

DIKTAT

EVALUASI PEMBELAJARAN FISIKA

Oleh:

Dr. Indrawati, M. Pd.

NIP. 195906101986012001

**Staf Pengajar Prodi Pendidikan Fisika
Jurusan Pendidikan Matematika dan Ilmu Pengetahuan Alam**



**KEMENTERIAN PENDIDIKAN NASIONAL
UNIVERSITAS JEMBER
FAKULTAS KEGURUAN DAN ILMU PENDIDIKAN**

Januari, 2010

KATA PENGANTAR

Mata Kuliah Evaluasi Hasil Belajar Fisika merupakan salah satu mata kuliah dalam kelompok Mata Kuliah Proses Belajar Mengajar (MKPBM). Mata kuliah ini diberikan pada mahasiswa Program Studi Pendidikan Fisika semester IV dengan bobot 4 sks.

Tujuan Mata Kuliah Evaluasi Hasil Belajar Fisika ini diberikan pada mahasiswa adalah agar memiliki mereka memiliki wawasan dan keterampilan tentang evaluasi pembelajaran fisika sekolah menengah. Untuk mencapai tujuan ini perkuliahan dilaksanakan dengan pendekatan kontekstual (Contextual Teaching and Learning) dan berpusat pada mahasiswa (*Student Centered learning*). Pendekatan ini mengajak mahasiswa untuk berpartisipasi secara aktif dalam perkuliahan dan berorientasi pada praktek yang ada di sekolah. Untuk itu, diperlukan bahan ajar yang bisa mempermudah mahasiswa dalam belajar.

Bahan ajar yang dimaksud antara lain berupa diktat kuliah. Agar diktat ini dapat digunakan mahasiswa untuk penunjang belajar evaluasi pembelajaran fisika, maka dalam penyajiannya dilengkapi dengan contoh-contoh yang berkaitan dengan pembelajaran di sekolah menengah.

Dalam penyajiannya, penulis menyadari bahwa tulisan ini jauh dari sempurna. Oleh karena itu, kritik dan saran dari semua pihak yang bersifat membangun sangat diharapkan. Trimakasih.

Jember, Januari 2010

Penulis;

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
KATA PENGANTAR	ii
DAFTAR ISI	iii
BAB I: EVALUASI PERFORMANSI KELAS	1
1.1 Beberapa Istilah dalam Evaluasi.....	1
1.2 Orientasi Pembelajaran dan Evaluasi	6
1.3 Sifat Evaluasi	7
BAB II : TUJUAN PEMBELAJARAN DAN KOMPETENSI	9
1.1 Tujuan Pendidikan	9
1.2 Kompetensi	13
BAB III: TUJUAN PEMBELAJARAN DAN EVALUASI	22
3.1 Merencanakan Tes	23
3.2 Jenjang-jenjang Tes	24
BAB IV: MENGONSTRUKSI DAN MENGGUNAKAN TES	27
4.1 Bentuk/Tipe Tes	27
4.2 Pembetulan Tes Siswa	32
4.3 Evaluasi Diri	32
4.4 Pedoman Umum Menggunakan Tes	33
BAB V : TES-TES PRAKTEK	35
5.1 Evaluasi Kerja Laboratorium	35
5.2 Laporan Laboratorium	37
5.3 Mencatat Performansi Siswa	25
BAB VI : EVALUASI BELAJAR TUNTAS DAN EVALUASI TRADISIONIL	42
6.1 Evaluasi Belajar Tuntas	42
6.2 Penilaian Tradisionil	46

BAB VII : KARAKTERISTIK DAN SPESIFIKASI BUTIR SOAL ...	48
7.1 Karakteristik Butir Soal	48
7.2 Spesifikasi Butir Soal	52
7.3 Karakteristik Perangkat Tes	53
DAFTAR PUSTAKA	57

1 EVALUASI PERFORMANSI KELAS

Banyak guru, termasuk guru fisika mempersepsikan evaluasi terbatas hanya pada pemberian tes. Aspek yang lebih luas dari evaluasi meliputi evaluasi diri oleh siswa, evaluasi kerja laboratorium, evaluasi diagnostik, formatif, dan sumatif, dan aspek-aspek lain dari total proses evaluasi kadang-kadang tidak dipahami oleh guru. Hal ini terjadi karena mereka hanya sedikit menerima latihan formal dalam evaluasi dan metode evaluasinya secara tradisional yang cenderung formal, mengkonsentrasikan pada kuis dan berakhir dengan tes.

Karena hasil pembelajaran dan evaluasi dipikirkan sebagai kesatuan yang terpisah, maka suatu hubungan yang berlawanan tampak ada antara yang mengevaluasi dan yang dievaluasi. Hal ini memperkecil efektivitas dan gangguan dari salah satu yang sangat beralasan untuk melakukannya (asesmen yang valid dari hasil belajar, sikap, perkembangan keterampilan, dan kemajuannya). Evaluasi harus dipertimbangkan sebagai bagian yang vital dan tidak dapat dipisahkan dari dari pembelajaran.

1.1 Beberapa Istilah dalam Evaluasi

Berkaitan dengan kegiatan evaluasi, ada beberapa istilah yang perlu dipahami oleh guru. Istilah-istilah tersebut adalah: *measurement and*

instrument, data, a test and examination, assessment and evaluation.
Semua istilah ini akan diuraikan seperti di bawah ini.

a. *Measurement and Instrument* (Pengukuran dan alat ukur)

Measurement refers to the process of obtaining a numerical description of the degree to which an individual possesses a particular characteristic (Sadler, R. 1998). Pengukuran adalah proses perolehan sebuah penggambaran dengan angka mengenai sejauhmana seseorang individu memproses karakteristik tertentu. Karakteristik ini dalam pembelajaran antara lain adalah: minat belajar, aktivitas belajar, hasil belajar, retensi hasil belajar, ketuntasan hasil belajar, kecerdasan, dan lain-lain.

Dalam belajar fisika, istilah pengukuran sudah tidak asing lagi. Misalnya pengukuran: suhu, massa, panjang, waktu, kuat arus, dan pengukuran besaran-besaran lain dalam fisika. Untuk menentukan atau mengukur setiap besaran tersebut memerlukan alat ukur (**instrument**). Mistar digunakan untuk mengukur panjang, termometer untuk mengukur suhu, dan neraca untuk mengukur massa. Hal ini juga terjadi dalam lingkungan psikologi pendidikan dan pembelajaran. Misalnya perangkat tes IQ digunakan untuk mengukur IQ seseorang, perangkat tes hasil belajar untuk mengukur hasil belajar siswa, perangkat tes bakat untuk mengukur bakat seseorang, perangkat lembar observasi untuk mengukur aktivitas belajar siswa, angket untuk mengukur minat belajar siswa, dan lain-lain. Jadi perangkat: tes IQ, tes hasil belajar, tes bakat, lembar observasi, angket, dan lain-lain adalah contoh-contoh instrumen yang digunakan dalam pembelajaran.

b. **Data**

Data kuantitatif

Data bisa dipetoleh melalui proses pengukuran. Pengukuran, instrumen dan data merupakan tiga istilah yang tidak bisa dipisahkan. Dalam

proses pengukuran diperlukan instrumen. Data adalah sesuatu yang diperoleh berdasarkan hasil dari proses pengukuran. Data hasil pengukuran dapat ditunjukkan dalam dua bentuk, yaitu data dalam bentuk kuantitatif dan dalam bentuk kualitatif. Data kuantitatif dapat ditunjukkan dalam empat bentuk skala, yaitu skala nominal, ordinal, interval, dan rasio. **Skala nominal** adalah suatu ukuran yang mana angka-angka merepresentasikan kategori yang tidak mempunyai urutan atau nilai kuantitatif (misalnya siswa dari Jawa sebagai "1", siswa Sunda sebagai "2", dan siswa Madura sebagai "3"). **Skala ordinal** adalah suatu ukuran dalam bentuk angka yang merepresentasikan suatu ranking yang mengurutkan individu atau obyek dalam variabel (misalnya siswa A ranking 1, siswa B ranking 2, siswa C ranking 3). Angka 1, 2, dan 3 menunjukkan urutan dari kategori tinggi ke kategori-kategori yang lebih rendah. **Skala interval** adalah suatu ukuran yang tidak mempunyai angka nol benar dan jarak antardua titik yang berurutan adalah sama (misalnya: skala termometer). **Skala rasio** adalah suatu ukuran yang mempunyai titik nol benar dan jarak antardua titik yang berurutan adalah sama (misalnya skala mistar).

Data kualitatif

Data selain berupa angka, dapat pula disajikan dalam bentuk deskriptif, yang biasa diistilahkan dengan data kualitatif. Data yang berbentuk kualitatif adalah data yang disajikan dalam bentuk deskripsi tentang fenomena, misalnya siswa kelas IE SMA X sulit menangkap informasi fisika yang sifatnya logikomatematik, siswa tidak aktif ketika mengikuti pelajaran fisika, siswa tidak mau bertanya ketika mempunyai kesulitan, dan lain-lain.

c. A Test and Examination (Tes dan Ujian)

A test refers to a particular situation set up for the purpose of making an assessment. The pencil-and-paper test in the classroom is an

obvious example but an essay subject set for homework or classwork also provides a test situation, though this may not be its only function.

Tes merujuk pada situasi khusus untuk tujuan pembuatan suatu asesmen. Tes tulis di kelas merupakan contoh yang jelas tetapi kumpulan subyek esai untuk pekerjaan rumah atau pekerjaan kelas juga merupakan situasi tes, meskipun fungsinya tidak sama.

Examination *refers to a larger scale test, or, more commonly, a combination of several tests, and perhaps other assessment procedures, whether within the school or conducted by an external examining board.* Ujian merujuk pada tes dalam skala lebih besar atau lebih umum, merupakan kombinasi dari beberapa tes, dan mungkin prosedur asesmen lain, apakah di dalam kelas atau dilakukan oleh penguji luar.

d. Assessment dan Evaluation

Assesment *is any of variety of procedures used to obtain information about student performance. Includes traditional paper-pencil tests as well as extended responses (e.g., essays) and performance of authentic tasks (e.g. Laboratory experiments). Assessment answers the question "How well does the individual perform?"* Asesmen merupakan salah satu dari prosedur yang digunakan untuk memperoleh informasi mengenai penampilan siswa. Mencakup tes tertulis yang perlu jawaban bebas (seperti esai), tes perbuatan yang otentik (seperti percobaan laboratorium). Asesmen menjawab pertanyaan "seberapa baikkah seorang individu berprestasi/tampil?"

Evaluation *is the systematic process of collecting, analysing, and interpreting information to determine the extent to which pupils are achieving instructional objectives (answers the questions "how good").* Evaluasi adalah proses sistematis pengumpulan, penganalisisan, dan penafsiran informasi untuk menentukan sejauh mana siswa mencapai tujuan pembelajaran (menjawab pertanyaan "Seberapa baik?").

Berkaitan dengan tujuan penilaian, *assessment refers to the extent to which pupils have benefitted from a course of instruction* (asesmen berkenaan dengan sejauh mana siswa memperoleh manfaat dari suatu proses pembelajaran) *and evaluation as the effectiveness of methods of teaching* (evaluasi berkenaan dengan efektivitas metode mengajar). Untuk memberikan wawasan tentang asesmen dan evaluasi, berikut ini diberikan cuplikan dari Bob Kizlik (2007) seperti berikut.

*Getting Ready for Distance Education
Distance Education Aptitude and Readiness Scale (DEARS)
Dr. Bob Kizlik*

Copyright 2007 Robert Kizlik

Assessment is a process by which information is obtained relative to some known objective or goal. Assessment is a broad term that includes testing. A test is a special form of assessment. Tests are assessments made under contrived circumstances especially so that they may be administered. In other words, all tests are assessments, but not all assessments are tests. We test at the end of a lesson or unit. We assess progress at the end of a school year through testing, and we assess verbal and quantitative skills through such instruments as the SAT and GRE. Whether implicit or explicit, assessment is most usefully connected to some goal or objective for which the assessment is designed. A test or assessment yields information relative to an objective or goal. In that sense, we test or assess to determine whether or not an objective or goal has been obtained. Assessment of skill attainment is rather straightforward. Either the skill exists at some acceptable level or it doesn't. Skills are readily demonstrable. Assessment of understanding is much more difficult and complex. Skills can be practiced; understandings cannot. We can assess a person's knowledge in a variety of ways, but there is always a leap, an inference that we make about what a person does in relation to what it signifies about what he knows. In the section on this site on behavioral verbs, to assess means To stipulate the conditions by which the behavior specified in an objective may be ascertained. Such stipulations are usually in the form of written descriptions.

Evaluation is perhaps the most complex and least understood of the terms. Inherent in the idea of evaluation is "value." When we evaluate, what we are doing is engaging in some process that is designed to provide information that will help us make a judgment about a given situation.

