

PENILAIAN DALAM PEMBELAJARAN BAHASA

PENILAIAN DALAM PEMBELAJARAN BAHASA

Ujang Suparman

Suluh Media

The logo for Suluh Media features the text "Suluh Media" in a bold, sans-serif font. A curved line or swoosh starts under the "S" and extends to the right, ending under the "a".

Penilaian dalam Pembelajaran Bahasa

oleh Ujang Suparman

Hak Cipta © 2016 pada penulis



SULUH MEDIA

Jl. Dr. Setiabudi No. 71C
Pamulang Timur - Ciputat
Tangerang
Telp. 021-7412412 Fax. 021-7494272

Hak Cipta dilindungi undang-undang. Dilarang memperbanyak atau memindahkan sebagian atau seluruh isi buku ini dalam bentuk apa pun, secara elektronis maupun mekanis, termasuk memfotokopi, merekam, atau dengan teknik perekaman lainnya, tanpa izin tertulis dari penerbit.

Tajuk Entri Utama: Suparman, Ujang

Penilaian dalam Pembelajaran Bahasa/Ujang Suparman

- Edisi Pertama. Cet. Ke-1. - Yogyakarta: Suluh Akademi, 2016
x + 140 hlm.; 25 cm

Bibliografi.:

ISBN :
E-ISBN :

1.

II. Judul

....

Semua informasi tentang buku ini, silahkan scan QR Code di cover belakang buku ini



KATA PENGANTAR

Alhamdulillah penulisan buku yang berjudul Evaluasi Pembelajaran Dalam Pendidikan Bahasa bisa diselesaikan. Buku ini diharapkan laksana setetes embun di Padang Pasir, walaupun sangat kecil namun tetap bermakna dan bermanfaat. Buku ini dimaksudkan untuk memenuhi kekurangan sumber untuk bahan ajar Mata Kuliah Penilaian pada Pendidikan Bahasa.

Buku ini bisa dipakai oleh mahasiswa pada Jurusan Pendidikan Bahasa di Universitas, atau Program Studi Pendidikan Bahasa yang biasanya berada di bawah naungan Fakultas Keguruan dan Ilmu Pendidikan (FKIP) atau di Sekolah Tinggi Keguruan Dan Ilmu Pendidikan (STKIP), mahasiswa di Fakultas Tarbiyah IAIN atau STAIN, bisa juga digunakan oleh guru-guru Sekolah Dasar, Sekolah Menengah Pertama, Sekolah Menengah Atas, Sekolah Menengah Kejuruan, guru-guru di Madrasah Tsanawiyah, Madrasah Aliyah, dan tentu saja oleh para peneliti di bidang evaluasi bahasa.

Penulis mengucapkan terima kasih kepada Dekan FKIP Unila yang telah memfasilitasi terbitnya buku ini, dan ucapan terima kasih juga disampaikan kepada penerbit Graha Ilmu.

Akhirnya semoga buku ini menjadi bermanfaat untuk meningkatkan kualitas pendidikan di Tanah Air tercinta.

Penulis

APPROVAL



DAFTAR ISI

KATA PENGANTAR	v
DAFTAR ISI vii	
BAB 1 PERENCANAAN PENILAIAN	1
1.1 Taksonomi Sebagai Pengarah	2
1.2 Peranan Tujuan Pembelajaran	7
1.3 Petunjuk dalam Menyatakan Tujuan Sampel	9
1.4 Mengevaluasi Daftar Tujuan yang Final	11
1.5 Menyiapkan Penilaian	12
1.6 Jenis Prosedur Penilaian	14
1.7 Validitas dan Reliabilitas dalam Rencana Penilaian	16
1.8 Rangkuman	21
BAB 2 HAKIKAT PENILAIAN	23
2.1 Jenis-Jenis Metode Penilaian	24
2.2 <i>Realisme Tugas</i>	26
2.3 <i>Kompleksitas Tugas</i>	27
2.4 Waktu Penilaian yang Diperlukan	27
2.5 Penilaian dalam Penyekoran	28
2.6 Petunjuk Untuk Penilaian Siswa yang Efektif	29
2.7 Penilaian Acuan Norma (<i>Norm-Referenced</i>) dan Penilaian Acuan Kriteria (<i>Criterion-Referenced</i>)	35
2.8 Rangkuman	38

BAB 3	MERANCANG TES ESAI	41
3.1	Bentuk Tes	41
3.2	Cara Mengkonstruksi Tes Esai	43
BAB 4	ALAT UKUR NONTES	53
4.1	Alat Ukur Evaluasi Diri	54
4.1.1	Model <i>Scoring</i>	54
4.1.2	Model <i>Ranking</i>	55
4.2	Alat Ukur <i>Rating</i>	56
4.2.1	Daftar Cek (<i>Check List</i>)	57
4.2.2	Skala <i>Rating</i> (SR)	61
4.3	Kartu-Kartu Skor (kks)	64
4.3.1	Kelebihan dan Kelemahan Kartu Skor	65
4.4	Bentuk-Bentuk Laporan	67
4.4.1	Laporan proyek	68
4.4.2	Laporan Kegiatan dan Buku harian atau log	69
4.4.3	Evaluasi Menggunakan Wawancara	71
4.4.4	Kelebihan dan Kelemahan Teknik Wawancara	73
4.4.5	Evaluasi dengan Kuesioner	74
4.4.5	Evaluasi dengan Portofolio	76
BAB 5	PENILAIAN ALTERNATIF (<i>ALTERNATIVE ASSESMENT</i>)	79
5.1	Mengapa Penilaian Alternatif?	79
5.2	Apa Pengertian Penilaian Alternatif?	80
5.3	Landasan Rasional	82
5.4	Karakteristik Asessmen Alternatif	85
5.5	Pro Dan Kontra Terhadap Penilaian Alternatif	87
5.5.1	Dasar Pemikiran bagi yang pro	88
5.5.2	Argumentasi bagi yang Kontra	91
5.6	Teknik-teknik yang digunakan dalam penilaian alternatif	93
5.6.1	Portofolio	94
5.6.2	Rubrik	101
5.6.3	Jurnal belajar (<i>learning journal</i>) atau jurnal siswa (<i>student journal</i>)	104
5.6.4	Catatan ringkas tentang belajar (<i>learning log</i>)	105
5.6.5	Penilaian diri (<i>self assessment</i>)	105

5.6.6	Peta Konsep (<i>Concept of Map</i>)	106
5.6.7	Diskusi Kelompok Terfokus (<i>Focused Grup Discussion</i>)	108
PENILAIAN AUTENTIK (<i>AUTHENTIC ASSESSMENT</i>)		109
6.1	Pengertian Penilaian Autentik	109
6.2	Ciri-Ciri Penilaian Autentik	110
6.3	Teknik-Teknik yang Digunakan dalam Penilaian Autentik	112
BAB 8 PENENTUAN NILAI ATAU GRADE		115
7.1	Makna <i>Grade</i> Hasil Belajar	117
7.2	Mempertimbangkan Perbedaan Individual	119
7.3	Macam-Macam Sistem <i>Grade</i>	120
7.4	Menentukan Sistem Penilaian Terakhir	123
7.5	Menyelenggarakan Evaluasi Hasil Belajar di Kelas	124
BAB 8 PENILAIAN PERFORMA (<i>PERFORMANCE ASSESSMENT</i>)		127
8.1	Pengertian Penilaian Performa (Performa)	127
8.2	Karakteristik Penilaian Performa	128
8.3	Tes Performa	129
8.4.	Jenis-jenis Tes Performa	131
8.5	Inventori Performa	135
DAFTAR PUSTAKA		137



Bab 1

PERENCANAAN PENILAIAN

Tujuan bab ini adalah agar mahasiswa mampu:

1. Menguraikan peranan tujuan pembelajaran dalam kaitannya dengan penilaian prestasi siswa.
2. Menjelaskan Taksonomi Tujuan Pendidikan yang Telah Direvisi dan manfaatnya dalam perencanaan penilaian.
3. Menjelaskan bagaimana tujuan pengajaran harus dirumuskan untuk kepentingan penilaian.
4. Menguraikan pentingnya membuat rencana penilain sebagai bagian dari rencana pengajaran.
5. Menjelaskan berbagai jenis metode peniaian dan penggunaannya dalam mengevaluasi prestasi siswa.
6. Menjelaskan alasan pentingnya mahasiswa memhami jenis-jenis item tes dan instrumen penilaian yang akan digunakan.
7. Menjelaskan maksud validitas, reliabilitas dan peranannya dalam menyusun prosedur penilaian

Perencanaan penilain prestasi siswa meliputi tiga isu penting:

1. Siswa diharapkan menguasai apa?
2. Jenis penampilan siswa yang bagaimana yang akan diterima sebagai indikator bahwa mereka sudah belajar?
3. Metode penilaian yang manakah yang bisa digunakan untuk mengevaluasi penampilan siswa secara paling baik?

Pertanyaan pertama berkenaan dengan tujuan kurikulum yang telah disusun oleh sekolah dan dipengaruhi oleh kebijakan pemerintah Pusat maupun Daerah. Sebagian besar sekolah dan Pemerintah Daerah memiliki *standar isi* yang merupakan garis-garis besar untuk menentukan apa yang harus diketahui, dipelajari, dikuasai dan bisa dilakukan oleh siswa. Hal ini telah mempengaruhi pemerintah daerah dan sekolah untuk memodifikasi kurikulum sedemikian rupa dan lebih menekankan pada hasil yang lebih kompleks.

Pertanyaan kedua membahas pentingnya kejelasan tujuan pengajaran dalam istilah unjuk kerja dan mencakup semua jenis hasil pembelajaran yang diharapkan. Tujuan-tujuan tersebut harus selaras dengan tujuan kurikulum dan akan merefleksikan standar isi kurikulum sekolah yang bersangkutan. Dalam analisis terakhir, penilaian prestasi siswa harus didasarkan pada keterampilan yang harus dipahami dan dikuasai oleh siswa dan pada akhirnya hal ini akan ditentukan oleh kurikulum sekolah. Tujuan pembelajaran merupakan alat untuk menyatakan tujuan kurikulum dalam istilah yang lebih spesifik sehingga rumusan tujuan tersebut akan bermanfaat baik dalam pembelajaran maupun dalam penilaian.

Pertanyaan ketiga berkaitan dengan kesesuaian antara tujuan pembelajaran dan prosedur penilaian. Tujuan menjelaskan hasil pembelajaran dalam istilah unjuk kerja dan metode penilaian digunakan untuk menentukan tingkat penampilan siswa yang memuaskan. Untuk hasil unjuk kerja tertentu, sebuah tes akan memberikan bukti yang cukup. Dalam kasus lainnya, jenis teknis observasi tertentu akan diperlukan untuk mengevaluasi hakikat unjuk kerja yang tengah berlangsung, misalnya unjuk kerja kemampuan berpidato, atau untuk mengevaluasi produk yang dihasilkan, misalnya sebuah laporan tertulis. Kunci penilaian yang efektif terletak pada penggunaan metode yang paling langsung dan relevan.

1.1 TAKSONOMI SEBAGAI PENGARAH

Revisi Taksonomi Bloom tentang Tujuan Pendidikan oleh Anderson dan Krathwohl (2001) memberikan kerangka untuk mengidentifikasi dan menyiapkan tiga hal berikut ini: (a) tujuan pembelajaran, (b) kegiatan pembelajaran, serta (c) metode penilaian.

Taksonomi yang telah direvisi memiliki dua dimensi: (a) dimensi pertama mengandung enam kategori proses kognitif yang bermula dari yang paling mudah sampai dengan yang paling sulit; dan (b) dimensi kedua memuat empat kategori pengetahuan yang berkisar dari yang paling konkrit sampai dengan yang paling abstrak.

Kerangka taksonomi tersebut bisa dituangkan dalam sebuah table berikut ini:

Tabel 1.1 Tabel dua dimensi yang menggambarkan Taksonomi Kognitif Bloom yang direvisi

Dimensi Pengetahuan (<i>The Knowledge Dimension</i>)	Dimensi Proses Kognitif					
	1. Mengingat (<i>remember</i>)	2. Memahami (<i>Understand</i>)	3. Menerapkan (<i>Apply</i>)	4. Menganalisa (<i>Analyze</i>)	5. Mengevaluasi (<i>Evaluate</i>)	6. Menciptakan (<i>Create</i>)
Pengetahuan faktual (<i>Factual knowledge</i>)						
Pengetahuan konseptual (<i>Conceptual knowledge</i>)						
C. Pengetahuan prosedural (<i>Procedural knowledge</i>)						
Pengetahuan Metakognitif (<i>Metacognitive knowledge</i>)						

Sumber: Diadaptasi dari Loren . Anderson & David R. Krathwohl, eds., *A taxonomy for Learning Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives* © 2001. Published by Allyn and Bacon, Boston, M.A., Copyright © 2001 by Pearson Education.

Setiap sel, di mana dua dimensi bertautan, mengelompokkan jenis pembelajaran yang terjadi. Misalnya, Sel 1A menunjukkan jenis pembelajaran *Remembering Factual Knowledge*; Sel 2B menunjukkan jenis pembelajaran *Understanding Conceptual Knowledge*; dan Sel 3C menunjukkan jenis pembelajaran *Applying Procedural Knowledge*. Masing-masing dimensi proses kognitif dan dimensi pengetahuan didefinisikan dan diberikan contoh pada lampiran.

Jenis pengetahuan metakognitif sering terabaikan meskipun merupakan salah satu jenis pengetahuan yang penting. Pengetahuan metakognitif berhubungan dengan strategi umum dan khusus yang digunakan dalam pembelajaran, pengetahuan tentang tugas-tugas kognitif, pengetahuan tentang kekuatan serta kelemahan seseorang yang disadarinya dan strategi yang seseorang gunakan dalam menyelesaikan tugas-tugas kognitifnya. Penggunaan portofolio sangatlah bermanfaat untuk mengevaluasi pengetahuan metacognitif karena penggunaan portofolio mencakup contoh-contoh karya siswa yang terbaik, hasil penilaian mandiri dan renungan siswa sendiri tentang butir-butir portofolio. Di samping itu, penggunaan portofolio memberikan fakta tentang perubahan-perubahan penilaian dalam pola pikir siswa setiap saat.

Tabel taksonomi di atas bisa membantu guru dalam memperluas daftar tujuan pembelajaran yang sedang disiapkan. Bila tidak menggunakan Tabel Taksonomi tersebut, maka *mengingat materia faktual* terlalu sering digunakan. Oleh karena itu, penggunaan Table Taksonomi bisa membantu guru beranjak dari proses kognitif yang paling rendah ini kepada penggunaan pembelajaran yang lebih bermakna yang tercakup dalam proses kognitif yang lebih tinggi.

Ada beberapa langkah yang perlu guru lakukan untuk menggunakan Tabel Taksonomi tersebut. Pertama, perhatikan kategori-kategori berikut ini: *Mengingat*, *Memahami* dan *Menerapkan*. Sebagian besar standar isi meliputi tujuan-tujuan pada dimensi ini (*Mengingat*, *Memahami* dan *Menerapkan*). Sebenarnya, banyak hakikat standar isi menggunakan *Pengetahuan (Mengingat)*, *Memahami*, dan *Aplikasi* sebagai kategori utama untuk mengelompokkan standar. Sebaiknya sejak awal, Anda sebagai guru mempertimbangkan *Menganalisa*, *Mengevaluasi*, dan *Menciptakan*. Ketika menganalisis bahan ajar dan masalah pada suatu mata pelajaran, apakah siswa perlu membedakan antara mana yang relevan dan mana yang tidak relevan; mana yang tepat dan mana yang tidak tepat; sebab dan akibat; fakta dan opini; dsb? Jika demikian, maka *Menganalisa* harus dipertimbangkan sebagai salah satu tujuan pembelajaran. Sama seperti itu juga, apakah siswa Anda harus membuat penilaian tentang ketepatan suatu prosedur, tentang kualitas sebuah laporan tertulis, ketepatan suatu keterampilan, dsb.? Apakah Anda mengharapkan siswa Anda melakukan sebuah penelitian, sebuah laporan tertulis,

menyiapkan subah bagan, grafik, poster atau sebuah model? Jika demikian, maka tujuan pembelajaran dalam bidang *Mengevaluasi*, dan *Menciptakan* sangatlah diperlukan. Dimensi *Pengetahuan* akan membantu Anda dalam menyempurnakan tujuan pembelajaran. Dengan memasukkan capaian pembelajaran yang lebih kompleks dalam daftar tujuan pembelajaran, Anda bukan saja sekedar memberikan kontribusi untuk mentransfer pembelajaran, tetapi juga memperkuat ingatan tentang fakta, konsep, prosedur, dan strategi.

Selain itu, Taksonomi bisa juga membantu guru memilih aktivitas yang paling bisa digunakan untuk mencapai tujuan-tujuan pembelajaran. Taksonomi bisa digunakan untuk mengalihkan fokus Anda dari berbagai aktivitas ke hasil akhir tujuan pembelajaran. Perhatian utama dari seorang guru jangan ditujukan pada kegiatan apa yang siswa lakukan tetapi pada apa yang siswa bisa pelajari dari aktivitas itu. Falsafah kuno *learning by doing* 'belajar sambil bekerja' sering gagal menjelmakan harapan karena terlalu banyak penekanan pada *bekerja* dan tidak cukup banyak penekanan pada hasil akhir pembelajaran yang dimaksudkan.

Taksonomi juga berguna untuk mempersiapkan prosedur dan instrumen penilaian yang selaras dengan tujuan dan aktivitas pembelajaran. Tujuan pembelajaran menjelaskan hasil akhir pembelajaran, aktivitas pembelajaran dirancang untuk menyediakan proses pembelajaran yang relevan, dan instrumen pembelajaran memberikan bukti sampai tingkat mana tujuan pembelajaran telah tercapai. Agar berfungsi secara efektif, ketiga unsur tersebut harus sejalan. Sedangkan Tabel Taksonomi menyediakan dasar untuk mengecek keselarasan antara tujuan pembelajaran, aktivitas pembelajaran dan penilaian.

Perlu dipastikan bahwa dalam merencanakan penilaian, penilaian yang dirancang untuk mengukur hasil pembelajaran di luar kategori *mengingat* harus mengandung kebaruan. Misalnya, dalam menilai *pemahaman* siswa mungkin bisa diminta untuk mendefinisikan istilah-istilah dengan menggunakan kata-katanya sendiri, atau diminta untuk memberikan contoh-contoh yang bukan diambil dari buku, sehingga jawabannya itu menunjukkan bahwa mereka memahami maksud dari konsep-konsep yang diajarkan, bukan mengingat-ingat apa yang sudah dipejarinya. Demikian pula untuk *mengaplikasikan*, *menganalisis*, *mengevaluasi* dan *menciptakan*, penilaian itu harus menuntut tanggapan terhadap materi baru atau

masalah baru. Proses kognitif tingkat lebih tinggi ini berkenaan dengan belajar mentransfer bukan belajar mengingat yang termasuk dalam kategori *mengingat*.

Soal bisa lebih mudah dihubungkan dengan Tabel Taksonomi dari penilaian unjuk kerja. Sebuah butir soal biasanya berada pada satu atau dua sel dalam Tabel Taksonomi. Misalnya, sebuah butir soal yang mengukur *penerapan* bisa mengukur bukan hanya kemampuan *mengingat* tentang pengetahuan prosedural akan tetapi juga *menerapkan* pengetahuan konseptual. Akan tetapi penilaian unjuk kerja bisa menyebar di seluruh Table Taksonomi tersebut. Misalnya, sebuah penelitian kecil bisa menuntut siswa untuk memilih dan menjelaskan sebuah masalah untuk dikaji, membaca literatur yang relevan dan sumber-sumber lain untuk menjawab masalah tersebut, menganalisis temuan, dan menulis laporan yang meringkas informasi, menyiapkan sebuah model, poster, menyajikan temuan-temuannya di depan kelas, dan merefleksikan apa yang sudah dielajarinya dari penelitiannya. Meskipun penelitian itu sederhana, namun bisa meliputi semua prosedur kognitif dan semua jenis pengetahuan. Tabel Taksonomi itu juga bermanfaat untuk menyiapkan penelitian kecil dengan cara mengarahkan perhatian pada jenis-jenis proses dan hasil akhir pengetahuan yang diharapkan. Juga bermanfaat dalam menyiapkan kriteria untuk mengevaluasi berbagai fase dari penelitian tersebut. Selain itu Tabel Taksonomi tersebut bisa digunakan bersama oleh guru dan siswa pada awal suatu tugas atau kegiatan, sehingga bukan hanya guru tetapi siswa juga bisa lebih memfokuskan perhatian pada hasil yang diharapkan dan bisa menghindarkan diri dari kegiatan yang kurang berguna.

Taksonomi psikomotor (Bott, 1996) sangat bermanfaat untuk membuat keputusan pada saat merencanakan penilaian unjuk kerja. Taksonomi tersebut terdiri dari empat tataran: (1) observasi, (2) imitasi, (3) latihan dan (4) adaptasi. Pada tataran yang paling rendah, *observasi*, siswa hanya diminta untuk mengobserasi gurunya yang tengah menunjukkan cara penyelesaian suatu tugas, dan melaporkan apa yang sudah guru lakukan. Pada tataran kedua, *imitasi/menyontoh*, siswa mampu melakukan imitasi atau meniru setelah melihat orang lain menyelesaikan suatu tugas serta mengikuti instruksi lisan atau instruksi tertulis. Pada tataran ketiga, *latihan*, siswa bisa melakukan tugas secara mandiri, tanpa bantuan. Sedangkan pada tataran

terakhir, *adaptasi*, siswa sudah bisa mengkoordinasikan dan menyelesaikan berbagai tugas dengan menggunakan keterampilan yang sudah dipelajari sebelumnya untuk menyelesaikan tugas-tugas baru secara mandiri, tanpa bantuan. Oleh karena itu, taksonomi psikomotor bisa digunakan sebagai pemandu untuk membantu guru dalam membuat dasar-dasar keputusan tentang tingkat performa siswa yang akan dianggap memadai dan berterima. Misalnya, jika seorang siswa diminta untuk mengungkapkan ucapan terima kasih atau diminta untuk menuliskan alasan seseorang temannya yang tidak hadir, guru harus memutuskan apakah siswa itu akan diijinkan untuk menunjukkan kemampuannya dengan bantuan intruksi tertulis (yakni *imitasi/menyontoh*) atau apakah siswa itu harus menyelesaikan tugasnya tanpa bantuan (yakni *latihan*).

