan merencanakan dan melaksanakan serta melaporkan hasil investigasi. Jika digambarkan akan tampak model sebagai berikut:



Gambar 1. Kinerja yang dilakukan pada proses IPA (Sumber: Bryce dkk, 1990: 2)

Dalam proses pembelajaran IPA menurut Towle (1989: 16-31) harus bertumpu pada proses ilmiah. Proses ilmiah tersebut melibatkan berbagai keterampilan. Misal keterampilan melakukan pengamatan dan mengoleksi data, termasuk kemampuan menggunakan alat dan melaporkan informasi yang spesifik yang berkaitan dengan hasil-hasil pengamatannya (*observing and collecting data*). Keterampilan lainnya yaitu: (a) keterampilan melakukan pengukuran, termasuk kemampuan menggunakan berbagai alat ukur (*measuring*), juga keterampilan mengorganisasi data

dalam bentuk grafik, tabel, diagram, peta ataupun urutan dengan pola tertentu agar lebih bermakna/mudah dipahami (*organizing data*), dan keterampilan mengklasifikasi data ke dalam suatu pola tertentu untuk menghasilkan pola yang baru berdasarkan karakteristik dari obyek yang diamati, misal mengklasifikasi berdasarkan ciri-ciri morfologinya, berdasar cara hidup, berdasar habitat (*classifying*). Keterampilan merumuskan hipotesis sebagai suatu pernyataan yang siap diuji/*testable* (*hypothesizing*), memprediksi berbagai hal yang relevan dalam rangka menguji

hipotesis (*predicting*), melakukan percobaan untuk menguji hipotesis dengan menentukan variabel bebas dan terikatnya serta mengontrol variabel luaran (*extraneous variable*) dan (*experimenting*), menarik kesimpulan berdasar fakta dan pengetahuan atau hasil percobaan sebelumnya (*inferring*), membuat model dalam bentuk bagan alir atau membuat model matematik (*modeling*), mengkomunikasikan hasil pengamatan atau percobaan (*communicating*) juga merupakan bagian dari keterampilan proses IPA. Jika proses ilmiah disusun dalam suatu urutan tertentu dan digunakan untuk memecahkan suatu permasalahan yang dihadapi, maka rangkaian proses ilmiah itu menjadi suatu metode ilmiah. Towle memberikan contoh cara menyusun model kegiatan laboratorium dengan maksud melatih peserta didik untuk menguasai aspek tertentu dari proses ilmiah yang ada disesuaikan dengan materi dan tujuan kegiatan yang diselenggarakan.

Mengingat siswa SMU dilihat dari perkembangan kognitifnya sudah mencapai tataran kemampuan berpikir formal, maka pembelajaran Biologi di SMU hendaknya sudah mengenalkan siswa kepada kemampuan untuk investigasi walapun sifatnya masih sederhana. Setidaknya siswa sudah harus mampu merencanakan pengamatan/percobaan, menyusun hipotesis berdasar pustaka bukan sekedar menurut dugaan yang rasional menurut logika, mampu melakukan dan melaporkan percobaan/pengamatan baik tertulis maupun secara lisan. Jika hal seperti itu dibiasakan maka hasil belajar yang dapat dicapai benar-benar akan memuat unsur kognitif, unsur afektif dan unsur psikomotor.

### Sistem Evaluasi di SMU

Dalam buku Petunjuk Pelaksanaan Penilaian di SMU (Depdik-

bud,1994:6-10) dijelaskan pula tujuh prinsip penilaian-dalam arti evaluasi-yaitu prinsip menyeluruh (menyangkut aspek proses dan hasil belajar, serta menyangkut aspek pengetahuan, sikap, perilaku dan nilai, serta keterampilan), berkesinambungan, berorientasi pada tujuan (sesuai dengan rumusan tujuan dalam GBPP), obyektif, terbuka (proses dan hasil penilaian diterima oleh semua pihak), dan kebermaknaan (bagi pihak-pihak yang berkepentingan). Evaluasi juga harus mengacu pada prinsip berkesesuaian yaitu sesuai dengan pendekatan dalam kegiatan pembelajaran dalam arti jika menggunakan pendekatan eksperimen maka kegiatan melakukan percobaan harus menjadi salah satu aspek yang dinilai.