Generally, any evaluation process requires information about the situation in question. A situation is an umbrella term that takes into account such ideas as objectives, goals, standards, procedures, and so on. When we evaluate, we are saying that the process will yield information regarding the worthiness, appropriateness, goodness, validity, legality, etc., of something for which a reliable measurement or assessment has been made. For example, I often ask my students if they wanted to determine the temperature of the classroom they would need to get a thermometer and take several readings at different spots, and perhaps average the readings. That is simple measuring. The average temperature tells us nothing about whether or not it is appropriate for learning. In order to do that, students would have to be polled in some reliable and valid way. That polling process is what evaluation is all about. A classroom average temperature of 75 degrees is simply information. It is the context of the temperature for a particular purpose that provides the criteria for evaluation. A temperature of 75 degrees may not be very good for some students, while for others, it is ideal for learning. We evaluate every day. Teachers, in particular, are constantly evaluating students, and such evaluations are usually done in the context of comparisons between what was intended (learning, progress, behavior) and what was obtained. When used in a learning objective, the definition provided on the ADPRIMA site for the behavioral verb evaluate is: To classify objects, situations, people, conditions, etc., according to defined criteria of quality. Indication of quality must be given in the defined criteria of each class category. Evaluation differs from general classification only in this respect.

To sum up, we measure distance, we assess learning, and we evaluate results in terms of some set of criteria. These three terms are certainly connected, but it is useful to think of them as separate but connected ideas and processes.

1.2 Orientasi Pembelajaran dan Evaluasi Kelas Sains

Semua pembelajaran dan evaluasi di kelas sains harus diorientasikan ke arah kesuksesan bukan pada kegagalan. Siswa merespon dengan senang pada pengalaman-pengalaman yang berhasil dan sebaliknya mereka secara cepat akan frustrasi dengan kegagalan-kegagalan yang terulang. Hal ini guru perlu mengorbankan standar-standarnya dalam bekerja pada suatu tingkat yang tak bermakna. Hal ini secara sederhana mengakui bahwa siswa akan berfungsi lebih baik ketika mereka mempunyai konsep-konsep positif tentang nilai/harganya sendiri

dan performansinya di kelas. Jadi setiap tugas harus disediakan agar dapat memberikan potensi untuk keberhasilan prestasi (pencapaian) di beberapa tingkat pada semua siswa tanpa menghiraukan kemampuan-kemampuan akademiknya. Di hampir semua kasus hasil belajar akhir hubungannya lebih dekat pada kecepatan belajar bukan pada seberapa tingkat IQ yang dipertanyakan.

1.3 Sifat Evaluasi

Evaluasi melibatkan total asesmen belajar siswa, meliputi evaluasi pemahamannya terhadap proses sains, kemampuan dan prestasi mata pelajaran, bakat-bakat ganda, sikap-sikap ilmiah, keterampilan-keterampilan laboratorium, dan keinginan untuk bekerja. Kemajuan siswa yang mengarah pada tujuan-tujuan pelajaran dan tujuan sekolah seperti efektifitas pembelajaran harus dipertimbangkan. Evaluasi yang baik menunjukkan kekuatan dan kelemahan pembelajaran. Sekali seorang guru telah melakukan suatu kegiatan asesmen, dia mempunyai suatu indikasi bagaimana meningkatkan pembelajarannya. Evaluasi bekerja sebagai umpan balik dalam proses eksperimen mengajar. *Guru harus bereks-perimen jika dia ingin maju dan menjadi lebih terampil.* Dia harus berkeinginan untuk mencoba metode-metode baru dan teknik-teknik baru, dan dengan melakukan pengembangan yang mengarah pada mengajar tuntas.

Para guru sains harus diorientasikan pada eksperimen, tidak hanya di laboratrum tetapi juga dalam pendekatan-pendekatan sehari-harinya mengajar. Pendekatan eksperimen mengasumsikan pengumpulan data untuk membuktikan keberhasilan dari metode-metode yang digunakan. Data harus berasal dari teknik-teknik evaluasi. Semakin baik evaluasi pembelajaran, maka semakin banyak informasi yang tersedia untuk guru dalam meningkatkan pengajarannya.

Ada tiga tipe evaluasi yang dipertimbangkan dalam bidang pengajaran, yaitu: diagnostik, formatif, dan sumatif. Tiga tipe tersebut

utamanya berbeda oleh posisi kronologis dalam urutan pembelajaran. **Evaluasi diagnostik** secara normal mendahului pembelajaran tetapi dapat digunakan untuk lingkungan khusus untuk menemukan masalah-masalah belajar siswa. Evaluasi diagnostik dapat memberikan informasi untuk guru tentang pengetahuan, sikap, dan keterampilan siswa untuk masuk suatu mata pelajaran dan dapat digunakan sebagai suatu dasar untuk perbaikan individu atau pembelajaran khusus.

Evaluasi formatif dihasilkan selama periode pembelajaran untuk memberikan umpan balik pada siswa dan guru pada bagaimana baiknya bahan yang diajarkan dan dipelajari. Karena mengajar merupakan proses dinamik, maka evaluasi formatif dapat memberikan informasi yang bermanfaat bahwa guru dapat menggunakan untuk memodifikasi pembelajaran dan dapat meningkatkan efektifitas mengajar untuk individu dan kelompok.

Jenis evaluasi ketiga adalah **evaluasi sumatif**, yaitu: jenis evaluasi yang sering digunakan oleh guru dan utamanya ditujukan untuk memberikan peringkat siswa dan melaporkan prestasi (hasil) belajar. Hal ini sering didasarkan pada perolehan-perolehan kognitif dan jarang menggunakan pertimbangan bidang-bidang intelektual lain.

Saat sekarang para guru telah memberikan banyak informasi mengenai bidang-bidang pengembangan intelektual, meliputi ranah kognitif, psikomotor, dan afektif. Ranah kognitif secara tradisional telah merupakan wilayah perhatian utama oleh hampir semua guru. Hasil belajar dalam bentuk informasi mengingat konsep, keterampilan memecahkan masalah, dan aspek-aspek perolehan informasi telah diuji sepenuhnya dan dalam berbagai bentuk. Ranah psikomotor dan afektif, bagaimanapun juga, belum diberikan melalui studi yang sama.

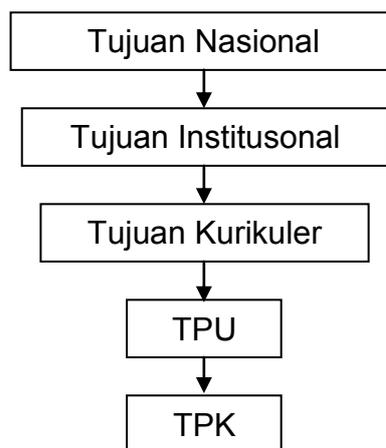
2 TUJUAN PEMBELAJARAN

2.1 Tujuan Pendidikan

Pendidikan merupakan salah satu program negara. Sekolah adalah lembaga pelaksana pendidikan yang utama. Sehingga, negara mengalokasikan sebagian anggarannya untuk pelaksanaan kegiatan sekolah. Guru merupakan faktor sumber daya manusia yang penting dalam kegiatan pendidikan di sekolah. Dia tidak hanya bertanggung jawab pada siswa atau sekolah, tetapi juga pada masyarakat dan negara. Oleh karena itu, sebagian alokasi dana pendidikan negara adalah untuk guru. Dengan demikian, guru mempunyai tanggung jawab moral yang paling berat dalam mewujudkan program pendidikan negara.

Sasaran pendidikan adalah manusia. Tujuan pendidikan nasional adalah untuk mencerdaskan kehidupan bangsa dan mengembangkan manusia seutuhnya. Untuk mewujudkan tujuan pendidikan nasional tersebut diserahkan pada lembaga-lembaga pendidikan, seperti SMP dan SMA. Setiap bentuk kegiatan pendidikan di setiap lembaga tersebut harus mengacu pada tujuan nasional, termasuk tujuan di setiap sekolah. Dalam pelaksanaan pendidikan melalui sekolah, tujuan tersebut diterjemahkan dalam kurikulum-kurikulum bidang studi atau mata pelajaran. Setiap bidang studi terdiri atas pokok-pokok bahasan dan sub-sub pokok bahasan. Tujuan untuk pokok-pokok bahasan dalam suatu bidang studi atau mata pelajaran dirumuskan dalam Tujuan Pembelajaran Umum (TPU) dan tujuan untuk sub-sub pokok bahasan dirumuskan dalam Tujuan

Pembelajaran Khusus (TPK). Secara hirarkis tujuan-tujuan pendidikan dapat dibagangkan seperti berikut.



Gambar 2.1 Bagan Hirarki Tujuan Pendidikan

Bagan di atas menunjukkan bahwa antara jenjang tujuan yang satu dan lainnya saling berkaitan. Jenjang yang bawah harus mengacu pada tujuan-tujuan jenjang atasnya atau sebaliknya tujuan yang lebih atas harus diterjemahkan oleh tujuan-tujuan di bawahnya.

Agar anda lebih jelas hubungan antarjenjang tersebut coba jelaskan secara singkat dalam selembar kertas apa yang dimaksud setiap jenjang tujuan tersebut (tujuan nasional, tujuan institusional, tujuan kurikuler, TPU, dan TPK) dan kumpulkan pada pertemuan mendatang.

Dari urutan jenjang tujuan tersebut, TPK merupakan jenjang tujuan yang terletak pada jenjang terbawah dan terbatas ruang lingkungannya, dan perumusannya ditugaskan oleh guru. Hal ini dapat dikatakan bahwa guru dapat dipandang sebagai ujung tombak penentu kualitas manusia. Oleh karena itu, perumusan TPK harus dipikirkan dengan cermat.

Dalam pelaksanaan pengajaran di sekolah, setiap kurikulum sekolah dijabarkan dalam Garis-garis Besar Program Pengajaran (GBPP) atau silabus mata pelajaran. Dalam silabus suatu kurikulum, tujuan pendidikan sudah dirumuskan sampai tingkat satuan/unit kegiatan belajar mengajar (KBM) yaitu tujuan Pembelajaran Umum (TPU). Setiap TPU ada yang beraspek kognitif, afektif, dan psikomotor. Tugas guru adalah menjabarkan TPU ke dalam TPK. Biasanya pada rentangan antara TPU dan TPK terdapat sejumlah *tujuan antara*. Tujuan antara ini berfungsi untuk menjembatani pencapaian tujuan umum dari sejumlah tujuan khusus.

Karena tugas guru adalah menerjemahkan TPU dalam silabus menjadi tujuan-tujuan pembelajaran khusus (TPK), maka perlu cara merumuskan TPK yang jelas dan mudah dievaluasi keberhasilannya. Mengajar pada dasarnya menciptakan/membantu siswa untuk belajar. Gagne mendefinisikan belajar sebagai perubahan tingkah laku karena pengalaman. Oleh karena itu, ia mengusulkan agar tujuan pembelajaran itu dirumuskan dengan menggunakan kata-kata yang menggambarkan kemampuan yang dapat diamati (*observable*), seperti: menghitung, menyebutkan, dan menjelaskan. Berikutnya, perumusan TPK harus berpusat pada pembelajar (siswa) bukan pada guru atau kemampuan yang diharapkan dimiliki siswa setelah belajar. Selain itu, rumusan TPK juga dapat memberi gambaran dalam kondisi apa kemampuan itu dapat dimunculkan. Tetapi sekarang orang menganggap bahwa syarat ketiga itu tidak perlu. Sehingga, perumusan TPK menjadi lebih sederhana, yaitu: hanya memuat pembelajar sebagai *audience* (A) dan memuat tingkah laku

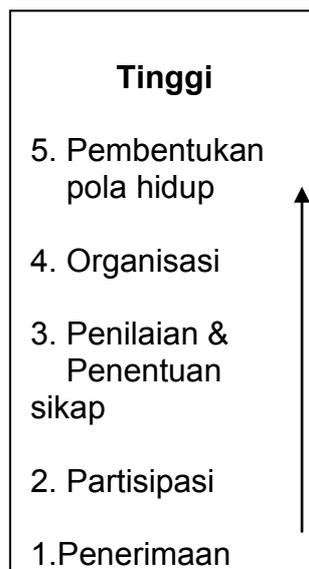
yang dapat diamati setelah akhir pembelajaran yang biasa disebut dengan *behaviour* (B).