Ringkasnya, taksonomi berdimensi ganda yang dibahas pada bagian ini memiliki empat manfaat:

- (1) Mengidentifikasi berbagai hasil belajar yang diharapkan (yakni tujuan pembelajaran) di luar kategori *mengingat*.
- (2) Merencanakan kegiatan pembelajaran yang memberikan sumbangan terhadap berbagai jenis pembelajaran bermakna yang relevan dengan tujuan pembelajaran.
- (3) Merencanakan metode penilaian yang sesuai dengan tujuan dan aktivitas pembelajaran. Dan
- (4) Mengecek keselarasan antara tujuan pembelajaran, kegiatan pembelajaran dan metode penilaian.

Jika tidak terdapat keselarasan dari ketiga unsur tersebut, maka penilaian yang valid tidak akan pernah bisa diperoleh.

1.2 PERANAN TUJUAN PEMBELAJARAN

Tujuan pembelajaran yang dirumuskan dengan baik memberikan penjelasan tentang hasil pembelajaran yang diharapkan, yakni, dalam arti jenis penampilan yang siswa bisa tunjukkan untuk membuktikan bahwa mereka sudah mencapai pengetahuan, pemahaman, atau keterampilan yang dijelaskan oleh tujuan. Dengan menjelaskan performa yang bisa diterima sebagai bukti sebuah prestasi, maka berarti kita sudah memberikan perhatian khusus pada pembelajaran, aktivitas belajar dan penilaian. Keselarasan dari ketiga unsur tersebut terpelihara oleh adanya tujuan pembelajaran. Misalnya,

untuk sebuah tujuan pembelajaran yang menekankan strategi pemecahan masalah (*problem solving*), maka siswa berlatih untuk menerapkan strategi tersebut, dan kita menilai keterampilan mereka dalam menerapkan strategi itu dengan masalah baru. Jadi tujuan pembelajaran, strategi belajar dan penilaian terhadap target merupakan tiga unsur yang sama bagaikan dua muka pada mata uang yang sama. Apabila siswa diarahkan untuk memahami tujuan pembelajaran sejak awal, maka baik guru maupun siswa akan bekerja ke arah tercapainya tujuan yang sama, di samping itu pembelajaran dan penilaian merupakan dua bagian dari proses yang sama.

Dalam merumuskan tujuan pembelajaran, sangatlah penting untuk dicamkan bahwa kita tidak menjelaskan prosedur pengajaran, aktivitas pembelajaran, atau proses belajar. Namun kita hanya menjelaskan performa siswa untuk ditampilkan pada akhir pengalaman belajar sebagai bukti bahwa dia telah belajar. Hal ini akan memungkinkan kita untuk menggunakan berbagai prosedur, materi ajar, dan aktivitas belajar untuk mencapai tujuan ahir yang dimaksudkan.

Prosedur yang bermanfaat adalah merumuskan tujuan umum yang difokuskan pada hasil belajar yang lebih luas kemudian mendefinisikan masing-masing tujuan umum ke dalam tujuan yang lebih khusus, sbb:

1. Memahami konsep-konsep ilmiah
 - 1.1 Menjelaskan konsep dengan menggunakan kata-katanya sendiri.
 - 1.2 Memberikan contoh tentang konsep yang baru.
 - 1.3 Membedakan konsep dari konsep ilmiah yang mirip.
 - 1.4 Menjelaskan peranan konsep dalam memahami even ilmiah.
 - 1.5 Menggunakan konsep dalam pemecahan masalah.

Hasil akhir belajar yang spesifik yang diurutkan di bawah tujuan umum tidak dimaksudkan sebagai tujuan yang lengkap, tetapi harus mengandung contoh-contoh yang representatif dari jenis performa yang akan kita terima sebagai bukti bahwa tujuan pembelajaran telah tercapai. Hasil akhir belajar yang bersifat khusus akan bervariasi juga dengan bidang prestasi. Misalnya, untuk membuat rincian hasil belajar yang bersifat khusus untuk memahami konsep membaca, maka penekanan akan diberikan lebih banyak pada menginterpretasikan, mengklasifikasikan dan pemecahan masalah. Contoh yang diberikan pada bagian ini berfungsi sebagai model, yang akan menjadi

lebih berorientasi pada isi ketika merencanakan sebuah unit pembelajaran. Dalam hal ini, tujuanyang bersifat umum dan hasil belajar yang bersifat khusus akan merujuk pada konsep yang tengah dipelajari. Dan hal tersebut akan mengubah seperrangkat hasil belajar yang bersifat khusus. Apa yang dikemukakan di sini merupakan prosedur yang diharapkan berguna dalam menyiapkan tujuan pembelajaran. Prosedur tersebut adalah sebagai mana bisa dilihat pada bagan berikut ini.

Bagan 1.1 Merumuskan Tujuan Pembelajaran

Merumuskan Tujuan Pembelajaran

Langkah berikut ini memberikan garis-garis besar untuk merumuskan tujuan pembelajaran yang bermanfaat dalam pembelajaran dan penilaian.

1. Rumuskan tujuan umum sebagai berikut:
 - 1.1 Tuliskan setiap tujuan sebagai hasil belajar yang dimaksudkan.
 - 1.2 Rumuskan setiap tujuan dengan menggunakan istilah performa.
 - 1.3 Mulailah masing-masing tujuan itu dengan kata kerjam(misalnya: “mengingat,” “memahami”).
 - 1.4 Tuliskan masing-masing tujuan secara umum agar bisa mengidentifikasi sebuah domain (bidang) dari beberapa hasil akhir yang bersifat khusus.
2. Urutkan dan rumuskan hasil akhir belajar yang bersifat khusus sebagai berikut:
 - 2.1 Jelaskan masing-masing tujuan umum dengan mengemukakan contoh hasil akhir belajar yang bersifat khusus yang representatif.
 - 2.2 Mulailah masing-masing rumusan dengan kata kerja yang menunjukkan performa siswa yang bisa tampak (seperti “pilihlah,” “jelaskan,”).
 - 2.3 Masukkan sejumlah hasil akhir belajar yang cukup banyak untuk menunjukkan capaian dari tujuan umum.
 - 2.4 Pastikan bahwa hasil akhir belajar yang berifat khusus itu relevan dengan tujuan umum dan pastikan pula bahwa tujuan umum itu selaras dengan tujaun sekolah.

1.3 PETUNJUK DALAM MENYATAKAN TUJUAN SAMPEL

Pada saat memulai identifikasi dan pemilihan tujuun pembelajaran, sangatlah penting untuk mengkaji dulu berbagai sumber ide. Berikut ini disajikan beberapa contoh sumber:

1. Rumuskan Standar Isi.

Standar isi memberikan pernyataan umum tentang hasil akhir belajar yang dimaksudkan yang diharapkan dikuasai oleh semua siswa. Rumusan standar isi itu biasanya terlalu umum untuk dijadikan sebagai tujuan pembelajaran, namun dalam program yang berdasar standar guru diharapkan merumuskan tujuan pembelajarannya berdasarkan pada standar isi. Jika ungkapan standar isi itu hanya bersifat saran, maka standar isi itu bisa ditinjau ulang untuk mempertimbangkan hasil akhir belajar, untuk dilengkapi dengan sumber-sumber lain). Tingkat di mana standar isi itu direfleksikan dalam kurikulum sekolah akan menentukan tingkat dimana standar isi itu harus direfleksikan dalam daftar tujuan pembelajaran.

2. Taksonomi Tujuan Pembelajaran yang Direvisi.

Tujuan pendidikan yang dijelaskan dalam Taksonomi itu berada dalam level menengah antara tujuan kurikulum dan tujuan pembelajaran. Tetapi seperti halnya standar isi, tujuan pendidikan tersebut menyediakan bidang penting dari hasil akhir belajar yang diharapkan untuk dipertimbangkan. Sub-kategori juga penting dalam merumuskan tujuan pembelajaran dengan cara menyarankan jenis performa yang merepresentasikan kategori utama. Misalnya, *pemahaman* bisa ditunjukkan dengan *menginterpretasikan, memberikan contoh, mengklasifikasikan, meringkas, menarik kesimpulan, membandingkan dan menjelaskan*. Nama dan contoh lain untuk sub-kategori ini bisa menjadi bantuan lebih lanjut dalam menyiapkan tujuan pembelajaran yang tepat. Mereview semua kategori dalam tabel (lihat lampiran) sangat berguna untuk memilih tujuan yang mencakup semua hasil akhir belajar yang penting dan untuk merumuskannya dengan bahasa yang jelas dan konsisten.

3. Publikasi Organisasi Pendidikan.

Ada berbagai organisasi dalam masing-masing bidang ajaran utama yang mempublikasikan materi yang bermanfaat tentang pengajaran yang mengandung tujuan-tujuan yang disarankan. Misalnya, *the National Council of Teachers of English, the National Science Teachers Association, dan the National Council for the Social Studies*, yang mempublikasikan buku tahunan secara periodik and publikasi lainnya tentang pengajaran dalam bidangnya masing-masing. Berbagai organisasi ini serta

organisasi lainnya telah menerbitkan berbagai perangkat standar yang mungkin bermanfaat dalam menyiapkan tujuan-tujuan pembelajaran, (lihat Bagan 1.2).

Bagan 1.2 Contoh Berbagai Standar oleh Organisasi Profesional

National Council of Teachers of English and the International Reading Association. 1996. *Standards for the English Language Arts*. Urbana, IL: National Council of Teachers of English.

National Council of Teachers of Mathematics. 1989. *Curriculum and Evaluation Standards for School Mathematics*. Reston, VA: National Council of Teachers of Mathematics.

National Council for Social Studies. 1994. *Expectations of Excellence; Curriculum Standards for Social Studies*. Washington, DC: National Council for Social Studies.

National Research Council. 1996. *National Science Education Standards*. Washington, DC: National Academy Press.

4. Buku Petunjuk Guru yang Melengkapi Buku Teks Siswa.

Buku-buku petunjuk semacam ini biasanya mengandung berbagai tujuan, tetapi sering kali difokuskan pada hasil akhir belajar tingkat yang lebih rendah, dirumuskan secara kurang baik, dan terlalu spesifik. Namun buku-buku petunjuk tersebut sangat baik untuk dikaji guna memperoleh hasil akhir belajar yang bisa dipertimbangkan. Mungkin juga untuk memodifikasi atau mengkombinasikannya untuk diubah menjadi sebuah bentuk yang lebih bisa digunakan.

Perlu dicamkan bahwa dalam memilih dan menyiapkan tujuan pembelajaran, semua tujuan pembelajaran harus selaras dengan tujuan program sekolah setempat.

1.4 MENGEVALUASI DAFTAR TUJUAN YANG FINAL

Ketika mengevaluasi daftar tujuan yang final dari suatu mata pelajaran atau beberapa unit mata pelajaran, maka perlu diajukan beberapa pertanyaan berkenaan dengan daftar tujuan tersebut:

- 1. Sudahkah semua hasil akhir pembelajaran dicantumkan?** Jangan mengabaikan hasil akhir pembelajaran yang lebih rumit hanya karena lebih sulit untuk dinyatakan.

2. **Apakah tujuan-tujuan pembelajaran itu selaras dengan tujuan kurikulum?** Daftar tujuan pembelajaran harus merepresentasikan standar isi yang sama dengan tujuan sekolah. Begitu juga, apabila tujuan sekolah menekankan penilaian mandiri dan pembelajaran independen, maka baik tujuan pembelajaran maupun rencana pengajaran harus memiliki penekanan yang sama.
3. **Apakah tujuan pembelajaran sesuai dengan tingkat capaian siswa?** Kalau tujuan itu harus dicapai oleh siswa, maka mereka harus memiliki latar belakang yang diperlukan untuk mengikuti proses pembelajaran. Hal ini mungkin menuntut guru untuk mereview tujuan-tujuan pembelajaran sebelumnya atau tujuan-tujuan pada level yang lebih rendah. Mengembangkan tujuan secara bersama-sama dengan guru-guru yang lain mungkin merupakan solusi yang paling baik.
4. **Apakah tujuan pembelajaran itu sesuai dengan waktu, materi dan fasilitas yang tersedia?** Pada saat kita bergerak menuju pada hasil belajar yang kompleks, maka perlu dicamkan bahwa hasil akhir belajar yang lebih kompleks memerlukan waktu yang relatif lebih lama untuk mencapainya dari pada hanya sekedar mengingat fakta. Begitu juga, materi ajar dan fasilitas
5. **Apakah tujuan itu sudah mengandung pembelajaran bermakna yang berguna dalam dunia nyata?** Kecenderungan dewasa ini terhadap proses pembelajaran yang lebih autentik bisa dimasukkan ke dalam tujuan pembelajaran dengan cara memasukkan paling sedikit beberapa hasil belajar yang dimaksudkan yang berkenaan dengan masalah-masalah dalam dunia nyata (misalnya peningkatan kualitas air). Dengan melibatkan konteks belajar yang realistis, siswa akan lebih termotivasi dan menjadi lebih sadar akan transfer nilai dari apa yang mereka tengah pelajari.

1.5 MENYIAPKAN PENILAIAN

Menyiapkan penilaian harus bersamaan dengan menyiapkan rencana program pengajaran (RPP) agar penilaian itu menjadi bagian integral dari proses belajar-mengajar. Beberapa pertanyaan berikut ini bisa digunakan sebagai panduan.

- 1. Apakah tes awal diperlukan?** Dengan memperhatikan isi dan metode pembelajaran, Anda bisa menilai apakah pembelajaran prasyarat diperlukan atau tidak. Hal ini mungkin cukup dengan cara memastikan apakah siswa bisa memahami ide yang terkandung dalam sebuah teks atau tidak. Apakah penguasaan kosakata para siswa sudah cukup memadai untuk memahami teks yang ditugaskan kepada mereka. Kalau diberi tugas menulis, apakah mereka sudah mengenal struktur sebuah paragraf – bagaimana menuliskan kalimat topik, bagaimana mengungkapkan kalimat pendukung dan bagaimana mengemukakan kalimat kesimpulan. Sebelum menyampaikan suatu program, mungkin akan sangat bermanfaat untuk memastikan pengetahuan dan keterampilan yang sudah dikuasai oleh siswa agar pembelajaran bisa disesuaikan dengan mudah.
- 2. Jenis penilain apa yang dibutuhkan selama proses pembelajaran?** Jika penillaian akan digunakan untuk memonitor serta mengembangkan kualitas belajar siswa, maka penilaian itu harus direncanakan bersamaan dengan pembuatan RPP. Mungkin perlu diadakan pengerjaan latihan, uji akhir setiap unit pelajaran, observasi dan menilai performa siswa, atau penilain produk yang dihasilkan siswa (misalnya, tulisan, laporan). Hakikat penilaian akan ditentukan oleh tujuan pembelajaran, aktivitas pembelajaran, dan kebutuhan akan *feedback* yang diberikan kepada para siswa. Jika ditangani secara baik, maka penilaian selama berlangsungnya pembelajaran itu bisa memberikan perkembangan yang baik bagi kualitas belajar siswa, yang sebenarnya merupakan fungsi utama dari penilaian.

Pada saat pembuatan rencana penilaian, Anda tidak perlu membuat instrumen penilain yang lengkap. Namun rencana penilian itu harus mencakup hal-hal yang bersifat global dan kapan penilain itu akan digunakan. Skala penilaian, ceklis, dan alat penilaian performa serta produk harus dibahas bersama para siswa sebelum proses pembelajaran dimulai agar instrumen penilaian itu bisa brguna sebagai panduan bagi mereka. Inilah salah satu contoh bagaimana proses pembelajaran dan penilain bisa digabung bersama untuk meningkatkan mutu belajar siswa.
- 3. Jenis penilain apa yang dibutuhkan pada akhir pembelajaran?** Guru perlu memastikan tingkat pencapaian tujuan pembelajaran agar nilai

bisa diberikan, keterampilan bisa disertifikasi. Penilaian ini hampir tidak memberikan kesempatan kepada siswa untuk memperbaiki kualitas belajarnya, namun umpan balik bisa membantu siswa memahami kekuatan serta kelemahannya. Dan informasi dari penilain itu bisa membantu Anda – guru – untuk memastikan pembelajaran berikutnya perlu dimodifikasi atau tidak.

1.6 JENIS PROSEDUR PENILAIAN

Program penilain yang efektif meliputi berbagai prosedur. Beberapa di antaranya berdasarkan pada hasil observasi guru selama berlangsung pembelajaran di ruang kelas, semntara lainnya memberikan alat yang lebih sistematis dan terstruktur untuk memperoleh data sebagai bukti capaian. Berikut ini disajikan lima buah contoh prosedur penilaian yang umum digunakan disertai dengan indikator fungsi utamanya.

- 1. Observasi informal pada saat berlangsungnya proses pembelajaran.**

Observasi di dalam kelas pda saat berlangsungnya proses belajar-mengajar memberikan panduan kepada guru tentang berbagai cara khusus untuk membimbing siswa. Pertanyaan siswa bisa menginsyaratkan adanya kebutuhan untuk mengulang kembali materi ajar yang sudah disampaikan, atau bisa juga memberikan kesempatan bagi guru untuk mengoreksi sedikit kesalahpahaman yang terjadi antara pemahaman siswa dengan apa yang dimaksudkan oleh guru. Observasi bisa menunjukkan bahwa tingkat kecepatan pembelajaran ada yang terlalu dan ada juga yang terlalu lambat bagi sebagian siswa. Beberapa siswa tertentu mungkin memerlukan bantuan khusus dalam memecahkan masalah, melakukan suatu keterampilan, atau menyelesaikan tugas. Observasi yang berlangsung pada saat proses pembelajaran terjadi itu serta keputusan yang guru ambil berdasarkan hasil observasi itu merupakan strategi mendasar untuk meningkatkan kualitas belajar siswa. Prosedur penilain yang lebih sistematis hanyalah merupakan cara untuk melengkapi observasi informal yang guru lakukan selama proses pembelajaran berlangsung.
- 2. Uji Kemampuan di kelas.** Tes yang dirancang untuk digunakan di kelas memberikn prosedur yang sistematis untuk menentukan tingkat sampai di mana hasil pembelajaran yang diharapkan bisa dicapai. Kalau

dirancang dengan baik, tes semacam ini bisa mengukur berbagai hasil belajar, termasuk *mengingat pengetahuan faktual, pemahaman, penerapan, kemampuan menalar, dan pemecahan masalah*. Sayangnya, tes kemampuan ini paling banyak digunakan hanya untuk mengukur *ingatan pengetahuan faktual* dan sudah banyak dikritik karena keterbatasannya dalam mengukur kemampuan yang tes itu berikan.

3. **Penilaian Performa.** Observasi serta penilaian langsung performa siswa selama berlangsungnya kegiatan biasanya dilakukan dengan menggunakan ceklis, skala penilaian, atau rubrik penilaian holistik. Tujuan alat ukur ini adalah untuk mengarahkan observasi pada unsur-unsur performa siswa yang paling penting dan bisa digunakan untuk memberikan ruang bagi pencatatan nilainya. Ujian tulis bisa digunakan untuk menentukan apakah siswa tahu apa yang mereka harus lakukan, namun observasi langsung terhadap performa siswa sangatlah penting untuk memastikan tingkat kualitas performa siswa (seperti performa dalam berbicara (*speaking*), membaca (*reading*), mengoperasikan komputer dsb.).
4. **Penilaian Produk.** Beberapa jenis performa menghasilkan suatu produk yang menjadi pusat penilaian. Beberapa contoh produk yang bisa dinilai secara langsung adalah: esai, grafik, peta, lukisan, poster, model, proyek kerja kayu, dan laporan laboratorium. Ceklis, skala penilaian, dan rubrik penilaian holistik juga sangat berguna dalam menilai produk. Instrumen-instrumen tersebut menyediakan fokus tentang fitur-fitur utama dari produk yang akan dinilai **serta** skala untuk menentukan penilaiannya.
5. **Penilaian Portofolio.** Portofolio bisa didefinisikan sebagai koleksi karya siswa yang digunakan untuk menunjukkan kemajuan belajarnya dalam periode waktu tertentu atau untuk menunjukkan karya terbaik siswa. Biasanya siswa berpartisipasi dalam memilih isi portofolio dalam istilah kemajuan belajar mereka. Isi portofolio biasanya meliputi hasil tes, contoh tulisan, gambar, laporan penelitian, performa dan nilai produk, proyek, laporan konferensi, dan bukti-bukti lain yang sesuai dengan hasil belajar yang dimaksudkan dalam suatu proses pembelajaran. Portofolio dianggap sebagai penilaian performa siswa yang sangat lengkap. Portofolio memberikan kesempatan kepada siswa untuk ikut serta

dalam proses penilaian dan dalam melakukan refleksi tentang apa yang mereka sedang pelajari.

Agar berhasil paling efektif, maka pemilihan dan penggunaan metode penilaian harus direncanakan pada waktu bersamaan dengan penyusunan rencana program penyajaran (RPP). Untuk memastikan metode penilaian apa yang akan digunakan, kapan metode penilain itu akan digunakan, dan bagaimana metode penilaian itu akan digunakan bisa dilakukan paling baik pada saat merencanakan kegiatan pembelajaran. Instrumen penilain mungkin tidak semuanya dibuat pada saat membuat RPP, namun tetap harus direncanakan agar penilaian itu bisa disesuaikan dengan jadwal belajar. Prosedur penilain itu memakan waktu cukup banyak, jika tidak direncanakan sejak awal, maka ada kemungkinan penilaian itu kurang atau tidak sesuai dengan yang semestinya dievaluasi.

Persiapan tes tulis dan metode evaluasi performa siswa harus dilakukan sedini mungkin agar contoh butir soal, seperti yang digunakan dalam tes tulis, dan instrumen penilaian bisa dibagikan kepada siswa jauh sebelum instrumen digunakan. Hal ini akan membuat siswa memahami dengan baik jenis performa yang diharapkan. Menyajikan contoh butir soal di hadapan siswa yang akan digunakan untuk mengukur berbagai hasil belajar, seperti *memahami*, *mengaplikasikan*, dan *mengevaluasi*, akan membuat siswa paham bahwa *mengingat informasi faktual* belum cukup. Membahas instrumen penilaian performa bersama siswa akan memberikan panduan kepada mereka tentang apa yang akan dipelajari dan penilaian secara mandiri.

1.7 VALIDITAS DAN RELIABILITAS DALAM RENCANA PENILAIAN

Ada dua karakteristik utama tentang prosedur penilain yang baik yaitu validitas dan reliabilitas. Kedua sifat utama ini merupakan fokus utama pada saat merencanakan penilaian dan sebagian besar saran yang dikemukakan pada buku ini diarahkan pada persiapan penilain yang memberikan interpretasi hasil penilaian yang sah dan terpercaya (*valid san reliabel*).