Dalam buku tersebut dijelaskan pula bahwa dalam menilai dalam arti melakukan *assessment* hasil belajar, aspek yang dinilai juga mencakup aspek kognitif sebagaimana yang disusun dalam taksonomi Bloom; aspek-aspek yang tercakup dalam sikap ilmiah dan nilai-nilai IPA seperti ketelitian, kecermatan, kejujuran, penghargaan terhadap pendapat orang lain, kemauan menerima saran; aspek-aspek kreatifitas, imajinasi serta tanggung jawab juga supaya diupayakan ikut dinilai dengan cara melakukan pengenalan secara individual kepada setiap subyek belajar agar dapat dibandingkan satu dengan yang lainnya.

### Penilaian Autentik sebagai Alternatif Pembaharuan

Dari apa yang digariskan dalam Kurikulum 1994, sebenarnya pengajaran Biologi di SMU beserta cara memberikan penilaian-dalam arti memberikan *assessment*- secara teoritik sudah memadai dari arah kebijakan yang digariskan karena tidak semata mengandalkan hasil tes. Namun demikian kenyataan guru hanya mengandalkan hasil tes dalam bentuk

ulangan harian dan ulangan umum. Namun demikian banyak kendala yang ditemukan di lapangan seperti banyak guru yang melebihi beban tugas yang diwajibkan sebanyak 18 jam per minggu, banyak guru yang *mismatch*, ukuran kelas yang terlalu besar.

Hasil penelitian di 20 propinsi menunjukkan banyak sekolah (hampir separoh SLTPN di Jawa Timur, juga sebagian besar propinsi di luar Jawa) belum memiliki laboratorium yang memadai dan semakin tinggi jenjang sekolah semakin banyak guru yang *mismatch* sehingga mata pelajaran IPA (Fisika, Kimia, Biologi) diampu oleh guru non-IPA (Jemari Mardapi dkk. 2001).

Hasil penelitian di DIY, Kalimantan Barat dan Sumatera Barat menunjukkan bahwa penyelenggaraan program perbaikan secara klasikal sulit dilaksanakan karena padatnya materi/bahan ajar. Penyelenggaraan program perbaikan di luar jam pelajaran pada pagi hari praktis tidak dapat dilaksanakan, sementara penyelenggaraan pada sore hari tidak didukung ketersediaan finansial (Toto Kuwato dan Djemari Mardapi 1999).

Berkait dengan pelaksanaan sistem evaluasi di SMU, hasil penelitian Toto Kuwato dan Djemari Mardapi (1999) menunjukkan bahwa:

1. Hasil sistem ujian yang ada selama ini belum seperti yang diharapkan. Masih banyak para guru yang belum secara rutin menyusun kisi-kisi ulangan, menelaah soal, menganalisis butir soal, menganalisis hasil ulangan, menginformasikan kegagalan siswa kepada orang tua, dan belum sepenuhnya menindaklanjuti kegagalan siswa dalam penguasaan konsep melalui program perbaikan. Guru belum diwajibkan untuk menyusun kisi-kisi ulangan. Dalam menyiapkan pelajaran pun guru

hanya mencontoh renaca pelajaran dan analisis materi pelajaran (AMP) yang disusun oleh MGMP.

2. Soal-soal ujian SMU belum dikalibrasi, sehingga sulit untuk membandingkan mutu sekolah baik antar wilayah maupun antar tahun.
3. Faktor finansial menjadi kendala pengembangan bank soal di tingkat wilayah.
4. Arus informasi hasil ujian yang sangat diperlukan oleh pihak-pihak terkait belum dapat diperoleh secara lengkap, walaupun ada juga belum dimanfaatkan secara optimal.
5. Mutu alat tes untuk ulangan harian yang dibuat para guru belum baik akibat belum adanya kisi-kisi dan tidak pernah ditelaah oleh guru lain yang sebidang.
6. Keterkaitan antara ulangan harian, ulangan cawu, ulangan kenaikan kelas belum baik akibat tidak adanya kisi-kisi ulangan.
7. Kurangnya dorongan dari pihak kepala sekolah kepada guru yang telah mengikuti pelatihan untuk menerapkan pengetahuannya di sekolah.