Dalam merumuskan tujuan pendidikan, Bloom dan kawan-kawan mengklasifikasi tujuan pendidikan dalam tiga ranah, yakni ranah: kognitif, psikomotor, dan afektif. Setiap ranah ini digolongkan dalam beberapa tingkat yang tersusun secara hirarkis. Susunan itu disebut dengan istilah taksonomi, artinya jenjang yang lebih atas dapat ditempuh kalau sudah melewati jenjang di bawahnya. Taksonomi untuk ranah kognitif, afektif, dan psikomotor dikemukakan berturut-turut oleh Bloom, dkk., Krathwohl dan Bloom, dkk., dan Simpson. Ketiga taksonomi tersebut dapat dibagikan sebagai berikut:



Bagan 2.1

Taksonomi Bloom, dkk.



Bagan 2.2

Taksonomi Krathwohl & Bloom, dkk



Bagan 2.3

Taksonomi Harrow

Setiap kategori dalam setiap ranah tujuan tersebut juga terdiri atas sub-sub kategori, seperti kategori memahami dalam ranah kognitif terdiri

atas mempertahankan, memperkirakan, menggeneralisasi, mengevaluasi, dan sebagainya. Untuk mengetahui lebih lengkap tentang sub-sub dari setiap kategori dalam setiap ranah dapat dilihat untuk ranah kognitif oleh Bloom, dkk. (1956:62), untuk ranah afektif oleh Krathwohl, dkk. (1981:95), dan untuk ranah psikomotor oleh Harrow (1972:100)

Dalam setiap kegiatan pembelajaran diharapkan ketiga ranah tersebut dapat dirumuskan oleh guru. Karena rumusan TPK harus memuat tingkah laku yang dapat diamati setelah pembelajaran, maka guru biasanya mengalami kesulitan dalam merumuskan TPK untuk ranah afektif bila dibandingkan untuk ranah kognitif dan psikomotor. Biasanya dampak pembelajaran pada sikap hanya dapat teramati setelah beberapa saat. Selain itu, sifat materi yang diajarkan juga menentukan faktor kesulitan dalam merumuskan TPK.

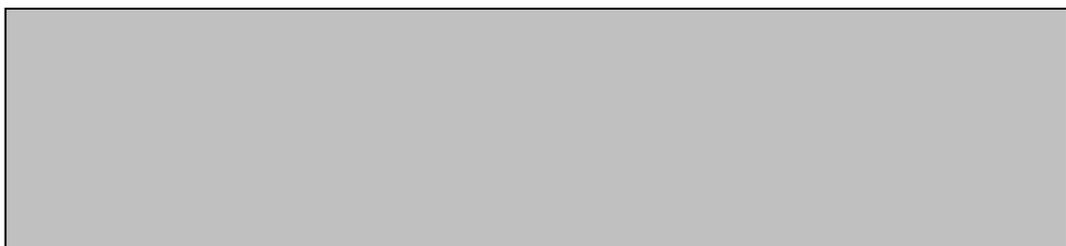
Dalam Rencana Pelaksanaan Pembelajaran (RPP), apakah TPK yang dirumuskan sudah memenuhi kriteria perumusan TPK?

Misalnya TPK 1: Setelah selesai pelajaran, siswa akan dapat menyebutkan ciri-ciri zat cair.

Pejelasan: Siswa (*audience*)

Akan dapat menyebutkan ciri-ciri zat cair (*behavior*)

Karena TPK 1 memuat *audience* dan perilaku siswa yang dapat diamati, maka TPK ini sudah memenuhi syarat. Dapat menyebutkan ciri-ciri zat cair termasuk dalam ranah kognitif pada kategori pemahaman syarat memuat audience dan behaviour dirumuskan untuk *ranah kognitif pada kategori pengetahuan, sub kategori pengetahuan tentang unsur universal dan abstraksi dalam suatu bidang, dan pada sub-sub kategori pengetahuan tentang prinsip dan generalisasi.*



Apakah TPK 2 dan TPK 3 juga sudah memenuhi syarat perumusan TPK? Jika ya, sebutkan pada ranah, kategori, dan sub kategori apa? Diskusikan dengan teman-teman Anda!

2.2 Kompetensi

Adanya perkembangan zaman, di dunia pendidikan juga mengalami perubahan. Perubahan ini diwujudkan dalam bentuk perubahan kurikulum. Kurikulum lama yang berbasis pada konten dipikirkan kurang bisa mengakomodasi kemampuan siswa dan juga kurang bisa membawa siswa yang mengarah pada keterampilan hidup (*life skill*). Adapun kurikulum sekarang yang berbasis kompetensi diharapkan dapat mengatasi kekurangan yang ada pada kurikulum sebelumnya. **Kompetensi** dalam pendidikan dapat dimaknai sebagai pengetahuan, sikap-perilaku, dan keterampilan yang tercermin dalam kebiasaan berpikir dan bertindak. Kebiasaan berpikir dan bertindak yang dilakukan secara konsisten dan terus menerus memungkinkan seseorang menjadi kompeten dalam bidang tertentu. Dengan demikian, dapat dikatakan bahwa seseorang dikatakan kompeten jika ia memiliki pengetahuan, keterampilan, dan nilai-nilai dasar yang tercermin dalam kebiasaan berpikir dan bertindak.

Kurikulum adalah seperangkat rencana dan pengaturan mengenai tujuan, isi, dan bahan pelajaran serta cara yang digunakan sebagai pedoman penyelenggaraan kegiatan pembelajaran untuk mencapai tujuan pendidikan tertentu. Saat ini kurikulum yang berlaku adalah kurikulum yang berbasis kompetensi adalah Kurikulum Tingkat satuan pendidikan (KTSP). KTSP adalah kurikulum operasional yang disusun oleh dan dilaksanakan di masing-masing satuan pendidikan. KTSP terdiri atas tujuan pendidikan tingkat satuan pendidikan, struktur dan muatan kurikulum tingkat satuan pendidikan, kalender pendidikan, dan silabus. Silabus adalah rencana pembelajaran pada suatu dan/atau kelompok

mata pelajaran/tema tertentu yang mencakup standar kompetensi, kompetensi dasar, materi pokok/pembelajaran, kegiatan pembelajaran, indikator, penilaian, alokasi waktu, dan sumber/bahan/alat belajar. Silabus merupakan penjabaran standar kompetensi dan kompetensi dasar ke dalam materi pokok/pembelajaran, kegiatan pembelajaran, dan indikator pencapaian kompetensi untuk penilaian.

Berkenaan dengan tujuan pendidikan seperti pada Gambar 2.1, dan KTSP, maka istilah TPU dan TPK tidak digunakan lagi, tetapi digunakan dengan istilah SK, KD, Indikator, dan Tujuan pembelajaran yang masing-masing dapat diuraikan sebagai berikut.

Standar kompetensi (SK)

Standar kompetensi (SK) adalah seperangkat kompetensi yg dibakukan dan harus dicapai siswa sebagai hasil belajarnya dalam setiap satuan pendidikan. SK digunakan untuk memandu penjabaran kompetensi dasar (KD) menjadi pengalaman belajar . Pengalaman belajar merupakan kegiatan fisik atau mental yg dilakukan siswa dalam mencapai KD dan materi pembelajaran. Urutan SK menggunakan menggunakan pendekatan prosedural dan hirarkis. Pendekatan prosedural digunakan apabila SK yg diajarkan berupa serangkaian langkah-langkah secara urut dalam mengerjakan tugas pembelajaran. Pendekatan hirarkis menunjukkan hubungan yang bersifat subordinat/berjenjang antara beberapa SK yang dicapai, sehingga ada yang mendahului dan ada yang kemudian. SK yang mendahului merupakan prasyarat SK berikutnya.

Kompetensi Dasar (KD)

Kompetensi Dasar (KD) adalah rincian dari SK, yang berisi pengetahuan, keterampilan, dan sikap yang secara minimal harus dikuasai siswa. Urutan (sekuens) menggunakan pendekatan prosedural, hirarkis, mudah-sukar, konkret-abstrak, spiral, tematik/terpadu, dan lain-lain.

Indikator

Indikator adalah kompetensi dasar secara lebih spesifik yang dapat dijadikan ukuran untuk menilai ketercapaian hasil pembelajaran. Indikator dikembangkan oleh guru sesuai kebutuhan dan potensi siswa, menggunakan kata kerja operasional yang dapat diukur dan cakupan materinya terbatas.

Tujuan Pembelajaran

Tujuan pembelajaran adalah target kompetensi spesifik yang ingin dicapai dalam pembelajaran, berupa perilaku siswa yang dapat diamati (*observable*), diukur (*measurable*), dan diuji (*testable*). Perilaku tersebut berupa kemampuan kognitif, afektif, dan psikomotorik siswa (Taksonomi Bloom) seperti yang diuraikan sebelumnya. Dalam RPP (Rencana Pelaksanaan Pembelajaran) ada yg hanya cukup menuliskan Indikator tanpa ada Tujuan Pembelajaran. Indikator sudah menggambarkan Tujuan Pembelajaran. Rumusan Tujuan Pembelajaran yang baik memuat unsur: *Audience* (A), *Behaviour* (B), *Condition* (C), *Degree* (D). Sekurang-kurangnya memuat A dan B (seperti yang diuraikan sebelumnya). Indikator identik dengan Tujuan pembelajaran tanpa *audience*.

2.3 Standar Proses

Peraturan Pemerintah Nomor 19 tahun 2005 tentang Standar Nasional Pendidikan salah satu standar yang harus dikembangkan adalah standar proses. Standar proses adalah standar nasional pendidikan yang berkaitan dengan pelaksanaan pembelajaran pada satuan pendidikan untuk mencapai kompetensi lulusan. Standar proses berisi kriteria minimal proses pembelajaran pada satuan pendidikan dasar dan menengah di seluruh wilayah hukum Negara Kesatuan Republik Indonesia. Standar proses ini berlaku untuk jenjang pendidikan dasar dan menengah pada jalur formal, baik pada sistem paket maupun pada sistem kredit semester. Standar proses untuk satuan pendidikan dasar dan menengah mencakup meliputi perencanaan proses pembelajaran, pelaksanaan

proses pembelajaran, penilaian hasil pembelajaran, dan pengawasan proses pembelajaran untuk terlaksananya proses pembelajaran yang efektif dan efisien (Permendiknas RI No 41 tahun 2007)

2.4 Pendidikan Karakter

Perkembangan zaman yang begitu pesat dan kompleks memungkinkan untuk merubah karakter bangsa. Hal ini sekolah harus bisa mengantisipasi dengan wajar. Untuk itu, pendidikan karakter perlu diberikan pada siswa melalui mata pelajaran-mata pelajaran. Dengan demikian, dalam penyusunan silabus dan RPP perlu memuat tujuan yang berorientasi pada pengembangan karakter siswa. suatu mata pelajaran.

- a. Berperilaku sesuai dengan ajaran agama yang dianut sesuai dengan perkembangan remaja
- b. Mengembangkan diri secara optimal dengan memanfaatkan kelebihan diri serta memperbaiki kekurangannya
- c. Menunjukkan sikap percaya diri dan bertanggung jawab atas perilaku, perbuatan, dan pekerjaannya
- d. Berpartisipasi dalam penegakan aturan-aturan sosial
- e. Menghargai keberagaman agama, bangsa, suku, ras, dan golongan sosial ekonomi dalam lingkup global
- f. Membangun dan menerapkan informasi dan pengetahuan secara logis, kritis, kreatif, dan inovatif
- g. Menunjukkan kemampuan berpikir logis, kritis, kreatif, dan inovatif dalam pengambilan keputusan
- h. Menunjukkan kemampuan mengembangkan budaya belajar untuk pemberdayaan diri
- i. Menunjukkan sikap kompetitif & sportif untuk mendapatkan hasil yang terbaik
- j. Menunjukkan kemampuan menganalisis dan memecahkan masalah kompleks
- k. Menunjukkan kemampuan menganalisis gejala alam dan sosial

- l. Berpartisipasi dalam kehidupan bermasyarakat, berbangsa, dan bernegara secara demokratis dalam wadah Negara Kesatuan Republik Indonesia
- m. Mengekspresikan diri melalui kegiatan seni dan budaya
- n. Mengapresiasi karya seni dan budaya
- o. Menghasilkan karya kreatif, baik individual maupun kelompok.
- p. Menjaga kesehatan dan keamanan diri, kebugaran jasmani, serta kebersihan lingkungan.
- q. Berkomunikasi lisan dan tulisan secara efektif dan santun.
- r. Memahami hak dan kewajiban diri dan orang lain dalam pergaulan di masyarakat.
- s. Menghargai adanya perbedaan pendapat dan berempati terhadap orang lain.
- t. Menunjukkan keterampilan membaca dan menulis naskah secara sistematis dan estetis.
- u. Menunjukkan keterampilan menyimak, membaca, menulis, dan berbicara dalam bahasa Indonesia dan Inggris.
- v. Menguasai pengetahuan yang diperlukan untuk mengikuti pendidikan tinggi

Contoh SK, KD, I, TP, dan Materi Pokok yang memuat Standar Proses dan Pendidikan Karakter

A. Standar Kompetensi

6. Memahami konsep dan prinsip GEM.

B. Kompetensi Dasar :

6.1 Mendiskripsikan spektrum GEM

6.2

C. Indikator

a. Kognitif

1. Produk:

- a) Mendeskripsikan konsep GEM.
- b) Mendeskripsikan spektrum GEM.
- c) Mendeskripsikan range spektrum GEM.
- d) Menyebutkan contoh-contoh penerapan GEM.
- e) Menentukan panjang gelombang GEM.
- f) Menentukan frekuensi GEM.
- g) Mendeskripsikan rumus energi GEM.