Validitas suatu instrumen penilaian bisa didefinisikan sebagai ketepatan dan kebermaknaan kesimpulan yang kita buat berdasarkan

hasil penilaian untuk penggunaan yang dimaksud. Misalnya, jika kita mengadakan tes kosa kata, kita bisa menginterpretasikan nilai yang siswa peroleh sebagai contoh yang representatif tentang istilah-istilah yang kita ajarkan. Jika mengerjakan dengan baik hal-hal berikut ini: (a) mendefinisikan dengan jelas bidang kosa kata yang akan dievaluasi, (b) menyiapkan kisi-kisi tes dengan cermat, dan (c) menyusun butir soal yang representatif dan relevan, maka interpretasi kita tentang hasil tes tersebut kemungkinan besar valid. Karena kita bisa menyimpulkan bahwa nilai kosa kata siswa yang tinggi menunjukkan pengetahuan yang baik tentang kosa kata yang sudah diajarkan. Perlu dicatat bahwa kesimpulan kitalah yang sangat penting. Jika kita menyimpulkan bahwa nilai tinggi itu menunjukkan kemampuan menulis yang baik atau kemampuan berbicara yang baik, maka kita telah membuat kesimpulan di luar aspek penilaian yang tengah dites, karena itu validitas dari kesimpulan yang kita buat sangatlah meragukan tanpa ada bukti lain yang cukup meyakinkan. Singkatnya, bukanlah skor sebuah tes yang valid atau tidak valid itu namun kesimpulan yang kita buat tentang hasil tes tersebut.

Penilaian performa siswa biasanya dipandang sebagai penilaian yang bisa memberikan kesimpulan yang lebih valid tentang hasil belajar dari pada tes tulis tradisional karena penilaian performa itu lebih difokuskan secara langsung terhadap jenis tugas performa yang kita ajarkan. Jika kita ingin mengevaluasi kemampuan siswa dalam membaca maka suruhlah mereka membaca suatu teks. Jika kita ingin mengetahui kemampuan siswa untuk berbicara atau berpidato, maka suruhlah dia berbicara atau berpidato. Jika kita ingin mengetahui apakah siswa bisa mengoperasikan komputer, maka suruhlah dia mengoperasikan komputer. Pada setiap kasus tersebut, setiap tugas yang diberikan itu memiliki validitas tampilan (disebut *face validity*). Namun validitas tampilan itu tidaklah sesederhana yang kita lihat. Dalam penilaian performa siswa, masalah khusus dihadapi pada saat mendefinisikan aspek yang dinilai, bagaimana performa yang dinilai, serta dalam menentukan sampel performa yang representatif. Misalnya, banyak jenis membaca yang berbeda-beda, banyak jenis berbicara yang berbeda-beda, dan banyak jenis masalah untuk diselesaikan pada komputer. Masing-masing memiliki rubrik dan sifat tersendiri. Mengingat penilaian performa itu memakan waktu lama, maka sampelnya cenderung terbatas.

Hal ini menyebabkan kesimpulan yang dibuat tentang performa suatu tugas diterapkan pula untuk menyimpulkan performa tugas lain dalam bidang yang sama.

Mengingat masalah serta keterbatasan di atas, maka kita perlu membuat rencana serta mengikuti prosedur yang paling mungkin menghasilkan kesimpulan yang valid berkenaan dengan pembelajaran siswa. Ada empat komponen yang harus kita lakukan pada saat membuat perencanaan untuk tes tulis atau penilaian performa: (a) memilih prosedur penilaian yang tepat, (b) menyiapkannya secara cermat, (c) menerapkannya secara efektif, dan (d) menginterpretasikan hasilnya dalam batasan aspek capaian tertentu yang dinilai. Pada saat menggunakan hasil penilaian, kita juga harus memperhatikan akibat dari penilaian tersebut, yakni apakah hasil penilaian itu meningkatkan kualitas belajar siswa sebagaimana yang dimaksud? Inilah pertanyaan yang sangat penting yang berkaitan dengan validitas penilaian karena tujuan utama kita dalam menilai prestasi siswa adalah untuk meningkatkan kualitas belajar siswa.

Karakteristik utama kedua adalah reliabilitas. *Reliabilitas* bisa didefinisikan sebagai konsistensi hasil penilaian. Misalnya, jika seorang siswa memperoleh nilai 70 pada suatu tes membaca, kita bisa mengatakan bahwa nilai 70 itu merepresentasikan kemampuan membaca siswa tersebut secara akurat pada tes tersebut. Jadi jika kita mengetes siswa tersebut pada waktu berbeda dengan menggunakan sampel tes berbeda dari butir yang mirip, maka kita bisa mengharapkan bahwa dia akan memperoleh skor yang serupa. Begitu juga jika seorang siswa memperoleh skor tinggi pada tes menulis laporan, kita bisa mengatakan bahwa skor tersebut merepresentasikan keterampilan menulis siswa tersebut yang sebenarnya, dan bahwa jika guru lain mengetes kemampuan menulis siswa tersebut, maka hasilnya akan serupa. Jadi, konsistensi hasil suatu tes itu akan menunjukkan bahwa hasil tersebut bebas dari kesalahan, oleh karena itu kita bisa memercayainya (yakni hasil tes tersebut memiliki *reliabilitas*).

Tentu saja kita tidak bisa mengharapkan hasil penilaian itu konsisten secara sempurna sepanjang waktu, dan bagi seluruh sampel pada bidang capaian yang sama. Faktor-faktor seperti ketaksamaan (kecenderungan memiliki dua interpretasi atau lebih), variasi sampel butir soal, naik turunnya

motivasi dan perhatian siswa serta faktor keberuntungan bisa menyebabkan terjadinya kesalahan yang mengakibatkan hasil penilaian berbeda. Demikian juga dalam menilai performa siswa, faktor pribadi si penilai bisa menyebabkan terjadinya kesalahan pada hasil penilaiannya. Tujuan yang terpenting dalam penilaian adalah menjaga jenis-jenis kesalahan ini seminimal mungkin agar hasil bisa sereliabel mungkin.

Selain itu, reliabilitas itu sangat penting untuk memperoleh kesimpulan yang valid dari hasil penilaian. Jika skor seseorang siswa berfluktuasi terlalu lebar pada suatu sampel butir soal, maka kita tidak bisa menghrapkan mampu mengambil kesimpulan yang valid tentang capaian siswa. Begitu juga jika penilaian tentang kemampuan menulis siswa bervariasi terlalu lebar, maka kesimpulan yang valid tidak bisa dicapai tentang kemampuan menulis siswa tersebut. Singkatnya, reliabilitas itu merupakan syarat yang diperlukan untuk membuat kesimpulan yang valid namun belum cukup.

Baik validitas maupun reliabilitas bisa dibuat pada saat menyiapkan prosedur penilaian. Apabila kita mendefinisikan dengan jelas hasil pembelajaran yang dimaksud, mendefinisikan bidang capaian yang akan dinilai, serta memilih seperangkat alat ukur penilaian yang representatif dan relevan, maka berarti kita sudah menyiapkan kesimpulan yang valid untuk pembelajaran. Apabila kita sudah menyusun sejumlah tugas yang cukup dalam penilaian yang kita siapkan, dan kita menggunakan prosedur yang terbebas dari ketaksaan, sumber kesulitan yang tidak relevan, kunci jawaban yang tidak disengaja, dan faktor lain yang bisa mempengaruhi hasil penilaian, berarti kita sudah menyiapkan penilaian yang valid dan reliabel. Sebenarnya, sebagian besar saran-saran untuk merancang tes kemampuan dan menyiapkan penilaian performa diarahkan untuk meningkatkan kualitas reliabilitas hasil penilain dan validitas kesimpulan yang akan kita buat berdasarkan pada hasil penilain tersebut.

Beberapa fitur yang sangat penting untuk meningkatkan kualitas validitas dan reliabilitas bisa dilihat pada Table 1.2. Fitur-fitur tersebut harus dipertimbangkan dengan sangat cermat sejak awal sebelum tes dibuat, bukan setelah tes digunakan atau sebelum datanya dianalisis.

Tabel 1.2 *Fitur untuk meningkatkan validitas dan reliabilitas hasil penilaian*

Fitur yang diinginkan	Prosedur untuk dilakukan
1. Hasil pembelajaran dirumuskan secara jelas	1. Rumuskan hasil pembelajaran yang dimaksudkan dalam istilah performa.
2. Sampel yang representatif tentang tugas pembelajaran dalam suatu bidang yang dedefinisikan	2. Siapkan deskripsi tentang capaian yang akan diukur dan contoh tugas yang akan digunakan.
3. Tugas yang sesuai dengan hasil pembelajaran yang diukur	3. Sesuaikan tugas penilaian dengan performa tertentu yang sudah dirumuskan dalam hasil pembelajaran.
4. Tugas yang memiliki tingkat kesulitan yang cocok.	4. Sesuaikan tingkat kesulitan pada penilaian dengan tugas pembelajaran, kemampuan siswa, dan penggunaan hasil penilaian
5. Tugas yang berfungsi secara efektif untuk membedakan antara siswa yang pintar dan yang tidak pintar	5. Ikuti petunjuk umum serta aturan khusus untuk mempersiapkan prosedur penilaian dan waspadalah terhadap faktor-faktor yang bisa mengganggu keabsahan hasilnya.
6. Tugas yang cukup banyak untuk mengukur contoh prestasi yang memadai, hasil yang bisa dipercaya, dan hasil penilaian yang memungkinkan untuk diinterpretasikan.	6. Manakala umur siswa atau waktu evaluasi yang tersedia bisa membatasi jumlah tugas, maka buatlah interpretasi sementara, berikan penilaian yang lebih sering, serta buktikan hasil penilain itu dengan menggunakan alat bukti yang lain.
7. Prosedur yang bisa memberikan kontribusi terhadap persiapan dan penggunaan yang efisien.	7. Tulislah petunjuk yang jelas dan susunlah prosedur untuk mempermudah administrasi, penyekoran, atau penilaian serta interpretasi.

1.8 RANGKUMAN

1. Rencana penilaian harus disesuaikan dengan apa yang siswa harus pelajari sebagaimana dijelaskan pada tujuan pembelajaran.
2. Tujuan pembelajaran harus selaras dengan tujuan sekolah dan harus merefleksikan standar bagi sekolah-sekolah yang sederajat.
3. Rencana penilain perlu mempertimbangkan semua jenis hasil belajar yang memungkinkan dalam bidang disiplin ilmu tertentu.
4. Taksonomi tujuan pendidikan yang direvisi memberikan kerangka yang bermanfaat untuk hal-hal berikut ini:
 - a. Mengidentifikasi hasil proses pembelajaran.
 - b. Merencanakan kegiatan pembelajaran.
 - c. Merencanakan metode penilaian.
 - d. Mengecek keselaran berbagai tujuan pembelajaran, pengajaran, dan penilian.
5. Perencanaan penilian harus merupakan bagian dari perencanaan pengajaran agar metode panilain sejalan dengan tujuan pengajaran, dan aktivitas pembelajaran.
6. Tujuan pembelajaran harus dirumuskan dalam istilah performa siswa yang harus didemonstrasikan di akhir pembelajaran sebagai bukti bahwa dia sudah belajar dan memahaminya.
7. Penilian di kelas bisa digunakan untuk mengukur semua jenis hasil belajar dari yang paling sederhana sampai dengan yang paling rumit. Performa yang diukur harus sesuai dengan performa yang dirumuskan pada tujuan pembelajaran.
8. Penilian performa berkaitan dengan keterampilan yang bisa dilihat, (misalnya berbicara) atau produk (musalnya menulis), biasanya menggunakan ceklis, skala penilian, atau rubrik penilaian (petunjuk penilian).
9. Penilian portofolio memberikan penilaian yang komprehensif tentang prestasi siswa serta memberikan kesempatan bagi siswa untuk merefleksikan apa yang sudah mereka pelajari.
10. Para siswa harus diberi contoh butir soal dengan baik sebelum mereka diuji dan harus diberi contoh instrumen performa sebelum kegiatan performa dimulai, agar mereka memiliki target capaian sebagai pemnadu proses belajarnya. Instrumen performa juga sangat bermanfaat bagi

siswa untuk mengembangkan keterampilan menilai capaian mereka secara mandiri.

11. Validitas dan reliabilitas merupakan dua karakteristik utama dari setiap metode penilaian dan harus dipertimbangkan pada saat membuat perencanaan dan persiapan prosedur penilaian.
12. Interpretasi hasil penilaian yang valid dan reliabel memerlukan hasil akhir proses pembelajaran yang dirumuskan dengan jelas, contoh tugas pembelajaran yang representatif, sistem penilaian yang bagus, serta trebebas dari faktor-faktor yang tidak relevan yang bisa menimbulkan kesalahan.

Inilah akhir dari Bab I. Pada bab berikutnya akan dibahas hakekat penilaian.

-oo0oo-



Bab 2

HAKIKAT PENILAIAN

Tujuan bab ini adalah agar mahasiswa mampu:

1. Mendeskripsikan situasi di mana pengujian dan penilaian performa diperlukan serta mampu menjelaskan alasannya.
2. Menjelaskan jenis-jenis metode penilaian utama dan memberikan contoh untuk masing-masing jenis.
3. Membedakan antara pengujian dan penilaian performa dalam kaitannya dengan realisasi tugas, tingkat kesulitan tugas, waktu penilaian yang diperlukan, serta penilaian dalam penyekoran.
4. Menyusun petunjuk untuk penilaian siswa secara efektif.
5. Membedakan antara penilaian beracuan norma dan penilaian beracuan kriteria.

Istilah *assessment* digunakan sebagai kategori yang sangat luas yang mencakup semua jenis metode yang digunakan untuk menentukan tingkat capaian siswa dalam hasil proses pembelajaran yang dimaksudkan. Hal ini meliputi bukan hanya pengujian tetapi juga penilaian performa (*performance assessment*). Misalnya, untuk menilai kemampuan mengemudi seorang siswa, tes objektif digunakan untuk mengukur pengetahuan siswa tentang cara mengemudikan mobil dan cara menggunakan aturan lalulintas. Sedangkan untuk memastikan kemampuannya dalam mengemudikan mobil, maka ia harus disuruh mengemudikan mobil dalam waktu tertentu dan dalam jarak tertentu (*performance assessment*). Tes tentang penguasaan aturan lalulintas mencakup sampel aturan-aturan lalulintas yang lebih

banyak dari pada yang mungkin ditemukan pada saat penilaian performa, tetapi keterampilan mengemudi mobil hanya bisa ditentukan dengan cara ia duduk di belakang setir dan kemudian mengemudikannya. Kedua jenis penilaian itu sangat penting. Tes pengetahuan menunjukkan tingkat pengetahuan siswa tentang apa yang harus ia lakukan sedangkan penilaian performa menunjukkan tingkat keterampilan yang ia bisa lakukan.

Para guru umumnya lebih menyukai tes pilihan seperti pilihan ganda, benar-salah dan menjodohkan. Hal ini terjadi karena mengingat ada tiga alasan: (a) banyak pertanyaan bisa diajukan dalam tempo yang relatif sangat singkat, (b) mudah diselenggarakan dan dinilai, dan (c) hasilnya bisa dinyatakan dalam bentuk angka yang mudah dicatat, dibandingkan serta dilaporkan kepada pihak lain. Sayangnya tes pilihan yang guru buat itu hampir semuanya terbatas hanya pada *mengingat* fakta dan istilah saja. Berbagai studi telah menunjukkan bahwa antara 80 - 90% tes yang guru buat itu dipusatkan pada unsur *mengingat* pengetahuan. Hampir bisa dipastikan bahwa penekanan yang berlebihan terhadap jenis tes pilihan dan pada pengetahuan yang sederhana telah menyebabkan berpindah kepada teknik penilaian yang mengukur hasil belajar yang lebih kompleks dalam setting yang lebih realistik. Fakta bahwa tes tertulis bisa juga dirancang untuk mengukur berbagai hasil belajar yang kompleks seringkali diabaikan setelah terjadi perpindahan ke penilaian performa. yang perlu kita perhatikan sekarang setelah berpindah dari tes tertulis kepada penilaian performa adalah bahwa pendidikan itu bisa dilayani secara paling baik dengan menggunakan kedua pendekatan tes tertulis dan penilaian performa aktual dengan menekankan kedua tes dan penilaian tersebut pada tugas-tugas belajar yang lebih kompleks dari pada apa yang terjadi pada masa lalu yang hanya menekankan pada unsur *mengingat* saja.

2.1 JENIS-JENIS METODE PENILAIAN

Metode penilaian bermacam-macam, namun secara garis besarnya bisa diklasifikasikan menjadi 4 kelompok (lihat Table 2.1.). *Tes pilihan* menuntut siswa untuk memilih jawaban yang paling benar, sebagaimana pada soal pilihan ganda, benar-salah dan menjodohkan. *Tes isian* menuntut siswa untuk menjawab dengan sebuah kata, frase atau sebuah esay yang lengkap. *Penilaian performa terbatas* berhubungan dengan penilaian tentang tugas terba-

tas yang sangat terstruktur dengan baik, seperti menulis sebuah paragraf lengkap tentang sesuatu topik, memilih perlengkapan laboratorium, mengukur tingkat kelembaban, atau mencari informasi dengan menggunakan komputer. *Penilaian performa yang diperluas* meliputi tugas performa yang lebih komprehensif yang kurang terstruktur seperti menulis cerita pendek, melakukan percobaan di laboratorium, meramal cuaca, atau menggunakan komputer untuk memecahkan masalah. Selain menuntut performa yang lebih diperluas, biasanya penilaian itu menuntut siswa untuk mengintegrasikan dan menerapkan pengetahuan serta keterampilan untuk melakukan tugas-tugas performa dalam setting yang realistis. Jika ada suatu produk yang dibuat (misalnya sebuah cerita pendek), siswa mungkin juga diharapkan untuk mereview dan merevisi produk itu sebelum diserahkan, guna menambah realitas pada tugas tersebut.

Tabel 2.1 Ringkasan Perbandingan metode Penilaian

Pengujian		Penilaian Performa	
Jawaban Pilihan	Jawaban Isian	Performa Terbatas	Performa Diperluas
Rendah	Realisme Tugas 		Tinggi
Rendah	Kompleksitas Tugas 		Tinggi
Rendah	Waktu Penilaian yang Diperlukan 		Tinggi
Rendah	Penilaian Dalam Penyekoran 		Tinggi

Sumber: Gronlund, dkk. (2009: 17)

Tabel 2.1 di atas menunjukkan 4 buah kategori penilaian berikut karakteristik dari masing-masing kategori. Keempat kategori penilaian itu terdiri dari:

1. Realisme Tugas
2. Kompleksitas Tugas
3. Waktu Penilaian yang Diperlukan dan
4. Penilaian dalam Penyekoran.

Masing-masing dari keempat kategori penilaian di atas akan dijelaskan secara rinci dalam paragraf berikut ini.

2.2 REALISME TUGAS

Realisme tugas bisa didefinisikan sebagai tingkat kemampuan para siswa mensimulasikan performa dalam dunia nyata. Tes tradisional *jenis pilihan* memiliki realisme yang rendah karena meliputi pemilihan satu jawaban dari seperangkat jawaban yang mungkin yang sudah ditentukan. Jawabannya terbatas hanya pada pilihan yang sudah disusun, dan masalah-masalah yang terstruktur sangat tinggi semacam itu jarang sekali terjadi dalam dunia nyata. *Penilaian performa yang diperluas* memiliki realisme yang tinggi karena penilaian tersebut berusaha untuk mensimulasikan performa dalam dunia nyata. Ketika menilai tingkat kecakapan siswa dalam mengemudikan kendaraan, mengoperasikan mesin, menyampaikan pidato, atau menerapkan pengetahuan serta pemahaman dalam menyelesaikan masalah-masalah dalam dunia nyata (seperti cara melindungi lingkungan), maka diperlukan seperangkat jawaban komprehensif yang mendekati apa yang terjadi dalam dunia nyata. Penilaian yang terletak di antara kedua ekstrim ini adalah *tes isian* (seperti tes dengan jawaban singkat atau esai) dan penilaian performa dengan respon terbatas yang memberikan struktur sedang tetapi memberikan kebebasan jawaban yang lebih besar dan oleh karenanya merupakan jenis masalah yang lebih realistik dari pada tes jenis pilihan.

Selain pindah dari tes tradisional pilihan menuju ke penilaian performa yang diperluas guna meningkatkan tingkat realisme, juga terdapat kecenderungan untuk membuat tes tertulis tradisional menjadi lebih autentik (yakni, memiliki realisme yang lebih tinggi). Hal ini telah menyebabkan rancangan tes sedemikian rupa sehingga cocok untuk mengukur hasil belajar yang lebih kompleks dan penggunaan soal serta prosedur lebih mirip dengan apa yang terjadi dalam dunia nyata. Misalnya, dalam masalah matematika siswa mungkin diberi fakta lebih banyak dari pada apa yang diperlukan untuk memecahkan masalah untuk menguji apakah mereka bisa memilih fakta yang diperlukan untuk memecahkan masalahnya. Dalam pemecahan masalah sains, para siswa mungkin diberi kebebasan untuk memilih prosedur guna memecahkan masalah dan dan diminta

untuk memberikan pembenaran tentang prosedur yang mereka gunakan. Dalam beberapa kasus, hal ini bisa diterapkan dengan cara mengubah dari penggunaan butir soal pilihan ke butir soal isian, tetapi dalam kasus lainnya mungkin menggunakan kombinasi dari kedua tipe butir soal tersebut (misalnya, mereka diminta menjelaskan mengapa mereka memilih jawaban yang mereka pilih).

2.3 KOMPLEKSITAS TUGAS

Butir soal dengan jawaban pilihan cenderung memiliki kompleksitas yang rendah baik dalam masalah yang disajikannya maupun hakikat jawaban yang diharapkannya. Meskipun butir soal bisa dirancang untuk mengukur pemahaman dan keterampilan berikir, namun soal-soal tersebut biasanya menyajikan satu masalah yang terbatas dan memerlukan pilihan jawaban yang paling benar. Sebaliknya, soal performa yang diperluas biasanya meliputi hasil belajar yang bermacam-macam, integrasi ide serta keterampilan dari berbagai sumber, tersedianya berbagai solusi yang memungkinkan, dan kebutuhan akan berbagai kriteria untuk mengevaluasi hasil (misalnya, merancang sebuah rencana untuk mengurangi peredaran narkoba di Indonesia). Begitu juga penilaian performa yang tengah berlangsung melibatkan pola gerakan yang kompleks yang dipandu oleh pengintegrasian informasi dan keterampilan khusus dari berbagai pengalaman belajar (misalnya, memainkan alat musik, mengoperasikan mesin, dan memperbaiki peralatan listrik). Sebagaimana pada kategori realisme, tes jenis isian berada di antara kedua ekstrim tersebut. Misalnya, tes esai bisa dirancang untuk mengukur kemampuan siswa untuk memilih, mengintegrasikan serta menyatakan ide, namun tugasnya biasanya lebih terbatas serta lebih terstruktur dari pada tugas-tugas pada penilaian performa.