Penelitian Toto Kuwato dan Djemari Mardapi (1999) juga menyertakan program tindakan untuk meningkatkan sistem evaluasi di sekolah. Hasilnya menunjukkan bahwa rendahnya taraf serap siswa dalam program tindakan lebih ditentukan oleh pandangan guru yang menganggap mudah suatu konsep atau memang karena kemampuan awal siswa yang rendah, jadi bukan karena rendahnya mutu soal yang dibuat guru.

Sejak diperbaikinya Kurikulum 1975 menjadi Kurikulum 1984, dimunculkan Evaluasi tahap Akhir nasional atau disingkat Ebtanas sebagai suatu *national assessment*. Tujuan diselenggarakannya Ebtanas antara lain sebagai berikut (Depdikbud, 1986).

1. Merintis terciptanya standar nasional mutu pendidikan dasar dan menengah.

2. Menyederhanakan prosedur penerimaan peserta didik baru pada sekolah yang lebih tinggi.
3. Mempercepat peningkatan dan pemerataan mutu pendidikan dasar dan menengah.
4. Menunjang tercapainya tujuan kurikulum.
5. Mendorong agar proses belajar mengajar dilaksanakan berdasar kurikulum, buku, dan alat peraga yang telah ditetapkan.

Kritik terhadap penyelenggaraan ujian yang sifatnya nasional, Nitko (1996) mengemukakan bahwa:

1. Hasil-hasil ujian tidak peka, baik terhadap perbaikan masukan (input) pendidikan, maupun terhadap persepsi guru dan orang tua perihal prestasi peserta didik.
2. Laporan hasil ujian tidak menerangkan tentang pengetahuan dan keterampilan yang dipelajari oleh peserta didik. Akibatnya pengambilan keputusan/pengembang kurikulum tidak mengetahui aspek kurikulum mana yang harus diperbaiki.
3. Hasil-hasil ujian memberikan dasar yang rapuh untuk membimbing peserta didik ke arah kejuruan dan pengembangan karir.
4. Kesesuaian antara tujuan belajar yang dinyatakan dalam kurikulum resmi dan pertanyaan-pertanyaan yang muncul dalam ujian seringkali tidak jelas bagi guru. Akibatnya para guru mengabaikan kurikulum resmi dan menggunakan soal-soal ujian yang telah lalu sebagai bahan ajar.
5. Para pendidik di semua jenjang pendidikan menunjukkan kelemahan-kelemahan tersebut bahkan ada yang menyandarkan pada hasil sekali ujian, sehingga beresiko tinggi karena mengabaikan kinerja peserta didik bertahun-tahun di kelas.
6. Keluasan dan kekayaan pemba-

haruan kurikulum diabaikan oleh para guru, yang atas kemauannya sendiri mempersempit kurikulum sehingga menjadi tugas-tugas yang bakal muncul dalam ujian.

Ebtanas sebagai instrumen untuk memantau mutu pendidikan secara makro ternyata hanya memfokuskan pada aspek kognitif. Sementara itu, hasil Ebtanas dijadikan alat untuk seleksi (di SLTP dan SMU). Hal ini memperkuat apa yang dikemukakan oleh Nitko, bahwa Ebtanas memacu guru untuk segera menyelesaikan kurikulum agar ia dapat mempersiapkan muridnya berhasil dalam menempuh Ebtanas.

Banyak guru yang menggunakan soal-soal Ebtanas sebagai acuan dalam penyelenggaraan kegiatan pembelajaran peserta didiknya (Wuryadi dan Bambang Subali 2000; Djemari Mardapi dkk 1999). Di sisi lain, dalam Petunjuk Teknis Mata Pelajaran Biologi di SMU (Depdikbud, 1995), guru tidak boleh mengubah tujuan pembelajaran yang sudah ditetapkan dalam GBPP. Oleh karena itu perlu menyertakan ujian praktik laboratorium agar lebih komprehensif.

