



**Prinsip Asesmen &
Evaluasi Pembelajaran**

Prof. Dr. Bambang Subali, M.S.

Prof. Dr. Bambang Subali, M.S.

**Prinsip Aesmen &
Evaluasi pembelajaran**

Prinsip Asesmen & Evaluasi Pembelajaran

Prof. Dr. Bambang Subali, M.S.

Undang-undang Republik Indonesia Nomor 19 Tahun 2002 tentang
Hak Cipta

Lingkup Hak Cipta

Pasal 2:

1. Hak Cipta merupakan hak eksklusif bagi Pencipta atau Pemegang Hak Cipta untuk mengumumkan atau memperbanyak ciptaannya, yang timbul secara otomatis setelah suatu ciptaan dilahirkan tanpa mengurangi pembatasan menurut peraturan perundang-undangan yang berlaku.

Ketentuan Pidana

Pasal 72:

1. Barangsiapa dengan sengaja atau tanpa hak melakukan perbuatan sebagaimana dimaksudkan dalam Pasal 2 ayat (1) atau Pasal 49 ayat (1) dan (2) dipidanakan dengan pidana penjara masing-masing paling singkat 1 (satu) bulan dan/atau denda paling sedikit Rp 1.000.000,00 (satu juta rupiah), atau pidana penjara paling lama 7 (tujuh) tahun dan/atau denda paling banyak Rp 5.000.000.000,00 (lima miliar rupiah).
2. Barangsiapa dengan sengaja menyiarkan, memamerkan, mengedarkan, atau menjual kepada umum suatu ciptaan atau barang hasil Pelanggaran Hak Cipta atau Hak Terkait sebagaimana dimaksudkan dalam ayat (1) dipidanakan dengan pidana penjara paling lama 5 (lima) tahun dan/atau denda paling banyak Rp 500.000.000,00 (lima ratus juta rupiah).

Prinsip Asesmen & Evaluasi Pembelajaran

Cetakan 1, November 2012

Penulis:

Prof. Dr. Bambang Subali, M.S.

Editor: Maman Suryaman

Tata Letak: Dhian Hapsari

Desain Cover: Kuncoro W.D

Dicetak dan diterbitkan oleh:

UNY Press

Jl. Affandi (Gejayan), Gg. Alamanda, Kompleks FT

Kampus Karangmalang, Yogyakarta

Telp. (0274) 589346

Email: redaksi.unypress@gmail.com

ISBN 978-979-8418-94-5

Perpustakaan Nasional: Katalog dalam Terbitan (KDT)
192 hlm; 16 x 23 cm

KATA PENGANTAR

Puji dan syukur saya panjatkan ke hadirat Allah sehingga buku Prinsip Asesmen dan Evaluasi Pembelajaran dapat saya terbitkan. Buku ini merupakan perbaikan dari buku *Penilaian, Evaluasi, dan Pembelajaran Biologi* yang diterbitkan oleh kerjasama FMIPA UNY dan JICA, dan sudah disesuaikan dengan konsep baru dan peraturan yang berlaku sejak adanya standar nasional pendidikan di tingkat pendidikan dasar dan menengah.

Sebagaimana diketahui bersama bahwa dengan adanya standar nasional pendidikan, maka standar penilaian juga menjadi salah satu produk hukum yang harus dipatuhi oleh guru. Dengan demikian, calon guru juga perlu memahaminya. Oleh karena itu, selain buku ini dijadikan pegangan bagi mahasiswa yang menempuh mata kuliah Penilaian, Evaluasi, dan Remedi Pembelajaran Biologi, diharapkan juga dapat dijadikan buku pegangan bagi para guru di lapangan dalam melaksanakan kegiatan profesinya.

Penggunaan istilah penilaian dalam buku ini yang dimaksudkan adalah *assessment*. Dalam buku-buku lama sering tidak jelas perbedaan antara penilaian dan evaluasi. Kedua istilah tersebut dianggap sama. Oleh karena itu, buku ini ditulis dengan judul *Prinsip Asesmen dan Evaluasi Pembelajaran*. Sementara, ada pula sebagian ahli yang menyamakan istilah asesmen dengan tes/ujian.

Kritik dan saran sangat saya perlukan untuk penyempurnaan buku ini ke depan.

Yogyakarta, Oktober 2012

Penulis

DAFTAR ISI

Kata Pengantar

Daftar Isi

Bab I	Konsep Dasar Aesmen dan Evaluasi Pendidikan—1
Bab II	Objek Aesmen—33
Bab III	Teknik Aesmen—51
Bab IV	Cara Penyiapan dan Penyelenggaraan Ujian—57
Bab V	Pengembangan Instrumen Aesmen—63
Bab VI	Telaah Instrumen—91
Bab VII	Validitas dan Reliabilitas—107
Bab VIII	Pengolahan Skor—129
Bab IX	Tindak Lanjut Hasil Aesmen—133
Bab X	Program Remedi dan Pengayaan—143
Daftar Pustaka—	183
Indeks	

I

Konsep Dasar Aesmen dan Evaluasi Pendidikan

Dalam dunia pendidikan, penilaian atau asesmen (*assessment*) diartikan sebagai prosedur yang digunakan untuk mendapatkan informasi untuk mengetahui taraf pengetahuan dan keterampilan peserta didik yang hasilnya akan digunakan untuk keperluan evaluasi. Informasi adalah data yang diperoleh melalui pengukuran dan nonpengukuran termasuk di dalamnya dengan melakukan observasi, kelas, menggunakan tes yang standar atau tes buatan guru, proyek, dan protfolio subjek belajar (*Assessment is a process of gathering information to measure a student's level of knowledge and/or skill. Assessment can incorporate many methods including classroom observation, standardized tests, teacher-developed tests, classroom projects, and portfolios of student work*).

Evaluasi merupakan suatu proses yang sistematis yang dilaksanakan untuk mengetahui tingkat keberhasilan dan efisiensi dari program yang bersangkutan. Dalam hal ini termasuk di dalamnya untuk mengetahui keberhasilan seluruh subjek belajar yang menem-puh suatu program.

Untuk memperoleh informasi yang berupa data kuantitatif dilakukan melalui pengukuran. Jadi, pengukuran (*measurement*) merupakan proses untuk memperoleh deskripsi numerik atau kuantitatif tentang tingkatan karakteristik yang dimiliki seseorang dengan aturan tertentu.

Untuk memperoleh data kuantitatif dapat dilakukan pengukuran melalui tes dan nontes. Tes merupakan metode pengukuran yang menggunakan alat ukur berbentuk satu set pertanyaan untuk mengukur sampel tingkah laku, dan jawabannya dapat dikategorikan benar dan salah. Nontes merupakan metode pengukuran yang menggunakan alat ukur untuk mengukur sampel tingkah laku, tetapi jawa-

bannya tidak dapat dikategorikan benar dan salah, misalnya positif dan negatif, setuju dan tidak setuju, suka dan tidak suka.

Berkait dengan kegiatan pengukuran, di dalam Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 tentang Standar Penilaian Pendidikan dinyatakan bahwa ulangan adalah proses yang dilakukan untuk mengukur pencapaian kompetensi peserta didik secara berkelanjutan dalam proses pembelajaran, untuk memantau kemajuan, melakukan perbaikan pembelajaran, dan menentukan keberhasilan belajar peserta didik. Ulangan harian adalah kegiatan yang dilakukan secara periodik untuk mengukur pencapaian kompetensi peserta didik setelah menyelesaikan satu Kompetensi Dasar (KD) atau lebih. Ulangan tengah semester adalah kegiatan yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik setelah melaksanakan 8 – 9 minggu kegiatan pembelajaran. Cakupan ulangan meliputi seluruh indikator yang merepresentasikan seluruh KD pada periode tersebut. Ulangan akhir semester adalah kegiatan yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik di akhir semester. Cakupan ulangan meliputi seluruh indikator yang merepresentasikan semua KD pada semester tersebut. Ulangan kenaikan kelas adalah kegiatan yang dilakukan oleh pendidik di akhir semester genap untuk mengukur pencapaian kompetensi peserta didik di akhir semester genap pada satuan pendidikan yang menggunakan sistem paket. Cakupan ulangan meliputi seluruh indikator yang merepresentasikan KD pada semester tersebut. Ujian sekolah/madrasah adalah kegiatan pengukuran pencapaian kompetensi peserta didik yang dilakukan oleh satuan pendidikan untuk memperoleh pengakuan atas prestasi belajar dan merupakan salah satu persyaratan kelulusan dari satuan pendidikan.

A. Makna Evaluasi

Suatu program, termasuk di dalamnya program pendidikan yang diwujudkan dalam bentuk kurikulum, adalah suatu kegiatan yang terencana yang lengkap dengan rincian tujuan beserta jenis-jenis komponen kegiatan pembelajarannya, seperti tujuan, strategi, materi/bahan ajar, sumber belajar, dan alokasi waktu. Oleh karena

itu, apakah suatu program yang diimplementasikan benar-benar berharga, diperlukan adanya evaluasi. Evaluasi yang dimaksud adalah suatu proses yang sistematis yang dilaksanakan untuk mengetahui tingkat keberhasilan dan efisiensi dari program yang bersangkutan.

Evaluasi terhadap tingkat efisiensi terutama ditujukan kepada program yang dilaksanakan berulang-ulang. Karena keberhasilan suatu program tidak dapat terlepas dari segi pelaksanaannya, evaluasi terhadap suatu program menyangkut berbagai hal yang terkait, baik yang menyangkut kualitas masukan, kualitas proses maupun kualitas hasil pelaksanaannya. Selain itu, evaluasi dapat dilaksanakan atas dasar sekuen implementasi program, dapat pula dilakukan terhadap setiap komponen program. Dalam program pendidikan misalnya, evaluasi menjadi sangat kompleks karena dapat dilakukan mulai dari evaluasi terhadap perencanaannya yakni mencakup evaluasi terhadap kurikulum, ketersediaan sarana dan prasarana pendukung, ketersediaan dan profesionalitas tenaga yang terlibat baik tenaga edukatif maupun administratif.

Evaluasi terhadap implementasi program ditujukan untuk mengetahui kelancaran pelaksanaan program dan kesesuaian antara implementasi dengan perencanaan program yang telah disusun. Dengan kata lain, apakah kurikulum beserta silabus dan rencana pelaksanaan pembelajaran (RPP) terimplementasi dengan baik. Bila silabus dan RPP tidak dapat terimplementasikan dengan baik, perlu ditelusuri apa saja yang menjadi penyebabnya. Selama implementasi dimungkinkan juga pendidik sebagai ujung tombak di lapangan memodifikasi kegiatan pembelajaran bila implementasi RPP tidak dapat terlaksana sehingga RPP disempurnakan lagi untuk tahun berikutnya.

Evaluasi terhadap hasil ditunjukkan untuk mengetahui efisiensi waktu penyelenggaraan program, dan yang tidak kalah pentingnya adalah seberapa jauh keefektifan program yang telah diselenggarakan. Meskipun suatu program dinyatakan efektif, jika tidak sesuai dengan alokasi waktu yang direncanakan, program tersebut perlu ditinjau ulang untuk diterapkan lagi pada tahun-tahun berikutnya.