2.4 WAKTU PENILAIAN YANG DIPERLUKAN

Sejumlah butir soal pilihan bisa diberikan kepada sekelompok siswa dalam waktu yang relatif singkat dan hasilnya bisa diskor dengan cepat baik secara manual maupun dengan menggunakan mesin. Keefisienan inilah yang tidak bisa diragukan lagi merupakan faktor utama penggunaannya secara sangat luas. Penilaian performa cenderung memerlukan waktu sangat lama.

Beberapa tugas bisa memerlukan waktu sehari-hari atau bahkan berminggu-minggu untuk menyelesaikannya (misalnya, melakukan studi percobaan) dan beberapa tugas lainnya mungkin memerlukan penilaian siswa satu per satu, (contohnya, menyampaikan pidato, mengoperasikan mesin). Pada sebagian besar kasus, mengevaluasi proses atau produk dari suatu performa sangatlah sulit dan memerlukan waktu lama. Tes isian, seperti halnya tes esai, memerlukan waktu lebih lama untuk mengoreksinya dari pada tes pilihan, tetapi relatif lebih singkat dari pada penilaian performa.

Sejumlah waktu yang lebih lama diperlukan untuk penilaian performa bisa mengakibatkan hilangnya cakupan isi karena keterbatasan jumlah soal penilaian yang bisa dimasukkan ke dalam program pengajaran. Hal ini meningkatkan pertanyaan apakah hasil penilaian bisa digeneralisasikan tugas-tugas lainnya yang serupa. Masalah “dunia nyata” tidak bisa dikemukakan kepada siswa, tetapi masalah itu biasanya unik bagi setting konteks tertentu. Oleh karena itu, transfer hasil belajar merupakan pertimbangan utama dalam penilaian performa. Kita bisa membenarkan jumlah waktu yang lebih banyak diperlukan hanya jika penilaian itu merupakan bagian integral dari pengajaran dan hasil belajar yang bisa ditransfer ditekankan (misalnya, penalaran, berpikir kritis, keterampilan psikomotorik).

2.5 PENILAIAN DALAM PENYEKORAN

Penilaian yang terlibat dalam penyekoran sangat berbeda-beda. Setiap butir soal pilihanditandai dengan benar atau salah sehingga penyekorannya benar-benar objektif (yakni penyekor yang berbeda-beda akan menghasilkan skor yang sama). Tes esai memberikan kebebasan yang lebih banyak bagi jawabannya, dan hal ini menyebabkan subjektivitas yang besar dalam penyekoran. Penyekor yang berbeda-beda akan menghasilkan skor yang berbeda-beda karena mereka mengukur unsur-unsur jawaban secara berbeda-beda (misalnya kelengkapan, organisasi, kejelasan tulisan dan semacamnya) dan melahirkan bias perorangan lainnya ke dalam penilaiannya. Ketika tugas menjadi lebih komprehensif dan kompleks, seperti dalam penilaian performa, tuntutan penilaian guru menjadi lebih besar. Tugas performa yang kompleks yang melibatkan integrasi dari berbagai jenis informasi dan keterampilan serta mungkin memiliki berbagai solusi menyebabkan sulit dan dalam beberapa kasus tidak diharapkan

untuk memiliki model jawaban seperti yang terjadi pada penilaian tulisan dalam sebuah esai. Dengan penilaian performa, kita sangat mungkin harus tergantung pada identifikasi kriteria mengenai performa yang berkualitas, kemudian menerapkan kriteria tersebut melalui skala penilaian atau seperangkat rubrik penilaian. Masing-masing langkah ini berdasarkan penilaian subjektif.

Berdasarkan telaahan terhadap masing-masing kategori ini maka menjadi jelaslah bahwa setiap metode penilaian itu memiliki kekuatan serta kelemahannya masing-masing. Apabila kita menggunakan tes pilihan, kita bisa memperoleh cakupan yang luas tentang domain isinya, dan bisa menyelenggarakan tes, menyekornya, dan menginterpretasikannya dengan mudah tetapi kita mengorbankan realisme dan beberapa jenis kompleksitas. Apabila kita menggunakan penilaian performa yang diperluas, kita bisa mendapatkan tingkat realisme yang tinggi, dan meningkatkan kompleksitas tugas yang bisa kita nilai, namun waktu yang diperlukan untuk penilaian itu sangatlah lama dan penilaian performanya sangat subjektif. Mungkin aturan yang bermanfaat adalah gunakanlah metode yang paling efisien selama metode tersebut cocok untuk menilai hasil belajar yang dimaksudkan, tetapi jangan sampai mengabaikan hasil belajar yang kompleks hanya karena metode penilaiannya memerlukan waktu lama serta hasil tesnya sulit untuk disekor atau dinilai.

2.6 PETUNJUK UNTUK PENILAIAN SISWA YANG EFEKTIF

Tujuan utama dari program penilaian siswa adalah untuk meningkatkan kualitas belajar siswa. Hal ini bisa sangat mungkin terjadi manakala penilaian itu diintegrasikan dengan pembelajaran dan dipandu oleh seperangkat kondisi dasar. Berikut ini adalah panduan sebagai kerangka umum untuk menggunakan penilaian siswa yang efektif.

1. *Penilaian yang efektif memerlukan konsepsi yang jelas tentang semua hasil belajar yang dimaksudkan.* Pada saat membuat rencana pengajaran dan penilaian, kita perlu bertanya pada diri kita sendiri: Apa hasil belajar yang dimaksudkan melalui pembelajaran itu? Jenis pengetahuan, pemahaman, aplikasi, dan keterampilan performa apakah yang akan kita terima sebagai bukti bahwa pengajaran itu telah berhasil? Dalam hal ini selalu ada kecenderungan yang sangat berbahaya yakni memouskan

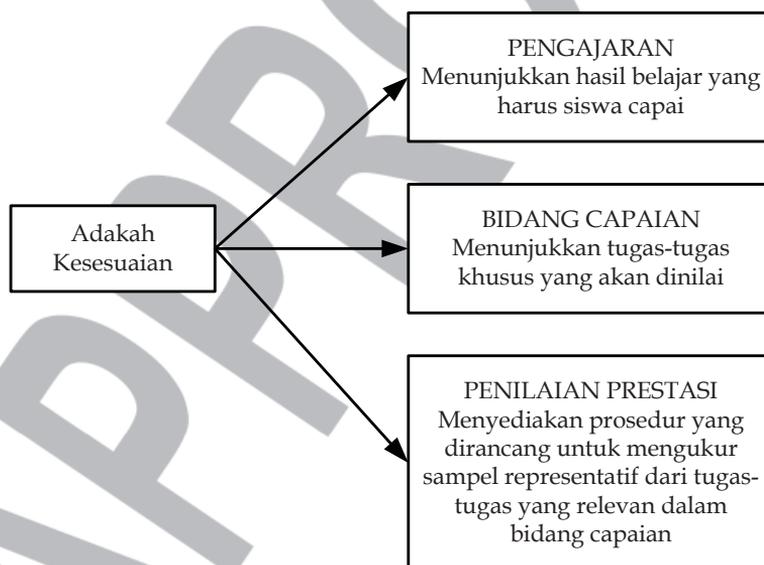
terlalu sempit pada hasil belajar yang sederhana karena pemokusannya semacam itu sangatlah mudah diidentifikasi, dirumuskan, dan diukur. Jika kita tidak melibatkan tujuan belajar yang lebih kompleks pada tahap perencanaannya, maka tujuan belajar yang lebih kompleks tersebut akan cenderung terabaikan pada saat penilaian. Oleh karena itu, kita perlu menjelaskan semua hasil belajar yang dimaksudkan dalam istilah performa siswa, serta harus membuat rencana untuk menilai semua tujuan tersebut.

2. *Penilaian yang efektif memerlukan bahwa beraneka macam prosedur penilaian digunakan.* Berbagai hasil belajar yang mungkin dicapai dalam pembelajarannya pada suatu bidang tertentu menunjukkan bahwa berbagai jenis prosedur penilaian harus dipertimbangkan pada saat membuat perencanaan penilaian. Tes pilihan mungkin bisa digunakan untuk beberapa hasil belajar dan tes esai untuk menilai tujuan belajar lainnya di mana kemampuan menyatakan ide dianggap sangat penting. Dalam menilai keterampilan performa, di mana kita harus sangat tergantung pada *judgement* (penilaian subjektif), maka skala penilaian atau ceklis mungkin diperlukan.

Dalam menilai hasil belajar yang lebih kompleks, maka kombinasi berbagai metode penilaian mungkin paling cocok. Misalnya, menyelesaikan masalah yang kompleks mungkin melibatkan pengumpulan informasi dari berbagai sumber, menganalisisnya, mengintegrasikannya, menuliskan solusi yang disarankan, serta menyajikannya secara lisan di hadapan sekelompok orang. Begitu juga, dalam mengidentifikasi dan memperbaiki bagian yang rusak pada sebuah mesin mungkin diperlukan untuk membaca buku manual, mengidentifikasi suara mesin, memilih alat yang tepat, menguji bagian-bagian mesin, dan membuat perbaikan yang diperlukan. Selain memperbaiki bagian mesin yang rusak, kecepatan performa pada saat melakukan langkah-langkah yang tepat dan faktor-faktor yang serupa mungkin merupakan bagian penting dalam proses penilaian. Dalam mengevaluasi keterampilan performa, berbagai penilaian sangatlah mungkin menjadi aturan yang harus diikuti dan bukannya pengecualian.

3. *Penilaian yang efektif memerlukan bahwa relevansi prosedur pengajaran harus dipertimbangkan.* Penilaian yang relevan dengan proses pengajaran

mengandung arti bahwa hasil pembelajaran yang dimaksudkan, domain tugas pembelajaran, serta prosedur penilaian semuanya harus selaras, (lihat Gambar 2.1). Juga mengandung arti bahwa perencanaan untuk menggunakan hasil penilaian dalam program pengajaran harus dipertimbangkan. Apakah tes di kelas akan memberikan informasi yang cukup untuk melakukan program remedial? Bisakah penilaian keterampilan performa yang kompleks dirancang sebagai kegiatan yang berlangsung sepanjang pembelajaran agar bisa memberikan kontribusi langsung terhadap pembelajaran? Pertanyaan-pertanyaan ini dan semacamnya sangatlah diperlukan untuk memperoleh integrasi maksimal dari pengajaran dan penilaian. Perlu dicamkan bahwa tujuan utama pengajaran dan penilaian adalah untuk meningkatkan kualitas proses pembelajaran siswa. Dengan rancangan program penilaian yang baik, maka kegiatan penilaian akan tampak hampir tidak berbeda dengan kegiatan pengajaran.



Gambar 2.1 Rangkaian dalam mempersiapkan penilaian yang relevan

4. Penilaian yang efektif memerlukan sampel yang cukup dari performa siswa. Penilaian selalu merupakan hal yang berkaitan dengan pengambilan sampel. Pengajaran kita biasanya meliputi berbagai capaian pengetahuan dan keterampilan, namun karena keterbatasan waktu untuk penilaian

dan karena hambatan-hambatan lainnya, maka kita hanya bisa mengukur atau menilai sampel terbatas dari performa siswa dalam bidang tertentu. Misalnya, dalam menyiapkan tes tertulis di kelas, mungkin terdapat 100 istilah yang harus diketahui oleh siswa, namun kita hanya menggunakan 20 istilah saja dalam tes tertulis tersebut. Oleh karena itu, kita harus memilih sampel yang representatif dari 100 istilah tersebut karena kita ingin membuat generalisasi dari performa siswa pada 20 item itu mengenai seberapa hebat siswa mengetahui 100 istilah tersebut. Jika sampel kita cukup, maka kita bisa memperkirakan bahwa jawaban siswa misalnya benar 18 dari 20 istilah berarti 90% dari 100 istilah itu sudah dikuasai oleh siswa.

Sampel juga merupakan masalah dalam penilaian performa. Misalnya, ketika menguji kemampuan mengemudi, tidaklah mungkin untuk memasukkan semua masalah yang berkaitan dengan kemampuan mengemudi dalam uji kemudi yang singkat, karena itu kita harus menetapkan sampel yang representatif dari kemampuan yang berkaitan dengan uji kemudi.

Dalam penilaian keterampilan performa di dalam kelas, biasanya terdapat dua masalah yang berkaitan dengan pengambilan sampel. Misalnya dalam penilaian kemampuan menulis, mungkin kita akan mengajukan pertanyaan-pertanyaan berikut ini: (a) Apakah tugas menulis yang diberikan kepada para siswa mengandung sampel yang representatif dari kemampuan menulis mereka yang ditekankan pada pengajaran itu? Dan (b) Apakah performa siswa dalam tugas menulis itu merepresentasikan seperti apa performa tulisan siswa pada tugas serupa? Mengingat penilaian performa yang diperluas (misalnya, kemampuan berpidato, melakukan sebuah percobaan, atau menerapkan prinsip matematika untuk memecahkan masalah dalam dunia nyata) memakan waktu cukup lama serta terikat oleh konteks, maka penggunaan sampel yang cukup merupakan pertimbangan yang sangat penting dalam merencanakan suatu penilaian.

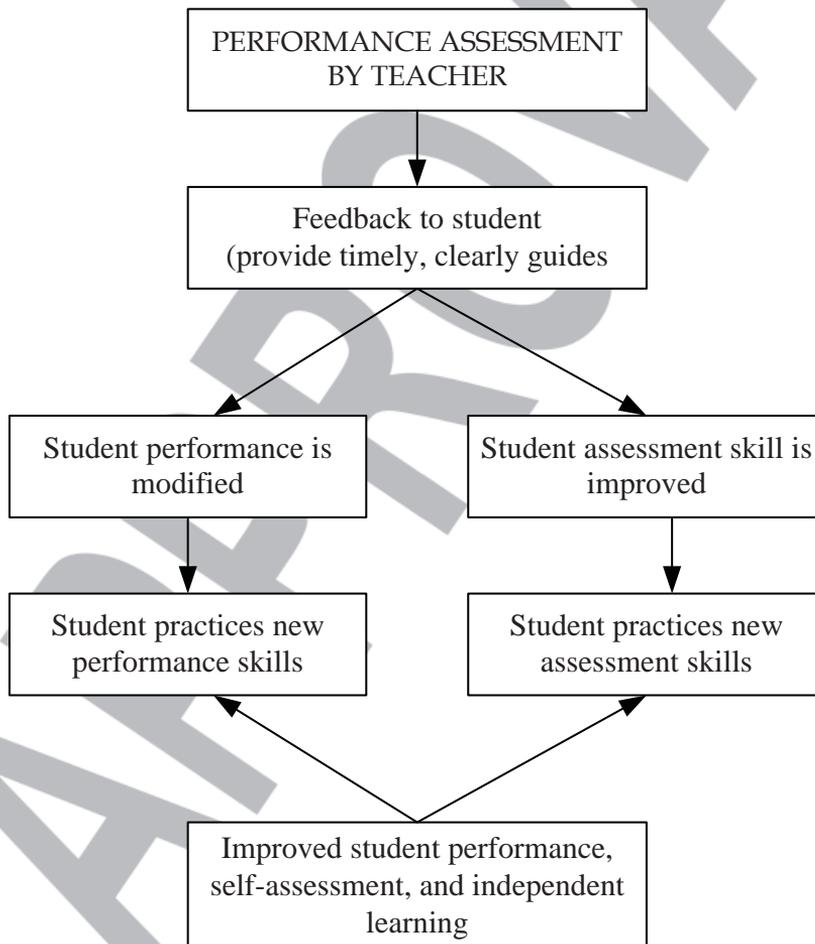
5. *Penilaian yang efektif memerlukan prosedur yang adil untuk setiap siswa.* Program penilaian yang sejalan dengan penciptaan atmosfer yang adil memiliki karakteristik berikut ini: (a) membuat hasil belajar lebih jelas bagi para siswa, (b) menggunakan prosedur penilaian yang sesuai

dengan pembelajaran, (c) merepresentasikan performa siswa yang cukup, dan (d) menggunakan hasil penilaian untuk mengembangkan kualitas proses pembelajaran.

6. *Penilaian yang efektif memerlukan spesifikasikriteria untuk menilai performa yang berhasil.* Pada masa lalu, keberhasilan itu ditentukan dengan cara membandingkan pefroom siswa dengan performa siswa lainnya (interpretasi berdasarkan patokan norma). Jika performa seorang siswa melebihi performa siswa lain, maka peforma itu dianggap bagus sekali. Sebaliknya jika performa seorang siswa ternyata lebih rendah dari performa siswa lainnya, maka performa itu dianggap rendah dan kurang berhasil. Meskipun metode penilaian keberhasilan performa semacam ini bermanfaat dalam situasi tertentu, namun tidak bisa memuaskan sebagai penilaian seberapa bagus siswa menuasai tujuan pembelajaran yang telah ditetapkan. maka untuk tujuan inilah kita memerlukan kriteria yang bisa menjelaskan apa yang bisa siswa lakukan manakala dia melakukan suartu performa dengan berhasil (misalnya bisa membaca satu paragraf teks berbahasa Inggris dengan ucapan yang tepat tanpa satu kesalahanpun, atau bisa menuliskan sebuah kalimat dengan benar secara gramatikal, tanpa ada kesalahan sedikitpun).
7. *Penilaian yang efektif memerlukan umpan balik (feedback) bagi siswa yang menekankan kehebatan performanya dan kelemahan-kelemahan yang harus dikoreksinya.* Hasil umpan balik satu penilaian bagi siswa merupakan faktor yang sangat penting dalam suatu program penilaian. Agar umpan balik itu sangat efektif, maka perlu memenuhi kriteria berikut ini:
 - a. Harus diberikan segera setelah atau pada saat penilaian dilakukan.
 - b. Harus rinci dan bisa difahami oleh siswa.
 - c. Harus difokuskan pada unsur performa yang dianggap berhasil dan pada kesalahan-kesalahan yang harus dikoreksi.
 - d. Harus memberikan saran-saran remedial untuk memeperbaiki kesalahan-kesalahannya.
 - e. Dan harus positif serta memberikan petunjuk untuk mengembangkan performa maupun penilaian oleh diri sendiri siswa.

Dalam penilaian performa, tujuan yang paling utama adalah untuk mengembangkan kualitas performa. Namun jika kita ingin mengembangkan juga kemampuan siswa untuk melakukan penilaian sendiri, maka kita perlu membantu mengembangkan keterampilan untuk me-

nilaia performa diri sendiri (*self-assessment skills*). Misalnya, kita menyarankan siswa bukan hanya bagaimana memodifikasi performa tetapi juga cara-cara untuk mengecek hasil modifikasi terhadap performa serta cara untuk menentukan perbaikan di masa-masa yang akan datang. Pertanyaan penting bagi siswa tentang pengaruh umpan balik (*feedback*) ini adalah *Akankah feedback ini membuat siswa menjadi lebih bergantung pada guru atau lebih mandiri?* Pilihan yang diinginkan tentu saja sangat jelas yani menciptakan siswa lebih mandiri (*independent*) seperti ditunjukkan pada Gambar 2.2.



Gambar 2.2 Peranan Umpan Balik Penilaian
(Diadaptasi dari Gronlund, dkk 2009: 24)

8. *Penilaian yang efektif harus didukung oleh penyekoran serta sistem pelaporan yang komprehensif.* Semua guru sering sekali menggunakan berbagai jenis prosedur penilaian kemudian menentukan nilai berdasarkan skor berdasarkan pada tes pilihan ganda. Sebenarnya, kita memerlukan sistem penilaian dan pelaporan yang merefleksikan penekanan pada penilaian kita. Jika separo dari hasil belajar dinilai dengan tes dan separohnya lagi dinilai dengan penilaian performa, dan sebuah nilai kita gunakan, maka kedua jenis penilaian itu harus memiliki bobot yang sama dan penentuan nilai. Dengan penilaian yang akhir-akhir ini lebih menekankan pada penilaian performa, maka kita memerlukan sistem penilaian dan pelaporan yang lebih rinci untuk menjelaskan performa siswa yang cukup memadai. Pelaporan berdasarkan hasil belajar yang diinginkan dan portofolio hasil kerja siswa semakin lebih penting untuk melaporkan hasil capaian belajar kepada siswa, orang tua dan pihak lainnya. Pelaporan yang terinci semacam ini sejalan dengan prosedur penilaian yang lebih rinci yang tengah digunakan.

Mengingat nilai huruf masih diperlukan untuk tujuan-ujuan tertentu (misalnya di universitas atau perguruan tinggi), maka kita perlu menggunakan bukan hanya nilai huruf akan tetapi juga penjelasan yang lebih rinci. Dalam setiap kesempatan, prosedur penentuan penilaian akhir dan pelaporan harus merefleksikan dan mendukung prosedur penilaian yang digunakan, harus dijelaskan kepada semua siswa pada awal pembelajaran, dan harus memberikan umpan balik secara berkala kepada para siswa berkenaan dengan perkembangan belajar mereka.

2.7 PENILAIAN ACUAN NORMA (*NORM-REFERENCED*) DAN PENILAIAN ACUAN KRITERIA (*CRITERION-REFERENCED*)

Ada dua manfaat penilaian prestasi: (1) untuk menentukan ranking yang relatif (tidak mutlak) kepada siswa atau (2) untuk memberikan deskripsi tentang tugas-tugas belajar yang bisa siswa lakukan atau tidak bisa lakukan. Hasil dari penilaian jenis pertama diinterpretasikan dalam posisi masing-masing siswa di antara siswa-siswa lainnya (misalnya, "Dia berada pada posisi tertinggi ketiga dari 40 siswa.") Metode untuk menginterpretasikan performa siswa macam ini dinamakan interpretasi acuan norma (*Norm-Referenced interpretation*). Hasil dari penilaian jenis kedua dinyatakan dalam istilah pengetahuan dan keterampilan khusus yang bisa ditunjukkan

oleh setiap siswa (misalnya, “Dia bisa mengidentifikasi ide utama dalam setiap paragraf; dia bisa membedakan fakta dan opini.”) Metode untuk menginterpretasikan hasil penilaian semacam ini dikenal dengan istilah interpretasi acuan kriteria (*Criterion-Referenced*). (lihat Bagan 2.1). Kedua metode untuk menjelaskan hasil penilaian ini sangat berguna. Metode yang pertama menunjukkan bagaimana performa seseorang siswa dibandingkan dengan performa siswa lainnya. Sedangkan metode yang kedua menunjukkan dengan istilah performa khusus apa yang bisa dilakukan oleh seseorang siswa tanpa merujuk kepada performa siswa lainnya.