Ebtanas juga menimbulkan perbedaan cara siswa dalam menyikapi suatu mata pelajaran. Siswa lebih memberikan respons yang positif pada mata pelajaran yang diujikan dalam Ebtanas dibanding yang tidak (Djemari Mardapi dkk. 2001).

Dari kenyataan di atas, sudah sewajarnya jika Ebtanas hanya dipakai untuk memantau mutu pendidikan, dan pelaksanaannya justru pada kelas V SD dan kelas II SLTP/ SLTA. Dengan demikian hasilnya dapat dipakai untuk memperbaiki mutu (dengan catatan hasil analisis hasil Ebtanas disampaikan ke setiap sekolah dan guru yang bersangkutan), dan kelulusan serta selski masuk diserahkan kepada masing-masing sekolah. Peta kualitas sekolah pada setiap wilayah berdasar hasil Ebtanas dapat diumumkan dan diketahui oleh masyarakat luas, sehingga masyarakat juga dapat menilai

secara obyektif, mengingat Ebtanas hanya mengukur aspek kognitif.

Dalam skala mikro, guru pada umumnya mengandalkan hasil ulangan harian dan ulangan umum untuk menilai prestasi siswa, termasuk dalam pelajaran Biologi di SMU. Jika diperhatikan, baik ulangan harian, ulangan catur wulan, ulangan kerjakan kelas, maupun Ebtanas menggunakan tes tertulis. Teknik evaluasi ini dipandang mengandung banyak kelemahan karena umumnya hanya mengukur sebagian kecil saja dari aspek (domain) prestasi yang dicapai oleh peserta didik. Aspek/domain yang lebih banyak diukur adalah aspek kognitif. Itupun hanya beberapa jenjang yang rendah saja, seperti pengetahuan (*knowledge*), pemahaman (*comprehension*) dan sebagian kecil aplikasi, sedangkan jenjang yang lebih tinggi, seperti analisis, sintesis dan evaluasi, kurang terukur dengan baik. Akibatnya, hasil yang terukur kurang mencerminkan pencapaian hasil belajar yang sesungguhnya.

Sebagai alternatif, penilaian di SMU dititik beratkan kepada penampilan atau kinerja yang ditunjukkan oleh siswa, yaitu melalui *authentic assesment*. *Authentic assesment* merupakan suatu penilaian yang dilakukan melalui penyajian atau penampilan oleh peserta didik dalam bentuk pengerjaan tugas-tugas atau berbagai aktivitas tertentu, yang secara lang-sung mempunyai makna pendidikan. Karena banyak menekankan pada aspek penampilan, maka alat penilaian ini sering disebut juga dengan istilah *performance assesment*.

Istilah *authentic assesment* menurut kepustakaan memang bervariasi. Ada yang menyebutnya *performance assesment*, *outcomes-based assesment*, dan *alternative assesment*. Namun demikian, menurut Marzano dkk (1993: 9-13) semua itu mengandung tiga unsur inovasi dalam bidang penilaian, yaitu (a) tidak

mengukur ketercapaian tujuan pembelajaran yang tradisional, tetapi lebih menekankan pada kemampuan nyata subyek belajar, (b) bersifat menyeluruh, mengembangkan seluruh kemampuan subyek melalui kegiatan pembelajaran menurut paham konstruktivisme, dan (c) tidak menggunakan sistem tes tradisional tetapi menggunakan berbagai cara.