Kurikulum SD/MI mengalokasikan waktu selama enam tahun bagi peserta didik untuk menempuhnya. Hal ini mengandung maksud bahwa peserta didik yang tidak mengalami kesulitan belajar yang sangat parah akan dapat menempuhnya selama enam tahun. Kurikulum SMP/MTs dan SMA/MA mengalokasikan waktu tiga tahun bagi peserta didik untuk menempuhnya. Hal ini juga mengandung maksud bahwa peserta didik yang tidak mengalami kesulitan belajar yang sangat parah akan dapat menempuhnya selama tiga tahun.

Kurikulum S-1 yang mengalokasikan waktu selama empat tahun bagi mahasiswa untuk menempuhnya umumnya diberi ketetapan daat ditempuh paling lama selama tujuh tahun. Kurikulum yang demikian menunjukkan keadaan yang mendua arti. Di satu sisi kurikulum S-1 dirancang agar dapat ditempuh mahasiswa selama empat tahun. Di sisi yang lain, mahasiswa masih dapat dinyatakan lulus asalkan tidak lebih dari tujuh tahun. Jadi, sebagai suatu program, kurikulum tersebut dinyatakan benar-benar efisien jika semua mahasiswa dapat menyelesaikannya selama empat tahun. Namun, di lapangan hanya sedikit sekali mahasiswa yang dapat menyelesaikannya selama empat tahun. Jika ditinjau dari segi efisiensi program maka semakin lama waktu yang ditempuh mahasiswa menyelesaikannya akan menjadi semakin kurang efisien.

Evaluasi terhadap dampak program pendidikan ditujukan untuk mengetahui seberapa jauh para lulusan dapat menjalankan tugas profesinya sebagai tenaga ahli yang profesional dalam bidangnya. Dalam hal ini, evaluasi dampak dapat dipantau melalui *tracer study*, baik melalui jaringan *on-line* ataupun dengan memberikan kuesioner kepada pihak pengguna lulusan. Hal yang sering diabaikan adalah bahwa dampak program bagi program yang bersifat terminal adalah dinyatakan positif tinggi bila banyak lulusan dapat bekerja sesuai dengan bidang keahliannya dan menunjukkan prestasi kinerja yang tinggi. Misalnya, lulusan SMK dapat tersebar dalam dunia kerja di bidangnya dan dapat bekerja secara profesional. Demikian pula dengan lulusan Program S-1 Pendidikan dinyatakan positif tinggi bila banyak lulusan dapat bekerja sesuai dengan bidang profesinya dan

menunjukkan prestasi kinerja yang tinggi. Namun, karena minimnya lapangan kerja, dunia kerja sering kali digelar para lulusan tidak lagi sesuai dengan bidang profesinya. Akibatnya, akan menjadi sangat sulit untuk mengevaluasi apakah mereka mampu meraih prestasi yang tinggi karena dampak dari program pendidikan yang ditempuhnya atau karena ia memang memiliki potensi yang tinggi. Hal yang harus disadari bahwa orang yang memiliki potensi yang tinggi akan dapat belajar hal-hal baru dalam waktu yang singkat dan akhirnya dapat meraih prestasi yang tinggi pula.

B. Manfaat Evaluasi

Suatu program ditunjukkan agar sasaran-sasaran yang telah ditetapkan dapat tercapai sesuai harapan. Oleh karena itu, manfaat atau kegunaan evaluasi berupa pengambilan keputusan atau untuk pertanggungjawaban terhadap kegiatan yang telah dilaksanakan. Pertanggungjawaban ditujukan kepada pihak yang menghendaki atau pihak sponsor. Dalam hal program pendidikan, pihak yang terkait adalah pemerintah dan masyarakat, khususnya pihak orangtua subjek belajar.

Pengambilan keputusan didasarkan pada hasil penilaian/assessment dan dilakukan dalam upaya untuk pengendalian kegiatan jika program masih berlangsung atau untuk upaya penyempurnaan untuk pelaksanaan selanjutnya. Dengan demikian, dengan adanya hasil evaluasi akan dapat diambil kebijaksanaan apakah suatu program akan dilaksanakan lagi pada periode berikutnya ataupun perlu direvisi terlebih dahulu. Bahkan, jika perlu, digantikan dengan program lain. Jadi, setiap program yang diimplementasikan hendaknya dilengkapi dengan kegiatan pengukuran dan penilaian yang dapat dijadikan dasar dilakukannya evaluasi agar dapat diambil suatu kebijaksanaan untuk menentukan apakah program yang dimaksud dilaksanakan ulang, direvisi, atau diubah.

Dalam dunia pendidikan, program yang ada dapat berbeda-beda tingkatannya. Ada program pendidikan yang berlaku pada tingkat kementerian. Dalam hal ini, program tersebut menjadi tang-

gung jawab Menteri Pendidikan dan Kebudayaan. Ada program pendidikan yang berlaku pada tingkat wilayah provinsi dan kabupaten/kota. Dalam hal ini program pendidikan tersebut menjadi tanggung jawab Kepala Dinas Pendidikan Provinsi dan Kepala Dinas Pendidikan Kabupaten/Kota. Ada program pendidikan yang berlaku pada tingkat sekolah atau satuan pendidikan. Dalam hal ini, program tersebut menjadi tanggung jawab Kepala Sekolah. Program yang paling rendah tingkatannya adalah program pendidikan yang berlaku pada tiap kelas dalam suatu sekolah. Dalam hal ini, program pendidikan tersebut menjadi tanggung jawab guru sebagai pendidik. Dilihat dari segi program pendidikan yang penyelenggaraannya dilakukan oleh pemerintah dan ada pula yang diselenggarakan oleh pihak swasta sehingga dikenal dengan sekolah negeri dan sekolah swasta serta perguruan tinggi negeri dan perguruan tinggi swasta.

Penanggung jawab program di masing-masing tingkatan harus melakukan berbagai pengukuran untuk menghimpun data yang diperlukan guna melakukan penilaian, dan dari hasil penilaian yang diperoleh dijadikan dasar untuk melakukan evaluasi terhadap kelancaran, ketepatan, keefektifan, dan efisiensi program yang menjadi tanggung jawabnya. Dengan kata lain setiap penanggung jawab program di setiap tingkatan harus melakukan evaluasi diri (*self evaluation*) atas penyelenggaraan program.

Idealnya, secara berjenjang, guru harus memantau/memonitor dan mengevaluasi kinerja peserta didik selama menempuh program pembelajaran. Kepala sekolah harus memantau/memonitor dan mengevaluasi kinerja guru/pendidik dalam merancang dan mengimplementasikan program pembelajaran di setiap kelas. Kepala Dinas Pendidikan Kabupaten/Kota harus memantau/memonitor dan mengevaluasi kinerja Kepala Sekolah di wilayahnya dalam merancang dan mengimplementasikan program pendidikan di masing-masing sekolah dalam bentuk Kurikulum Tingkat satuan pendidikan (KTSP). Kepala Dinas Provinsi harus memantau/memonitor dan mengevaluasi kinerja Kepala Dinas pendidikan Kabupaten/Kota dalam merancang dan mengimplementasikan program pendidikan masing-masing kabupaten/kota. Menteri Pendidikan dan Kebudayaan harus memantau/

memonitor dan mengevaluasi kinerja Kepala Dinas Pendidikan provinsi dalam merancang dan mengimplementasikan program pendidikan di masing-masing provinsi. Namun, dengan adanya otonomi daerah, otoritas penyelenggaraan program di tingkat Kabupaten/Kota, termasuk program pendidikan, tidak semua aspeknya dapat dikoordinasikan oleh Dinas Pendidikan Provinsi.

Dari segi pelaksanaannya, evaluasi terhadap suatu program pendidikan dapat dilakukan oleh perencanaan dan pelaksanaan program yang bersangkutan melalui evaluasi diri, dapat pula diserahkan kepada pihak lain/pihak eksternal yang dianggap ahli dan tidak terlibat dalam pelaksanaan program. Jika evaluasi dilakukan terhadap setiap satuan kecil dari suatu program pendidikan yang lebih besar yang masih berjalan dalam upaya untuk pengendalian pelaksanaan program, maka evaluasi dilakukan oleh pihak pelaksana program. Dalam hal ini dikenal dengan evaluasi pada skala mikro.

Sebagai contoh evaluasi pada skala mikro adalah evaluasi program pembelajaran di kelas. Evaluasi program dalam skala mikro ini pada umumnya dilaksanakan secara periodik dalam waktu yang relatif singkat selama berlangsungnya program pembelajaran di kelas. Dalam hal ini, dikenal dengan istilah evaluasi formatif. Evaluasi formatif adalah evaluasi yang dilaksanakan selama berlangsungnya program pembelajaran. Oleh karena itu, dari sisi program disebut pula dengan istilah evaluasi proses pembelajaran. Hal yang dijadikan dasar untuk melakukan evaluasi formatif ini adalah hasil penilaian/asesmen formatif yang datanya dihimpun melalui pengukurur dan pengamatan oleh guru yang bersangkutan. Berdasarkan hasil penilaian selama berlangsungnya proses pembelajaran oleh guru yang bersangkutan Guru dapat menghimpun data dan informasi yang berkaitan dengan kemajuan belajar peserta didik, juga jika ada kesulitan yang dihadapi peserta didik. Atas dasar hasil evaluasi formatif selama berlangsungnya proses pembelajaran itulah guru dapat menyempurnakan program pembelajarannya agar sasaran/target pembelajaran yang dapat tercapai sesuai harapan. Dalam skala mikro/semipit, orientasi utama evaluasi program ditujukan kepada masalah metode pembe-

lajaran. Sebaliknya, evaluasi juga dilakukan pada skala makro/luas yang dititikberatkan pada masalah efisiensi pelaksanaan, yaitu berkenaan dengan strategi dan pelaksanaan program. Oleh karena itu, evaluasi pada skala makro akan lebih baik jika dilakukan oleh pihak luar/pihak independen.

C. Evaluasi Proses dan Hasil Belajar

Pencapaian prestasi atau hasil belajar diperoleh setelah dilaksanakannya suatu program pembelajaran. Evaluasi pencapaian hasil belajar merupakan langkah mengetahui seberapa jauh target pembelajaran dapat dicapai. Kumulatif pencapaian hasil belajar dari seluruh bidang studi akan mencerminkan keberhasilan suatu program pendidikan atau suatu kurikulum sekolah. Dengan demikian, kedudukan evaluasi hasil belajar suatu bidang studi hanyalah sebagai bagian kecil dari evaluasi pencapaian program pendidikan. Oleh karena itu, setiap pengampu bidang studi dituntut kesadarannya bahwa dirinya sebagai bagian dari sistem pendidikan. Seorang guru bukanlah pribadi yang lepas, ia tidak dapat dilepaskan dari sistem pendidikan. Oleh karena itu, apa yang diambil sebagai keputusan saat ia melakukan evaluasi atas program yang diselenggarakan merupakan bagian dari sistem pendidikan.