Tabel 2.1 Istilah yang Mirip dengan Interpretasi Acuan Kriteria

<i>Domain-Referenced Interpretation</i>	: Hasil penilaian diinterpretasikan dalam istilah seperangkat tugas-tugas yang berkaitan dan didefinisikan secara jelas (yang disebut <i>domain</i>). Maknanya mirip dengan <i>criterion-referenced interpretation</i> tetapi istilah itu jarang digunakan, meskipun istilah itu lebih deskriptif.
<i>Content-Referenced Interpretation</i>	: Secara esensi maknanya sama dengan <i>Domain-Referenced Interpretation</i> apabila isi domain didefinisikan secara luas yang mencakup tugas-tugas yang merepresentasikan baik isi maupun proses (yakni reaksi terhadap isi). Penggunaan istilah ini menurun dan digeser oleh penggunaan istilah <i>criterion-referenced interpretation</i> .
<i>Objective-Referenced Interpretation</i>	: Hasil penilaian diinterpretasikan dengan istilah masing-masing tujuan khusus yang direpresentasikan oleh seperangkat istilah tes. Metode ini sering disebut <i>criterion-referenced interpretation</i> , tetapi penunjukkan yang lebih terbatas lebih disukai di mana interpretasinya terbatas hanya pada masing-masing tujuan terpisah.

Pada dasarnya, istilah *norm-referenced* (acuan norma) dan *criterion-referenced* (acuan kriteria) hanyalah menunjukkan metode untuk menginterpretasikan hasil penilaian. Jadi, kedua metode interpretasi hasil penilaian tersebut bisa digunakan untuk penilaian yang sama. Misalnya, kita bisa mengatakan, “Yeni mengungguli 90% siswa (interpretasi acuan norma) dengan cara menjawab dengan benar 25 dari 30 soal tentang *Reading Comprehension* (interpretasi acuan kriteria). Namun kedua jenis interpretasi itu akan lebih bermakna manakala suatu penilaian dirancang secara khusus

untuk salah satu jenis interpretasi. Secara umum, interpretasi acuan norma akan lebih dipermudah dengan rentangan nilai yang cukup luas sehingga tingkat perbedaan reliabilitas bisa dibuat di antara para siswa yang memiliki tingkat prestasi yang bervariasi. Sedangkan interpretasi acuan kriteria akan sangat dipermudah oleh tugas-tugas penilaian yang memberikan deskripsi performa siswa yang cukup rinci. Dalam pengujian, hal ini bisa mengandung arti setiap tugas mengandung sejumlah butir soal. Sedangkan dalam penilaian performa, hal ini menunjukkan tugas-tugas mana yang bisa dikerjakan oleh seorang siswa dan tugas-tugas mana yang tidak bisa dikerjakan.

Tabel 2.2 Ringkasan perbedaan antara dua pendekatan uji capaian pembelajaran

	Penilaian acuan norma (Norm-referenced testing)	Penilaian berpatokan Kriteria (Criterion-referenced testing)
Penggunaan utama	Uji survey	Uji penguasaan
Penekanan utama	Mengukur perbedaan capaian individu	Menjelaskan tugas yang bisa dikerjakan oleh siswa
Interpretasi hasil	Membandingkan performa dengan performa siswa lain	Membandingkan performa dengan domain capaian yang sudah ditentukan
Cakupan isi	Biasanya mencakup bidang capaian yang luas	Difokuskan pada tugas-tugas belajar yang terbatas
Sifat rencana pengujian	Umumnya menggunakan tabel kisi-kisi	Lebih mengutamakan kisi-kisi domain yang rinci
Prosedur pemilihan butir soal	Butir soal yang dipilih adalah yang bisa memberikan perbedaan maksimal di antara para siswa (untuk memperoleh ranking yang reliabel). Butir soal yang sangat mudah biasanya dibuang dari tes.	Mencakup semua butir soal yang diperlukan untuk menjelaskan performa secara cukup. Tidak ada perubahan terhadap soal yang sulit juga tidak mengeliminasi butir soal yang mudah untuk meningkatkan rentangan skor.

	Penilaian acuan norma (Norm-referenced testing)	Penilaian berpatokan Kriteria (Criterion-referenced testing)
Standar Performa	Tingkat performa ditentukan oleh <i>posisi relatif</i> dalam kelompok tertentu yang sudah ditentukan (misalnya tingkat ke 5 dalam kelompok yang terdiri dari 20 siswa)	Tingkat performa umumnya ditentukan oleh standar <i>mutlak</i> (misalnya menunjukkan penguasaan dengan cara mendefinisikan 90% dari istilah teknis).

Tabel di atas mengandung ringkasan perbedaan antara dua pendekatan dasar terhadap uji capaian pembelajaran.

Meskipun interpretasi acuan norma dan interpretasi acuan kriteria berlaku untuk semua jenis penilaian, namun perbedaannya akan tampak dengan sangat nyata manakala diterapkan pada pengujian. Ringkasan karakteristik umum dari tes yang dirancang khusus untuk masing-masing jenis interpretasi itu bisa dilihat pada Tabel 2.2. Tetapi harus dicamkan bahwa ini adalah masalah penekanan. Misalnya penilaian acuan norma biasanya, tetapi tidak selalu, digunakan untuk menyurvei capaian hasil belajar yang sangat luas, sedangkan penilaian acuan kriteria umumnya, tetapi tidak selalu, digunakan untuk uji penguasaan. Secara sepintas, perbedaan karakteristik dari masing-masing penilaian tersebut bisa dilihat pada Tabel 2.2.

2.8 RANGKUMAN

1. Metode penilaian sangat bervariasi namun bisa dikelompokkan menjadi tes pilihan, tes isian, penilaian performa dengan jawaban terbatas, atau penilaian performa dengan jawaban yang diperluas.
2. Penilaian siswa perlu diperluas sehingga bisa mencakup penilaian hasil belajar yang lebih kompleks dari pada penilaian yang terjadi di masa lalu.
3. Tes pilihan (misalnya pilihan ganda) adalah jenis tes yang terendah dalam hal realisme dan kompleksitas dari tugas yang nilai, tetapi memerlukan waktu yang sangat singkat untuk menyelenggarakannya serta bisa diseekor dengan cepat dan objektif.
4. Tes isian (misalnya tes esai) merupakan jenis tes yang lebih tinggi dalam hal realisme dan kompleksitas tugas yang akan diukur (misalnya kemampuan untuk melacak asal usul suatu peristiwa, kemampuan

untuk mengintegrasikan, serta kemampuan untuk menyatakan gagasan) dari pada tes pilihan, namun tes macam ini memerlukan waktu lebih lama untuk menerapkannya serta lebih sulit untuk disekor.

5. Penilaian performa, baik yang memerlukan jawaban terbatas maupun jawaban yang diperluas, bisa dirancang dengan tingkat realisme yang tinggi (misalnya masalah-masalah dalam dunia nyata) yang difokuskan pada tugas-tugas belajar yang sangat kompleks, namun penilaian semacam ini memerlukan waktu yang sangat lama untuk menerapkannya, dan penilaiannya bersifat subjektif.
6. Kebijakan penilaian yang bagus biasanya menggunakan metode yang paling efisien untuk menilai masing-masing hasil belajar yang diinginkan selama penilaian itu cocok, tetapi jangan mengabaikan hasil belajar yang kompleks hanya karena metode penilaian memerlukan waktu lebih lama untuk menyelenggarakannya dan lebih sulit untuk disekor atau dinilai.
7. Hasil penilaian bisa diinterpretasikan dengan cara membandingkan performa siswa dengan performa siswa lainnya (*normed-referenced*) atau dengan menjelaskan performa siswa tentang seperangkat tugas yang didefinisikan dengan jelas (*criterion-referenced*).
8. Penilaian efektif memerlukan: (a) konsep yang jelas tentang hasil belajar yang diharapkan, (b) berbagai prosedur penilaian yang sesuai dengan pembelajaran, (c) sampel tugas yang cukup, (d) prosedur yang adil untuk setiap orang, (e) kriteria untuk menilai keberhasilan, (f) umpan balik yang tepat waktu dan rinci terhadap siswa, dan (g) sistem pengambilan nilai akhir dan sistem pelaporan yang sesuai dengan program penilaian.
9. Interpretasi acuan kriteria (*criterion-referenced*) terutama sekali penting untuk penggunaan hasil penilaian dalam pembelajaran, tetapi interpretasi acuan norma (*normed-referenced*) mungkin sangat diperlukan untuk pemilihan dan pengelompokan keputusan.
10. Dalam beberapa kasus, baik interpretasi acuan norma maupun interpretasi acuan kriteria mungkin digunakan untuk penilaian yang sama.

Bab 3

MERANCANG TES ESAI

Tujuan bab ini adalah agar mahasiswa:

1. bisa mengidentifikasi perbedaan antara tes tulis dan tes lisan
2. bisa membedakan contoh tes tulis yang dan yang jelek.
3. mampu merancang tes esai,
4. mampu menginterpretasikan data,
5. bisa mengidentifikasi kelebihan dan kelemahan tes esai,
6. bisa menyekor tes esai,
7. dan mengatasi kelemahan yang terdapat pada tes esai.

3.1 BENTUK TES

Bentuk tes dapat dibedakan menjadi dua macam: (1) tes tertulis dan (2) tes lisan.

- 1) Yang dimaksud dengan *Tes Tertulis* pada bab ini adalah sekumpulan item pertanyaan atau pernyataan yang dirancang baik oleh guru maupun evaluator secara sistematis, guna memperoleh informasi tentang siswa. Tes tertulis pada dasarnya tidak bisa digunakan secara efektif untuk mengevaluasi keterampilan psikomotor siswa. Akan tetapi, tes tertulis dapat digunakan untuk mengevaluasi prinsip-prinsip yang menyertai keterampilan termasuk keterampilan kognitif, afektif, dan psikomotorik.
- 2) *Tes lisan* merupakan sekumpulan butir pertanyaan atau pernyataan yang disusun secara terencana, diberikan oleh seorang guru kepada para siswanya tanpa melalui media tulis. Menurut Kusuma (2016: 127),

tes lisan dapat digunakan secara efektif pada kondisi tertentu, misalnya apabila jumlah siswa kecil (kelompok siswa yang praktik laboratorium) atau sekelompok kecil siswa yang memerlukan tes remedi. Menurutnya, tes lisan ini sebaiknya berfungsi sebagai tes pelengkap, setelah tes utama dalam tes bentuk tertulis dilakukan. Sebetulnya tes lisan sangat penting dan sangat utama untuk menguji kemampuan berbicara (*speaking*) siswa dalam mata pelajaran keterampilan berbahasa, terutama keterampilan berbicara baik di tingkat sekolah menengah maupun di perguruan tinggi.

Tes tertulis termasuk sebagai kelompok evaluasi yang menggunakan *papers and pencils*. Berdasarkan bentuk butir soalnya, tes tulis dapat dibedakan menjadi dua macam, yaitu tes esai dan tes objektif. Pada bab ini pembahasan akan difokuskan pada tes esai.

Cross (2013) menyatakan bahwa paling sedikit ada tujuh langkah persiapan yang perlu guru lakukan agar tujuan evaluasi yang telah ia rancang bisa tercapai dengan baik. Ketujuh langkah tersebut terdiri dari: 1) mengidentifikasi tujuan pembelajaran yang hendak dievaluasi, 2) mengembangkan kisi-kisi kerja atau tabel spesifikasi yang menunjukkan presentase butir-butir untuk setiap tujuan dan cakupan isi, 3) mendaftar semua isi pelajaran yang tercakup dalam silabus dan yang telah diberikan selama proses pembelajaran, 4) memilih atau mengkonstruksi butir-butir soal dan menyusunnya dalam sebuah tes, 5) menyelenggarakan tes kepada siswa dengan menggunakan butir-butir soal yang telah disusun, 6) menganalisis hasil tes yang telah dilakukan, dan 7) membuat laporan sebagai masukan bagi para pengambil keputusan.

Tes esai secara ontologi dapat didefinisikan sebagai salah satu bentuk tes tertulis, yang susunannya terdiri atas butir-butir pertanyaan yang masing-masing mengandung permasalahan dan memerlukan jawaban siswa melalui uraian-uraian kata yang merefleksikan kemampuan berpikir mereka. Tes esai dapat juga disebut sebagai tes yang menggunakan pertanyaan terbuka, di mana dalam tes tersebut siswa diminta menjawabnya sesuai dengan pengetahuan yang mereka miliki. Jika dilihat dari segi jawaban yang siswa berikan, tes esai menurut Gronlund (2010) dapat dibedakan menjadi dua macam, yaitu: pertama tes esai dengan jawaban panjang, dan kedua tes esai dengan jawaban singkat.

Tes esai memerlukan jawaban panjang, manakala dalam aplikasinya tes tersebut memerlukan jawaban siswa secara luas. Evaluator dalam hal ini, memberikan kesempatan kepada para siswa untuk memberikan jawaban secara tuntas, lengkap dan jelas. Sedangkan pada sisi *siswa*, mereka juga diberikan kesempatan untuk mengorganisasi penuangan gagasan tersebut menjadi satu kesatuan yang utuh sehingga mudah dipahami oleh pembaca.

Tes esai dengan jawaban panjang semacam ini dirancang oleh para evaluator untuk mengidentifikasi kemampuan siswa dalam menuangkan gagasan dalam satu kesatuan yang komprehensif, koheren, dan sistematis sehingga memberikan jawaban yang jelas. Jawaban tes esai yang tidak membatasi ide-ide yang dituangkan oleh siswa merupakan tes yang disusun untuk tujuan tertentu. Misalnya pada tes tertulis ujian tahap akhir, yakni ujian skripsi (untuk mahasiswa S1), ujian tesis (untuk mahasiswa S2), dan ujian disertasi (untuk mahasiswa S3), di perguruan tinggi, di mana mahasiswa dituntut untuk menjawab pertanyaan secara komprehensif dan mendalam. Di samping itu, tes ini juga diberikan kepada siswa di sekolah untuk mengetahui penguasaan pengetahuan yang di dalamnya mengandung unsur-unsur kedalaman dan keluasan pengetahuan yang siswa miliki. Di samping itu tes esai digunakan juga untuk mengetes kemampuan manulis (*writing*) pada jurusan bahasa baik di Sekolah Menengah Atas maupun di Jurusan Bahasa.

Tes esai memerlukan jawaban terbatas adalah manakala siswa dalam menjawab pertanyaan hanya diminta menguraikan ide-idenya secara singkat dan tepat sesuai dengan *space* atau ruang yang disediakan. Jawaban pertanyaan tes esai terbatas ini biasanya mengarah kepada jawaban yang lebih spesifik dan lebih pasti seperti kunci jawaban yang telah dibuat evaluator.

3.2 CARA MENGKONSTRUKSI TES ESAI

Bentuk pertanyaan atau butir tes esai dapat dikonstruksi dengan menggunakan kata bantu pertanyaan tertentu yang mengandung unsur singkatan 4 W + 1 H, yakni *where* (di mana), *who* (siapa), *what* (apa), *why* (mengapa), dan *how* (bagaimana). Di samping itu, pertanyaan esai direncanakan secara sistematis untuk mendorong para siswa agar memiliki kemampuan untuk mengekspresikan gagasan mereka dengan menggunakan bahasa atau kata-

kata mereka sendiri, menggunakan informasi dari pengetahuan mereka sendiri, kemudian menuangkannya secara bebas dalam lembaran jawaban yang ada. Kemampuan mengekspresikan gagasan siswa sendiri itulah yang sebenarnya merupakan kelebihan dari tes esai.

Beberapa contoh tes esai di antaranya dapat dilihat berikut ini.

- a. Sebutkan empat alasan mengapa merokok tidak sehat bagi kesehatan!
- b. Jelaskan apa yang dimaksud dengan *intrinsic* dan *extrinsic motivation*?

Kegagalan yang sering terjadi, yang berasal dari para evaluator adalah guru atau evaluator dalam mengkonstruksi pertanyaan tidak dapat memotivasi para siswa agar mengungkapkan kemampuannya dalam menuangkan ide-ide mereka. Berikut adalah beberapa contoh item pertanyaan esai yang dirasakan masih kurang baik.

Contoh jelek : Mengapa para siswa lebih menyukai bersekolah di kota?

Contoh lebih baik : Nyatakan tiga macam alasan yang dapat menerangkan, mengapa para siswa lebih menyukai bersekolah di kota?

Berkaitan dengan tes esai ini ada lima hal yang perlu diperhatikan oleh setiap guru atau pembuat naskah soal tes esai: Merancang tes esai, menginterpretasikan data, kelebihan dan kelemahan tes esai, menyekor tes esai dan mengatasi kelemahannya. Masing-masing dari kelima hal tersebut akan diuraikan secara singkat pada bagian berikut ini.

1. Merancang Pertanyaan Esai

Ada dua hal penting yang perlu diperhatikan oleh para evaluator guna meningkatkan kualitas pertanyaan esai sebagai alat ukur hasil belajar yang kompleks. Kedua hal tersebut adalah: a) bagaimana merancang pertanyaan esai yang mengukur perilaku yang direncanakan, dan b) bagaimana menyekor jawaban yang diberikan oleh siswa. Berikut adalah paparan tentang kiat dalam menyusun tes esai yang dimaksud.

- (1) Pertanyaan esai hendaknya difokuskan pada materi pembelajaran yang tidak dapat diungkap dengan bentuk tes lain, misalnya dengan tes objektif. Ada beberapa faktor penting dalam proses belajar mengajar, yang hanya bisa diungkapkan oleh tes esai, di antaranya pembelajaran

- yang kompleks, organisasi materi, integrasi penyusunan jawaban, dan ekspresi penuangan ide dari pemikiran siswa ke dalam bentuk jawaban soal. Itu semua menjadikan tes esai tetap menjadi pilihan para guru atau para evaluator.
- (2) Butir pertanyaan hendaknya diformulasikan sedemikian rupa sehingga bisa mengungkap perilaku spesifik yang diperoleh dari pengalaman hasil belajar. Tes yang direncanakan oleh guru, baik tes objektif maupun tes esai perlu tetap mengukur penilaian tujuan instruksional. Pertanyaan yang tidak mengarah pada tujuan instruksional sebaiknya dikesampingkan lebih dahulu. Contoh: Tulis jawaban secara lengkap, jelas, dan komprehensif tentang alasan mengapa seorang siswa perlu menguasai dasar-dasar untuk berkomunikasi dengan bahasa Inggris.
 - (3) Butir-butir pertanyaan tes esai sebaiknya jelas dan tidak menimbulkan kebingungan sehingga para siswa dapat menjawab dengan tidak ragu-ragu. Menggunakan kata-kata yang spesifik, seperti *terangkan*, *bandingkan*, *buktikan*, *nyatakan*, *jelaskan*, *gunakan*, dan sebagainya.
 - (4) Cantumkan dengan jelas petunjuk waktu pengerjaan untuk setiap pertanyaan, agar para siswa dapat memperhitungkan kecepatan berfikir, menulis, dan menuangkan ide sesuai dengan waktu tersebut; pertimbangan waktu hendaknya didasarkan pada tingkat kesulitan setiap pertanyaan.
 - (5) Pada saat merancang sejumlah pertanyaan esai, guru hendaknya menghindari penggunaan pertanyaan pilihan. Pertanyaan pilihan biasanya terletak pada kalimat instruksi pengerjaan pada awal tes. Misalnya "Pilih empat soal dari lima pertanyaan yang tersedia". Penggunaan pertanyaan pilihan mungkin akan mempengaruhi reliabilitas tes esai yang direncanakan.

2. Menginterpretasikan Data

Tes esai memiliki banyak kelebihan, salah satunya dalam evaluasi pendidikan adalah bahwa tes esai merupakan tes yang memiliki kemampuan dalam menginterpretasi data melalui jawaban yang diberikan oleh para siswa. Hal ini sesuai dengan kehidupan manusia termasuk siswa yang tidak lepas dari kegiatan interpretasi, terutama yang berkaitan dengan kegiatan manusia pada umumnya, seperti membaca, melihat, dan mendengar. Oleh karena itu, guru mempunyai tanggung jawab moral kepada para siswanya

dalam memberikan kemampuan, baik melalui materi pelajaran maupun pengalaman evaluasi para siswa yang akan membantu meningkatkan keterampilan dalam menyusun jawaban terhadap pertanyaan esai. Butir soal esai direncanakan mempunyai tingkat yang bervariasi. Siswa diharapkan mampu mengakomodasi beberapa situasi. Pertanyaan yang kompleks membantu para siswa secara kelompok mengembangkan kemampuan intelektual, sedangkan pertanyaan yang sederhana, diharapkan dapat membantu mereka yang kurang pandai.

Pada umumnya, pertanyaan esai memberikan cara untuk mengetes penerapan asas ilmu pengetahuan. Penguasaan asas pengetahuan memberikan dasar-dasar yang diperlukan peserta didik ketika berinteraksi dengan lingkungannya. Oleh karena itu, pertanyaan esai sebaiknya dirancang dari tahap yang paling sederhana sampai tahap aplikasi atau tahap yang lebih tinggi termasuk sintesis dan evaluasi. Pada bidang pendidikan bahasa, pendidikan umum, maupun di bidang perilaku, tes esai telah digunakan untuk mengungkapkan keterampilan menginterpretasikan data dari berbagai kegiatan praktik maupun kegiatan eksperimen di laboratorium.

Berikut ini beberapa contoh pertanyaan tes esai yang sangat berhubungan dengan pemahaman konsep-konsep psikologi dalam pendidikan bahasa.

- a. Terangkan secara jelas peranan kepribadian siswa dalam meningkatkan kemampuan belajar mereka!
- b. Identifikasi macam-macam karakteristik siswa yang berkepribadian *embroidery* dan *skeleton*?
- c. Berikan tiga alternatif kegiatan agar siswa yang berkepribadian *phelgmatic* yang cenderung kurang percaya diri berubah menjadi siswa yang percaya diri !
- d. Jelaskan macam-macam kegiatan yang bisa meningkatkan motivasi siswa untuk berbicara dalam Bahasa Inggris dengan penuh percaya diri!

Mendeskripsikan atau memberikan ilustrasi merupakan salah satu aplikasi prinsip pengetahuan. Pada proses pengajaran, ilustrasi merupakan bagian yang tidak dapat dipisahkan pada tingkat komprehensif dari proses kognitif yang sering tumpang tindih (*overlapping*) dengan tingkat aplikasi.

Para guru sekolah menengah sering menggunakan ilustrasi dalam proses pengajaran untuk meningkatkan usahanya dalam pencapaian tujuan intruksional yang telah ditetapkan.

- a. Berikan contoh pendidikan karakter yang bermanfaat dalam meningkatkan kepemimpinan seseorang anak!
- b. Gambarkan secara diagram kaitan antara populasi dan sampel dalam kaitannya dengan sebuah penelitian yang dilakukan di sekolah menengah!
- c. Terangkan konsekuensi apakah yang akan terjadi apabila mahasiswa Pendidikan Bahasa Inggris jarang berpraktik berbicara dengan menggunakan Bahasa Inggris!
- d. Berikan ilustrasi yang memberikan penjelasan antara konsep validasi dan reliabilitas suatu instrumen!