Gronlund (1998: 14-15) menambahkan perihal kelebihan *performance assesment*. Melalui *performance assesment* yang diperluas (*extended performance assesment*) guru dapat mengetahui berbagai kemampuan yang lebih kompleks yang dicapai siswa yang tidak dapat diukur dengan menggunakan tes tertulis dalam bentuk uraian sekalipun. Untuk mengetahui prestasi dalam bidang matematika misalnya, melalui *paper-and-pencil test* sebagai salah satu bentuk tes perbuatan dapat dipakai untuk mengetahui kemampuan siswa dalam memilih diantara fakta-fakta yang tersedia untuk memecahkan permasalahan matematika yang dihadapi. Demikian pula dalam bidang IPA, melalui *paper-and-pencil test* dapat dipakai untuk mengetahui kemampuan testi dalam memilih prosedur untuk memecahkan permasalahan. Namun demikian dalam beberapa hal guru dapat mengkombinasikan tes perbuatan dengan tes tertulis. Menurut Gronlund (1998: 2) melalui *performance assesment* akan dapat diketahui penampilan yang aktual dari siswa dalam menguasai keterampilan yang telah dipelajarinya seperti kemampuan memakai peralatan laboratorium, kemampuan melaksanakan eksperimen, kemampuan menjalankan mesin, dan sebagainya. Sementara istilah *authentic assesment* dipakai untuk suatu *performance assesment* yang difokuskan pada aplikasi atau dari

pengetahuan yang dikuasai siswa atau untuk mengetahui keterampilan siswa untuk memecahkan problem-problem yang dihadapi dalam dunia nyata.

Menurut Newman dan Wehlage (1993: 12) *authentic assessment* adalah proses pengumpulan data dimana mahasiswa memahami dan menghasilkan pengetahuan yang berarti/bermakna. *Authentic assessment* disebut juga *performance assessment* karena didasarkan atas apa yang dapat dilakukan oleh subyek belajar. Marzano dkk. (1993: 13) mendeskripsikan kegiatan *authentic assessment* sebagai berikut:

- a. Mahasiswa diberi kesempatan untuk mendemonstrasikan kebolehannya, pemahamannya, keterampilannya secara kontekstual dan variatif.
- b. Dilakukan secara kontinyu dan terstruktur menurut tujuan instruksional
- c. Menghasilkan karya nyata (*tangible product*) dan penampilan yang dapat diamati (*observable performance*).
- d. Memacu mahasiswa untuk melakukan penilaian diri (*self-assessment*),
- e. menyadari kelebihan dan kelemahannya dan mampu mengembangkan kelebihannya tersebut dan memperbaiki kelemahannya.
- f. Mengungkap kemampuan mahasiswa berdasarkan kriteria yang telah ditentukan.

O'Neil (1992: 14-19) menambahkan bahwa *authentic assessment* memberi data yang lebih lengkap tentang kemampuan mahasiswa dan didasarkan atas kegiatan pembelajaran, menghargai produk dan proses sama baiknya.

*Authentic assessment* identik pula dengan *outcomes-based education* seperti yang diungkap oleh Spady (1993: 4). Menurut Spady, program studi harus memiliki standar lulusan. Karena berbasis sekolah, maka siswa dituntut untuk dapat melakukan

berbagai kegiatan pembelajaran di sekolah. Oleh karena itu *authentic assessment*, *performance assessment*, dan *outcomes-based education* dikembangkan dengan baik. Pro-gram IPA di SMU mestinya juga telah mengembangkan standar kemampuan atau kapasitas lulusannya sebagai calon mahasiswa yang nantinya akan berkiprah dalam bidang IPA.. Dengan demikian, setiap mata pelajaran dalam program tersebut ditujukan untuk mengembangkan kemampuan menuju standar yang telah ditetapkan. Fungsi *authentic assessment* ialah untuk melacak kemampuan standar mana yang telah dikuasai siswa dan kemampuan mana yang belum dikuasai. Dengan demikian maka kualitas lulusan akan memenuhi standar mutu yang telah ditetapkan.