Jika dalam evaluasi hasil belajar pada masa lalu hanya untuk mengetahui sejauh mana isi/materi atau bahan ajar sudah dikuasai oleh subjek belajar, sekarang evaluasi hasil belajar lebih diorientasikan kepada seberapa jauh sasaran-sasaran pendidikan telah dapat dicapai dan disertai pula dengan pelacakan peran berbagai faktor penentu aktualisasi proses pembelajaran. Sebagaimana telah diketahui bersama bahwa ada banyak faktor penentu aktualisasi proses pembelajaran. Faktor-faktor tersebut antara lain berupa faktor tujuan pendidikan, tenaga pengajar (minat, sikap, pandangan, langgam mengajar, stabilitas emosional, dll.), isi/materi pelajaran atau bahan ajar, organisasi materi, metodologi atau sistem metodologi yang dikembangkan, peningkatan yang digunakan, pengelolaan kelas yang diselenggarakan, bimbingan dan penyuluhan yang diberikan, sarana dan prasarana

pendidikan, sistem administrasi, kondisi subjek belajar, maupun sistem penilaian, dan sistem evaluasi yang digunakan.

Selain itu, aktualisasi proses pembelajaran tidak dapat dilepaskan dari faktor lingkungan yang melingkupinya, baik yang bersifat fisik maupun sosial. Jika dikaitkan dengan fungsi guru, maka aktualisasi proses pembelajaran sebagai ujung nyata implementasi program, tidak dapat lepas dari kualitas program itu sendiri. Kualitas program tidak dapat lepas dari kualitas pendidik dan tenaga pendidikan lainnya yang terlibat dalam perancangannya. Dalam skala mikro, baik buruknya penyelenggaraan program pembelajaran yang dirancang oleh guru sebagai seorang pendidik sangat tergantung kepada sejauh mana ia memahami kurikulum yang sedang berlaku, karena dalam hal ini dituntut kemampuannya dalam melakukan *restatement* kurikulum. Sementara, banyak guru yang sekedar melaksanakan kurikulum dengan asal menyusun tanpa memperhatikan potensi subjek belajar dan spesifikasi lingkungan dimana satuan pendidikan berada. Dalam implementasi kurikulum operasional sebagai tuntutan Kurikulum Tingkat Satuan Pendidikan, sudah saatnya dikembangkan program pembelajaran yang menjadi karakteristik satuan pendidikan yang bersangkutan tanpa meninggalkan standar isi yang menjadi tuntutan minimal dalam implementasi kurikulum di tingkat satuan pendidikan.

Karena tujuan belajar adalah terjadinya perubahan tingkah laku ke arah yang diinginkan, maka evaluasi pencapaian hasil belajar juga merupakan usaha untuk menetapkan tercapainya perubahan perilaku, kecakapan dan status subjek belajar selama dan setelah melakukan kegiatan belajar dari peserta program yang diselenggarakan. Perubahan tingkah laku tersebut sudah dirumuskan dalam tujuan pembelajaran, yang dalam kurikulum berbasis kompetensi tujuan pembelajaran dirumuskan dalam bentuk Standar Kompetensi Lulusan (SKL), Standar Kompetensi (SK), dan Kompetensi Dasar.

Berkait dengan kurikulum berbasis kompetensi, evaluasi pencapaian hasil belajar bukan semata-mata ditujukan untuk menetapkan apa yang telah berhasil diketahui peserta program, tetapi lebih diorien-

tasikan kepada apa yang dapat dikerjakan dalam bentuk kinerja (*performance*), apa sikap positif yang tumbuh di dalam diri subjek belajar, juga bagaimana kemampuan peserta program mengaktualisasikan diri dalam hidup bersama sesuai dengan hakekat belajar yakni *learning to know, learning to do, learning to be dan learning to live together*. Oleh karena itu, dewasa ini di banyak negara, penilaian menjadi bergeser ke arah asesmen yang otentik yakni penilaian yang didasarkan pada pemecahan permasalahan dalam kehidupan sehari-hari dalam bentuk *performance assessment* sebagai suatu *alternative assessment*.

Selama proses evaluasi formatif berlangsung guru sebagai perancang dan pelaksana program dapat melakukan penyempurnaan program pembelajarannya. Dengan demikian, diharapkan akan dapat meningkatkan prestasi belajar tiap peserta didik. Hasil penilaian yang formatif sebagai dasar melakukan evaluasi formatif bukanlah hasil yang final. Oleh karena itu, nilai formatif merupakan "nilai antara". Nilai formatif seorang subjek belajar digunakan untuk memantau ada tidaknya kesulitan yang dialaminya selama terlibat dalam proses pembelajaran. Atas dasar hasil penilaian formatif itulah guru melakukan evaluasi formatif atas program yang dirancang agar ia dapat memutuskan perlu tidaknya ia memperbaiki program pembelajarannya.

Pada akhir program, guru melakukan penilaian sumatif untuk mengetahui keberhasilan atau prestasi akhir setiap peserta didik. Evaluasi sumatif merupakan keputusan untuk menentukan hasil akhir keberhasilan subjek belajar dalam menempuh program belajar. Evaluasi akhir semester untuk memutuskan keberhasilan subjek belajar menempuh program selama satu semester. Evaluasi akhir tahun untuk memutuskan apakah subjek belajar harus gagal sehingga tinggal kelas atau dapat dinyatakan sukses untuk naik kelas.

Dari uraian di atas tampak bahwa demikian evaluasi pencapaian belajar atau evaluasi hasil belajar meliputi evaluasi sumatif (*summative evaluation*) yang dikenakan pada subjek belajar sebagai penempuh program pembelajaran, guru sebagai perancang dan pelaksana program (yang rancangannya diwujudkan dalam bentuk kurikulum, silabus sampai dengan RPP), dalam hubungannya dengan

efektivitas kegiatan belajar dalam suatu program tertentu, sedangkan evaluasi formatif (*formative evaluation*) yang dilakukan untuk mengetahui keberhasilan subjek belajar sebagai peserta program pembelajaran selama berlangsungnya proses pembelajaran.

Jika benar-benar dapat didudukkan sebagai evaluasi proses belajar, maka akan ada upaya-upaya yang melekat atau yang menyertainya yaitu upaya untuk meningkatkan efektivitas proses belajar sehingga dapat mencapai hasil yang lebih baik. Akan sangat ideal pula jika sebelum pelaksanaan proses pembelajaran dilaksanakan evaluasi penempatan (*placement evaluation*) dengan menghimpun data melalui penempatan/asesmen penempatan (*placement assessment*) untuk menjawab pertanyaan

- 1) sejauh mana subjek belajar sebagai peserta program menguasai kemampuan prasyarat yang diperlukan untuk proses pembelajaran yang akan diselenggarakan;
- 2) seberapa jauh subjek belajar sebagai peserta program telah menguasai kompetensi yang ditargetkan,
- 3) seberapa jauh subjek belajar sebagai peserta program bernilai terhadap program pembelajaran yang akan diselenggarakan.

Dari ragam evaluasi yang ada, maka tujuan evaluasi menjadi spesifik. Berikut ini adalah tujuan masing-masing evaluasi.

1. **Evaluasi sumatif** didasarkan pada kumulatif hasil **penilaian sumatif** subjek belajar dalam menempuh program. Dalam hal ini pengertian penilaian sumatif adalah hasil final dari subjek belajar menempuh suatu program. Misalnya, nilai sumatif dalam suatu program semester diperoleh melalui ulangan akhir suatu pokok bahasan (setelah dilakukan proses remedi bagi yang mengalami kegagalan dan dilakukan program pengayaan bagi yang sudah menguasai target pembelajaran berdasarkan penilaian formatif), ulangan

tengah semester, dan ulangan akhir semester. Adapun tujuan evaluasi sumatif adalah untuk:

- a. menentukan nilai akhir seluruh peserta penempuh program pembelajaran, agar dapat dinyatakan berhasil atau gagal. Bila berhasil maka akan dapat diberi sertifikat karena ia telah menguasai kecakapan ataupun keterampilan tertentu yang ditargetkan dalam program pembelajaran yang dirancang;
- b. meramalkan kecakapan subjek belajar dalam menyelesaikan program/semester berikutnya;
- c. menetapkan efektivitas dan efisiensi penyelenggaraan suatu program pembelajaran;
- d. dalam konteks untuk seleksi, seperti seleksi masuk, berarti menetapkan yang layak lolos seleksi. Bila untuk seleksi menetapkan juara mewakili satuan pendidikan yang bersangkutan, berarti menetapkan siapa yang layak lolos menjadi wakil satuan pendidikan yang bersangkutan.

2. **Evaluasi formatif** didasarkan pada hasil penilaian formatif selama subjek belajar sebagai penempuh program pembelajaran mengikuti proses pembelajaran dalam kaitannya dengan penyelenggaraan program, dan tujuannya untuk:

- a. menetapkan langkah-langkah/urutan kegiatan belajar selanjutnya agar supaya lebih efektif dan efisien;
- b. pendalaman dan pemantapan penguasaan perilaku yang ditargetkan;
- c. mendiagnosis kesulitan belajar, dalam arti bahwa subjek belajar yang mendapat nilai jelek identik belum menguasai perilaku yang ditargetkan;
- d. mencari cara mengatasi kesulitan belajar jika subjek belajar dinyatakan gagal, berdasar kegiatan belajar yang telah dilakukan;

- e. umpan balik bagi guru dalam mengelola kegiatan pembelajaran sehingga mengetahui seberapa jauh tujuan yang ditetapkan sudah dapat dicapai;
- f. meramalkan seberapa jauh keberhasilan peserta program belajar dalam mengikuti penilaian sumatif;
- g. mengetahui seberapa jauh seluruh subjek belajar sebagai penempuh program pembelajaran akan berhasil dalam mengikuti proses pembelajaran sampai akhir program, berdasarkan kecakapan dan keterampilan yang dikuasainya sekarang, dalam konteks bahwa subjek belajar sebagai masukan;
- h. mengetahui subjek belajar yang mana yang harus dibantu melalui program remedi agar ia dapat berhasil menempuh program yang ditempuh;
- i. mendiagnosis penyebab kegagalan subjek belajar dalam dalam menguasai kemampuan yang ditargetkan dari program pembelajaran yang diselenggarakan.

3. **Evaluasi penempatan** didasarkan pada hasil penilaian terhadap subjek sebelum menempun program pembelajaran, dan bertujuan untuk:

- a. mengetahui penguasaan kemampuan prasyarat yang diperlukan dalam KBM yang akan diselenggarakan;
- b. menajagi penguasaan subjek belajar sebagai program terhadap kemampuan yang ditargetkan;
- c. meneliti interest, langgam belajar, ataupun karakteristik personal subjek belajar sebagai peserta program pembelajaran;
- d. mendiagnosis kemampuan subjek belajar yang mengalami kegagalan dalam menguasai kemampuan prasyarat yang diperlukan.