Melakukan perbandingan (komparasi) antara dua situasi merupakan sarana penggunaan asas pengetahuan. Komparasi merupakan proses yang positif dan mudah dipahami, baik oleh guru maupun peserta didik di kelas. Dengan komparasi para guru dapat mengidentifikasi bagaimana dua subjek atau objek memiliki persamaan, di samping itu dengan komparasi seorang guru dan para siswanya juga dapat melihat perbedaan di antara subjek atau objek yang sejenis. Tes esai dapat mengungkapkan usaha komparasi dua objek tersebut. Berikut ini ditampilkan beberapa contoh tugas dalam tes esai.

- a. Bandingkan apa perbedaan antara siswa yang berkepribadian *extroversion* dengan *introversion*!
- b. Bandingkan manfaat keuangan bantuan operasional sekolah (BOS) di sekolah perkotaan dengan sekolah yang ada di pedesaan!
- c. Identifikasi perbedaan kesejahteraan guru yang berada di daerah kaya dan guru yang berada di daerah miskin!
- d. Komparasikan sikap orang tua dalam hal bimbingan terhadap anaknya di sekolah favorit dengan sekolah biasa!

Manfaat lain dari tes esai adalah bahwa pertanyaan esai dapat digunakan juga untuk menganalisis hubungan sebab akibat. Hubungan suatu kejadian sebagian besar di lingkungan kita memiliki faktor penyebab terjadinya hubungan dua atau lebih fenomena. Faktor fenomena yang memiliki hubungan sebab-akibat dalam konteks penelitian disebut variabel.

Hubungan sebab-akibat dari variabel pendidikan dapat digunakan sebagai basis penyusunan butir pertanyaan esai. Permasalahan dua variabel yang berpasangan dalam konteks hubungan sebab-akibat dapat digunakan untuk mengembangkan konteks evaluasi pendidikan yang baik. Berikut beberapa contoh pertanyaan esai yang berkaitan dengan hubungan sebab-akibat.

- a. Bagaimanakah hubungan tingginya tingkat pengangguran dengan meningkatnya kriminalitas masyarakat dapat terjadi di suatu daerah?
- b. Berikan beberapa kemungkinan alasan mengapa terjadi antara dua anak yang sama-sama berumur 6 tahun berbeda tinggi dan berat badannya!
- c. Buatlah daftar tiga faktor penyebab mengapa mahasiswa pada Program Studi Pendidikan Bahasa Inggris enggan untuk berpraktik berbicara dengan menggunakan Bahasa Inggris!
- d. Bagaimanakah kaitan antara anggaran rendah dengan hasil pendidikan yang kurang bermutu?

Di samping itu, pertanyaan esai dapat juga memberikan sarana atau *means* untuk mengorganisasi fakta dalam urutan yang logis. Mendeskripsikan proses secara berurutan merupakan tugas yang sering dilakukan oleh seorang guru kepada peserta didiknya di kelas. Ilustrasi fakta secara sekuensial ini bertujuan agar para siswa tidak hanya menghafal, tetapi juga dapat mengurutkan secara logis sehingga dapat digunakan untuk dasar pengembangan pengetahuan sejenis di masa mendatang. Selain itu, ilustrasi fakta secara berurutan juga membawa implikasi tertibnya seorang siswa dalam menghadapi pekerjaan yang berkaitan dengan teori yang diberikan. yang lebih penting lagi, tes esai mampu menyajikan deskripsi berurutan yang mendukung tercapainya kaitan antar fakta secara sistematis. Berikut adalah beberapa contoh soal esai yang direncanakan untuk tujuan tersebut.

- 1) Terangkan secara jelas, bagaimana seorang mahasiswa perlu mengatur waktu belajar, selain kegiatan belajar di kampus!
- 2) Langkah-langkah apakah yang perlu diperhatikan agar seorang mahasiswa berhasil dalam kuliah secara tepat waktu?
- 3) Langkah-langkah apakah yang perlu dilakukan untuk merencanakan pengajaran dan proses belajar dengan baik?

- 4) Langkah antisipatif apakah yang perlu dilakukan seseorang yang hendak melakukan pendidikan lanjut di luar negeri agar bisa selesai tepat waktu?

Tidak kalah pentingnya dari paparan yang sudah dibahas di atas adalah meringkas (*summarization*). Meringkas merupakan penggunaan utama dari suatu pertanyaan esai. Untuk dapat meringkas dengan baik, seseorang siswa memerlukan keterampilan yang perlu dikembangkan sejak awal. Seorang guru dapat membantu mengembangkan kemampuan siswanya melalui latihan. Latihan tersebut dimulai dari hal-hal yang paling sederhana, yaitu bertanya secara lisan sampai dengan bentuk bertanya secara tertulis yang butir-butir pertanyaannya disesuaikan dengan bentuk pertanyaan esai.

Kegiatan meringkas pada umumnya berkaitan erat dengan membuat kesimpulan yang selalu diperlukan dalam setiap kegiatan pembelajaran. Dengan membuat kesimpulan dan melakukan generalisasi pada permasalahan yang sejenis, siswa dapat memiliki pengetahuan yang komprehensif tentang materi pembelajaran yang telah guru sampaikan. Untuk mencapai tujuan tersebut guru dapat mengembangkan kemampuan ini melalui tes esai seperti berikut.

- 1) Ungkapkan dengan menggunakan bahasa sendiri tentang apa yang telah anda pelajari berkaitan dengan hubungan antara olah raga dengan kesehatan seorang siswa!
- 2) Jelaskan kesimpulan apakah yang dapat ditarik dari cerita rakyat Loro Jongrang.
- 3) Ringkas secara naratif maksimal tiga alenia, artikel tentang manfaat workshop Penelitian Tindakan Kelas bagi guru-guru di sekolah!
- 4) Kesimpulan akhir apakah yang bisa ditarik dari tampilan film tentang bahaya merokok bagi kesehatan?

3. Kelebihan Dan Kelemahan Esai

Sama seperti jenis tes yang lain, tes esai memiliki keunggulan-keunggulan di samping kelemahan-kelemahan. Dalam proses belajar mengajar di kelas, tes esai memiliki keunggulan yakni dapat digunakan untuk menilai hal-hal yang berkaitan erat dengan beberapa butir positif berikut ini:

- a. Mengukur proses mental para siswa dalam menuangkan ide ke dalam jawaban item secara tepat.
- b. Mengukur kemampuan siswa dalam menjawab pertanyaan yang diajukan melalui kata dan bahasa mereka sendiri.
- c. Mendorong siswa untuk mempelajari, menyusun, merangkai, dan menyatakan pemikirannya secara aktif.
- d. Mendorong siswa untuk berani mengemukakan pendapat serta menyusunnya dalam bentuk kalimat mereka sendiri.
- e. Mengetahui seberapa jauh siswa telah memahami dan mendalami suatu permasalahan atas dasar pengetahuan yang diajarkan di dalam kelas.

Selain kelebihan-kelebihan seperti yang telah diuraikan di atas, ternyata tes esai juga memiliki beberapa kelemahan yang perlu diperhatikan dan diatasi oleh seorang guru, di antaranya sebagai berikut:

- a. Ada kecenderungan pengaruh subjektif yang selalu muncul dalam pribadi seorang guru ketika memeriksa jawaban pertanyaan tes esai. Ini terjadi, terutama ketika telah terjadi hubungan moral yang baik antara para siswa dengan guru.
- b. Pertanyaan esai yang guru atau evaluator susun cenderung kurang mencakup seluruh materi yang telah diberikan.
- c. Bentuk pertanyaan yang memiliki arti ganda sering menyulitkan para siswa sehingga memunculkan asal menjawab, dan jawabannya tidak mengandung kepastian, namun ragu-ragu.

4. Menskor Tes Esai

Memberikan skor tes esai dapat dikatakan mudah dan juga dapat dikatakan sukar. Dikatakan mudah, karena setiap guru pasti merasa bisa menilai jawaban yang diberikan oleh para siswanya termasuk penggunaan jawaban yang berasal dari tes esai, karena dalam pemberian skor pada tes esai tidak ada penjelasan penilaian angka diberikan secara pasti. Sebaliknya, sebagian guru merasakan sukar dalam memberikan skor pada tes esai, karena banyak faktor selalu muncul yang sedikit banyak dapat memengaruhi pengambilan keputusan pada penilaian siswa. Faktor-faktor tersebut di antaranya subjektivitas, pertimbangan, dan pengaruh interaksi antara guru dengan siswa selama proses belajar mengajar berlangsung. Untuk mengatasi adanya pengaruh tiga faktor di atas, berikut ini beberapa petunjuk yang dapat

digunakan sebagai acuan para guru. Dalam pemberian tes esai, seorang guru sebaiknya:

- a. Menyusun kunci jawaban untuk setiap pertanyaan yang mengandung materi penting yang dapat digunakan sebagai acuan dasar ketika melakukan penilaian.
- b. Menentukan nilai dari setiap pertanyaan berdasarkan bobot permasalahan, kompleksitas jawaban, dan waktu yang diperlukan untuk menyelesaikan jawaban.
- c. Memutuskan beberapa butir pengurangan skor penilaian apabila siswa melakukan kesalahan kecil, misalnya kesalahan ejaan, tanda baca, dan penggunaan kata.
- d. Mengevaluasi satu pertanyaan pada semua lembar jawaban, sebelum pindah ke pertanyaan lainnya.
- e. Guna mengecek kesamaan kualitas jawaban, kelompokkan lembar jawaban siswa ke dalam 3-5 tumpukan dengan memerhatikan ranking dari yang tertinggi sampai terendah dan menempatkan lembar jawaban siswa ke dalam tumpukan yang ada atas dasar nilai yang dicapai.
- f. Usahakan guru tidak melihat nama siswa penjawabnya dalam proses penilaian jawaban soal.
- g. Disarankan untuk sering beristirahat untuk mencegah kelelahan dan kejenuhan yang dapat mengakibatkan pemberian skor berubah secara signifikan.

5. Mengatasi Kelemahan Tes Esai

Untuk meminimalkan hal-hal yang membuat lemahnya tes esai, dalam mempersiapkan soal-soal esai, guru hendaknya memerhatikan beberapa pertimbangan berikut:

- a. Sediakanlah waktu yang cukup untuk menyusun pertanyaan dalam setiap soal. Walaupun banyak anggapan bahwa merancang tes esai adalah sangat mudah, karena setiap guru bisa membuat soal-soal esai yang baik, soal esai yang baik memerlukan kecermatan dalam pembuatannya. Selain dilihat dari unsur bahasa juga perlu dilihat aspek substansi dari setiap butir pertanyaannya.

Contoh jelek : Bandingkan siswa yang introspektif dengan siswa retrospektif!

Ontoh lebih baik : Dari segi keterampilan berbahasa (*language skills*), bandingkan kelebihan dan kekurangan siswa yang berkepribadian *choleric* dengan siswa yang *phelgmatic!*

- b. Butir pertanyaan yang direncanakan hendaknya memuat persoalan penting yang telah diajarkan dalam proses belajar mengajar.
- c. Permasalahan yang hendak dirumuskan memiliki arti yang dinyatakan secara eksplisit dalam tujuan instruksional.
- d. Kata-kata yang digunakan dalam pertanyaan hendaknya tidak diambil secara langsung dari buku atau catatan. Para guru atau evaluator hendaknya memodifikasi atau menggunakan parafrase dengan cara menggunakan kata-kata lain yang memiliki arti yang sama agar siswa tidak semata-mata menghafal.
- e. Setiap butir soal sebaiknya dilengkapi dengan kunci jawaban. Membuat kunci jawaban sebaiknya menjadi bagian yang tidak terpisahkan dalam pembuatan pertanyaan esai. Kegagalan membuat kunci jawaban akan membuat reliabilitas jawaban bisa berubah secara signifikan.
- f. Pertanyaan esai yang direncanakan sebaiknya dibuat bervariasi dan bisa mencakup unit-unit mata pelajaran yang telah diajarkan di kelas.

oo0oo-



Bab 4

ALAT UKUR NONTES

Tujuan Bab ini:

Setelah selesai mempelajari bab ini, mahasiswa dirapkan bisa memahami dan mengaplikasikan:

1. Penilaian atau pengukuran yang orientasinya pada evaluasi diri
2. Alat ukur yang tergolong nontes
3. Alat ukur bentuk-bentuk laporan,
4. Alat ukur dengan menggunakan wawancara,
5. Kuesioner,
6. dan portofolio

Pada bab ini, pembahasan akan difokuskan pada penilaian atau pengukuran yang orientasinya pada evaluasi diri, dimana pada kesempatan tersebut, para siswa diberikan kesempatan untuk memberikan jawaban atas dasar relevansi dan kecocokan pertimbangan mereka. Di samping itu, pada subbagian ini juga dibahas alat ukur, yang masih tergolong alat ukur nontes, yaitu alat ukur bentuk-bentuk laporan, alat ukur dengan menggunakan wawancara, kuesioner, dan portofolio. Pada alat ukur nontes ini, prinsipnya adalah pemberian jawaban atas dasar relevansi dan bentuk laporan yang berasal dari pendapat pribadi siswa setelah mereka mengerjakan tugas yang diberikan.

4.1 ALAT UKUR EVALUASI DIRI

Beberapa alat ukur yang diuraikan pada bab ini adalah termasuk alat ukur nontes (Kusuma, 2016). Di antara alat ukur nontes termasuk skala rating, bentuk laporan, dan sosiometrik. Manfaat yang paling utama dari alat ukur nontes ini terletak pada evaluasi hasil pembelajaran yang erat hubungannya dengan kualitas pribadi, dan keterampilan. Kualitas pribadi dan keterampilan ini hanya bisa dievaluasi dengan tepat melalui penampilan sebagai efek penguasaan domain keterampilan. Alat ukur nontes ini juga dapat digunakan dengan tepat untuk menilai, bukan saja kegiatan yang berkaitan dengan proses belajar mengajar, tetapi juga banyak dipakai dalam kegiatan di luar kelas, seperti kegiatan penelitian atau bentuk proyek lain yang dilakukan dalam kaitannya dengan manajemen lembaga pendidikan. Sebagaimana alat ukur lainnya, alat ukur nontes ini, selain dapat menghasilkan nilai atau skor, mereka juga dapat dipergunakan untuk mengevaluasi pengetahuan dalam bentuk:

- 1) produksi,
- 2) kualitas personal yang mencakup nilai, dan
- 3) sikap yang dihasilkan sebagai keterkaitan antara kedua unsur di atas.

Alat ukur *rating* mendeskripsikan dengan jelas setiap derajat karakter objek yang hendak dievaluasi. Derajat tersebut pada umumnya diidentifikasi dalam bentuk angka. Pada alat ukur *rating* ini, siswa yang dirangking mengobservasi dan memberikan *check* pada pilihan *rating* yang tepat bagi dirinya.

Alat ukur *rating*, menurut Freeman (2012) dan Linn & Grondlund (2000) dapat dibedakan menjadi dua, yaitu 1) model numerik atau *scoring*, dan 2) model *ranking*. Masing-masing model akan dijelaskan lebih rinci pada paragraf berikut ini.

4.1.1 Model *Scoring*

Alat ukur nontes jenis *scoring*, pada umumnya digunakan oleh para guru atau para evaluator untuk mengevaluasi siswa dengan model titik, tingkat, atau pada skala dengan acuan langsung. Para siswa, dalam hal ini tanpa dibandingkan dengan siswa lain dalam kelasnya, mendapatkan hasil penilaian mereka. Jika alat *rating* digunakan untuk *me-rating* para pekerja

dalam suatu perusahaan, maka posisi pekerja adalah juga secara individu tanpa membandingkan dengan anggota pekerja dalam kelompoknya. Gambaran skala *rating* dengan model skor dapat dilihat pada gambar berikut, yakni gambaran penampilan dalam model *scoring* siswa dalam berpidato.

Penampilan Anda Dalam Berpidato (1) (2) (3) (4) (5)

Gambar 4.1 Skala Rating Model Skor Atau Numerik

Keterangan

- (1) = Sangat tidak memuaskan
- (2) = Tidak memuaskan
- (3) = Sesuai standar
- (4) = Memuaskan
- (5) = Sangat memuaskan

4.1.2 Model Ranking

Alat ukur nontes dengan model *rating* dikatakan menggunakan tipe *ranking*, jika alat ukur *rating* tersebut mengukur karakteristik siswa yang diasosiasikan dalam grup tunggal atau di-*rating* dengan membandingkan satu siswa dengan siswa lain dalam kelompoknya. Gambaran alat ukur nontes dengan menggunakan *ranking* dapat dilihat seperti pada gambar berikut yang me-*rating* kinerja siswa dalam kerja kelompok.

Kinerja Anda Dalam Kerja Kelompok

Tidak memuaskan	memuaskan	sangat memuaskan
Sesuai standar	super memuaskan	

Gambar 4.2 Skala Rating dengan Ranking

Pada alat ukur *rating* dengan dengan jenis *ranking*, para guru sebagai evaluator mengatur nama siswa secara berurutan, dengan mempertimbangkan status atau posisi siswa dalam karakter spesifik yang diperlukan, misalnya tertinggi, rerata, dan terendah. Evaluator kemudian menempatkan siswa pertama sebagai individu yang memiliki *ranking* tertinggi, dan menempatkan siswa lain dengan cara berurutan, tertinggi, rerata, rendah, dan seterusnya.

Alat ukur *rating* sering digunakan dalam evaluasi pendidikan. Bila dipandang dari segi fisik, alat ukur *rating* dapat dikelompokkan ke dalam beberapa bentuk: bentuk daftar, skala *rating*, dan kartu skor. Alat ukur *rating* tersebut sangat berguna terutama untuk mengevaluasi mutu pribadi. Sedangkan skala *rating* dan kartu skor dapat digunakan untuk mengevaluasi macam-macam keterampilan. Skala *rating* dan kartu skor dibedakan dalam indikator produksi dan penampilan yang disebabkan oleh adanya proses serta perkembangan keterampilan kerja, setelah mereka memperoleh pembinaan dari guru.

Bentuk-bentuk laporan, seperti laporan proyek, *diaries*, *log* atau catatan harian, angket, *anecdotal records*, dan autobiografi juga termasuk sebagai alat ukur nontes. Laporan proyek sering digunakan para guru untuk mengevaluasi pribadi, keterampilan dan pengetahuan tentang *diaries* dan *log* bersama angket dan autobiografi. Bila dicermati secara teliti, laporan proyek akan memberikan informasi yang menunjukkan kualitas pribadi siswa. Oleh karena itu, tidak aneh jika dari alat ukur ini lebih banyak dipakai dari pada alat ukur nontes lainnya.

Teknik audio visual biasanya digunakan bersama dengan teknik lain. Alat ini merupakan alat kelengkapan yang mampu mengangkat alat ukur nontes, terutama dalam mengatasi unsur penampilan yang semula tidak dapat dilihat ketika seseorang yang dinilai menampilkan kemampuannya. Dengan menggunakan teknik audio visual, peristiwa yang terdiri atas beberapa rangkaian gerakan yang merupakan unsur dari penguasaan keterampilan dapat direkam dan diulang kembali. Bahkan, dengan menggunakan *pause* dan diputar secara perlahan, gerakan dalam kesatuan penampilan dapat dilihat kembali dengan gerakan lebih pelan, sehingga dapat diamati kembali dan dievaluasi benar tidaknya unsur gerak dari siswa yang dievaluasi tersebut.

4.2 ALAT UKUR RATING

Secara umum, fungsi utama alat *rating* dapat digunakan untuk:

- 1) mengevaluasi kualitas pribadi siswa, dan
- 2) memberikan penghargaan proses serta produk yang dihasilkan dari hasil pembelajaran siswa.

Pada umumnya, hampir semua alat evaluasi *rating* digunakan untuk melengkapi alat eksplorasi data dengan menggunakan pendekatan observasi. Untuk mencapai tujuan tersebut, seorang guru atau evaluator perlu melakukan pengamatan secara intensif. Alat evaluasi *rating* dengan menggunakan model observasi dapat digunakan untuk memperoleh informasi yang relevan, yang dilakukan baik oleh guru maupun oleh siswa. Alat *rating* ini sangat bagus untuk penilaian diri (*self evaluation*), yaitu pemberian skor atas dasar pilihan yang sesuai dengan diri sendiri siswa.

Pada dasarnya, alat ukur *rating* dapat dikelompokkan menjadi tiga macam bentuk, yaitu daftar cek, skala *rating*, dan kartu skor dengan penjelasan sebagai berikut.

4.2.1 Daftar Cek (*Check List*)

Daftar cek adalah salah satu alat evaluasi yang termasuk alat ukur *rating*. Alat ini banyak digunakan oleh guru karena dua keunggulan, yaitu cara pembuatannya paling sederhana, selain itu penggunaannya juga mudah sehingga dengan sedikit *training*, guru sudah bisa menggunakannya. Secara definitif, *check list is a list of characteristics or kriteria which requires a simple check of whether or not each item on the list has been met* (Cross, 2013 dan Grondlund, 2010) yang berarti "Daftar cek adalah seperangkat karakteristik atau kriteria yang memerlukan jawaban sederhana, misalnya dengan membubuhkan tanda cek (√), apakah setiap butir dalam daftar telah terpenuhi atau belum".

Daftar cek, pada prinsipnya, adalah metode mencatat apakah suatu karakteristik ada atau tidak ada pada suatu subjek atau objek yang dievaluasi. Pada proses penggunaannya, guru menyusun daftar kriteria, kemudian memberikannya kepada siswa. Para siswa selanjutnya mengisi, apakah kriteria atau karakteristik yang diinginkan guru ada atau tidak ada. Jika ada, maka ia memberikan tanda cek ("√"), sebaliknya jika tidak ada, maka siswa akan membiarkannya kosong.

Untuk mengkoordinasikan ada atau tidak adanya karakteristik atau kriteria, guru biasanya memberikan dua pilihan jawaban, yaitu *ya* atau *tidak*. Dalam kondisi tertentu, daftar cek juga dapat dirancang dengan tiga jawaban pilihan: *ya*; *kadang-kadang*; dan *tidak*.

Daftar cek juga memiliki kelebihan dan kelemahan sebagaimana alat ukur evaluasi pembelajaran lainnya. Kelebihan alat ini di antaranya adalah tidak mengharuskan seorang evaluator menggunakan daftar cek atau tidak, tetapi alat ini memberikan kesempatan kepada para guru sebagai perencana untuk memilih secara bijaksana atas penggunaannya. Faktor penentu kapan penggunaan daftar cek dianjurkan pemakaiannya, minimal ada tiga pertimbangan, yaitu apabila penilaian mengandung unsure-unsur seperti berikut:

- a) Prilaku seperti apakah yang hendak dievaluasi dari para siswa.
- b) Apakah informasi yang diperoleh siswa memberi manfaat.
- c) Apakah tepat atau tidak penggunaan daftar cek tersebut.