Tugas-tugas siswa yang dikembangkan melalui *authentic assessment* bervariasi namun tidak terlepas dari tiga prinsip dasar. Pertama, tugas harus sangat bermakna bagi mahasiswa (*meaningful*). Kedua, senantiasa disertai dengan kriteria penilaian. Ketiga, didasarkan atas apa yang dapat dilakukan oleh siswa (Marsh 1996: 224). Bentuk dari tugas-tugas tersebut meliputi: (a) portfolio, (b) pembuatan jurnal/ paper, (c) simulasi, (d) membuat desain dan presentasi, (e) observasi kritis, (f) mengerjakan proyek individu dan kelompok, (g) melaporkan hasil studi lapangan, (h) melakukan kegiatan pemecahan masalah, (i) membuat peta konsep, dan sebagainya. Melalui penilaian portofolio setiap siswa dapat diikuti perkembangan kemampuannya berdasar program yang dipilih dan disusunnya. Dalam hal ini kemampuan siswa menjadi sangat individual dan prestasi setiap siswa dapat diikuti secara individual pula.

Dari pengalaman penerapan *performance assessment* dalam mata

kuliah Penilaian Pencapaian Hasil Belajar Biologi yang dicoba oleh Bambang Subali, dkk (2000), juga penerapan *authentic assessment* dalam mata kuliah Mikrobiologi (Bambang subali dkk. 2001:35) ternyata banyak tantangan yang dihadapi, diantaranya yaitu:

1. sukar mendesain tugas/kegiatan yang mampu menghasilkan kinerja yang terukur;
2. sukar membuat desain tugas dengan tujuan pembelajaran kompleks tetapi dapat dikerjakan dengan baik oleh peserta didik;
3. untuk dapat mengerjakan tugas/kegiatan memerlukan pengetahuan awal dan keterampilan prasyarat yang menghabiskan banyak waktu sebelum tugas/kegiatan yang pokok dapat dilaksanakan;
4. sukar menentukan teknik penskoran terhadap kinerja peserta didik agar benar-benar merupakan penghargaan yang setimpal dengan usaha yang dilakukannya;
5. penyelesaian tugas/kegiatan yang lengkap dan baik banyak memakan waktu, sementara setiap jenis tugas tidak mampu mencerminkan kebulatan tujuan yang ditargetkan;
6. penskoran terhadap hasil kinerja berupa jawaban tertulis memerlukan banyak waktu;
7. skor terhadap kinerja peserta didik tidak terjamin kesahihan dan keandalannya;
8. penyelesaian tugas/kegiatan yang lengkap dan bernilai hanya akan dapat dicapai oleh peserta didik yang mampu dan yang mau melibatkan diri secara aktif, sehingga hanya sedikit peserta didik yang benar-benar dapat berhasil dengan baik;
9. dengan ukuran kelas yang sedang (45 orang) menjadikan peserta didik yang berperan aktif lebih sedikit jika waktu persentasi sangat terbatas.

Dari kenyataan di atas penerapan penilaian terhadap kinerja/penampilan siswa perlu persiapan yang serius dan harus dapat mengubah perilaku guru yang selama ini mengandalkan teknik evaluasi berupa tes. Hasil yang perlu dipertimbangkan adalah bahwa setelah dianalisis, hasil kinerja yang bersifat kognitif dengan hasil tes kognitif masih menunjukkan korelasi yang signifikan pada taraf kesalahan 5% , sementara hasil kinerja yang lebih bersifat psikomotor dengan hasil tes kognitif tidak menunjukkan korelasi yang signifikan. Dengan demikian memadukan hasil tes dan hasil penilaian terhadap kinerja/penampilan akan lebih mampu menggambarkan prestasi siswa.

### Kesimpulan

Penilaian skala makro pada pengajaran Biologi di SMU dengan menggunakan Ebtanas yang hanya mengukur kemampuan kognitif hendaknya hanya sebagai alat untuk memetakan mutu pendidikan sebagai upaya perbaikan kebijakan, bukan untuk menentukan kelulusan ataupun untuk seleksi masuk. Hasil analisis hasil Ebtanas harus sampai ke meja guru agar guru dapat memperbaiki kualitas pembelajaran dari aspek kognitif. Kelulusan ataupun prestasi siswa dalam mata pelajaran Biologi sesuai dengan hakekat IPA yang memuat produk dan keterampilan proses IPA lebih baik dinilai melalui penilaian autentik sehingga kinerja siswa menjadi orientasi pencapaian pembelajaran.