E. Prinsip Penilaian dan Evaluasi dalam Pendidikan

Ada beberapa prinsip penilaian dan prinsip evaluasi pencapaian hasil belajar yang harus dijadikan landasannya. Prinsip-prinsip evaluasi program pembelajaran antara lain:

1. Prinsip integralitas/menyeluruh

Bahwa evaluasi harus didasarkan pada penilaian terhadap setiap subjek belajar secara komprehensif, mencakup seluruh aspek, baik yang menyangkut kemampuan (*ability*) dan personalitas (*apthitude*), atau menyangkut aspek pengetahuan, keterampilan, sikap, dan perilaku. Untuk itu diperlukan banyak teknik evaluasi yang harus diterapkan karena setiap macam penilaian memerlukan teknik tersendiri dan setiap teknik penilaian memiliki kelemahan.

2. Prinsip kontinuitas/berkesinambungan

Pelaksanaan evaluasi harus didasarkan pada proses penilaian yang dilakukan secara kontinyu dan periodik dengan harapan adanya kegiatan penilaian berfungsi untuk membimbing perkembangan subjek belajar.

3. Prinsip objektivitas

Evaluasi hasil belajar harus didasarkan pada penilaian terhadap setiap subjek belajar yang harus bebas dari unsur yang bersifat subjektif, harus dapat dimaknakan/ditafsirkan dengan jelas dan tegas. Semakin banyak data yang dijadikan dasar penilaian, maka hasil penilaian akan semakin objektif.

4. Prinsip berorientasi pada tujuan

Evaluasi hasil belajar harus ditunjukkan untuk mengetahui keberhasilan peserta didik mednempuh program, sehingga penilaian hasil belajar terhadap setiap subjek belajar hendaknya dimaksudkan untuk mengetahui seberapa jauh subjek belajar menguasai tujuan yang ditargetkan.

5. Prinsip terbuka

Proses evaluasi didasarkan pada proses penilaian yang dilakukan terhadap setiap subjek belajar perlu diketahui oleh semua pihak. Oleh karena itu hasil penilaian harus disebarluaskan (dapat diketahui dan diterima) oleh pihak-pihak yang terkait (siswa, orang tua, sekolah, pemerintah dan masyarakat).

6. Prinsip kebermaknaan

Tujuan evaluasi adalah untuk mengetahui keberhasilan program. Oleh karena itu evaluasi harus didasarkan pada hasil penilaian terhadap setiap subjek belajar dan harus memiliki kebermaknaan bagi orang yang menggunakan. Bagi guru, selain harus berguna untuk meningkatkan hasil belajar siswa juga untuk umpan balik dalam upaya memperbaiki proses pembelajaran. Bagi siswa juga harus berguna untuk memperbaiki diri dalam hal cara belajarnya agar pada penilaian berikutnya hasilnya akan lebih baik..

7. Prinsip kesesuaian

Evaluasi terhadap keberhasilan peserta didik sebagai peserta program harus didasarkan pada seberapa jauh hasil penilaian terhadap setiap subjek belajar bersesuaian dengan pendekatan/strategi/metode kegiatan pembelajaran yang diterapkan dalam rangka pelaksanaan kurikulum. Apabila dalam pelaksanaan kurikulumnya menggunakan pendekatan induktif, penilaiannya juga harus menjadikan pendekatan induktif menjadi salah satu aspek yang dinilai. Kalau dalam pembelajarannya menerapkan metode eksperimen, maka kemampuan bereksperimen harus menjadi salah satu aspek yang dinilai.

8. Prinsip determinasi dan klarifikasi

Evaluasi hasil belajar harus didasarkan pada kejelasan apa yang dinilai dalam kegiatan penilaian terhadap setiap subjek belajar sebagai pesewrta program, apakah tentang kemajuan belajarnya ataukah pencapaian hasil akhirnya.

9. Prinsip mendidik

Evaluasi diharapkan menerapkan prinsip mendidik. Hasil penilaian terhadap setiap subjek belajar hendaknya dapat digunakan untuk membina dan memberikan motivasi pada subjek belajar agar dapat meningkatkan hasil belajarnya. Hasil penilaian harus dinyatakan dan dapat dirasakan sebagai suatu penghargaan bagi subjek belajar yang berhasil dan merupakan peringatan bagi subjek belajar yang gagal Hasil penilaian yang dicantumkan dalam rapor hendaknya merupakan pertanggungjawaban subjek belajar yang bersangkutan kepada orang tuanya yang telah mempercayakan pendidikan anaknya kepada pihak sekolah/guru. Dengan demikian, penilaian dapat memperkuat perilaku dan sikap subjek belajar.

Prinsip evaluasi program pendidikan menurut Cronbach menyangkut hal-hal sebagai berikut.

1. **Kebijaksanaan pendidikan adalah kebijaksanaan pemerintahan, sehingga evaluator harus membantu pemerintah.** Konsekuensi lanjut adalah bahwa guru harus mengikuti kebijaksanaan sekolah dalam melakukan evaluasi.
2. **Evaluasi adalah seni, evaluasi yang baik bukan mengandalkan hasil penilaian yang tunggal.** Dalam hal ini semakin banyak data dan informasi tentangnya akan dapat dilakukan evaluasi yang semakin akurat. Pendekatan humanistik kualitatif dan positifistik kuantitatif akan dapat dipadukan secara harmonis mengingat peserta didik adalah manusia dengan segala variasi yang dimilikinya.
3. **Evaluasi bukan keputusan yang absolut, tujuannya memilihkan alternatif bagi pengambil kebijaksanaan untuk mengambil keputusan.** Dalam hal ini evaluasi yang hanya didasarkan pada data kuantitatif boleh jadi akan memiliki kelemahan tanpa disertai dengan penggalan lebih jauh apa dibalik angka-angka yang ada.
4. **Tidak ada orang yang mampu membuat seluruh pertimbangan dalam merancang evaluasi dan menafsirkan hasil evaluasi.** Dalam hal ini harus disadari bahwa harus diingat bahwa aktualisasi proses pembelajaran sebagai bentuk implementasi kurikulum sebagai program suatu program pembelajaran sangat dipengaruhi oleh banyak faktor seperti "raw in-put", ketersediaan sarana-prasarana, pandangan dan kompetensi serta komitmen guru, motivasi peserta didik sebagai peserta program, dan faktor lingkungan baik lingkungan fisik maupun lingkungan sosial peserta didik.
5. **Evaluator tidak hanya mengacu pada salah satu aliran evaluasi dalam melaksanakan evaluasi. Metode objektif kuantitatif dan humanistik kualitatif harus saling melengkapi.** Dalam hal ini harus disadari bahwa fungsi penyelenggaraan pendidikan tidak dapat dipisahkan antara fungsi edukasi dan fungsi pembelajaran.
6. **Evaluasi sebagai suatu proses harus bersifat kontinyu dan luwes.** Dalam hal ini harus disadari bahwa data dan informasi haruslah dihimpun setiap waktu, bukan hanya data yang dihimpun melalui tes akhir seperti ujian sekolah ataupun ujian yang diselenggarakan oleh pemerintah sebagai suatu ujian yang berisiko tinggi (*high stake test*), tetapi evaluasi yang didasarkan pada data dan informasi harian menjadi tidak kalah penting bagi keberhasilan pelaksanaan program.
7. **Evaluator harus mengidentifikasi setiap permasalahan yang berkaitan dengan evaluasi.** Dalam hal ini, evaluasi harus dirancang dengan baik, mulai dengan apa tujuan evaluasi, jenis data dan informasi apa saja yang diperlukan, seberapa jauh data dan informasi tersebut dapat diperoleh secara tepat dan akurat, faktor-faktor apa saja yang berpengaruh menjadikan penyebab ketidakepatan dan ketidakakuratan.
8. **Program pendidikan bukan perlakuan tunggal, banyak faktor yang berpengaruh.** Oleh karena itu, evaluator harus melihat bagian dalam proses, antar perlakuan, dan dalam

Populasi unftuk melihat besarnya pengaruh tindakan yang diberikan. Dalam hal ini hendaknya harus disadari bahwa keberhasilan suatu program tidak dapat terlepas dari banyak faktor penentu aktualisasinya yang semuanya itu diharapkan dapat dipenuhi ketika tindakan evaluasi hasil program akan ditetapkan.

9. **Dalam melakukan evaluasi, aspek psikomotor dan afektif tidak boleh dikesampingkan.** Dalam hal ini, bahwa evaluasi keberhasilan dengan kedudukan peserta sebagai peserta program hendaklah dikembalikan kepada target pembelajaran secara menyeluruh. Dettmer (2006) bahkan menekankan bahwa semua domain hendaknya menjadi target apembelajaran dalam bentuk satu kesatuan (*unity*). Pembelajaran tidak hanya menekankan pada salah satu domain/ranah, melainkan kesatuan ranah antara ranah kognitif, afektif, sensorimotor (pengganti istilah psikomotor), dan sosial (sebagai domain yang fokusnya pada interaksi dengan orang lain sehingga dipisahkan dari domain afektif yang lebih bersifat personal).
10. **Evaluasi formatif dan sumatif harus menjadi satu kesatuan yang utuh, jadi harus melihat seluruh hasil dari pelaksanaan program, jangan hanya melihat hasil evaluasi sumatif.** Dalam hal ini evaluasi harus dilakukan sejak perencanaan, implementasi program yang di dalamnya termasuk pelaksanaan keseharian beserta hasil-hasil antara selama terjadinya proses pembelajaran, serta hasil akhir program.
11. **Analisis keseluruhan lebih dapat dipertanggungjawabkan daripada hanya berdasar pada satu pengukuran.** Data dan informasi serta hasil penilaian terhadap setiap komponen perlu keberhasilan program harus menjadi pertimbangan ketika akan dilakukan pengambilan keputusan.
12. **Evaluasi sebagai suatu simpulan berdasarkan hasil penilaian harus inferensial, bukan hanya berdasar data tetapi juga berdasarakan asumsi.** Sebagai probadi, setiap peserta didik yang berkedudukan sebagai peserta program adalah manusia yang sedang mengalami perkembangan. Kedua,

bahwa implementasi program pembelajaran tidak dapat terlepas dari adanya keterbatasan dari pihak-pihak yang terlibat di dalamnya juga keterbatasan faktor pendukung dalam bentuk sarana-prasarana. Oleh karena itu, akan menjadi bijaksana ketika evaluator melihat jauh kedepan atas ketetapan yang diputuskan kepada tingkat prestasi peserta didik sebagai peserta program.

Evaluasi program pembelajaran didasarkan pada hasil penilaian terhadap setiap peserta didik orang per orang. Adapun prinsip penilaian menurut Gronlund dan Linn adalah

1. harus ada spesifikasi yang jelas apa yang mau dinilai: penempatan, formatif, ataukah sumatif;
2. harus komprehensif: afektif, psikomotor, dan kognitif;
3. butuh berbagai ragam teknik/metode penilaian, baik metode tes maupun nontes;
4. harus dapat memilih instrumen penilaian yang sesuai;
5. harus jelas apa maksud dan tujuan diadakan penilaian, jadi akan jelas pula apa tindak lanjutnya.