2. Proses:

- a) Merumuskan masalah,
- b) Merumuskan hipotesis,
- c) Mengidentifikasi variabel manipulasi,
- d) Mengidentifikasi variabel respon,
- e) Mengidentifikasi variabel kontrol,
- f) Membuat tabel,
- g) Melakukan analisis data, dan
- h) Menyimpulkan.

Psikomotor (Tidak ada.....)

b. Afektif

- 1. Mengembangkan perilaku berkarakter, meliputi:
 - a) kritis
 - b) jujur,
 - c) peduli,
 - d) tanggung jawab,
- 2. Mengembangkan keterampilan sosial, meliputi:

- a) bertanya,
- b) menyumbang ide atau berpendapat,
- c) menjadi pendengar yang baik,
- d) berkomunikasi
- e) bekerjasama

D. Tujuan Pembelajaran:

a. Kognitif

1. Produk:

Dengan mendiskusikan bahan ajar (*handout*) secara kelompok, siswa dapat:

- a) Menjelaskan pengertian GEM.
- b) Menjelaskan bentuk GEM.
- c) Menjelaskan spektrum GEM.
- d) Menjelaskan range gelombang dalam spektrum GEM.
- e) Memberikan contoh-contoh penerapan GEM.
- f) Menghitung besarnya panjang gelombang GEM.
- g) Menghitung besarnya frekuensi GEM.
- h) Merumuskan besarnya energy GEM.
- i) Menghitung besarnya energy GEM.

2. Proses:

Disediakan Daftar spektrum GEM dan LKS, dengan memodifikasi beberapa macam nama gelombang elektromagnetik siswa dapat menentukan hubungan antara panjang gelombang, frekuensi, dan kecepatan GEM, sesuai rincian tugas kinerja (RTK) dalam “LKS Proses” meliputi: **merumuskan masalah, merumuskan hipotesis, mengidentifikasi variabel manipulasi, mengidentifikasi variabel respon, mengidentifikasi variabel**

kontrol, membuat tabel, melakukan analisis data, dan menyimpulkan.

b. Afektif

2. Karakter

Terlibat dalam proses belajar mengajar berpusat pada siswa, paling tidak siswa dinilai pengamat **Membuat kemajuan** dalam menunjukkan perilaku berkarakter meliputi: *kritis, jujur, peduli, dan tanggung jawab*.

3. **Keterampilan sosial:**

Terlibat dalam proses belajar mengajar berpusat pada siswa, paling tidak siswa dinilai pengamat **Membuat kemajuan** dalam menunjukkan keterampilan sosial bertanya, menyumbang ide atau berpendapat, menjadi pendengar yang baik, berkomunikasi dan kerjasama

3 TUJUAN PEMBELAJARAN DAN EVALUASI

Ada tingkat-tingkat (jenjang) dalam tujuan, ada juga tingkat-tingkat pengujian (*testing*). Banyak evaluasi yang belum menguji untuk semua tujuan yang penting dalam sains. Biasanya guru hanya mengevaluasi kualitas pembelajarannya dari ranah kognitif saja belum menemukan bagaimana baiknya dia mengajar. Ranah afektif dan psikomotor diberikan untuk memberikan umpan balik yang lengkap pada guru sehingga guru dapat memodifikasi secara efektif teknik-teknik pembelajarannya.

Aturan pertama sebagai petunjuk dalam merencanakan tes adalah pada penggunaan tujuan-tujuan dan prinsip-prinsip sains. Biasanya siswa cenderung belajar tentang apa yang akan diujikan pada mereka. Jika penekanannya adalah pada mengingat fakta, maka untuk memuaskan guru mereka akan mengingat fakta. Jika penekanannya pada pemahaman prinsip, pengembangan keterampilan proses, kreativitas, aspek-aspek inkuiri lain, atau tipe-tipe mengajar investigasi, tes-tes ini seterusnya akan membawa tipe-tipe jawaban tersebut dari siswa. Hal ini penting bagi guru untuk mempertimbangkan apakah tes menekankan pada semua tujuan-tujuan penting dari pelajaran. Jika tes-tes diberikan, mereka harus mengevaluasi bagaimana baiknya siswa untuk mencapai tujuan-tujuan ini, tidak memperhatikan apakah mereka dimaknai sebagai siswa atau sebagai guru.

Untuk menulis tujuan-tujuan pembelajaran relatif mudah, tetapi untuk mengevaluasi hasil belajarnya sering sulit. Benyamin Bloom

mengidentifikasi enam tingkat tujuan kognitif, yaitu: pengetahuan, pemahaman, aplikasi, analisis, sintesis, dan evaluasi (seperti yang dijelaskan pada bab sebelumnya). Sehingga, pertanyaan-pertanyaan harus direncanakan untuk mengevaluasi setiap tingkat ini. Jika kita menulis tujuan perilaku dengan baik dan menggunakannya sebagai petunjuk untuk mengkonstruksi tes kita, maka memungkinkan kita akan memberikan ujian-ujian lebih baik daripada tidak disiapkan sebelumnya.

3.1 Merencanakan Tes

Satu cara untuk meyakinkan tes kita mengikuti tujuan dan memberikan perhatian yang memadai untuk setiap bidang yang diinginkan adalah merencanakan Tabel Spesifikasi atau biasa disebut dengan kisi-kisi tes. Tabel tersebut dapat ditunjukkan pada Tabel 3.1 berikut.

Tabel 3.1: Tabel Spesifikasi Perencanaan Tes

TPK	Materi (<i>content</i>)					Total (%)
	A	B	C	D	
1. Pengetahuan						
2. Pemahaman						
3. Aplikasi						
4. Analisis						
5. Sintesis						
6. Evaluasi						
Total (%)						

Dengan menggunakan tabel spesifikasi selama pembelajaran dan mencatat jumlah waktu yang dibutuhkan untuk setiap bidang dan tingkat-tingkat tujuan yang diharapkan, dan mengikuti tabel secara hati-hati ketika membuat tes (khususnya untuk sumatif), kita akan membuat suatu tes yang mempunyai validitas isi lebih baik daripada yang diambil bersama-

sama secara tergesa-gesa dari mengingat atau didasarkan pada ingatan guru dari waktu luang pada setiap bidang.

Meskipun Tabel 3.1 dirancang untuk menganalisis ranah kognitif, tabel tersebut dapat dengan mudah dimodifikasi untuk memasukkan ranah afektif dengan menambah kategori lain pada kolom sama dengan kolom kognitif dengan judul “ranah afektif.” Begitupula untuk ranah psikomotor.

Mengkonstruksi dan menganalisis tes dengan menggunakan kisi ini membantu untuk meyakinkan kemajuan kita dalam menulis tes yang mengevaluasi tingkat-tingkat belajar lebih tinggi. Guru yang belum mengevaluasi semua tujuannya atau mengklasifikasi pertanyaan-pertanyaan tesnya dengan beberapa cara yang mirip dengan kisi-kisi tersebut cenderung mengevaluasi pada tingkat-tingkat paling bawah dari taksonomi Bloom. Penelitian yang dilaporkan pada tahun 1960 menghasilkan bahwa penekanan utama pada tes-tes kimia adalah pada pengetahuan fakta. Aplikasi, analisis, dan sintesis secara nyata diabaikan.

3.2 Jenjang-jenjang Tes

Jenjang terbawah dalam hirarki taksonomi Bloom adalah jenjang pengetahuan. Penekanan-penekanan ini menyederhanakan hafalan (*recall*) atau pengenalan (*recognition*) dan menggambarkan jenjang terbawah dari belajar, atau hanya memerlukan mengingat informasi. Mengajar hafalan merupakan jenjang terbawah dari pembelajaran, tetapi guru sering mencurahkan jumlah waktu banyak untuk ini karena pertanyaan-pertanyaan yang sifatnya hafalan dapat ditulis dengan mudah. Ada suatu aturan bahwa tes yang sifatnya pertanyaannya hafalan harus tidak lebih dari 20%.

Jenjang berikutnya adalah jenjang pemahaman. Kadang-kadang ini dipertimbangkan pada tingkat pertama pemahaman. Siswa dapat mengetahui sesuatu tentang suatu topik, dapat mengikuti suatu proses, atau mengetahui bagaimana menulis pertanyaan tanpa memahaminya

secara utuh. Pertanyaan pemahaman memerlukan siswa untuk (1) menginterpretasi pernyataan, (2) menerjemahkan atau menjelaskan proses atau ide dengan kata-kata mereka sendiri, (3) mengekstrapolasi, atau (4) menginterpolasi. Pertanyaan-pertanyaan pemahaman tidak sulit untuk menulis tetapi memerlukan beberapa pemikiran guru untuk mengungkapkan dalam suatu bentuk yang memerlukan tingkat berpikir lebih tinggi dari sekedar mengingat hal yang sifatnya sederhana.

Jenjang ketiga adalah aplikasi. Jenjang ini dapat dijelaskan paling baik adalah membandingkannya dengan pemahaman. Masalah pemahaman memerlukan siswa mengetahui suatu abstraksi yang cukup baik untuk menunjukkan secara tepat penggunaannya ketika ditanyakan secara khusus dalam mengerjakannya. Aplikasi, di sisi lain, memerlukan suatu tahap di luar ini. Misalnya, siswa diberi suatu masalah baru, siswa dapat menerapkan abstraksi dengan tepat tanpa ditunjukkan cara menggunakannya dalam situasi itu. Dalam pemahaman, harus dapat menunjukkan bahwa siswa dapat menggunakan abstraksi ketika penggunaannya dikhususkan. Dalam aplikasi, siswa menunjukkan bahwa dapat secara spontan menggunakan abstraksi secara tepat.

Jenjang keempat taksonomi Bloom adalah analisis. Ini dihubungkan pada pemahaman dan evaluasi. Analisis menekankan pemerincian bahan (materi) ke dalam bagian-bagiannya dan mendeteksi hubungan bagian-bagian ini dalam organisasinya. Kadang-kadang hal ini juga diarahkan pada teknik-teknik atau peralatan-peralatan yang digunakan untuk memberikan suatu pengertian atau membuat suatu kesimpulan.

Jenjang kelima, sintesis, melibatkan pemilihan unsur-unsur atau bagian-bagian secara bersama untuk menguraikan pola atau struktur yang didefinisikan kurang baik sebelumnya. Prosedur ini biasanya melibatkan pengombinasian pengalaman sebelumnya dengan materi baru, konstruksi keseluruhan yang terintegrasi kurang atau lebih baik. Kegiatan ini merupakan bentuk kegiatan kreatif. Kategori ini adalah dalam ranah kognitif dan tidak dapat selalu melibatkan ekspresi kreatif bebas karena

siswa biasanya diharapkan bekerja di dalam batas-batas yang disesuaikan dengan masalah, materi, atau kerangka metodologi khusus.

Jenjang paling tinggi dari taksonomi adalah evaluasi. Jenjang ini merupakan proses penilaian jenjang tentang ide-ide, penyelesaian, metode, dan bahan-bahan yang memenuhi kriteria. Ini juga melibatkan penggunaan kriteria sebagai standar untuk menilai tingkat pada item-item khusus akurat, efektif, ekonomis, atau memuaskan. Penilaian seperti ini dapat bersifat kuantitatif atau kualitatif dan kriteria ini dapat ditentukan oleh siswa atau dapat disediakan oleh guru.

Seperti telah dijelaskan sebelumnya dan berkaitan dengan tiga ranah tujuan pembelajaran beserta hirarki dalam setiap ranah, maka pengkonstruksian tes juga berlaku untuk ranah afektif dan psikomotor.

4 MENGKONSTRUKSI & MENGGUNAKAN TES

Pada dasarnya tes dapat dibedakan berdasarkan bentuknya, pelaksanaannya, dan tujuannya. Berdasarkan bentuknya tes dapat berbentuk obyektif dan subyektif (esai). Berdasarkan pelaksanaannya bisa dilakukan secara tertulis atau lisan. Ditinjau dari tujuannya, tes dapat berupa tes hasil belajar, tes diagnostik, dan tes kecerdasan. Berikut ini dibahas tentang berbagai jenis tes yang biasa digunakan untuk mengukur hasil belajar (tes hasil belajar).