### Daftar Pustaka

- Bambang Subali, Slamet Suyanto dan Paidi. 2000. *Peningkatan Kualitas Perkuliahan Melalui Authentic Assessment*. Laporan Penelitian Yogyakarta: Jurdik Biologi, FMIPA UNY.

- Bambang Subali, Paidi, Siti Umniyati, and Bernadeta Octavia. 2001. *Upaya Peningkatan Kualitas Penilaian Mengarah Ke Model Authentic Assessment. (dalam Mata Kuliah Mikrobiologi)*. Laporan Penelitian. Yogyakarta: Jurdik Biologi, FMIPA UNY.
- Bryce, T.G.K, McCall, J., MacGregor, J., Robertson, I.J., dan Weston, R.A.J. 1990. *Techniques for Assessing Process Skills in Practical Science: Teaching Guide*. Oxford: Heinemann Educational Books.
- Depdikbud. 1997. *Bahan Penataran Pengujian pendidikan*. Jakarta: Pusat Penelitian dan Pengembangan Sistem Pengujian Balitbang, Depdikbud.
- Depdikbud. 1995. *Kurikulum Sekolah menengah Umum (SMU): Petunjuk Teknis Mata pelajaran Biologi*. Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah, Depdikbud.
- Depdikbud. 1994. *Petunjuk Pelaksanaan Penilaian*. Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah, Depdikbud.
- Depdikbud. 1994. *Pedoman Program Perbaikan dan Pengayaan*. Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah, Depdikbud.
- Depdikbud. 1986. *Petunjuk Pelaksanaan Ebtanas*. Jakarta: Direktorat Jenderal Pendidikan Dasar dan Menengah, Depdikbud.
- Djemari Mardapi, Bambang Subali, Badrun Kartowagiran, Nukron dan satunggalno. 2001. *Sistem Ujian Akhir dalam Otonomi daerah*. Laporan penelitian sementara. Yogyakarta: Fakultas Teknik UNY.
- Djemari Mardapi, Badrun Kartowagiran, Satunggalno, Nurcholis, dan M. Fauzan. 1999. *Evaluasi penyelenggaraan Ebtanas*. Laporan penelitian. Kerjasama Pusiasjian Balitbang Dikbud dan Lembaga Penelitian IKIP Yogyakarta.
- Djemari, Bambang Subali, Agus Widiyantoro, dan Nukron 1999. *Survei Evaluasi Hasil Belajar di Kelas*. Laporan penelitian. Kerjasama Pusiasjian Balitbang Dikbud dan Lembaga Penelitian IKIP Yogyakarta.
- Gronlund, N.E. 1998. *Assessment of Student Achievement*. Boston: Allyn and Bacon.
- Marsh, C.J. 1996. *Handbook for Beginning Teachers*. Melbourne, Australia: Longman.
- Nitko, A.S. 1996. *Workshop Papers No. 2*. IKIP Yogyakarta, 22-24 Agustus 1996.
- Patric, G. dan Nix, P. 1991. *Educational Assessment and Reporting*. Sidney: Harcourt Brace Publisher.
- Toto Kuwato dan Djemari Mardapi. 1999. *Studi Pengembangan Sistem Ujian Berkesinambungan Sekolah Menengah Umum*. Laporan Penelitian. Kerjasama Direktorat Jenderal Pendidikan Dasar dan Menengah Depdik-bud bekerjasama dengan Fakultas Psikologi UGM.
- Towle, A. 1989. *Modern Biology*. Austin: Holt, Rinehart and Winston.
- Wuryadi dan Bambang Subali. 2000. *Profil Penyelenggaraan Kegiatan penilaian Prestasi Belajar IPA-Biologi/Biologi oleh Guru SLTP SMU di Propinsi DIY Ditinjau dari Latar Belakang Akademik Guru*. Laporan Penelitian. Yogyakarta: Jurdik Biologi, FMIPA UNY.