Prinsip penilaian hasil belajar peserta didik pada jenjang pendidikan dasar dan menengah menurut Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 adalah sebagai berikut.

1. sah, berarti penilaian didasarkan pada data yang mencerminkan kemampuan yang diukur.
2. objektif, berarti penilaian didasarkan pada prosedur dan kriteria yang jelas, tidak dipengaruhi subjektivitas penilai.
3. adil, berarti penilaian tidak menguntungkan atau merugikan peserta didik karena berkebutuhan khusus serta perbedaan latar belakang agama, suku, budaya, adat istiadat, status sosial ekonomi, dan gender.

4. terpadu, berarti penilaian oleh pendidik merupakan salah satu komponen yang tak terpisahkan dari kegiatan pembelajaran.
5. terbuka, berarti prosedur penilaian, kriteria penilaian, dan dasar pengambilan keputusan dapat diketahui oleh pihak yang berkepentingan.
6. menyeluruh dan berkesinambungan, berarti penilaian oleh pendidik mencakup semua aspek kompetensi dengan menggunakan berbagai teknik penilaian yang sesuai, untuk memantau perkembangan kemampuan peserta didik.
7. sistematis, berarti penilaian dilakukan secara berencana dan bertahap dengan mengikuti langkah-langkah baku.
8. beracuan kriteria, berarti penilaian didasarkan pada ukuran pencapaian kompetensi yang ditetapkan.
9. akuntabel, berarti penilaian dapat dipertanggungjawabkan, baik dari segi teknik, prosedur, maupun hasilnya.

Prinsip penilaian menurut BSNP (2007) untuk pendidikan dasar dan menengah mengacu kepada standar penilaian pendidikan jenjang pendidikan dasar dan menengah. Prinsip tersebut mencakup:

1. *Sahih*, yakni penilaian didasarkan pada data yang mencerminkan kemampuan yang diukur. Oleh karena itu, instrumen yang digunakan perlu disusun melalui prosedur sebagaimana dijelaskan dalam panduan agar memiliki bukti kesahihan dan keandalan.
2. *Objektif*, yakni penilaian didasarkan pada prosedur dan kriteria yang jelas tanpa dipengaruhi oleh subjektivitas penilai. Oleh karena itu, dalam rangka meningkatkan objektivitas penilaian, pendidik menggunakan rubrik atau pedoman dalam memberikan skor terhadap jawaban peserta didik atas item uraian dan tes praktik atau kinerja.
3. *Adil*, yakni penilaian tidak menguntungkan atau merugikan peserta didik karena berkebutuhan khusus serta perbedaan latar belakang agama, suku, budaya, adat istiadat, status sosial ekonomi, dan gender. Faktor-faktor tersebut tidak relevan di

- dalam penilaian, sehingga perlu dihindari agar tidak berpengaruh terhadap hasil penilaian.
4. *Terpadu*, yakni penilaian oleh pendidik merupakan salah satu komponen kegiatan pembelajaran. Dalam hal ini hasil penilaian benar-benar dijadikan dasar untuk memperbaiki proses pembelajaran yang diselenggarakan oleh peserta didik. Jika hasil penilaian menunjukkan banyak peserta didik yang gagal, sementara instrumen yang digunakan sudah memenuhi persyaratan secara kualitatif, berarti proses pembelajaran kurang baik. Dalam hal demikian, pendidik harus memperbaiki rencana dan/atau pelaksanaan pembelajarannya.
 5. *Terbuka*, yakni prosedur penilaian, kriteria penilaian, dan dasar pengambilan keputusan dapat diketahui oleh pihak yang berkepentingan. Oleh karena itu, pendidik menginformasikan prosedur dan kriteria penilaian kepada peserta didik. Selain itu, pihak yang berkepentingan dapat mengakses prosedur dan kriteria penilaian serta dasar penilaian yang digunakan.
 6. *Menyeluruh dan berkesinambungan*, yakni penilaian mencakup semua aspek kompetensi dengan menggunakan berbagai teknik penilaian yang sesuai, untuk memantau perkembangan kemampuan peserta didik. Oleh karena itu, penilaian bukan semata-mata untuk menilai prestasi peserta didik melainkan harus mencakup semua aspek hasil belajar untuk tujuan pembimbingan dan pembinaan.
 7. *Sistematis*, yakni penilaian dilakukan secara berencana dan bertahap dengan mengikuti langkah-langkah baku. Oleh karena itu, penilaian dirancang dan dilakukan dengan mengikuti prosedur dan prinsip-prinsip yang ditetapkan. Dalam penilaian kelas, misalnya, guru mata pelajaran agama menyiapkan rencana penilaian bersamaan dengan menyusun silabus dan RPP.
 8. *Beracuan kriteria*, yakni penilaian didasarkan pada ukuran pencapaian kompetensi yang ditetapkan. Oleh karena itu,

instrumen penilaian disusun dengan merujuk pada kompetensi (SKL, SK, dan KD). Selain itu, pengambilan keputusan didasarkan pada kriteria pencapaian yang telah ditetapkan.

9. *Akuntabel*, yakni penilaian dapat dipertanggungjawabkan, baik dari segi teknik, prosedur, maupun hasilnya. Oleh karena itu, penilaian dilakukan dengan mengikuti prinsip-prinsip keilmuan dalam penilaian dan keputusan yang diambil memiliki dasar yang objektif.

E. Fungsi dan Tindak Lanjut Evaluasi

Evaluasi hasil belajar merupakan serangkaian kegiatan yang dilaksanakan tahap demi tahap berdasarkan keseluruhan hasil penilaian yang dilakukan. Adapun manfaat hasil evaluasi bagi subjek belajar adalah untuk bimbingan belajar, bimbingan pribadi, dan kebutuhan subjek belajar yang berkaitan dengan studinya. Jadi, meliputi aspek bimbingan dan aspek pembelajaran. Dengan demikian evaluasi proses dan hasil belajar akan berfungsi untuk memberi

1. arah dan petunjuk dalam pelaksanaan pembelajaran, baik guru maupun subjek belajar;
2. gambaran tentang diri subjek belajar mengenai perkembangan baik kemampuan maupun personalitasnya, sehingga mereka mampu mengenali diri/mawas diri serta seberapa jauh produktivitasnya, sehingga mampu menentukan langkah/keputusan lebih lanjut guna peningkatan prestasi.
3. dorongan/motivasi subjek belajar agar mampu berusaha untuk meraih prestasi yang lebih baik.
4. masukan untuk perbaikan dan pelaksanaan program guna memperbaiki proses pembelajaran yang akan diselenggarakan saat berikutnya.

Dengan mengetahui hasil evaluasi yang diperoleh, maka dapat diambil keputusan sebagai tindak lanjut. Keputusan tersebut sangat tergantung kepada jenis evaluasi yang telah dilaksanakan, yakni:

1. Keputusan atas dasar evaluasi sumatif

Jika sistemnya menggunakan sistem kenaikan, kegagalan subjek belajar dalam evaluasi sumatif dapat dinyatakan tinggal kelas. Di perguruan tinggi misalnya, setelah empat semester pertama dilakukan peninjauan terhadap hasil belajarnya, jika tidak mencapai jatah yang ditetapkan, subjek belajar disuruh mengundurkan diri.

2. Keputusan atas dasar evaluasi formatif

Karena evaluasi formatif berfungsi untuk memberikan umpan balik, baik kepada subjek belajar ataupun guru, maka kegagalan kelas berarti merupakan masukan baik bagi guru maupun subjek belajar.

- a. Bagi guru, kegagalan kelas berarti masukan untuk merevisi program pembelajaran yang telah disusun, baik dari segi metode, media, urutan kegiatan, guru, kegiatan pembelajaran yang dilakukan subjek, belajar, serta alokasi waktunya, sehingga dalam proses pembelajaran selanjutnya akan ada perbaikan hasil. Dengan demikian, akan dapat disusun suatu program pembelajaran yang lebih efektif yang lebih sesuai dengan kondisi, minat, bakat dan kecerdasan subjek belajar sebagai peserta program.
- b. Bagi subjek belajar maka kegagalan dalam evaluasi formatif hendaknya digunakan untuk mawas diri dan mengambil langkah-langkah guna mengatasi kegagalan. Bagi yang belum memuaskan prestasinya dapat mengambil langkah-langkah baru guna meningkatkan prestasi.

3. Keputusan yang diambil berdasarkan evaluasi diagnostik

a. Jika evaluasi diagnostik dilakukan sebelum proses pembelajaran, maka tujuannya guru dapat mengambil sikap perlu tidaknya subjek belajar diberi pelajaran ekstra agar mampu menguasai prasyarat yang belum dikuasainya. Jika tanpa penguasaan prasyarat subjek belajar akan mengalami kesu-

litan dalam mengikuti proses pembelajaran, maka mau tidak mau guru harus meningkatkan penguasaan subjek belajar terhadap prasyarat tersebut.

b. Jika dilakukan selama proses, maka tugas guru untuk melakukan remedi terhadap subjek belajar yang mengalami kesulitan belajar, sehingga tidak terjadi kegagalan pada evaluasi sumatifnya. Dalam hal ini dapat dilakukan dengan berbagai cara, seperti remedi akademik dilakukan dengan menggunakan sistem modul ataupun tugas terstruktur kalau tidak ada waktu khusus untuk tatap muka di kelas. Kendala utama adalah kalau beban mengajar guru terlalu banyak.

4. Keputusan yang harus diambil berdasar evaluasi penempatan

- a. Subjek belajar yang tidak menguasai kemampuan prasyarat harus diremedi (sebagai penilaian diagnosis sebelum proses pembelajaran).
- b. Kemampuan yang sudah dikuasai subjek belajar tanpa harus melalui proses pembelajaran jangan ditargetkan sebagai kemampuan yang harus dikuasai melalui proses pembelajaran yang akan diselenggarakan.
- c. Subjek belajar yang tidak tertarik atau berbeda karakteristik personalnya perlu mendapat perhatian khusus.

F. Kedudukan Pengukuran dalam Evaluasi

Evaluasi merupakan suatu proses penetapan untuk menyatakan baik atau buruk, berhasil atau gagal, sukses atau tidaknya sesuatu. Dengan kata lain evaluasi adalah proses untuk menentukan kualitas atau mutu sesuatu. Agar dapat dilakukan proses tersebut, maka harus ada data yang dijadikan dasar untuk penentuannya. Data harus benar-benar dapat dipercaya/akurat agar ketetapan yang diambil tidak salah. Jika salah datanya salah pula hasil penilaiannya dan akibatnya salah pula keputusannya. Data yang diolah untuk penilaian dan yang selanjutnya untuk dasar evaluasi guna menetapkan apakah berhasil atau gagal atau apakah baik atau buruk itu harus data kuantitatif yang diperoleh melalui proses pengukuran. Jadi pengukuran merupakan

proses untuk memperoleh kuantitas tentang sesuatu. Oleh karena itu diperlukan adanya alat atau instrumen penilaian yang dipakai dalam kegiatan pengukuran atau dapat pula disebut sebagai alat ukur. Ada alat ukur yang berupa tes dan ada pula alat ukur yang tergolong nontes. Agar dapat diperoleh alat penilaian atau alat ukur yang baik perlu dikembangkan suatu prosedur yang benar, dan disebut dengan prosedur pengembangan alat penilaian yang meliputi:

1. perencanaan penilaian yang memuat maksud dan tujuan penilaian;
2. penyusunan kisi-kisi;
3. penyusunan instrumen/alat ukur;
4. penelaahan (*review*) untuk menilai kualitas alat ukur/instrumen secara kualitatif, yakni sebelum digunakan;
5. uji coba alat ukur, untuk menyelidiki kesahihan dan keandalan secara empiris;
6. pelaksanaan pengukuran;
7. penilaian yang merupakan interpretasi hasil pengukuran;
8. pemanfaatan hasil penilaian.