4.1 Bentuk/Tipe Tes

4.1.1 Tes Benar-Salah

Jika ujian dibatasi pada pertanyaan-pertanyaan benar-salah, statistik menunjukkan bahwa 75 item atau lebih adalah perlu untuk mengatasi faktor menebak (*guessing*). Dalam tes 100 pertanyaan benar-salah, siswa harus dapat menjawab sekitar 50 pertanyaan tepat dengan menebak. Beberapa petunjuk untuk mengeliminasi masalah ini adalah dengan mengurangi jumlah jawaban-jawaban salah dari jumlah jawaban benar dalam menentukan skor; mereka dapat menghukum untuk menebak. Prosedur ini tidak direkomendasi karena siswa biasanya berpikir bahwa instruktur/guru menggunakan teknik ini secara dendam/dengki. Hal ini juga tidak dapat diinginkan karena siswa dihukum untuk menebak; dalam sains kita ingin mempunyai siswa membuat

hipotesis, oleh sebab itu kemampuan dalam tebakan-tebakan juga diperlukan.

Dalam membuat tes, hindari ketidak-seimbangan tes dengan cukup banyak pertanyaan benar-salah, usahakan membuat soal rata dalam jumlah, sehingga siswa yang mengetahui sedikit tentang materi tes tidak dapat memperoleh skor tinggi dengan mudah, dengan berasumsi bahwa lebih banyak pertanyaan benar (atau salah). Hindari menggunakan pernyataan-pernyataan yang dapat menipu siswa. Jangan menggunakan bahasa yang sama seperti dalam teks atau siswa cenderung untuk mengingat. Hindari pernyataan-pernyataan yang mendua-arti. Misalnya, jangan menulis, "Erosi dicegah dengan penyemaian." Hindari menggunakan kalimat-kalimat kompleks dalam pernyataan-pernyataan anda. Jangan menggunakan bahasa kualitatif jika anda dapat memungkinkan menghindarinya. Jangan menulis, misalnya, logam-logam yang lebih baik menghantarkan listrik lebih cepat. Susun pernyataan-pernyataan anda dalam 10 sampai dengan 20 pertanyaan. Prosedur ini menggantikan terlalu banyak ketegangan pada siswa. Ambil blok/tempat jawaban ada dalam satu margin sehingga mereka dapat diperiksa dengan mudah dengan menggunakan suatu kunci.

4.1.2 Tes Pilihan Ganda

Tes pilihan-ganda terdiri atas item-item yang mempunyai lebih dari tiga jawaban. Jika tidak ada sekurang-kurangnya empat jawaban yang mungkin untuk setiap pertanyaan, rumusan jawaban benar harus digunakan. Tes pilihan ganda berbeda dari tes jawaban ganda yaitu hanya ada satu jawaban benar untuk setiap pertanyaan dalam tipe pertama tes.

- a. Ada beberapa petunjuk untuk mengkonstruksi soal pilihan ganda.
- b. Semua jawaban harus masuk akal.
- c. Semua jawaban harus konsisten dari sudut tata bahasa.
- d. Cobalah untuk menjaga semua jawaban panjangnya hampir sama.
- e. Acak jawaban-jawaban benar sehingga tidak ada pola dalam ujian.

- f. Ingat bahwa jawaban benar sering dapat ditentukan oleh suatu proses eliminasi boleh dikatakan dengan mengetahui jawaban-jawaban benar.
- g. Cobalah untuk mencegah hal ini dalam penyusunan jawaban.
- h. Sajikan pertama istilah atau konsep yang anda inginkan untuk mengetes apa.
- i. Ujikan tingkat-tingkat pemahaman yang tinggi sebanyak mungkin.
- j. Gunakan metode sederhana untuk jawaban.
- k. Sediakan garis-garis pendek untuk jawaban sepanjang satu margin dari satu halaman sehingga mereka dapat dicocokkan dengan mudah.
- l. Kelompokkan item-item anda dalam bagian-bagian. Sistem ini membuatnya mudah untuk merujuk pada berbagai bagian dari tes dan membantu memecahkan monotomi dalam mengerjakan tes.
- m. Kelompokkan bersama semua pertanyaan dengan jumlah pilihan sama.

4.1.3 Tes Melengkapi dan Menjodohkan

Oleh karena tes melengkapi dan menjodohkan biasanya menekankan pada hafalan dan sering secara verbal rumit, maka tes tipe ini jumlahnya harus diperkecil. Beberapa petunjuk yang dapat digunakan dalam tipe tes ini.

- a. Jika pertanyaan-pertanyaan menjodohkan digunakan, mereka harus dikelompokkan. Jika ada lebih dari 15 item menjodohkan dalam kelompok, tes menjadi tidak praktis.
- b. Nomori pertanyaan-pertanyaan anda dan gunakan huruf untuk jawaban anda, atau sebaliknya, tetapi konsisten.
- c. Buat pilihan-pilihan menjodohkan lebih banyak dari jumlah pertanyaan untuk memperkecil mendapatkan jawaban-jawaban dengan penyisihan.

Meskipun tes menjodohkan menekankan secara tradisional pada hafalan, mereka dapat digunakan untuk menguji pengenalan atau penerapan prinsip-prinsip. Berikut ini diberikan contoh tipe tes menjodohkan dan melengkapi.

Tes Menjodohkan:

- 1. Besaran yang mempunyai besar dan arah A. $[LT^{-2}]$
B. Skalar
- 2. Dimensi percepatan adalah C. $[LT^2]$
- 3. Perbandingan antara gaya tarik atau gaya tekan dan pemanjangan atau pemendekan pegas adalah konstan. D. Vektor
E. Hukum Newton II
F. Hukum Hooke

Contoh Tes Melengkapi:

1. Besaran yang hanya mempunyai besar saja disebut besaran
2. Kecepatan termasuk besaran
3. Dimensi dari besaran energi adalah
4. Hambatan-hambatan yang disusun seri, hambatan ekivalennya akan lebih dibandingkan jika hambatan tersebut disusun paralel.

4.1.4 Tes Diri

Tes diri konstruksinya mirip dengan tes lain tetapi diberikan siswa utamanya sebagai suatu alat belajar. Biasanya pertanyaan di satu sisi dan jawabannya di sisi lain. Siswa mengerjakan tes, kemudian membalik halaman dan mencocokkan jawabannya. Instruktur dapat menggunakan tes melengkapi sebagai alat merangsang diskusi.

Guru biasanya maset tes-tes diri untuk materi sama dan mencetak cukup kopi untuk siswanya. Format yang disarankan seperti pada Gambar 1. Halaman balik harus memuat penjelasan lengkap (ditail) untuk setiap jawaban sehingga siswa belajar dari tes. Siswa melipat jawaban di sisi

kanan halamannya. Jawabannya kemudian didekatkan dengan jawaban-jawaban benar dan penjelasannya di margin kiri halaman balik yang membuatnya mudah untuk memeriksa tes.

1. Pertanyaan	Jawaban: -----	Jawaban	1. Jawaban dan penje-lasan benar dari ja-waban.
2.	-----	-----	2.
3.	-----	-----	3.
.			
.			

Gambar 4.1: Format Tes-Diri

4.1.5 Tes Masalah (*Problem*)

Jenis tes ini menyajikan suatu masalah dan menanyakan siswa untuk mengerjakannya. Ini mirip dengan ajakan untuk berinkuiri kecuali bahwa ini dilakukan oleh siswa individu. Tes biasanya memuat sederetan pertanyaan yang harus siswa jawab untuk memecahkan masalah. Tes masalah dapat dikonstruksi dengan relatif mudah jika didasarkan pada suatu masalah yang telah secara nyata bertetangan dengan saintis. Masalah ini dapat diperoleh dengan mudah dari jurnal saintifik. Berikan siswa informasi tentang masalah dan ijin membuat hipotesis-hipotesis, rancangan penelitian, atau metode mengumpulkan dan mencatat data mereka sendiri. Tes masalah paling baik dapat digunakan untuk memperkenalkan siswa dengan proses-proses ilmiah. Tes ini mempunyai lembar jawaban mirip dengan tes diri atau mereka dapat digunakan untuk mendorong/merangsang diskusi. Masalah-masalah sains terbuka lebih baik untuk jenis tes ini.

Beberapa contoh masalah yang dapat digunakan untuk mengkonstruksi suatu tes dari sifat ini adalah:

- a. Bandul sederhana berosilasi dengan frekuensi f . Berapa frekuensinya jika percepatan bandul $\frac{1}{2}g$ (a) ke atas, dan (b) ke bawah?
- b. Jika anda memuati sisir dengan menggosoknya dengan selendang sutera, bagaimana anda menentukan jika sisir termuati positif atau negatif?

4.1.6 Tipe Tes Diagram atau Gambar

Banyak tes sains mengevaluasi lebih banyak untuk membaca daripada untuk mengerjakan sains. Banyak siswa memahami prinsip-prinsip sains tetapi karena mereka memiliki kesulitan verbal, seperti membaca dan menginterpretasikan apa yang telah mereka baca, mereka cenderung mengerjakan tes kurang baik. Studi-studi menunjukkan banyak siswa mengerjakan secara signifikan lebih baik pada tes utamanya yang terdiri atas gambar dan hanya memerlukan untuk memeriksa jawaban-jawaban yang benar.

4.2 Pembetulan Tes Siswa

Pembetulan (koreksi) tes siswa dapat dilakukan dalam beberapa cara. Siswa dapat bekerja sebagai pembantu dengan menggunakan kunci yang disiapkan oleh instruktur/guru. Instruktur/guru dapat memberikan tes itu pada siswa secara random, membaca jawaban di kelas, dan mempunyai jawaban-jawaban benar dan menghitung skor-skor. Siswa dapat diberi kunci setelah mereka menerima tes; mereka kemudian memeriksa tes-nya sendiri. Untuk jawaban-jawaban salah, mereka harus menulis di balik tes atau pada kertas terpisah tentang penjelasan jawaban-jawaban salahnya. Penjelasan ini dapat dianalisis untuk menentukan pertanyaan-pertanyaan yang harus dimodifikasi atau dibuang pada ujian selanjutnya.

4.3 Evaluasi Diri

Salah satu nilai penting dari evaluasi pembelajaran adalah memberikan umpan balik pada siswa. Yang termasuk dalam aspek ini harus menjadi komponen evaluasi diri siswa dengan perhatian tidak hanya pada perasaan dan perhatiannya mengenai kelas tetapi juga mengenai pemahamannya tentang kemajuannya sendiri dalam pelajaran. Salah satu tujuan yang kita perjuangkan dalam pendidikan adalah mengajar siswa untuk menerapkan tujuan-tujuan untuk mereka sendiri dan kemudian menghasilkan prosedur yang perlu untuk mencapai tujuan ini. Evaluasi diri merupakan bagian yang alami dan proses ini perlu, dan siswa harus diyakinkan untuk memperoleh kebiasaan lebih awal. Semua individu dari waktu ke waktu melihat kemajuannya dan hasil belajarnya sendiri dengan berintrospeksi. Hal ini wajar bahwa kesempatan-kesempatan dapat diberikan di kelas dengan harapan pembelajarannya menjadi lebih baik.

4.4 Pedoman Umum Menggunakan Tes

Jika kita dihadapkan dengan kemungkinan menyiapkan suatu tes, ada beberapa metode yang akan menyumbangkan untuk evaluasi agar tepat dan efektif. Beberapa pedoman untuk mengetes dapat ditunjukkan di bawah ini.

- a. Gunakan tes-tes secara ramah sebagai alat belajar dan diagnostik. Berikan siswa kesempatan untuk menunjukkan bahwa mereka telah belajar tentang apa yang salah dalam suatu tes.
- b. Jangan pernah menggunakan tes sebagai hukuman.
- c. Perkecil penggunaan pertanyaan-pertanyaan melengkapi dan menjodohkan.
- d. Gunakan tes-tes atau infentarisasi evaluasi diri untuk mengevaluasi semua tujuan-tujuan perilaku anda (TPK), termasuk proses dan sikap sains.
- e. Luangkan waktu dengan setiap siswa memeriksa pertanyaan-pertanyaan yang dia salah. Hal ini dapat dilakukan selama kelas terlibat dalam kera laboratorium.