Apabila ternyata penggunaan daftar cek memberikan informasi yang lebih maksimal, maka dianjurkan untuk dipergunakan. Sebaliknya, apabila dipertimbangkan bahwa ada alat ukur lain yang dapat digunakan secara lebih tepat, maka dianjurkan untuk menggunakan alat ukur lain yang dimaksud.

Daftar cek, pada prinsipnya, dapat dikembangkan dan digunakan sebagai bagian dari kegiatan proses belajar mengajar. Siswa pada kesempatan ini dapat membantu mengembangkan daftar cek, dengan mengidentifikasi kriteria yang hendak dimasukkan. Ini berarti bahwa daftar cek tersebut secara simultan juga dapat dikatakan lebih siap diterima oleh para siswa yang dievaluasi. Karena mereka sejak awal telah dilibatkan dalam mengidentifikasi kriteria-kriteria yang diperlukan.

Manfaat lain dari daftar cek yang pembuatannya melibatkan para siswa, adalah bahwa ternyata juga membawa konsekuensi positif bagi mereka. Mereka telah mengetahui kriteria apa yang dimaksudkan, sehingga para siswa merasakan terlalu mudah dalam menjawab dan dimasukkan pertimbangan diri ke dalam kriteria yang telah disediakan. Implikasi lain yang perlu dipertimbangkan, yaitu bahwa petunjuk dalam daftar cek pada umumnya dibuat dengan bahasa yang lebih sederhana. Sehingga para siswa juga lebih cepat memberikan keputusan untuk menjawabnya dibandingkan dengan alat ukur rating lainnya.

Selain itu, daftar cek mudah dan tidak memerlukan waktu dalam merancanginya. Pada keadaan ini, para guru dapat melakukan dengan

cara mengambil daftar kriteria yang diberikan siswa; mengedit kriteria yang ada untuk memperoleh bentuk yang konsisten; menulis petunjuk bagaimana cara mengisi, kemudian guru dapat melengkapi dengan kolom untuk memberi kesempatan *checking* di antara *ya* atau *tidak*.

Yang lebih penting lagi, daftar cek sangat fleksibel penggunaannya, yakni dapat digunakan dalam situasi yang bervariasi. Misalnya, dalam hampir semua bidang pendidikan, daftar cek yang paling sederhana pun dapat digunakan untuk tingkat umur yang bervariasi, yaitu sejak dari siswa pada sekolah dasar, sampai pada sekolah menengah dengan kemampuan yang variasi.

Contoh:

Tujuan : Menunjukkan kemampuan siswa dalam pembelajaran matematika.

Petunjuk : berilah tanda silang (X) pada kolom yang sesuai dengan kenyataan di kelas.

No	Keterampilan Matematika untuk Tingkat Dasar	Ada	Tidak
1.	Mengidentifikasi nilai pecahan		
2.	Membedakan antara pembilang dan penyebut		
3.	Menjumlahkan dua bilangan pecahan atau lebih		
4.	Mengurangkan dua bilangan pecahan atau lebih		
5.	Menyamakan nilai pecahan ke dalam bentuk decimal		

Daftar cek dapat digunakan untuk mengevaluasi siswa secara berkelompok atau secara individual. Di samping itu, daftar cek juga dapat digunakan secara berulang-ulang dengan situasi yang serupa.

Di samping kelebihan yang sudah diuraikan di atas, daftar cek ternyata juga memiliki beberapa kelemahan yang perlu diperhatikan oleh para guru atau para evaluator. Beberapa kelemahan tersebut, di antaranya sebagai berikut.

- a. Daftar cek tidak langsung memberikan skor atau nilai. Untuk penilaian pribadi atau personal yang tidak memerlukan pengangkatan. Penggunaan daftar cek tidak ada permasalahan, tetapi apabila diperlukan pengangkatan, akan timbul permasalahan yang serius.

- b. Daftar cek hanya member informasi evaluasi yang dangkal. Reliabilitasnya masih dipertanyakan, ketepatan daftar cek pada umumnya tergantung pada tingkat intensif karakteristik tersebut direncanakan dalam mengevaluasi kualitas diri.
- c. Daftar lus hanya tepat untuk mengevaluasi mutu pribadi, proses, dan produk, ketika guru menginginkan siswa yang terlibat langsung dalam evaluasi diri.
- d. Tidak terlalu memerhatikan kesepakatan siswa dalam mengisi jawaban yang ada pada daftar cek.

Daftar cek hanya tepat untuk mengevaluasi mutu personal, partisipasi proses dan proyek. Daftar cek dapat digunakan untuk mengevaluasi, baik oleh siswa maupun guru. Siswa sebagai individual atau dalam kelompok secara bersama. Mutu personal yang dapat dievaluasi dengan menggunakan daftar cek termasuk di antaranya informasi yang ada hubungannya dengan personal, yaitu interes, pendapat, sikap, dan penghargaan.

Contoh:

Tujuan : Menunjukkan kemampuan siswa dalam bekerja sama dengan siswa lainnya.

Petunjuk : Berilah tanda silang (X) pada kolom yang menggambarkan perilaku Anda.

Semacam Apakah Anda?	Kadang-kadang	Tidak
Apakah Anda merasa senang dengan kehadiran siswa lain?		
Apakah Anda ramah pada orang lain yang mendatangi Anda?		
Apakah Anda menyediakan sedikit waktu untuk menceritakan kesulitan pada orang lain?		
Apakah Anda pendengar yang baik?		
Apakah Anda jarang bertengkar dengan teman Anda?		
Apakah Anda mempunyai lebih dari 1 orang teman?		
Apakah Anda mempercayai teman Anda?		

Untuk mendapatkan daftar cek yang baik, seorang guru perlu memperhatikan beberapa faktor penting dalam menyusun daftar cek. Berikut adalah beberapa langkah penting yang dianjurkan untuk mengembangkan dan mendapat daftar cek yang baik.

- (1) Identifikasi dan catat semua kriteria atau karakteristik yang diutamakan; kriteria harus berkaitan langsung dengan kualitas subjek atau objek yang dievaluasi.
- (2) Edit semua item karakteristik secara teliti untuk memberikan konsistensi secara gramatik. Ini berarti bahwa setiap pernyataan. Jika satu item dalam bentuk kalimat atau pernyataan maka semuanya juga dalam bentuk kalimat.
- (3) Kriteria atau perilaku diatur dalam sistematika logis. Item sejenis dikumpulkan sesuai dengan urutan atau sequensial yang dapat diobservasi sesuai dengan keadaan nyata.
- (4) Semua item sebaiknya dinyatakan dalam cara yang sama, misalnya dalam kalimat negative. Kombinasi antara kalimat positif dan negatif. Kombinasi antara kalimat positif dan negative dengan jawaban ya atau tidak akan membingungkan siswa yang dievaluasi.
- (5) Petunjuk atau arahan sebaiknya dinyatakan secara singkat atau ringkas, dan ditempatkan pada posisi sebelum item daftar cek.
- (6) Proses pengecekan pada setiap tindakan/urutan nomor sedapat mungkin dalam sequensial atau berurutan dan dapat dilakukan dengan cara sederhana.

4.2.2 Skala *Rating* (SR)

Skala rating merupakan alat ukur keterampilan yang masih juga tergolong alat ukur nontes. Seperti alat ukur daftar cek, alat ukur ini juga sudah lama digunakan di bidang evaluasi pendidikan. Pada umumnya, alat ukur rating terdiri atas dua bagian, yaitu:

- a) Satu rangkaian karakteristik atau kualitas yang hendak dinilai, dan
- b) Beberapa tipe skala ukur yang menunjukkan tingkat atau derajat atribut subjek atau objek yang ada (Grondlund dan Linn, 2015: 383).

Skala rating merupakan sarana untuk mencapai pertimbangan responden dengan referensi atau acuan pada beberapa karakteristik yang

dimilikinya. Secara historis, alat ini, sudah digunakan pada awal Perang Dunia 1 (Freeman, 1962),

Sebelumnya mereka memformalisasikan dan menggunakan skala dengan metode statistic. Di samping untuk bidang pendidikan, skala rating juga digunakan dalam bidang social dan psikologi yang berusaha mengungkap presepsi atas dasar pilihan responden yang diselidiki. Pada umumnya, skala rating digunakan untuk mengevaluasi kualitas pribadi, produksi, proses, dan proyek yang dihasilkan para siswa. Kualitas pribadi yang dievaluasi dengan penggunaan skala rating, pada umumnya mutu pribadi yang berkaitan erat dengan penguatan hubungan antarpribadi. Skala rating juga dapat digunakan untuk mengevaluasi sikap dan penghargaan yang tidak memerhatikan adanya tingkatan skor atau *grade*.

Skala rating bukan saja sebuah daftar karakteristik, tetapi juga usaha evaluator dalam mendiskripsikan siswa atau responnden dengan karakteristik multitingkat. Ketika beberapa aspek kualitas dinyatakan secara abstrak dan diperkirakan bahwa mereka bervariasi dalam beberapa tingkatan, skala rating merupakan salah satu metode yang tepat untuk digunakan Roid dan Halayana (1982). Kemampuan mengungkap tingkat deskriptif yang multilevel inilah yang membuat skala rating berbeda dengan daftar lain sejenisnya, seperti kartukartu skor. Skor-skor pada daftar cek dengan tanda *check* ("") digunakan untuk menunjukkan deskriptif yang telah diidentifikasi dengan cermat.

Dibandingkan dengan alat ukur sejenis lainnya, skala rating memiliki beberapa kelebihan, diantaranya bahwa perilaku yang diinginkan didefinisikan secara pasti dengan tingkat deskriptif yang menunjukkan dengan jelas tentang harapan pada tingkat tertinggi medium, dan rendah. Ini berarti bahwa baik para guru maupun para siswa dapat dengan cara mudah dan cepat dalam mengidentifikasi perilaku yang paling diinginkan sehingga memberikan keadaan dan motivasi bagi mereka dalam belajar mencapai harapan yang terbaik. Di samping itu, siswa yang mengetahui kriteria yang paling tidak diharapkan maupun yang paling diharapkan sehingga memberikan kesadaran dan motivasi bagi mereka dalam usaha mencapai hasil belajar dengan skor penilaian yang terbaik. Kelebihan lain menurut Roid dan Haladyna (2012: 69) ada tiga macam, yaitu:

- (1) Rancangan dan penggunaan skala rating sederhana,
- (2) Mudah diinterpretasikan, dan
- (3) Hasil rating dapat diadaptasi pada analisis statistika.

Walaupun demikian, alat ukur ini juga memiliki kekurangan. Kekurangan skala rating yang utama adalah pada waktu mengonstruksi item-item memerlukan waktu lebih banyak. Hal ini biasa terjadi, karena penulisan tingkat deskriptif tidak dapat dikerjakan dengan tergesa-gesa. Mereka atau para evaluator maupun para guru harus dapat membuat kalimat pernyataan atau pertanyaan yang tepat dan singkat. Ditambah lagi setiap level harus bervariasi antara satu dengan yang lainnya. Jika skala kontinu perlu ada maka skala rating juga dianjurkan dilakukan *pilot study* sebelum digunakan. Tujuan pilot studi adalah untuk mengurangi semaksimal mungkin kesalahan dan koreksi kebahasaan. Ini semua memerlukan waktu. Karena perlu sinkronisasi antara pernyataan dengan tingkat perilaku yang ingin dipresentasikan dalam skala.

Kelemahan lain dari skala *rating* adalah tumbuhnya harapan siswa dalam menggunakan secara mudah, semudah mencocokkannya dengan alat yang sebenarnya mudah diingat, ketika para siswa atau responden dilibatkan dalam mengkonstruksi skala *rating*. Akibatnya, mereka dapat mengisi secara cepat tanpa melalui pengamatan secara cermat atau teliti.

Ada beberapa cara untuk mengatasi skala *rating* yang fokusnya untuk eksplorasi informasi siswa atas dasar pilihan mereka. Ketiga cara mengatasi tersebut adalah sebagai berikut.

- a. Pada kolom diskripsi perlu diberi tabel dengan angka, tingkat paling rendah diberi angka 1 dan tingkat paling tinggi 5. Sedangkan antara 1-5 dapat disisipkan tingkat-tingkat skala, misalnya 2, 3, dan 4. Skala diberikan agar para siswa dapat lebih menempatkan pilihannya sesuai dengan tingkat yang ada.
- b. Pada setiap karakteristik sebaiknya berdiri bagian kesatuan yang lain. Pengelompokan *karakteristik* yang tidak diatur dengan baik akan berdampak pada siswa yang dievaluasi.
- c. Setiap level sebaiknya berisi kata-kata yang jelas, yang menunjukkan kondisi subyek yang diukur dan tidak membuat kebingungan para siswa.

- d. Tingkatan deskriptif sebaiknya dibuat secara singkat, kurangi pengulangan kata yang sama, disamping itu juga hilangkan kata-kata yang tidak diperlukan dalam membuat item skala *rating*.

Contoh:

Tujuan : Menentukan bagaimana seorang siswa me-*rating* sikap atau perilaku siswa lainnya yang masih berada dalam satu kelas.

Petunjuk : Pada skala *rating* ini terdapat skor dari 1 sampai 5. Skor 1 = sangat jelek (terendah), 2 = jelek, 3 = sedang, 4 = baik, dan 5 = sangat baik atau paling tinggi.

1	2	3	4	5	Skor rating
Sedih, muram	Gembira		Biasanya gembira di setiap waktu		
Beda dengan interes siswa lainnya	Memerhatikan sebagian dengan interes siswa lain		Menghormati/mendukung interes siswa lainnya		
Tidak berusaha memahami nasihat guru	Berusaha memahami nasihat guru atau orang tuanya		Memahami semua nasihat guru atau orang tua lainnya		
Tidak pernah diskusi dengan siswa lainnya	Kadang mendiskusikan hal penting dengan siswa lainnya		Mendiskusikan semua hal penting dengan sesame kawannya		
Mandiri (tidak tergantung orang lain)	Umumnya tergantung siswa lain		Biasa tergantung dengan siswa lainnya		

4.3 KARTU-KARTU SKOR (KKS)

Jenis alat ukur nontes yang ketiga adalah kartu skor. Kartu skor banyak digunakan dalam pendidikan pada umumnya, dan juga pada pendidikan kejuruan. Alat ini sebenarnya sudah lama digunakan oleh para pendidik maupun para evaluator program guna mengevaluasi keterampilan yang mencakup proses dan produk atau kerja proyek yang diselesaikan para siswa di bengkel atau laboratorium. Alat ini muncul untuk penilaian kegiatan yang

berkaitan dengan proses belajar dan mengajar maupun dengan produk yang diajarkan oleh guru selama satu semester atau satu kuartal.

4.3.1 Kelebihan dan Kelemahan Kartu Skor

Seperti alat ukur nontes lainnya, penggunaan kartu skor mempunyai beberapa kelebihan, diantaranya mudah dalam cara pembuatannya. Para guru dengan sedikit *training* tentang penyusunan kartu skor, akan dengan mudah dan cepat dapat menyusun daftar karakteristik yang diinginkan dan menunjukkan skor standar. Ini merupakan kelebihan, sehingga kartu skor sampai sekarang masih digunakan dalam proses evaluasi dengan jenis alat ukur nontes. Setiap kartu skor, sebaiknya direncanakan terdiri atas butir pertanyaan yang jumlahnya tidak lebih dari 12 butir pertanyaan.

Kartu skor dapat digunakan untuk mengevaluasi proses dan produk. Kartu skor akan lebih tepat jika guru yang sama mengerjakan semua proses evaluasi. Kartu skor juga dapat digunakan sebagai bagian dari kegiatan proses belajar mengajar. Dalam hal ini, siswa dilibatkan dalam menyusun kartu-kartu skor dengan memberikan kesempatan untuk membantu mengidentifikasi karakteristik yang diperlukan. Sedangkan kaitannya dengan buku-buku acuan dapat digunakan untuk mendapatkan kata-kata kunci yang bermanfaat dalam menyusun karakteristik. Pada kesempatan lain, siswa juga dapat dilibatkan untuk membantu dalam memutuskan skor baku yang akan ditempatkan pada setiap karakteristik atau item sesuai dengan prioritas yang digunakan.

Keutamaan lainnya adalah bahwa kartu skor dapat menghasilkan skor atau nilai. Karena pada kartu skor diberi kolom yang diisi dengan skor yang ditetapkan oleh guru pengampu atau penilai. Dengan dicantumkan skor ini, orangtua atau siswa dapat mengetahui hasil penelitian dari produk atau proses yang dapat dicapai anaknya di sekolah, sehingga mereka dapat memberikan masukan pada para guru pengampu atau memberikan nasihat seperlunya pada putra-putrinya.

Sebenarnya, hampir semua aspek keterampilan, khususnya produksi dapat dievaluasi dengan penggunaan kartu-kartu skor. Kelemahan kartu skor yang paling menonjol adalah kartu skor kurang tepat untuk menilai kualitas pribadi. Di samping itu, kartu skor kadang-kadang dipertanyakan

objektivitas dan reliabilitasnya. Kesulitan penggunaan kartu skor adalah untuk sampai pada kesepakatan pada setiap butir skor yang harus diberikan pada siswa. Sebagai contoh, jika dalam satu item kompetensi memiliki skor 8, maka akan ada variasi delapan alternatif penilaian yang dapat diberikan oleh guru penilai. Ketidak-sepakatan ini dialami baik oleh guru junior maupun senior. Di samping itu, jika dalam penilaian digunakan secara tim, guru yang belum berpengalaman merasa kesulitan, karena tidak ada informasi yang dapat digunakan untuk acuan dalam menentukan nilai yang telah diberikan. Sebaliknya, bagi guru yang berpengalaman, penggunaan kartu skor juga masih memerlukan acuan tambahan agar memiliki dasar yang lebih rasional, ketika mereka memberikan skor kepada para siswanya.

Seperti halnya menyusun skala rating, merancang kartu skor sangat sederhana. Walaupun demikian, perlu ada tambahan tips atau masukan agar menjadi lebih baik. Tips tersebut antara lain sebagai berikut. *Pertama*, daftar karakter kartu skor sebaiknya dinyatakan secara sederhana dalam frasa atau kalimat yang singkat dan jelas agar tidak menyulitkan siswa. *Kedua*, skor baku perlu dibangun untuk setiap item karakter. Kadang-kadang setiap karakter dalam kartu-kartu skor diberi kartu skor sama, misalnya 10 dari sepuluh daftar item karakteristik. Sebaliknya, sering terjadi bahwa pada setiap item karakteristik dibuat secara profesional, sesuai dengan prioritas dan tingkat kesulitan yang dialami siswa ketika menyelesaikan proyek yang hendak dievaluasi. *Ketiga*, space atau ruang untuk nilai perlu disediakan untuk setiap karakteristik. Dalam hal ini, nilai yang disediakan untuk setiap item karakteristik dan nilai yang dapat dicapai oleh para siswa. Ruang juga dapat disediakan bagi penilaian siswa dan dari para guru. Cara ini digunakan terutama untuk meningkatkan kewajaran atau *fairness* yang dapat diterima, baik oleh guru maupun oleh para siswa sebagai subjek yang dievaluasi.

Tabel 4.1 Contoh Kartu Skor Penelitian Pendidikan

No	Komponen	Kriteria	Skor baku	Skor yang dicapai
1.	Judul	Maksimal 15 kata	5	
2.	Pendahuluan	Masalah penelitian nyata	10	
3.	Rumusan dan pemecahan masalah	Rumusan masalah dalam bentuk pertanyaan	15	

Tabel 4.1 Contoh Kartu Skor Penelitian Pendidikan (Lanjutan)

No	Komponen	Kriteria	Skor baku	Skor yang dicapai
4.	Tujuan	Sesuai dengan rumusan masalah	5	
5.	Manfaat penelitian	Jelas hasilnya	5	
6.	Kajian pustaka	Antara konsep, teori, dan permasalahan memiliki relevansi	15	
7.	Prosedur penelitian	Subyek, tempat dan waktu penelitian jelas	20	
8.	Hasil penelitian	Bermanfaat bagi masyarakat	20	
9	Lain-lain	Jadwal penelitian dan daftar kepustakaan yang relevan	5	
Total			100	

4.4 BENTUK-BENTUK LAPORAN

Para siswa sekolah menengah dan mahasiswa tidak jarang mendapat tugas kegiatan lapangan dan kerja proyek. Kegiatan lapangan dan kerja proyek tersebut mungkin pelaksanaannya sebagian besar di luar kampus dengan diawasi para instruktur yang ditunjuk. Kerja praktik lapangan ini, pada umumnya memerlukan laporan tentang hasil kerja tersebut. Laporan kerja itu menjadi salah satu dasar autentik yang merefleksikan apa yang telah siswa lakukan selama kegiatan lapangan atau kerja proyek. Agar dapat mengevaluasi hasil laporan yang dibuat oleh para siswa, guru perlu memahami beberapa prinsip evaluasi yang berkaitan erat dengan kegiatan lapangan dan kerja proyek tersebut. Prinsip penilaian tersebut perlu tetap mendasarkan pada laporan hasil kerja mereka di lapangan.

Ada tiga macam teknik evaluasi yang sering digunakan untuk menilai bentuk laporan kerja proyek, yaitu laporan proyek, log, dan laporan kegiatan. Pada evaluasi dengan bentuk-bentuk laporan, pada umumnya guru dihadapkan pada informasi seorang siswa yang bersangkutan dibanding dengan penilaian diatas.

4.4.1 Laporan proyek

Praktik industri, magang, dan kerja lapangan merupakan kegiatan penting yang sering dilakukan oleh para siswa di sekolah menengah, terutama mereka yang menekuni bidang keahlian pendidikan kejuruan. Praktik kerja dan semacamnya dapat dilakukan, baik di dalam kampus, misalnya di bengkel, laboratorium, atau di luar kampus, seperti di industri, perusahaan, atau lembaga-lembaga pemerintah maupun swasta yang relevan dengan bidang keahlian siswa. Mereka membuat laporan proyek atau laporan kerja setelah periode waktu tertentu untuk menyelesaikan pekerjaannya. Laporan proyek tersebut biasanya dinilai, baik oleh pihak industri/perusahaan maupun oleh guru pembimbing siswa.