Dalam hal ini, jika pengukuran dilakukan dengan menggunakan instrumen yang berupa tes, maka penyusunan instrumen harus mengacu pada cara-cara penyusunan suatu tes, demikian pula apabila instrumen yang digunakan berupa instrumen nontes maka cara penyusunan instrumen harus mengacu pada cara-cara penyusunan instrument nontes.

G. Macam Skala Pengukuran

1. Skala nominal

Skala nominal yaitu ukuran yang hanya berdasarkan pada prinsip tak tumpang tindih (*mutually exclusive*) dan tuntas (*exhaustive*), serta tidak ada asumsi tentang jarak maupun urutan/jenjang antar kategori yang ada di dalamnya. Angka-angka yang ada sekedar label untuk memisahkan kategori-kategori yang ada. Misalnya skala 1 : SD Negeri, skala 2 : SD Swasta atau sebaliknya; skala 1 : Islam, skala 2 : Kristen, skala 3: Katolik, skala 4: Hindu,

dan skala 5: Budha, dapat pula sebaliknya; skala 1 : laki-laki dan skala 2 : perempuan atau sebaliknya, dan sebagainya.

2. Skala ordianal

Skala ordianal yaitu ukuran yang sudah mengurutkan objek yang diukur dari jenjang atau peringkat terendah sampai dengan yang tertinggi pada suatu atribut tertentu, tetapi tanpa petunjuk yang jelas tentang berapa jumlah/nilai absolut yang dimiliki oleh objek pada tiap atribut, dan jarak atau interval antara satu objek dengan yang lainnya juga tidak diketahui. Misalnya, untuk variabel tingkat pendidikan digunakan skala 1: SD, 2: SLTP, 3: SLTA, 4: PT Diploma sampai S1, 5: PT S2, dan 6: PT S3, tak berarti bahwa responden yang memiliki ijazah SD sekaligus MI sama dengan memiliki ijazah SLTP. Seseorang yang memiliki ijazah SMP sekaligus MTs setara dengan memiliki ijazah S-1. Tingkat pendidikan SLTA juga tidak sama tiga kalinya tingkat pendidikan SD.

3. Skala interval

Skala interval yaitu ukuran yang menunjukkan kedudukan tiap subjek yang diukur, tetapi titik nolnya arbitrer, misalnya skor hasil ulangan/ujian dalam penguasaan hasil belajar seperti penguasaan konsep untuk aspek kognitif.

4. Skala rasio

Skala rasio yaitu ukuran yang menunjukkan kedudukan tiap subjek yang diukur dan titik nolnya absolut, misal berat benda, panjang benda, volume benda dan sebagainya.

H. Kesalahan Pengukuran

Adanya ketidaktepatan data atau data yang tidak dapat dipercaya kebenarannya dapat ditimbulkan oleh beberapa faktor berikut.

1. Kesalahan Pada Alat Ukur

Seperti halnya dalam Ilmu Sosial pada umumnya, variabel-variabel dalam bidang pendidikan banyak yang bersifat abstrak fenomenanya. Oleh karena itu terhadap variabel yang demikian, sukar dilakukan pengukuran secara langsung. Jadi, tidak seperti mengukur tinggi peserta didik, panjang jangkannya atau berat tubuhnya, karena fenomenanya benar-benar kongkrit. Guru tidak dapat mengukur kecakapan subjek belajar dalam mata pelajaran Biologi secara langsung misalnya. Kita juga tidak dapat secara langsung mengukur intelegensi subjek belajar. Dengan demikian jika fenomenanya sendiri abstrak, maka hanya dapat dilakukan pengukuran secara tak langsung. Persoalannya adalah bagaimana kita mampu menyusun alat ukur yang sah (*valid*) dalam arti bahwa hasil pengukuran mencerminkan apa yang diukur. Kedua, bagaimana kita dapat menghasilkan alat ukur yang dapat memberikan data yang ajeg atau konsisten (*reliable*), yakni yang menghasilkan skor yang sama jika dilakukan pengukuran berulang-ulang. Variabel yang memiliki fenomena abstrak diukur secara tidak langsung yakni dengan menetapkan indkan-indkan/ indikator-indikator dari variabel yang bersangkutan. Agar mampu mencari indkan yang tepat, maka variabel yang akan diukur harus didefinisikan dengan tegas. Misalnya, apa yang dimaksud subjek belajar telah menguasai Perindungan Khusus dalam bidang Perawatan Gigi? Apakah subjek belajar harus hafal semua teks tentang Perindungan Khusus atau mampu memecahkan berbagai persoalan perindungan khusus terhadap gigi, atau terampil menguasai kemampuan dalam melakukan perindungan khusus terhadap gigi, atau ketiganya? Demikian pula apa indkan-indikannya siswa berminat terhadap Biologi? Adanya definisi akan menegaskan apa sebenarnya indkan-indkan dari variabel tersebut. Dengan adanya indkan yang tepat kemudian disusun pertanyaan-pertanyaan yang sesuai. Persoalan yang muncul adalah apakah pertanyaan yang dibuat sudah benar-benar mengukur indkan yang dimaksud. Kadang-kadang variabel yang

akan kita ukur bersifat multidimensi. Sementara, dalam melakukan pengukuran seharusnya dilakukan terhadap satu dimensi atau harus bersifat unidimensional. Selain itu, suatu variabel harus bersifat unidimensional, dalam arti jika dilakukan pengukuran maka setiap hasil pengukuran dapat ditentukan letaknya pada garis yang menggambarkan keseluruhan harga dari variabel yang bersangkutan. Dengan kata lain hasil pengukuran harus dapat diplotkan pada garis abstrak (*abstract continuum*) yang merupakan garis variabel tersebut. Namun demikian, pada kenyataannya sering suatu variabel dapat mengandung banyak dimensi, sehingga pada saat melakukan pengukuran terhadap variabel tersebut harus mencakup pengukuran terhadap seluruh dimensinya. Untuk itu perlu dicari terlebih dahulu apa saja dimensi dari variabel yang akan kita ukur. Setelah diperoleh dimensinya baru kita jabarkan ke dalam indikan-indikan/indikator-indikatornya.

2. Kesalahan dalam Proses Pengukuran

Kesalahan ini ditimbulkan ketidaktehitan fihak yang melakukan pengukuran, seperti kesalahan dalam mengoreksi jawaban, kesalahan dalam memberikan skor, kesalahan dalam melakukan konversi skor menjadi nilai, dan kesalahan dalam menulis nilai sebagai hasil akhir peserta didik. Kesalahan tersebut dalam proses pengukuran juga dapat diakibatkan karena kesalahan dalam menerapkan alat ukur terhadap pihak yang akan diukur; maksudnya, apakah pertanyaan-pertanyaan yang ada dalam alat ukur tersebut benar-benar cocok untuk pihak yang akan diukur. Suatu alat ukur pengukur prestasi hasil belajar siswa harus sesuai dengan apa yang dipelajari dan kompetensi apa yang ditargetkan. Jika soal-soal yang diujikan tidak sesuai dengan rumusan indikator hasil belajar maka secara otomatis tidak mengukur penguasaan kompetensi dasar yang harus dikuasai sebagai target pembelajaran.

3. Kondisi Pihak yang Diukur

Karena yang diukur adalah manusia, maka situasi pada saat dilakukan pengukuran akan sangat menentukan skor hasil pengukuran. Hasil pengukuran dalam kondisi stress akan lain dengan dalam kondisi yang bugar. Anak yang penakut, dalam kondisi cemas, akan memberikan hasil pengukuran yang lain jika dilakukan pengukuran terhadapnya pada situasi yang berbeda. Oleh karena itu, peserta didik akan lebih berhasil dalam ujian nasional jika sering berlatih. Dengan berlatih maka siswa akan terbiasa menghadapi pola maupun prosedur pelaksanaannya. Namun, guru yang hanya mengajar atas apa yang sering diujikan dalam ujian nasional menjadi lupa akan kurikulum dan mengarah kepada pola *teaching to the test*.

4. Kondisi/situasi selama dilakukan pengukuran

Karena yang diukur seseorang, maka situasi yang gaduh akan menghasilkan skor yang berbeda dibandingkan dengan situasi yang tertib dan tenang pada waktu dilakukan pengukuran, khususnya untuk anak-anak yang mudah terganggu oleh situasi tersebut.

5. Pengaruh pengulangan pengukuran

Untuk keperluan tertentu maka pengukuran dengan menggunakan alat ukur yang sama dilakukan secara berulang. Jika jarak antara pengukuran pertama dengan pengukuran ulang terlalu dekat, maka pihak ada efek *testing*. Efek *testing* merupakan efek yang ditimbulkan akibat subjek yang diukur masih memiliki ingatan yang baik terhadap alat ukur tersebut, sehingga ia masih ingat apa jawaban-jawaban yang diberikan pada pengukuran pertama. Sebaliknya, jika pengulangan pengukuran terlampau lama selang waktunya, maka kemampuan seseorang akan menurun kalau ia tidak memiliki kesempatan untuk berlatih kembali, sehingga akan terjadi hasil yang berbeda antara pengukuran pertama dengan pengukuran ulangnya.

I. Kesalahan Penilaian

Setelah diperoleh data, kemudian dilakukan proses pengolahan data untuk mengambil keputusan akhir dalam menilai subjek belajar. Dalam hal ini, meskipun datanya benar dapat terjadi kesalahan pengambilan keputusan akibat oleh beberapa faktor. Faktor-faktor tersebut adalah sebagai berikut.

1. Kekeliruan dalam mengolah data

Hal ini akibat kehidakteritian dalam mengubah skor menjadi nilai atau karena kesalahan cara pengolahan datanya.

2. Pengaruh penilaian sebelumnya

Penilai dapat terpengaruh oleh keadaan-keadaan yang mendahului. Misalnya, seorang anak biasanya mendapat nilai jelek, kemudian suatu saat mendapat nilai baik. Keadaan ini dapat mempengaruhi penilai untuk menurunkan nilai anak tersebut.

3. Menilai mahal atau murah

Kecenderungan seorang penilai untuk memberi nilai lebih rendah atau lebih tinggi, sehingga keluar istilah ada guru yang murah ada guru yang mahal tanpa ada dasar yang dapat percaya.