- f. Ingat bawa tes-tes hanya merupakan suatu contoh yang telah dipelajari dan mungkin bukan merupakan satu contoh yang sangat baik. Jadi mereka harus tidak digunakan hanya sebagai alat evaluasi. Gunakan dalam pertimbangan untuk semua hal yang telah siswa lakukan di kelas untuk mengembangkan bakat-gandanya.
- g. Teskan untuk semua jenjang taksonomi Bloom dan gunakan kisi-kisi analisis tes untuk melihat bahwa ini dilakukan.
- h. Tanyakan pertanyaan-pertanyaan untuk menentukan bagaimana perasaan siswa tentang materi yang digunakan.
- i. Tempatkan pertanyaan-pertanyaan yang lebih mudah di awal tes sehingga siswa memperoleh keyakinan dan memperkecil kecewa dan gelisah/cemas.
- j. Pertimbangkan faktor waktu, berapa lama waktu yang disediakan siswa untuk melengkapi tes? Beberapa siswa akan selesai lebih cepat dari yang lain. Apa yang akan anda lakukan dengan mereka? Jika tidak mempunyai pedoman kerja, mereka mungkin menimbulkan masalah-masalah kedisiplinan.
- k. Rancang tes untuk diskor dengan mudah. Sediakan ruang untuk semua jawaban pada satu margin.
- l. Daripada siswa menulis jawaban dalam tes, sediakan jawaban-jawaban mereka pada suatu lembar jawab. Prosedur ini meyakinkan untuk mudah mencatat dan mengamankan kertas jika tes mungkin digunakan untuk lebih dari satu kelas.

5

TES-TES PRAKTEK

Dalam tes praktek, siswa dapat diarahkan pada penampilan tugas laboratorium tertentu. Guru dapat mengamati teknik-teknik yang digunakan, kebenaran prosedur, dan hasil yang diperoleh. Prosedur seperti identifikasi zat-zat kimia yang tidak diketahui, praktek umum dalam analisis kualitatif dalam kimia, dapat diperluas untuk sains bumi, biologi, dan fisika. Tes-tes hasil belajar, yang dirancang untuk mengases pemahaman konten (isi) pelajaran, merupakan teknik evaluasi yang penting untuk kerja laboratorium karena perhatian bahwa pengetahuan benar diperoleh melalui metode-metode laboratorium. Tes-tes pengenalan (*recognition*) dan hafalan murni adalah bentuk yang tidak cocok untuk tes hasil belajar pengalaman laboratorium. Tes-tes yang bergantung pada pengamatan yang akurat, pengenalan (*recognition*) data yang berkaitan, dan kemampuan untuk benalar secara logis lebih cocok untuk mengukur kerja laboratorium.

Kerja laboratorium melibatkan pengujian bahan-bahan, penentuan zat-zat yang tidak diketahui, mengenal dan mengklasifikasi organisme, menguraikan prosedur percobaan yang penting. Masalah dengan ujian-ujian laboratorium praktek adalah membutuhkan banyak waktu untuk menset. Asisten lab siswa atau salah satu dari dua siswa A dapat sangat membantu dalam menset ujian sebelum atau setelah sekolah.

5.1 Evaluasi Kerja Laboratorium

Konsep yang lebih tentang evaluasi meliputi pemberian perhatian pada semua kegiatan yang digunakan siswa selama bertahun-tahun di sekolah menengahnya. Dengan peningkatan waktu dan penekanan dicurahkan pada kerja laboratorium, menjadikan perlu untuk merencanakan metode-metode yang cocok untuk mengevaluasi kegiatan ini. Seperti dengan semua evaluasi, tujuan kegiatan atau metode mengajar harus diidentifikasi sebelum prosedur digunakan secara nyata dalam evaluasi ditentukan. Untuk kerja laboratoium, daftar tujuan umum (TPU) yang disarankan adalah:

- a. untuk mengembangkan keterampilan dalam pemecahan masalah melalui identifikasi masalah, mengumpulkan dan menginterpretasikan data, dan menarik kesimpulan
- b. untuk mengembangkan keterampilan dalam memanipulasi alat-alat laboratorium
- c. untuk membentuk kebiasaan-kebiasaan sistematis menjaga catatan
- d. untuk mengembangkan sikap ilmiah
- e. untuk belajar metode-metode ilmiah untuk memecahkan masalah
- f. untuk mengembangkan kepercayaan diri dan dapat dipertanggungjawabkan
- g. untuk menemukan kesempatan-kesempatan yang tidak tergalikan dan penyelidikan
- h. untuk mempromosikan kegairahan pada mata pelajaran sains

Kegiatan-kegiatan khusus yang biasanya siswa terlibat dalam kerja laboratorium meliputi:

- a. Merencanakan eksperimen dan membentuk hipotesis
- b. Merencanakan suatu darmawisata
- c. Menset alat
- d. Menyusun bahan dan alat

- e. Mengamati gejala alam
- f. Mengamati suatu proses di dalam laboratorium
- g. Meneliti informasi dokumenter yang berwenang dalam topik itu
- h. Mengumpulkan dan mencatat data
- i. Mengumpulkan bahan-bahan percobaan
- j. Mengklasifikasi dan mengorganisasi bahan-bahan
- k. Memodifikasi alat
- l. Membaca alat
- m. Mengkalibrasi alat
- n. *Menggambar diagram dan grafik*
- o. Menganalisis data
- p. Menarik kesimpulan dari data
- q. Menulis laporan percobaan
- r. Menggambarkan dan menjelaskan eksperimen pada siapa saja
- s. Mengidentifikasi masalah lanjut untuk penelitian
- t. Membongkar, membersihkan, menyimpan, dan memperbaiki alat.

Kegiatan-kegiatan dan hasil umum kerja laboratorium ini dapat dievaluasi dalam beberapa cara, seperti tes-tes praktek, menggunakan hal-hal yang tidak diketahui, tes-tes hasil belajar, pengamatan langsung dari teknik-teknik laboratorium, laporan-laporan tertulis, konferensi-konferensi individual dan kelompok.

5.2 Laporan Laboratorium

Laporan tertulis, merupakan suatu metode yang sering digunakan untuk mengevaluasi kemajuan dan pemahaman di lab, dan laporan ini harus diperiksa dengan cermat. Sering laporan tertulis menjadi suatu bentuk klise (*stereotyped*) yang menghilangkan nilainya sebagai suatu instrumen evaluasi. Setiap siswa diminta untuk menggunakan suatu bentuk baku yang meninggalkan sedikit kesempatan pada kreatifitas dan fleksibilitas.

Karakteristik-karakteristik berikut dari laporan eksperimen baik mengarah pada item-item pokok yang diperlukan setelah meninggalkan ruangan untuk berinisiatif dan berkreasi pada diri siswa. Tipe laporan ini, siswa memulai dengan lembar kertas kosong. Jika dia menulis, dia menjaga di dalam criteria pemikiran seperti berikut.

- a. Pembaca dapat menceritakan dengan tepat apa yang siswa coba tentukan.
- b. Pembaca dapat melihat prosedur yang siswa gunakan untuk memperoleh jawaban pada masalah. Menggambarkan dengan jelas, ringkas, dan lengkap.
- c. Data yang dikumpulkan terorganisasi baik dan dapat dipahami dengan mudah.
- d. Semua pengukuran ditunjukkan dengan satuan-satuan yang tepat.
- e. Diagram, jika digunakan, dilabelkan dengan jelas dan hati-hati. Diagram hanya bermanfaat untuk membuat percobaan lebih jelas untuk pembaca.
- f. Grafik, jika digunakan, diberi judul, diberi label, dan digambarkan dengan rapi. Tujuan grafik adalah untuk menunjukkan hubungan antara data yang diperoleh, sehingga kesimpulan dapat ditarik.
- g. Kesimpulan harus menjawab masalah, menggunakan data yang diperoleh dari percobaan.
- h. Laporan harus membantu siswa, dalam mereviu pelajaran, menghafal dengan tepat tentang percobaan apa dan kesimpulan apa yang dicapai.
- i. Kriteria utama untuk evaluasi laporan percobaan ini adalah: apakah laporan ditulis cukup jelas orang yang tidak diberitahu dapat membacanya, mengetahui dengan tepat apa yang dicoba, bagaimana percobaan dilakukan, dan kesimpulan apa yang dicapai, dan jika perlu, dapatkah dia menduplikasi percobaannya sendiri, menggunakan laporan sendiri sebagai suatu pedoman?

KETERAMPILAN. PROSES										
Mengorganisir kerja Mencatat data dengan tepat										

Komentar:

Biasanya pekerjaan lab dapat diperiksa dengan teliti dan hati-hati. Pada saat-saat yang lain pekerjaan tersebut dapat diseleksi. Bentuk dapat diacu untuk akhir kwartal atau semester sebagai bantuan dalam menentukan nilai untuk aspek lab pada kelas-kelas sains.

6 EVALUASI BELAJAR TUNTAS & EVALUASI TRADISIONAL

6.1 Evaluasi Belajar Tuntas

Belajar tuntas merupakan suatu bidang strategi pembelajaran yang menghubungkan pada proses evaluasi. Hal ini mengacu pada ketuntasan siswa dari tujuan-tujuan tertentu yang ditentukan untuk jenjang-jenjang yang diidentifikasi.

Pada tahun 1950-an sederetan studi menunjukkan bahwa di US, perbedaan antara cara hasil belajar tertinggi dan terendah yang diperoleh siswa di sekolah-sekolah terhadap kebangsaan kira-kira satu standar deviasi (1 SD). Hal ini menyatakan bahwa beberapa negara bagian membutuhkan 12 tahun sekolah untuk mencapai tingkat yang sama dan skor-skor hasil belajar diperlukan hanya 8 tahun di negara-negara bagian lain. Implikasi dari petunjuk ini dari sejumlah peneliti mempertimbangkan ide-ide belajar tuntas. Orang yang pertama mengemukakan ide ini adalah John B. Carroll dari Universitas Havard, yang pada tahun 1960 mengembangkan suatu model yang menggunakan *unsur waktu sebagai variabel pembelajaran inti dalam proses belajar*. Ini merupakan pendapatnya bahwa, waktu yang cukup diberikan dan ketekunan untuk melengkapi tugas, siswa harus mencapai kriteria tingkat hasil belajar yang diinginkan. Sementara kecepatan belajar akan bervariasi dari siswa ke

siswa, hasil akhir akan menjadi sama. Carroll mengidentifikasi 5 faktor yang mempengaruhi belajar: (1) sikap, (2) kemampuan untuk memahami pembelajaran, (3) kualitas pembelajaran, (4) waktu yang diijinkan untuk belajar, dan (5) ketekunan.

Pada tahun 1974, Benyamin Bloom mengembangkan suatu perencanaan untuk belajar tuntas didasarkan pada penemuan Carroll. Dia meneorikan bahwa dalam kondisi yang menyenangkan 95% siswa di sekolahnya akan mampu mencapai tingkat hasil belajar yang sekarang dicapai hanya oleh 20 % kelompok atas. *Bloom mengkritik metode-metode evaluasi pendidikan saat ini ketika dia memperingatkan bahwa salah satu efek yang paling mengganggu pendidikan modern adalah mendasarkan standar performansi pada distribusi nilai tradisinal yang didasarkan pada kurve normal.* Hasilnya, dia menyatakan bahwa hanya proporsi kecil siswa yang diberi dengan suatu pengalaman pendidikan yang berhasil. Dia juga menyatakan bahwa usaha-usaha pendidikan dapat diakui tidak berhasil pada tingkat bahwa hasil belajar siswa terdistribusi secara normal. Hal ini menceritakan pada kita bahwa secara praktik semua siswa mempunyai kemampuan untuk belajar apa yang kita usulkan untuk mengajarnya di sekolah-sekolah menengah. Perbedaan berada pada jumlah waktu yang diperlukan dan dalam kesulitan untuk mengakomodasi nilai-nilai belajar yang bervariasi ini di dalam struktur waktu tradisional lamanya bersekolah kita.

Bloom telah mengamati bahwa kerja awal yang dilakukan pada belajar tuntas adalah pada mata-mata pelajaran yang ada beberapa prasyarat, seperti aljabar dan sains. Jika mata-mata pelajaran mengandalkan secara berat pada belajar pelajaran sebelumnya, hal ini tak mungkin bahwa banyak siswa akan mencapai tuntas di dalam jadwal waktu pelajaran normal. Sains merupakan salah satu bidang yang mana konsep belajar tuntas dapat digunakan dengan menyenangkan. Bidang-bidang isi sains mempertemukan persyaratan-persyaratan tertutup dan urutan (rentetan). Hal ini berarti bahwa mereka perlu suatu minimum

belajar sebelumnya, dipelajari secara urut, menekankan berpikir konvergen, dan tertutup selama mereka memiliki kumpulan ide-ide terbatas dan perilaku-perilaku kognitif yang biasanya guru dapat setuju.

Bloom telah mendefinisikan lima ketentuan bahwa instruktur dalam suatu kelas belajar tuntas harus memenuhi kriteria berhasil untuk menjadi efektif, yaitu:

- a. Siswa harus diberi informasi tentang harapan-harapan pelajaran, biasanya melalui penggunaan tujuan-tujuan belajar.
- b. Standar-standar untuk tuntas yang disusun dalam kemajuan dan nilai-nilai diberikan dengan istilah-istilah performansi.
- c. Tes-tes diagnostik pendek, disebut tes-tes formatif, digunakan untuk setiap satuan pelajaran.
- d. Belajar tambahan untuk membantu siswa mencapai tingkat kriteria ditentukan dengan angka-angka yang tepat.
- e. Waktu belajar tambahan disediakan untuk siswa yang membutuhkannya.