4. Pengaruh kesan luar

Penampilan pihak yang dinilai dapat berpengaruh terhadap si penilai. Seorang subjek belajar yang "suka mendebat" oleh guru tertentu dinilai jelek, walaupun sebenarnya skor yang diperoleh selalu tinggi, karena guru tersebut tidak suka didebat. Sebaliknya, anak yang berpenampilan manis, penurut, akan diberi nilai yang lebih tinggi walaupun skornya rendah.

5. Pengaruh hallo effect

Kesan jelek terhadap seseorang yang dinilai akan dapat mempengaruhi pihak penilai. Seorang guru yang merangkap dua mata pelajaran, dapat terpengaruh kesan pada nilai dari salah satu mata pelajaran. Jika dalam mata pelajaran A yang satu anak tersebut baik,

maka ia dapat terpengaruh sehingga akan memberikan nilai yang baik pula dalam mata pelajaran B walaupun skor yang diperoleh dalam mata pelajaran B selalu rendah.

LATIHAN

1. Jelaskan hubungan antara pengertian pengukuran, penilaian, dan evaluasi?
2. Apa perbedaan utama antara penilaian dan evaluasi dalam kaitannya dengan kegiatan pembelajaran?
3. Apa makna diselenggarakannya evaluasi terhadap penyelenggaraan suatu program pembelajaran yang dilaksanakan oleh sekolah?
4. Mengapa evaluasi perlu dilaksanakan terhadap semua komponen penentu aktualisasi belajar yang terselenggara?
5. Apakah evaluasi perlu dilakukan terhadap kurikulum sekolah?
6. Samakah antara prinsip-prinsip evaluasi dan prinsip-prinsip penilaian?
7. Mengapa perpaduan pendekatan positifistik kuantitatif dan humanistic kualitatif sangat ditekankan dalam melakukan evaluasi pembelajaran?
8. Mengapa kesalahan pengukuran hasil belajar dapat terjadi?
9. Apa saja sumber kesalahan pengukuran?
10. Apakah sama antara sumber kesalahan pengukuran dan sumber kesalahan penilaian?

II Objek Penilaian

Dalam penilaian pencapaian hasil belajar, kedudukan subjek belajar merupakan subjek yang dinilai. Adapun sebagai objek penilaian adalah apa-apa yang melekat pada diri subjek belajar yang dijadikan sasaran penilaian. Adapun yang melekat pada diri subjek belajar selain variabel-variabel yang berhubungan dengan fisik subjek belajar juga yang berhubungan dengan nonfisik. Dalam hal ini berupa kemampuan (*ability*) dan personalitas (*aptitude*). Menurut Gorman kemampuan seseorang meliputi kemampuan intelektual dan kemampuan psikomotor. Kemampuan intelektual mencakup kemampuan penguasaan produk dan proses. Adapun yang termasuk produk adalah fakta, konsep dan struktur ilmu pengetahuan, sedangkan yang termasuk proses adalah kreativitas, pemecahan masalah, dan komprehensif. Kemampuan psikomotor berupa keterampilan (*skill*), sementara personalitas mencakup temperamen yang berupa tingkah laku (*behavior*), karakter, moral dan sikap (*attitude*).

A. Ranah Kognitif

Menurut taksonomi Bloom, kemampuan intelektual atau kognitif meliputi jenjang sebagai berikut:

1. Ingatan (*knowledge*):
 - a. Ingatan tentang hal yang spesifik, baik ingatan tentang peristilahan (terminologi) maupun kejadian yang spesifik, misal menyebutkan bagian-bagian, menyebutkan istilah, nama, sifat, contoh, dan sebagainya; mengingat definisi, bagian-bagian, kejadian, tempat, dan sebagainya.

sebesar 0,70. Seperti yang tersaji pada Tabel 3, suatu tes yang memiliki nilai $r = 0,70$ akan memiliki nilai koefisien kappa mendekati daerah 0,35 sampai 0,50 tergantung pada lokasi skor penggalan terhadap skor rata-rata atau ukuran yang nisbi antara kelompok yang menguasai dan yang tidak menguasai.

Sebagai contoh, jika koefisien keandalan tes $r = 0,70$ dan suatu kelas diharapkan hanya memiliki 15 % siswa yang tidak menguasai ($|z| = 1,00$) maka besarnya koefisien kappa yang diharapkan sekitar 0,45. Tentu saja, nilai kappa yang lebih tinggi dapat diharapkan bila suatu tes diharapkan memiliki koefisien keandalan yang lebih tinggi pula.

Dari diskusi di atas aturan berikut ini dapat dijadikan pegangan: Suatu tes yang diterapkan dalam suatu kelas pada periode yang penuh (katakanlah 1 semester) dapat menggunakan nilai koefisien kappa sebesar 0,35 sampai 0,50. Nilai yang lebih tinggi dapat diterapkan disesuaikan dengan proporsi kelompok yang menguasai dan yang gagal. Nilai kappa mendekati 0,45 untuk tes yang terdiri dari 10 item jika siswa yang gagal diharapkan hanya 15%.

Di dalam menerapkan kriteria seperti yang diusulkan di atas, harus melihat kembali bagaimana perhitungan kedua koefisien itu diperoleh. Sebagai contoh, tes yang terdiri dari 10 item dengan koefisien keandalan $r = 0,50$ dan penggalan $|z| = 1,60$, akan menghasilkan koefisien persetujuan $p_o = 0,92$ berarti melebihi batas minimal $p_o = 0,85$ karena lokasi skor penggalan menghasilkan proporsi antara siswa yang berhasil dan yang gagal cukup besar. Namun, besarnya koefisien kappa $k = .21$ untuk tes yang sependek itu dibawah patokan minimal yang yang diusulkan (yakni $= 0,35$). Dalam kasus tertentu guru dimungkinkan membuat perjanjian yang berkebalikan. Artinya patokan minimal untuk koefisien kappa ditetapkan terlebih dahulu.

DAFTAR PUSTAKA

- Abrams, L.M. (2007). "Implications of high-stakes testing for the use formative classroom assessment". In: J.H. McMillan. *Formative classroom assessment: Theory into practice*. New York: Teachers College Columbia University.
- Allen, M.J. & Yen, W.M. (1979). *Introduction to measurement theory*. California: Brooks/Cole Publishing Company.
- Ary, D. & Jacobe, L.Ch., & Razavieh, A. (2010). *Introduction to research in education*, 8th ed. New York: Holt, Rinehart and Winston.
- Badan Standar Nasional Pendidikan (2007). *Panduan penilaian kelompok mata pelajaran ilmu pengetahuan dan teknologi*. Jakarta: Badan Standar Nasional Pembelajaran.
- Bambang Subali & Pejiyati Suyata, (2011). Standardisasi penilaian hasil belajar pola konvergen dan divergen berbasis satuan pendidikan pada jenjang SMA. Laporan Penelitian. Tidak dipublikasikan. Yogyakarta: Lembaga Penelitian dan PPM Universitas Negeri Yogyakarta.
- _____. (2010) Bias item tes keterampilan proses sains pola divergen dan modifikasinya sebagai tes kreativitas. *Jurnal Penelitian dan Evaluasi Pendidikan*. Tahun 14. No. 2. Desember 2010. hlm: 309-334.
- _____. (2011). Pengukuran Kreativitas Keterampilan Proses Sains dalam Konteks *Assessment For Learning*. Cakrawala Pendidikan, Th. XXX, No. 1, Februari 2011, hlm. 130-144.
- _____. (2009). Pengukuran Keterampilan Proses Sains Pola Divergen Dalam Mata Pelajaran Biologi SMA di Provinsi DIY dan Jawa Tengah. Disertasi. Tidak dipublikasikan. Yogyakarta: Program Pascasarjana Universitas Negeri Yogyakarta.

- Banicky, L.A. & Foss, H.K. (2000). *Assessing student learning* [Versi elektronik]. Delaware: Delaware Educational Research and Development Center, University of Delaware. Diunduh pada tanggal 8 Agustus 2008.
- Black, P., Harrison, Ch., Lee, Cl., Marshall, B., & William, D. (2005). *Assessment for Learning: Putting it into practice* [Versi elektronik]. New York: Open University Press. Diunduh pada tanggal 11 Januari 2012.
- Brennan, R.L. (2006). *Educational measurement*, 4th ed. Westport: American Council on Education and Praeger Publishers.
- Brooks, V. (2002). *Assessment in secondary schools: The new teacher's guide to monitoring, assessment, recording, reporting, and accountability* [Versi elektronik]. Buckingham: Open University Press. Diunduh pada tanggal 12 Desember 2011.
- Bruun & McKane. (1989). *Study guide biology: Exploring life*. New York: John Wiley & Sons.
- 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: Teacher's guide*. Oxford: Heinemann Instructional Books.
- 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: Teacher's guide*. Oxford: Heinemann Educational Books.
- Burke, A.A. (2007). The Benefits of Equalizing Standards and Creativity: Discovering a Balance in Instruction [Versi elektronik]. *Gifted Child Today*, 30, 1, 58-63.
- Carin, A.A. dan Sund, R.B. (1989). *Teaching science through discovery*. Columbus: Merrill Publishing Company.

- Cochran, S.M. & Lytle, S.L. (2006). Troubling images of teaching in no child left behind [Versi elektronik]. *Harvard Educational Review*, 76, 4, 668-699.
- Collete, A.T. & Chiappetta, EL. (1994). *Science instruction in the middle and secondary schools*. 3rd ed. New York: Macmillan Publishing Company.
- Dettmer, P. (2006). New Blooms in Established Fields: Four Domains of Learning and Doing [Versi elektronik]. *Roeper Review*, 28, 2, 70-78.
- Djemari Mardapi. (2007). *Teknik penyusunan instrumen tes dan non tes*. Yogyakarta: Mitra Cendekia Press.
- Drake, S.M. (2007). *Creating standards-based integrated curriculum: Aligning curriculum, content, assessment, and Instructions*. Thousand Oaks: Corwin Press.
- Ebel, R.L. & Frisbie, D.A. (1986). *Essentials of educational measurement*, 4th ed. New Jersey: Prentice-Hall, Inc.
- Frisbie, D.A. (2005). Measurement 101: Some fundamentals revisited. *Educational Measurement: Issues and Practice* [Versi elektronik]. Fall 2005. Vol. 24. No.3. pp.21-28. Diunduh pada tanggal 19 Agustus 2007.
- Garrison, M.J. (2009). *A measure of failure: The political origins of standardized testing*. Albany: State University of New York Press. [Versi elektronik]. Diunduh tanggal 6 Desember 2011.
- Glencoe. (tt.). *Performance assessment in the science classroom*. Professional Glencoe Science series. New York: McGraw-Hill.
- Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context*. Washington, DC: National Center for Education Statistics, Institute of Education Sciences

- Gronlund, N.E. (1973). *Preparing criterion-referenced tests for classroom instruction*. New York: The Macmillan Company. Hart, D. (1994). *Authentic assessment: A handbook for educators*. California: Addison-Wiley Publishing Company.
- Gronlund, N.E. (1998). *Assessment of student achievement*, 9th ed. Boston: Allyn and Bacon.
- Guskey, Th. R. (2007). "Formative classroom assessment and Benjamin S. Bloom: Theory, research, and practice". In: J.H. McMillan. *Formative classroom assessment: Theory into practice*. New York: Teachers College Columbia University.
- Harlen, W. (2007). "Formative classroom in science and mathematics". In: J.H. McMillan. *Formative classroom assessment: Theory into practice*. New York: Teachers College Columbia University.
- Hedges, W.D. (1969). *Testing and evaluation for the science*. Belmont, California: Wadsworth Publishing Company, Inc.
- Hegge, Th.G., Kirk, S.A., & Kirk, W.D. (1965). *Remedial reading drills*. Novato, CA.: Academic Therapy Publications. Versi Elektronik. www.academitherapy.com. Diakses tanggal 2 Agustus 2010.
- Heubert, J.P. & Hauser, R.M. (Editors). (1999). *High Stakes: Testing for Tracking, Promotion, and Graduation* [Versi elektronik]. Washington, D.C.: National Academies Press. <http://www.nap.edu/catalog/6336.html>. Diunduh tanggal 29 Agustus 2009.
- Hilbard, K.M. (t.t.). *Performance assessment in the science classroom*. New York: McGraw-Hill Companies.
- IEA. (2004). *TIMSS 2003 technical report: Findings from IEA's trends in international mathematics and science study at fourth and eighth grades*. Boston: TIMSS & PIRLS International Study Center.