Bloom selalu mencela ide-ide tentang kompetisi yang kadang-kadang digunakan sebagai jastifikasi untuk sistem evaluasi tradisional dan “persiapan hidup.” Dia memperingatkan bahwa banyak belajar dan pengembangan dapat terganggu jika kompetisi merupakan dasar utama untuk motivasi. Malahan siswa harus diberi suatu standar performansi dan nilai-nilai yang menunjukkan keberhasilan akan dihargai/dihadiahi untuk semua yang mencapai tingkat kriteria. Pada waktu dulu biasanya kita telah mendasarkan pada standar tingkat tuntas yang secara sejarah adalah tingkat hasil belajar A dan B di kelas-kelas kita. Khusus hal ini merujuk tuntas pada 75-85% dari tugas-tugas yang diperlukan dalam sederetan tujuan-tujuan khusus (TPK).

Unsur-unsur yang paling penting dalam strategi belajar tuntas adalah *prosedur umpan balik/koreksi, yang merupakan produk dari tes formatif*. Tes-tes formatif diberikan secara periodik di seluruh pembelajaran dan digunakan untuk memberikan umpan balik (*feedback*)

untuk siswa dan juga untuk guru. Mereka tidak dinilai, dan siswa dapat mengulang tes dengan waktu yang cukup untuk mencapai tingkat tuntas. Dasar pemikiran ini adalah bahwa pola-pola untuk belajar harus disusun sebagai bagian dari proses berputar, yaitu siswa dapat menguji pengetahuan mereka dan kemudian diberi umpan balik yang berguna untuk mengarahkan usaha-usaha belajarnya dalam upayanya untuk menuntaskan pelajarannya.

Metode-metode tradisional untuk memberikan tes sumatif pada akhir bab tidak melayani kebutuhan-kebutuhan siswa untuk belajar karena mereka tidak diberi tahu tentang hasil-hasil dan tidak ada usaha yang dibuat mengulang belajar bahan pelajaran yang masih salah. Tes-tes sumatif hanya mengukur hasil belajar dan kegagalan untuk mengidentifikasi bidang-bidang kelemahan. Kita didesak untuk memberikan nilai di hampir semua sistem sekolah, dan tes-tes sumatif akan dibatasi pada tujuan itu, sebagai kalimat/kata keterangan kuantitatif untuk evaluasi siswa lebih menyeluruh.

Salah satu cara untuk melihat pengajaran tuntas adalah untuk mempertimbangkan bahwa suatu pelajaran terdiri atas dua tipe konsep, satu tipe menguraikan tujuan-tujuan inti dari mata pelajaran dan menunjukkan harapan-harapan minimal yang diharapkan untuk semua, dan tipe kedua adalah menguraikan tujuan-tujuan yang tidak inti di luar persyaratan minimal. Bloom menjaga untuk meluangkan ekstra waktu di dalam periode kalender sama untuk mencapai tingkat hasil belajar yang sama dengan teman-teman sekelas memberikan suatu perasaan yang dia kerjakan dengan baik. Hal ini merupakan harapan yang beralasan, selama waktu yang dibutuhkan, waktu yang banyak membuat siswa mulai mengalami kecewa (frustasi) dan motivasi untuk sekolah menurun.

Secara ringkas, dapat dinyatakan bahwa belajar tuntas dapat diimplementasikan dengan praktis pada setiap isi pelajaran. Keterbatasan umum adalah waktu dan kebutuhan dedikasi untuk mengkhususkan hasil-hasil belajar dan untuk mengembangkan ukuran-ukuran acuan kriteria

untuk mengases hasil belajar. Guru dihadapkan dengan tugas memperhalus item-item tes dan TPK-TPK dan mengembangkan bahan-bahan baru sebagai perubahan situasi.

Belajar yang didasarkan tuntas bukan merupakan suatu obat mujarab dalam pendidikan. Hal ini membutuhkan jumlah waktu yang sama dan mungkin lebih besar pada sisi guru dan pembelajar dibandingkan dengan kerja kelas tradisional. Namun demikian, beberapa studi menunjukkan bahwa sikap secara umum tetap positif, tingkat-tingkat hasil belajar meningkat, dan semangat untuk proses belajar tidak menurun. Hal ini merupakan hasil yang kokoh dan terkenal untuk pengalaman dalam proses belajar di sistem-sistem sekolah kita.

6.2 Penilaian Tradisional

Proses mengevaluasi bisa negatif dan bisa juga positif. Fungsinya tergantung pada bagaimana instruktur/guru menggunakannya. Dia dapat menggunakannya untuk mendiagnosis pengajarnya dan hasil belajar siswa atau untuk membandingkan satu siswa dengan yang lain.

Sayangnya, hampir semua guru menggunakan tes untuk mendapatkan rangking dari beberapa siswa untuk diidentifikasi tinggi dan lainnya rendah yang dapat ditentukan melalui skor-skor hasil belajar. Kemudian skor-skor ini digunakan untuk meranking siswa. Semakin meningkat, pendidik menanyakan pada menyukai penggunaan tes untuk tujuan ini dan bahkan akan digunakan untuk meranking. Mereka berpendapat bahwa meranking menimbulkan sistem kompetisi yang terikat pada penurunan harga dan merendahkan diri pada konsep-konsep diri dari banyak siswa. Hal ini benar bahwa kompetisi dapat memberikan motivasi pada beberapa orang; bagaimanapun juga individu akan berkompetisi hanya jika dia tahu bahwa dia mempunyai kesempatan untuk memenangkan. "Permainan urutan," menempatkan satu posisi di atas yang lain, terikat dengan kontribusi pada keyakinan terhadap beberapa siswa yang terjebak dalam suatu sistem yang mereka miliki dari beberapa

kesempatan untuk sukses. Jika ini benar, tidak mengagumkan bahwa surat-surat kabar sering membawakan cerita-cerita anak muda perusak, misalnya mendobrak sekolah dan merusak alat yang bernilai jutaan rupiah. Reaksi umum dari masyarakat adalah bahwa anak-anak ini “jelek” dan harus dihukum. Tetapi apa yang dibuatnya jelek? Mengapa mereka benci dan mogok sekolah? Tentunya ini sifat manusia untuk menghindari bahaya dengan hal-hal yang anda sayangi. Pada kenyataannya bahwa sistem sekolah di Amerika mempunyai persentase *droupout* tinggi. Individu yang droup out sering meninggalkan karena mereka tidak mempunyai perasaan sukses dari lingkungan sekolah.

Guru menyadari masalah-masalah dalam menggunakan evaluasi untuk menentukan hirarki hasil belajar siswa yang telah merencanakan berbagai pendekatan untuk tidak menekankan pada pentingnya tes dan ranking. Beberapa metode ini adalah:

- a. motivasi-diri, dengan cara siswa mengevaluasi dirinya sendiri dan kemudian memutuskan, dalam bekerja sama dengan instruktur/guru, dimana dia butuhkan untuk meningkatkan.
- b. Kriteria performansi, dengan cara guru membuat pedoman bahan-bahan apa yang harus dilengkapi untuk suatu ranking tertentu. Guru juga dapat menggunakan kriteria performansi untuk melihat jika siswa telah mencapai suatu tingkat tertentu.
- c. Evaluasi tertulis, instruktur/guru menulis ringkasan hasil-hasil belajar dan kelemahan-kelemahan siswa, siswa lebih menyukai.
- d. Lulus dan gagal, siswa diberi keputusan lulus atau gagal karena suatu ranking. Sistem ini mengurangi kompetisi tetapi tidak mengenal keunggulan (*excellence*).
- e. Memberikan semua siswa peringkat yang sama. Sistem ini menghilangkan kompetisi tetapi biasanya tidak baik diterima oleh semua siswa, khususnya jika peringkat yang diberikan di bawah A.
- f. Merangking secara rahasia, guru memberi peringkat tetapi tidak menceritakan pada siswa peringkat berapa kecuali dia berada di atas

atau di bawah rata-rata. Kerugian pendekatan ini adalah bahwa siswa dapat mengalami kecemasan di luar perasaan instruktur/guru tentang dia.

- g. Perjanjian siswa-guru, guru dalam bekerja sama dengan siswa, membuat sejumlah kegiatan yang harus dikerjakan untuk menerima peringkat A, B, C, dan seterusnya.

7 KARAKTERISTIK & SPESIFIKASI BUTIR SOAL

Yang dimaksud karakteristik butir soal adalah parameter kuantitatif dari butir soal. Sebaliknya, parameter kualitatif butir soal disebut dengan spesifikasi butir soal. Spesifikasi butir soal ditentukan atas dasar penilaian ahli (*expert judgment*). Kedua hal ini akan diuraikan seperti berikut.

7.1 Karakteristik Butir Soal

Karakteristik butir soal untuk tes hasil belajar dipertimbangkan berdasarkan tingkat kesukaran (p), daya beda (D), dan berfungsi atau tidaknya pilihan.

7.1.1 Tingkat Kesukaran

Tingkat kesukaran butir soal dimaknai sebagai proporsi peserta tes menjawab benar terhadap butir soal tersebut, yang dirumuskan dengan:

$$b = \frac{\sum \text{peserta yang menjawab benar}}{\sum \text{peserta keseluruhan}}$$

Misalnya soal nomor 1, jumlah peserta tes yang menjawab benar 3 orang dan jumlah seluruh peserta tes 15 orang, maka tingkat kesukaran butir soal nomor satu adalah 0,2. Butir soal ini dikatakan sukar. Jika dalam

suatu perangkat soal terdapat 10 butir soal, dengan tingkat kesukaran setiap butirnya berturut-turut 0,2; 0,6; 0,4; 0,5; 0,7; 0,5; 0,35; 0,45, 0,8; dan 1,0, maka tingkat kesukaran perangkat soal ($p_{\text{naskah ujian}}$) tersebut adalah:

$$\begin{aligned} p_{\text{naskah ujian}} &= (0,2 + 0,6 + 0,4 + 0,5 + 0,7 + 0,5 + 0,35 + 0,45 + 0,8 + 1,0)/10 \\ &= 5,5/10 \\ &= 0,55 \end{aligned}$$

atau $p_{\text{naskah ujian}} = \frac{\sum b}{\sum \text{butir soal}}$

Tingkat kesukaran butir soal atau perangkat soal biasanya dikategorikan mudah, sedang, dan sukar. Penentuan ketiga kategori tersebut dapat menggunakan pedoman:

Tingkat kesukaran	Kategori
0,00 - 0,25	Sukar
0,26 - 0,75	Sedang
0,76 - 1,00	Mudah

Contoh perangkat soal tersebut termasuk kategori sedang. Untuk menyusun suatu naskah ujian atau tes hasil belajar sebaiknya digunakan butir soal yang tingkat kesukarannya berimbang, yaitu: sukar (25%), sedang (50%), dan mudah (25%). Komposisi ini dapat diterapkan pada PAN dan PAP. Jika komposisi butir soal tidak seimbang maka penggunaan PAN tidak tepat, sebab informasi kemampuan yang dihasilkan tidak berdistribusi normal. Jadi ukuran butir soal atau perangkat soal yang baik tidak ditentukan oleh tinggi atau rendahnya tingkat kesukaran tetapi pada komposisi tingkat kesukarannya.

7.1.2 Daya Beda

Dalam suatu kelompok peserta tes, biasanya kita jumpai kelompok yang berprestasi tinggi (kelompok atas) dan kelompok yang berprestasi rendah (kelompok bawah). Indeks yang menunjukkan tingkat kemampuan butir soal yang dapat membedakan kelompok atas dan bawah disebut daya beda butir soal. Daya beda biasanya disimbolkan dengan D dan dirumuskan:

$$D = (Ba - Bb)/1/2T$$

Dimana D = daya beda

Ba = Σ kelompok atas yang menjawab benar

Bb = Σ kelompok bawah yang menjawab benar

T = Σ peserta tes (jika jumlah ganjil = T-1)

Jika jumlah peserta banyak, maka kelompok atas dan bawah masing-masing diambil 27%. Untuk memudahkan analisis, apabila jumlah peserta besar maka kelompok dibuat menjadi tiga, yakni: kelompok atas, tengah dan bawah.