- Jehlen, A. (2007). Testing how the sausage is made [Versi elektronik]. *NEA Today*, 25, 7, 29-34.
- Jumadi, Bambang Subali, & Das Salirawati. (2011). Pengembangan bahan ajar mata pelajaran ipa berbasis ideational learning untuk mengembangkan kreativitas bagi SMP Ritisian sekolah Berstandar internasional di provinsi DIY. Laporan Penelitian. Tidak dipublikasikan. Yogyakarta: Lembaga Penelitian dan PPM Universitas Negeri Yogyakarta.
- Karlan, J.W. (2000). The biosphere challenge: Developing ecological literacy [Versi elektronik]. *Green Teacher*, 62, 13-18.
- Keeley, P. (2009). *Science formative assessment: 75 practical strategies for linking assessment, instruction, and learning*. [Versi elektronik]. California: Corwin Press. Diunduh pada tanggal 20 Desember 2011.
- Kementerian Pendidikan Nasional. (2011). *Panduan pemanfaatan hasil ujian nasional untuk perbaikan mutu pendidikan ujian nasional 2010/2011*. Jakarta: Badan Penelitian dan Pengembangan Pendidikan
- Kind, P. M. & Kind, V. (2007). Creativity in science education: Perspectives and challenges for developing school science [Versi elektronik]. *Studies in Science Education*, 43, 1-37.
- Ludwig, J.A. dan Reynold. J.F. (1988). *Statistical ecology*. New York: John Wiley and Sons
- Marwaha, Pr. (2009). *Role of remedial teaching* [Versi elektronik]. [Http://www.hicow.com/education/teacher/classroom-173494.html](http://www.hicow.com/education/teacher/classroom-173494.html). Diunduh pada tanggal 2 Agustus 2010.
- Masri Singarimbun. (1982). *Metode penelitian survei* (Edt). Jakarta: LP3ES.
- McMillan, J.H. (Ed). (2007). *Formative classroom assessment: Theory into practice*. New York: Teacher College, Columbia University.

McNeil, L.M. (2002). *Contradictions of school reform: Educational costs of standardized testing*. [Versi elektronik]. New York: Taylor & Francis e-Library. Diunduh pada tanggal 11 Desember 2011.

Menteri Pendidikan Nasional Republik Indonesia. (2007). *Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 20 Tahun 2007 Tentang Standar Penilaian Pendidikan untuk Satuan Pendidikan dasar dan Menengah*.

(2006).

Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 24 Tahun 2006 Tanggal 2 Juni 2006 tentang Pelaksanaan Peraturan Menteri Pendidikan Nasional Nomor 22 Tahun 2006 Tentang Standar Isi untuk Satuan Pendidikan Dasar dan Menengah dan Peraturan Menteri Pendidikan Nasional Nomor 23 Tahun 2006 Tentang Standar Kompetensi Lulusan untuk Satuan Pendidikan Dasar dan Menengah.

(2006).

Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 23 Tahun 2006 Tanggal 23 Mei 2006 tentang Standar Kompetensi Lulusan untuk Satuan Pendidikan Dasar dan Menengah.

(2006).

Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 22 Tahun 2006 Tanggal 23 Mei 2006 tentang Standar Isi untuk Satuan Pendidikan Dasar dan Menengah.

(2007).

Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 6 Tahun 2007 Tanggal 13 Februari 2007 tentang Perubahan Peraturan Menteri Pendidikan Nasional Nomor 24 Tahun 2006 tentang Pelaksanaan Peraturan Menteri Pendidikan Nasional Nomor 22 Tahun 2006 tentang Standar Isi untuk Satuan Pendidikan Dasar dan Menengah dan Peraturan Menteri Pendidikan Nasional Nomor 23 Tahun 2006 tentang Standar Kompetensi Lulusan Untuk Satuan Pendidikan Dasar dan Menengah.

Michalko, M. (2000). Four steps toward creative thinking [Versi elektronik]. *The Futurist*; May/June 2000; 34, 3; ProQuest Education Journals. Diunduh tanggal 19 Agustus 2007.

Miller, P.W. (2008). *Measurement and teaching*. Munster: Patric W. Miller & Associates.

Muijs, D. & Reynolds, D. (2008). *Effective teaching: Teori dan aplikasi*. (Terjemahan Helly Prajitno Soeftipito & Sri Mulyantini Soecipta). London: Sage Publications Ltd. (Buku asli diterbitkan tahun 2008).

Popham, W.J. (2005). *Classroom assessment: What teachers need to know (4thed)*. Boston: Pearson Education, Inc.

Puckett, M.B. & Black, J.K. (1994). *Authentic assessment of the young child: Celebrating development and learning*. New York: Merrill, and imprint of Macmillan College Publishing Company.

Rezba, R.J., Sparague, C.S., Fiel, R.L., Funk, H.J., Okey, J.R., & Jaus, H.H. (1995). *Learning and assessing science process skills*. 3rd ed. Iowa: Kendall/Hunt Publishing Company.

Roid, G.H. & Haladyna, Th.M. (1982). *A technology for test-item writing*. Orlando: Academic Press, Inc.

Smith, J.K. (2003). *Reconsidering reliability in classroom assessment and grading [Versi elektronik]*. *Educational Measurement, Issues and Practice*, 22, 4, 26-33.

Stark, S., Chernyshenko, S., Chuah, D., Wayne Lee, & Willington, P. (2001). *IRT modeling lab: IRT tutorial [Versi elektronik]*. Urbana: University of Illinois.

Stark, S., Chernyshenko, S., Chuah, D., Wayne Lee, & Willington, P. (2001). *IRT modeling lab: Test Development Using Classical Test Theory [Versi elektronik]*. Urbana: University of Illinois.

Stiggins, R.J. (2002). *Assessment Crisis: The Absence of Assessment for Learning [Versi elektronik]*. *Kappan Professional Journal*.

Last updated 6 June 2002. URL: <http://www.pdkintl.org/kappan/k0206sti.html>. Copyright 2002 Phi Delta Kappa International. Diunduh tanggal 31 Desember 2011.

Subkoviak, M.J. (1988). A practitioner's guide to computation and interpretation of reliability indices for mastery tests. *Journal of Educational Measurement Spring 1988*. Vol.25, No. 1, pp. 47-55.

Tentang Penulis



Prof. Dr. Bambang Subali, M.S., terlahir di Wonosobo, 12 Januari, 1952, adalah Guru Besar Universitas Negeri Yogyakarta. Beliau menyelesaikan S1 Bidang Pendidikan Ilmu Hayat di IKIP Yogyakarta pada 1980, S2 di UGM spesialisasi Ilmu kehutanan, dan S3 di UNY Bidang Penelitian dan Evaluasi Pendidikan pada 2009.

Sebagai dosen yang memulai kariernya pada 1983, Prof. Bambang telah melakukan banyak penelitian, salah satunya Standardisasi Penilaian Hasil Belajar Pola Konvergen dan Divergen Berbasis Satuan Pendidikan Pada Jenjang Sma tahun III. Ia juga pernah menjabat sebagai Koordinator PPHP LPM UNY (1996 – 1999), Pembantu dekan II FMIPA UNY (1999 – 2004), Sekretaris Tim Adhoc BSNP untuk Penyusunan draf Permendiknas No 20 Tahun 2007 dan Panduan Penilaian Kelompok Mata Pelajaran, Ketua Tim Adhoc BSNP untuk Pemantauan dan Evaluasi Implementasi Standar Penilaian Pendidikan (2012), Ketua Kantor Penjaminan Mutu UNY (2010 – 2011), dan Ketua Pusat Penjaminan Mutu LPPMP UNY (2011 – 2012).

Bapak yang ramah ini juga aktif menulis buku diantaranya *Biometri (2011)*, *Pembelajaran Tematik untuk SD & MI Kelas 1 (2007)*, *Pembelajaran Tematik untuk SD & MI Kelas 2 (2007)*, dan *Pembelajaran Tematik untuk SD & MI Kelas 3 (2007)*.

Prinsip Asemen & Evaluasi Pembelajaran

Prof. Dr. Bambang Subali, M.S.



Dalam dunia pendidikan, penilaian atau asesmen (assessment) diartikan sebagai prosedur yang digunakan untuk mendapatkan informasi untuk mengetahui taraf pengetahuan dan keterampilan peserta didik yang hasilnya akan digunakan untuk keperluan evaluasi. Informasi adalah data yang diperoleh melalui pengukuran dan nonpengukuran termasuk di dalamnya dengan melakukan observasi kelas, menggunakan tes yang standar atau tes buatan guru, proyek, dan protfolio subjek belajar (*Assessment is a process of gathering information to measure a student's level of knowledge and/or skill. Assessment can incorporate many methods including classroom observation, standardized tests, teacher-developed tests, classroom projects, and portfolios of student work*).

Evaluasi merupakan suatu proses yang sistematis yang dilakukan untuk mengetahui tingkat keberhasilan dan efisiensi dari program yang bersangkutan. Dalam hal ini termasuk di dalamnya untuk mengetahui keberhasilan seluruh subjek belajar yang menempuh suatu program.

ISBN 979-6416-14-8



9 789796 161451

UNY
PRESS

Jl. H. Afrandi (Jl. Gelajaran)
Gg. Alamanda, Komplek FT-UNY
Kampus Karangmalang, Yogyakarta,
Kode Pos. 55281
Telp. (0274) 589346
redaksi.unypress@gmail.com
cedaik.unypress@gmail.com