Andaikan jumlah seluruh peserta tes 10 orang. Untuk soal nomor X misalnya, kelompok atas yang menjawab benar adalah 4 orang dan kelompok bawah yang menjawab benar hanya satu orang, maka proporsi kelompok atas yang menjawab benar adalah 0,8 dan proporsi kelompok bawah yang menjawab benar adalah 0,2. Jadi daya beda soal nomor X adalah:

$0,8 - 0,2 = 0,6$ atau dapat dihitung dengan rumus:

$$D = (4 - 1)/5 = 0,6$$

Koefisien atau indeks daya beda berkisar antara -1 sampai dengan $+1$. Daya beda berharga $+1$ berarti semua kelompok atas menjawab benar dan semua kelompok bawah menjawab salah terhadap suatu butir soal. Sebaliknya untuk daya beda yang berharga -1 . Harga daya beda yang

dianggap masih memadai untuk sebutir soal adalah $\geq 0,25$. Kurang dari 0,25, butir soal dianggap kurang mampu membedakan peserta tes yang siap menghadapi tes dari peserta yang tidak siap. Jika daya beda negatif, maka butir soal tidak dapat digunakan untuk mengukur hasil belajar siswa. Sehingga butir soal ini harus dibuang atau tidak dihitung dalam penentuan skor mahasiswa. Jadi, makin tinggi daya beda suatu butir soal, makin baik butir soal tersebut, dan sebaliknya makin rendah daya beda makin tidak baik butir soal tersebut.

7.1.3 Berlaku Tidaknya Pilihan

Dalam tipe soal obyektif, khususnya untuk soal pilihan ganda, untuk menentukan berfungsi tidaknya pengecoh suatu butir soal, maka butir soal tersebut perlu dianalisis. Untuk menganalisis setiap butir soal tersebut, lembar jawaban peserta kelompok atas dan bawah dijadikan sebagai sumber informasi. Distribusi dari jawaban kedua kelompok ini untuk setiap butir soal dimasukkan dalam satu tabel seperti berikut:

Andaikan butir soal nomor 12

Tabel : Distribusi jawaban soal nomor 12

Pilihan Kelompok	A*	B	C	D	E
Atas	3	0	1	0	0
Bawah	1	1	1	2	1
Jumlah	4	1	2	2	1

Jawaban yang benar adalah A (tanda *), jumlah peserta yang memilih A adalah banyak, khususnya untuk kelompok atas. Pengecoh B, C, D, dan E ada yang memilih terutama kelompok bawah. Jadi dapat disimpulkan bahwa pengecoh berfungsi sebagai jawaban yang salah. Sehingga semua pilihan pada soal nomor 12 sudah berfungsi.

Andaikan tabel di atas jumlah yang menjawab benar A lebih banyak kelompok bawah, maka pilihan ini harus diperbaiki. Apabila pada pengecoh ditemukan kelompok atas lebih banyak dibandingkan kelompok bawah, pilihan ini juga kurang baik dan perlu diperbaiki. Selain itu, bila pada pengecoh tidak ada satupun yang memilih, maka pilihan ini harus diperbaiki pula.

Dari uraian di atas, cocokkan secara bersama jawaban THB tentang besaran, satuan dan angka penting anda di kelas, kemudian kerjakan bersama tugas berikut di kelas.

Tugas

- a. Tentukan tingkat kesukaran setiap butir soal!
- b. Tentukan daya beda setiap butir soal dan tentukan pula kelayakannya untuk digunakan sebagai THB!
- c. Apakah pada butir-butir soal obyektif ada pilihan-pilihan yang perlu diperbaiki? Sebutkan!

7.2 Spesifikasi Butir Soal

Untuk menganalisis suatu butir soal ada dua spesifikasi yang harus dipertimbangkan, yakni: validitas isi dan keakuratan pengukuran tujuan yang ingin dicapai.

Validitas isi (konten) pelajaran sangat diperlukan untuk menentukan apakah suatu butir soal merupakan alat ukur yang baik untuk suatu hasil belajar tertentu. Analisis validitas isi ini hanya bisa dilakukan oleh seorang yang menguasai bidang studi tersebut dengan baik. Analisis dimulai dengan mengadakan kajian terhadap kisi-kisi soal. Dalam kisi-kisi itu ditentukan bahwa butir soal tertentu dimaksudkan untuk mengukur pokok bahasan atau sub pokok bahasan tertentu. Jadi kisi-kisi soal digunakan sebagai tolok ukur untuk memvalidasi butir soal.

Selain memvalidasi, aspek yang harus dianalisis secara kualitatif oleh seorang ahli bidang studi adalah apakah butir soal yang digunakan apakah mengukur tujuan pendidikan tertentu yang ditetapkan dalam kisi-

kisi. Untuk menganalisis ini perlu penguasaan tentang tujuan pendidikan. Yang perlu diperhatikan bahwa butir soal yang tidak secara akurat mengukur tujuan yang telah ditetapkan akan merupakan butir soal yang sia-sia. Berbahayanya, bila butir soal itu digunakan untuk menentukan keputusan bagi seseorang, hal ini akan berakibat jauh bagi siswa di masa yang akan datang. Di Indonesia, perumusan tujuan pendidikan masih cenderung mengacu pada tujuan pendidikan menurut Bloom dan kawan-kawan.

7.3 Karakteristik Perangkat Tes

Meskipun suatu tes terdiri atas butir-butir soal yang baik, belum tentu akan membuat perangkat tes (soal ujian) menjadi baik. Selain penilaian terhadap setiap butir soal, ada dua hal yang harus diperhatikan dalam menilai soal ujian, yakni: validitas dan reliabilitasnya. Kedua hal ini dapat dijelaskan sebagai berikut.

7.3.1 Reliabilitas Tes

Ketika kita mengamati skor hasil tes siswa, muncul pertanyaan atau keraguan pada diri kita, yaitu apakah skor itu benar-benar menggambarkan tingkat kemampuan siswa kita. Keraguan atau pertanyaan ini sulit dijawab, karena dalam setiap tes selalu akan terdapat unsur kekeliruan (*error*). Kekeliruan ini bisa bersumber pada alat ukurnya atau mungkin faktor yang lain. Untuk melihat apakah perangkat tes itu dapat dipercaya sebagai alat ukur yang dapat menggambarkan kemampuan peserta tes, maka dapat dilihat dari aspek reliabilitasnya. Secara umum, reliabilitas dimaknai sebagai sejauh mana suatu alat ukur dapat diyakini memberikan informasi yang konsisten (ajeg) dan tidak mendua tentang karakteristik peserta tes yang diujikan.

Skor yang diperoleh peserta tes pada dasarnya merupakan skor yang secara langsung berhubungan dengan alat ukur dan kondisi eksternal saat tes berlangsung. Kondisi eksternal tidak dapat didefinisikan

seungguhnya, begitupula alat ukur yang digunakan tidak dapat diketahui sepenuhnya kekuatan dan kelemahannya. Sehingga skor yang diperoleh peserta tes adalah skor yang kemungkinan besar mengandung kekeliruan yang tidak dapat diketahui. Andaikan skor peserta tes itu tidak mengandung unsur kekeliruan, maka skor itu merupakan skor yang sesungguhnya. Tetapi skor sesungguhnya itupun tidak kita ketahui. Untuk itu kita kenal adanya tiga bentuk skor dalam setiap hasil tes, yaitu: skor yang diperoleh (*obtained score*), skor sesungguhnya (*true score*), dan kekeliruan skor (*score error*). Hubungan ketiganya dinyatakan dengan:

Skor yang diperoleh = skor sesungguhnya – kekeliruan

Secara operasional reliabilitas tes didefinisikan sebagai koefisien korelasi antara dua perangkat skor yang dihasilkan oleh perangkat tes yang sama atau paralel yang diadministrasikan kepada sekelompok peserta tes yang sama. Karena reliabilitas merupakan salah satu bentuk khusus korelasi yang menggambarkan keajegan alat ukur (tes), maka ada beberapa prosedur untuk memperoleh koefisien korelasi yang menggambarkan reliabilitasnya. Reliabilitas tes dapat ditinjau dari unsur stabilitas, ekuivalensi dari dua tes yang paralel, dan konsistensi atau homogenitas tes.

Reliabilitas ditinjau dari stabilitas dapat ditentukan dengan mengkorelasikan anatardua skor dari satu tes yang diadministrasikan dua kali kepada kelompok peserta tes yang sama. Selang waktu antara dua pengadministrasian tes harus dekat, *mengapa?*

Reliabilitas dalam arti ekuivalensi dari dua tes yang paralel. Dalam hal ini, kita harus mengkonstruksi dua perangkat tes yang paralel. Kedua perangkat tes diadministrasikan pada kelompok peserta tes yang sama dalam waktu berurutan. Hasil tes dari dua perangkat tes tersebut dikorelasikan.

Reliabilitas dalam arti konsistensi tes merupakan koefisien korelasi yang menunjukkan seberapa jauh suatu perangkat tes homogen, dalam arti mengukur mata pelajaran atau bidang studi yang sama. Untuk menentukan koefisien korelasi ini dapat dilakukan dengan dua cara, yaitu: teknik split-half dan teknik Kuder-Richardson (KR). Teknik split-half dapat dilakukan dengan mengkorelasikan skor setengah pertama dengan skor setengah kedua dari suatu tes. Untuk memperoleh skor setengah pertama dan kedua dapat dilakukan dengan mengkorelasikan skor pada nomor ganjil dan nomor genap. Teknik Kuder-Richardson dikembangkan oleh Kuder dan Richardson, dengan rumus ke-20nya (KR-20):

$$KR-20 = \frac{n}{n-1} \left[\frac{SB^2 - \sum pq}{SD^2} \right]$$

dimana $n = \sum$ butir soal

SB = simpangan baku skor-skor tes

p = tingkat kesukaran tes (perangkat tes)

q = 1 - p

Setelah anda mempelajari reliabilitas tes di atas, diskusikan dan kerjakan secara kelompok tugas 10 berikut ini.

Tugas

- a. Apakah penskoran tes yang kurang obyektif berpengaruh terhadap reliabilitas tes? Jelaskan.
- b. Apakah peserta tes yang bervariasi berpengaruh terhadap reliabilitas tes? Jelaskan.
- c. Apakah jumlah butir soal dalam perangkat tes berpengaruh terhadap reliabilitas tes? Jelaskan.

7.3.2 Validitas Tes

Seperti halnya pada butir soal, perangkat tes yang baik juga harus memenuhi kriteria valid (tepat). Validitas tes didefinisikan sebagai

seberapa jauh perangkat tes itu berguna dalam mengambil keputusan yang relevan dengan tujuan yang telah ditentukan. Untuk tes hasil belajar, aspek validitas yang paling penting adalah validitas isi (*content validity*), yaitu: ukuran yang menunjukkan sejauh mana skor dalam tes berasosiasi dengan penguasaan peserta dalam bidang studi yang diuji melalui perangkat tes tersebut. Validitas isi ini ditentukan oleh ahli yang menguasai bidang studi tersebut. Jadi untuk validitas ini analisisnya lebih bersifat kualitatif. Oleh karena itu, yang bisa menganalisis tes harus orang mempunyai latar belakang bidang studi yang baik.

Selain validitas isi, juga kita kenal jenis validitas tes yang lain, yaitu: validitas prediktif, validitas serempak, dan validitas konstruk. Ketiga jenis validitas tersebut tidak dibahas di sini karena ketekaitannya dengan keperluan terhadap penilaian perangkat tes hasil belajar tidak terlalu kuat.

Setelah anda mempelajari validitas tes, diskusikan secara kelompok tugas 11 berikut.

Tugas

Apakah reliabilitas tes berpengaruh terhadap validitas tes? Jelaskan.

DAFTAR PUSTAKA

- Bloom B. S., Madaus G. F., dan Hastings, *Evaluation to Improve Learning*, McGraw-Hill Book Company, New York.
- Djamarah S. B. & Zain A., 1995, *Strategi Belajar Mengajar*, Rineka Cipta, Jakarta.
- Farmer W. A., Farrel M. A., (1980), *Systematic Instruction in Science for The Middle and High School Years*, Addison-Wesley Publishing Company, Inc., Sydney.
- Giancoli D., 1995, *Physics*, Fourth edition, Prentice-Hall International, Inc., Englewood Clifs, New Jersey.
- Kertiasa N, 1993, Fisika 1 untuk Sekolah Menengah Umum Kelas 1, Depdikbud., Jakarta.
- Permendiknas RI No. 15 tahun 2005 .tentang Standar Pendidikan Nasional.
- Permendiknas RI No. 41 tahun 2007.tentang Standar Proses untuk Satuan Pendidikan Dasar dan Menengah.
- Sadler, R. 1998. Formative Assessment: Revisiting the Territory'. *Assessment in Education: Principles, Policy and Practice*, 5, pp. 77-84.

Trowbridge L. W., Bybee R. W., 1990, *Becoming a Secondary School Science Teacher*, Merrill Publishing Company, Columbus.

Zainul A., Nasoetion N., 1996, *Program Pengembangan Keterampilan Teknik Instruksional (Pekerti): Penilaian Hasil Belajar*, Depdikbud, Jakarta.