Mengevaluasi kerja proyek dan semacamnya sebaiknya menjadi bagian keahlian bagi para pembimbing yang juga bersumber dari prinsip-prinsip evaluasi nontes. Hasil evaluasi bisa bervariasi, ada yang mendapat skor, dan ada yang tidak mendapatkannya. Untuk yang tidak dengan skor, pada umumnya, penilaian dapat berupa pernyataan *lulus*, *tidak lulus*, atau *memuaskan*, *tidak memuaskan*. Agar dapat mengevaluasi laporan proyek, diperlukan keahlian menilai dari guru maupun pembimbing. Ada beberapa komponen penting yang perlu diperhatikan ketika melakukan penilaian laporan proyek: a) sistematika laporan, b) kelengkapan, dan c) tata tulis laporan.

Di samping itu, pada butir b) berisi 1) keahlian yang diperoleh selama praktik industri atau kerja lapangan, dan 2) substansi laporan. Semua komponen tersebut juga memerlukan perhatian bagi guru pembimbing.

Permasalahan yang sering muncul, di antaranya adalah memberikan penilaian yang adil antara para siswa yang benar-benar melakukan pekerjaannya, dengan siswa yang hanya sekedar berpartisipasi dan tidak memberi kontribusi pada proyek yang dikerjakan. Ada juga siswa yang dalam teori di kelas tergolong lancar dan cepat menguasai materi pembelajaran, tetapi kurang baik dalam melakukan tugas proyek. Selain itu, sering terjadi antar siswa yang aktif dan yang tidak aktif mendapat nilai yang sama dari seorang guru pembimbing. Guru baru menyadari ketika mereka mendapatkan protes atau laporan dari anggota tim lainnya.

Tujuan penilaian laporan proyek, walaupun dikerjakan secara bersama atau tim, penilaian harus tetap berbasis pada kompetensi dan kontribusi siswa secara individual. Dalam kerja tim, sebaiknya guru tidak membandingkan siswa yang satu dengan siswa lain dalam kelompoknya. Agar guru dapat memberikan penilaian yang adil, seorang guru sebaiknya perlu hadir dan melakukan observasi terhadap para siswanya, terutama ketika mereka sedang melakukan tugas atau dalam kegiatan proyek. Berikut ini diberikan contoh komponen laporan proyek dari kegiatan praktik industri yang diterbitkan oleh salah satu Fakultas Teknik di Jakarta yang perlu diperhatikan, ketika para dosen pembimbing atau para instruktur menilai laporan proyek siswa atau mahasiswa.

Nama industri :						
Alamat industri :						
Nama siswa :						
Lama praktik :						
No.	Komponen	Penilaian				Ket
		100-80	79-66	65-56	55-40	
	Sistematika laporan Kelengkapan laporan Tata tulis					
	Keahlian yang diperoleh Substansi laporan					
	Nilai ujian					

Catatan : Butir 1 mempunyai bobot 40%
Butir 2 mempunyai bobot 60%

Kesimpulan nilai laporan dan ujian :

Jakarta, 15 Juni 2015

Dr. X

4.4.2 Laporan Kegiatan dan Buku harian atau log

Bentuk lain alat ukur yang juga tergolong sebagai keluarga alat penilaian laporan adalah alat ukur dalam bentuk laporan kegiatan dan buku harian atau *log*. Secara garis besar, dua macam alat ukur ini dapat digunakan secara efektif dalam tiga hal, yaitu:

- (1) Memberikan kesempatan pada siswa untuk melakukan evaluasi diri sendiri atau (*self evaluation*),
- (2) Mudah pembuatannya, dan
- (3) Mudah diadaptasi pada hampir semua kegiatan pendidikan, terutama pada tingkat menengah ke atas.

Laporan kegiatan digunakan untuk menilai tugas khusus yang dapat dikerjakan di rumah, di sekolah, atau di masyarakat yang menjadi tugas para siswa, sedangkan *log* atau buku harian merupakan catatan harian yang dilakukan secara kontinu yang mungkin dicatat dalam periode waktu tertentu. *Log* ini semula merupakan catatan harian yang digunakan dalam kegiatan dikapal laut, tetapi dalam perkembangan pendidikan, *log* juga digunakan untuk mencatat kegiatan sekolah yang dilakukan secara kontinu. Kedua tipe alat penilai ini dapat digunakan untuk mengevaluasi kegiatan siswa, dan keduanya merupakan alat ukur yang pelaksanaannya memerlukan penguasaan ide atau reaksi siswa sesuai dengan apa yang telah mereka lakukan selama mendapat tugas dari guru.

Contoh Laporan Kegiatan

Tujuan : Mengevaluasi penggunaan waktu praktik laboratorium bahasa siswa SMA.

Petunjuk : Buatlah daftar secara berurutan kegiatan yang telah dikerjakan selama berpraktik di lab bahasa.

Kegiatan yang dikerjakan	Waktu persiapan	Waktu memasak	Waktu dihidangkan	Waktu membersihkan
1.				
2.				
3.				
4.				
5.				

Desain laporan kegiatan dan buku harian dapat bervariasi. Pada umumnya, perbedaan yang direncanakan bertujuan untuk menyesuaikan dengan tujuan penggunaan kedua alat tersebut. Ada catatan laporan yang berisi garis besar komponen kegiatan, tetapi ada juga laporan kegiatan dan *log* yang berisi semua bagian penting dari tugas yang diperlukan.

Pada laporan kegiatan dan *log* bagian atas sering kali disediakan petunjuk laporan kegiatan. Petunjuk itu bermanfaat bagi para pembimbing atau guru lain yang terlibat agar dapat dengan mudah mengisinya. Buku harian atau *log* merupakan alat evaluasi yang sangat efektif untuk mencatat perkembangan anak, hubungan keluarga, dan perkembangan pengelolaan kelas atau sekolah. Laporan kegiatan sangat baik untuk kegiatan manajemen sekolah yang mencakup penggunaan keuangan dan penggunaan ruang dalam proses belajar mengajar.

Penggunaan laporan kegiatan dan *log* pada prinsipnya sangat fleksibel. Hal ini menguntungkan karena hampir semua kegiatan pendidikan yang berupa praktik industri, kerja lapangan, kegiatan darmawisata dapat dievaluasi dengan alat ukur nontes tersebut.

Contoh Buku Harian Guru

Tujuan : Mengetahui reaksi pribadi terhadap perilaku untuk anak di taman kanak-kanak.

Petunjuk : Catat kejadian signifikan dalam kolom 1 dan reaksi anda sebagai guru dalam kolom 2.

Nama guru :	Nama anak :
Tanggal :	Umur anak :tahun..... bulan
Apa yang terjadi	Reaksi anda

4.4.3 Evaluasi Menggunakan Wawancara

Teknik evaluasi lain yang harus guru miliki sebagai evaluator di bidang pendidikan adalah teknik evaluasi nontes dengan menggunakan interview dan kuesioner secara efektif. Dalam mengevaluasi dengan kuesioner, guru bertanya kepada siswa tentang informasi pribadi yang diperlukan. Informasi ini memiliki peranan yang sangat penting. Informasi yang terkait dengan pribadi siswa dan keluarganya dapat diperoleh melalui teknik nontes ini. Oleh karena itu, wawancara dengan mereka sebaiknya merupakan bagian kompetensi yang perlu bagi guru.

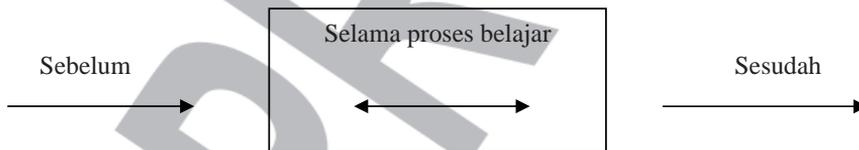
Mengenai apa yang dimaksud dengan wawancara dalam evaluasi nontes, Johnson dan Johnson (2001) menyatakan sebagai berikut, "An

interview is a personal interaction between interviewer (teacher) and one of more interviewees (students) in which verbal questions are asked." Wawancara adalah interaksi pribadi antara pewawancara (guru) dengan yang diwawancarai (siswa) di mana pertanyaan verbal diajukan kepada mereka.

Dalam wawancara ada beberapa persyaratan penting yang perlu diperhatikan:

- a. Adanya interaksi atau tatap muka guru dengan siswa.
- b. Ada percakapan verbal di antara mereka, dan
- c. Memiliki tujuan khusus.

Tujuan wawancara dalam butir ketiga merupakan aspek penting. Karena, berdasarkan tujuan wawancara yang telah ditentukan akan diperoleh makna wawancara yang bervariasi antara satu dengan yang lainnya. Guru harus menyadari pentingnya perencanaan dan tujuan wawancara. Secara umum, wawancara bukan hanya digunakan dalam konteks evaluasi pendidikan saja, namun secara lebih luas lagi, wawancara digunakan dalam konteks kegiatan penelitian, terutama sebagai alat untuk mengeksplorasi data dari para responden.



Gambar 4.1 Wawancara dalam Evaluasi Pendidikan

Dalam konteks evaluasi pendidikan, wawancara dapat dilakukan secara individual maupun secara berkelompok, di mana seorang guru bertatap muka dan melakukan tanya-jawab dengan para siswa. Di samping itu, wawancara dapat dilakukan, baik sebelum, selama, dan sesudah proses belajar mengajar berlangsung. Untuk tujuan evaluasi, wawancara bisa difokuskan pada sasaran yang bervariasi, termasuk buku wajib, materi pembelajaran yang telah diberikan oleh guru, proyek, penulisan skripsi, tesis, dan portofolio setelah diselesaikan, atau bahkan ketika pengerjaan proyek yang sedang berlangsung.

Secara garis besar, teknik wawancara dapat dibedakan menjadi dua macam bentuk, yaitu menggunakan model pertanyaan dengan jawaban

pasti (*fixed ended questions*) atau pertanyaan dengan jawaban terbuka (*open ended questions*), tergantung pada pertanyaan yang hendak dilakukan para guru.

Teknik wawancara dikatakan menggunakan pertanyaan *fixed ended* jika para evaluator merencanakan pertanyaan dilengkapi dengan jawaban alternatifnya. Bentuk pertanyaan *fixed ended* ini biasanya guru gunakan manakala jawaban pertanyaan tersebut sudah pasti, jelas, atau terbatas. Sebagai contoh:

Masih tepatkah pelajaran keterampilan hidup diberikan pada siswa kelas tiga SMP?

- a. Ya
- b. Tidak

Wawancara dikatakan menggunakan pertanyaan *open ended* jika isi yang hendak dievaluasi memiliki jawaban kompleksitas tinggi. Indikator jawaban yang kompleks di antaranya.

- a. Jawaban pasti tidak diketahui,
- b. Tujuan wawancara lebih menekankan agar siswa dapat mengeksplorasi pengetahuan mereka
- c. Jawaban memerlukan alasan tanpa proses, dan
- d. Jawaban masih bisa ditafsirkan sesuai dengan pengetahuan para siswa.

4.4.4 Kelebihan dan Kelemahan Teknik Wawancara

Pada umumnya, wawancara dengan menggunakan pertanyaan *fixed ended*, memiliki beberapa kelebihan sebagai berikut:

1. Jawaban siswa mudah dimengerti.
2. Jawaban siswa mudah diadministrasikan.
3. Siswa tidak perlu melakukan interpretasi jawaban.
4. Siswa diberikan kesempatan untuk menginterpretasikan hal-hal yang berbeda.

Kelemahan pertanyaan *fixed ended* di antaranya sebagai berikut:

1. Mendorong siswa memberikan jawaban bervariasi menurut versi pengetahuan mereka
2. Mendorong siswa mengabadikan jawaban alternatif penting lainnya

3. Memberikan kesempatan kepada siswa untuk menginterpretasikan hal yang berbeda.

Evaluasi dengan menggunakan pertanyaan *open ended* mempunyai kelebihan sebagai berikut.

1. Memberikan kemampuan kepada siswa untuk menyampaikan alasan atau *reasoning* terhadap pertanyaan guru.
2. Memberikan siswa keberanian dalam menjawab sesuai dengan pandangan mereka dengan memberikan jawaban alternatif.
3. Memberikan peluang kepada siswa untuk menyampaikan jawaban yang bertujuan penyelidikan (*probing*).

Di samping itu, evaluasi dengan pertanyaan *open ended* memiliki kelemahan sebagai berikut:

1. Jawaban sulit diadministrasikan.
2. Memerlukan pewawancara yang terlatih untuk dapat melakukan wawancara dengan baik.
3. Memerlukan pelatihan bagi para pewawancara.
4. Memancing jawaban yang kompleks.
5. Hasil wawancara sulit dianalisis.

4.4.5 Evaluasi dengan Kuesioner

Wawancara dan kuesioner memiliki hubungan yang sangat erat. Keterkaitan tersebut dapat ditunjukkan sebagai berikut. Untuk dapat melakukan wawancara dengan baik, seorang guru memerlukan alat, yaitu kuesioner, minimal dengan model pertanyaan *fixed ended*. Dalam evaluasi pendidikan, kuesioner direncanakan untuk memperoleh informasi dari para siswa. Pada umumnya, kuesioner tidak direncanakan untuk memperoleh skor atau nilai. Alat ini sangat membantu para guru dalam mendapatkan informasi pribadi tentang siswa. Informasi tersebut kemudian disimpan dalam bentuk map di kantor pusat informasi.

Informasi dalam kuesioner ini bisa berupa informasi keluarga, alamat keluarga, alamat siswa apabila belajar di kota lain, tingkat pendidikan orang tua, pekerjaan, serta penghasilan orang tua. Informasi dari keluarga tersebut bisa digunakan untuk merencanakan hal-hal yang berkaitan dengan pribadi

siswa, misalnya minat siswa, kebutuhan siswa, dan program belajar pilihan siswa.

Kuesioner untuk mengevaluasi siswa, sebaiknya menggunakan pertanyaan atau pernyataan pendek agar lebih mudah mereka mengerti. Pada pertanyaan tersebut sebaiknya diberi spasi yang cukup agar siswa dapat memberikan jawaban dengan baik.

Contoh Pertanyaan

1. Persiapan apa yang diperlukan sebelum siswa melakukan praktik lapangan?
2. Apa yang perlu dilakukan seorang siswa untuk melakukan praktik lapangan agar lebih berhasil?
3. Bagaimana merancang kuesioner yang baik untuk tujuan evaluasi siswa?

Berikut ini diuraikan beberapa kiat penting yang perlu guru perhatikan untuk mendapatkan kuesioner yang baik.

1. Kuesioner sebaiknya dibuat dengan kalimat pendek dengan pertanyaan jelas.
2. Susunan kuesioner sebaiknya menghindari kalimat majemuk dalam satu item.
3. Ruang jawaban perlu disediakan secukupnya untuk setiap butir pertanyaan, agar para siswa yang dievaluasi dapat memberikan jawaban dengan cukup longgar.
4. Jawaban sebaiknya direncanakan agar mudah ditabulasi.
5. Pertanyaan perlu direncanakan untuk dapat memenuhi butir tiga.
6. Setiap item pada kuesioner diberi nomor yang relevan.

Berikut ini diberikan contoh evaluasi menggunakan kuesioner.

Tujuan : Mengidentifikasi latar belakang siswa kelas I SMP. Kuesioner tentang Keluarga.

Petunjuk : Lengkapi setiap pertanyaan tentang keluarga Anda, jawaban Anda akan disimpan secara rahasia.

1. Siapakah nama Anda?
2. Dengan siapakah Anda tinggal?
3. Siapakah nama Ayah Anda?

4. Siapakah nama Ibu Anda?
5. Apakah kedua orang tua Anda masih bersama Anda?
6. Apakah pendidikan ayah Anda?
7. Apakah pendidikan Ibu Anda?

4.4.5 Evaluasi dengan Portofolio

Portofolio banyak dimanfaatkan oleh para guru maupun para pengguna lain, misalnya arsitek, artis, tim penilai akreditasi, dan sebagainya. Portofolio dapat digunakan untuk menggambarkan mutu kinerja siswa yang ingin dievaluasi. Di bidang pendidikan, portofolio juga digunakan untuk mengumpulkan data tentang kinerja siswa. Menurut Johnson (2012) portofolio bisa didefinisikan sebagai pengumpulan data secara terorganisasi yang dilakukan dalam periode tertentu tentang perkembangan siswa atau perkembangan program kelompok mahasiswa, pencapaian, keterampilan, atau sikap.

Portofolio dapat berupa sekumpulan *file* yang terdiri atas topik pilihan tugas atau pekerjaan yang telah dan akan siswa selesaikan dalam satu semester, satu tahun atau sejak masuk sampai selesai menjadi siswa atau mahasiswa di suatu institusi pendidikan tergantung tujuan penggunaan portofolio tersebut. Portofolio ini juga dapat menggambarkan kinerja siswa dalam satu atau beberapa mata pelajaran atau bahkan semua mata pelajaran yang telah dicapainya. Dilihat dari sasarannya, portofolio dapat bervariasi, misalnya satu portofolio bisa digunakan untuk satu orang siswa atau satu grup siswa. Di samping itu, portofolio juga dapat disimpan dalam map guru, dalam *note book*, kotak, atau dalam *compactdisk*. Apa isi portofolio? Pada prinsipnya, portofolio dapat berisi macam-macam informasi, misalnya untuk keperluan menyimpan data sekolah, maka dapat berisi semua mata pelajaran, atau bahkan semua nilai rapor siswa selama belajar di sekolah tersebut.

Berikut ini adalah beberapa alasan mengapa portofolio sampai sekarang masih banyak digunakan:

1. Memberikan kesempatan kepada siswa untuk dapat mengerahkan potensi belajar mereka sesuai dengan kemampuan.
2. Menentukan tingkat pencapaian hasil belajar
3. Mengetahui perkembangan usaha belajar siswa

4. Dapat digunakan untuk mengevaluasi dan meningkatkan efektivitas kurikulum dan instruksional.

Manfaat portofolio sebenarnya masih banyak sekali sehingga para guru atau para administrator dapat mengembangkannya sesuai dengan bidang keahlian dan tujuan penggunaan mereka.

-oo0oo-

APPROVAL

Bab 5

PENILAIAN ALTERNATIF (*ALTERNATIVE ASSESMENT*)

Setelah mempelajari bab ini, mahasiswa diharapkan:

1. Bisa mengidentifikasi alasan perlunya penilaian alternatif.
2. Bisa menjelaskan pengertian Pengertian Penilaian Alternatif.
3. Bisa mengemukakan landasan rasional penggunaan penilaian alternatif.
4. Bisa menyebutkan Karakteristik Penilaian Alternatif.
5. Bisa menjelaskan kelompok Pro Dan Kontra Terhadap Penilaian Alternatif.
6. Bisa menjelaskan Dasar Pemikiran bagi yang pro.
7. Bisa menjelaskan Dasar Pemikiran bagi yang kontra.
8. Bisa membedakan teknik-teknik yang digunakan dalam penilaian alternatif

5.1 MENGAPA PENILAIAN ALTERNATIF?

Dengan mempertimbangkan keterbatasan dan kelemahan tes (terutama tes objektif) dalam mengungkap pemahaman, kemampuan, dan keterampilan peserta didik yang sesungguhnya dalam proses dan hasil belajar, maka keterbatasan tersebut mendorong timbulnya upaya dan pola pikir untuk mencari alternatif lain. Sehingga, apa yang diungkapkan sebagai penilaian, benar-benar menggambarkan keadaan yang sesungguhnya. Dengan menggunakan instrumen dalam bentuk tes objektif saja, baik pada waktu peserta didik sedang belajar maupun dalam mengungkap hasil belajar, maka peserta didik hanya memilih satu respons dari daftar jawaban

yang disediakan. Peserta tidak menciptakan suatu jawaban terhadap satu pertanyaan, dan juga belum mencakup semua kemampuan yang seharusnya dikuasainya. Prioritas yang ada baru pada kognitif tingkat rendah dan apa yang dijawab dengan benar, belum tentu mereka dapat dan mampu mengaplikasikan kemampuan tersebut dalam situasi nyata yang sesungguhnya (dalam kehidupan sehari-hari).

Tes objektif memang baik, kalau sesuai dengan tujuannya. Artinya tidak berhenti di sana. Menurut Yusuf (2015), kata *baik* harus diikuti dengan kata *untuk*, jadi baik untuk menghimpun sebanyak mungkin informasi dari peserta didik sesuai dengan materi yang diberikan. Tes objektif lebih bersifat mengingat kembali (*recall*) pengetahuan dan pemahaman, namun terbatas sekali dalam menilai kemampuan mengaplikasikan, menganalisis, mensintesis, dan mengevaluasi. Selain itu, tes objektif tidak mungkin digunakan untuk menilai keterampilan (*skills*) maupun unjuk kerja (*performance*). Kadang-kadang para pendidik sering lupa dan menyusun butir soal dalam bentuk tes objektif pada ujian mid semester untuk satuan pelajaran yang tujuan pembelajarannya adalah “agar peserta didik mampu melakukan...” begitu juga dengan tes esai. Walaupun sering pula digunakan *untuk* mengetahui kemampuan peserta didik dalam menyatakan pendapatnya, menerangkan atau menjelaskan sesuatu, namun luasnya cakupan yang diujikan terbatas; dan tes esai juga tidak dapat menjawab pertanyaan: Apakah peserta didik mampu menerapkan apa yang diujikan dalam dunia nyata?

Masing-masing tes objektif dan tes esai mempunyai kekuatan dan kelemahan tersendiri. Kedua bentuk tes ini tidak dapat digunakan untuk mengungkapkan informasi tentang aspek-aspek tertentu dalam waktu yang cepat dan sesering mungkin, serta tidak ditujukan supaya mampu melakukannya dalam situasi nyata, dan/atau dalam dunia nyata di lingkungannya. Oleh karena itu maka dikembangkanlah Penilaian Alternatif (*alternative assessment*) dengan menggunakan berbagai jenis instrumen lain, selain tes esai dan tes objektif.

5.2 APA PENGERTIAN PENILAIAN ALTERNATIF?

Sebagaimana telah diuraikan di atas, bahwa penilaian hasil belajar dengan menggunakan tes objektif pada umumnya hanya mampu menilai aspek

kognitif tingkat rendah dan sangat terbatas dalam menilai sikap, dan yang lebih penting lagi tidak mampu sama sekali untuk menilai kemampuan dan keterampilan melakukan sesuatu aspek *skills* dan *performance* yang ingin diukur. Sementara, tes esai mampu menguraikan sebageian kemampuan kognitif tingkat tinggi peserta didik namun dalam jumlah materi yang sangat terbatas dan tidak mampu mewakili semua aspek yang telah dipelajari dalam mata pelajaran tertentu. Hal itu tidak dapat diabaikan, sebab apabila belajar hanya dinilai dan kemudian ditransfer ke dalam angka atau huruf, maka sesungguhnya bukanlah kemampuan peserta didik yang diharapkan sesuai target belajar (*learning targets*), yang telah dirumuskan sebelumnya. Pada target belajar, pendidik merumuskan “Peserta didik mampu menilai hubungan antara curah hujan dan banjir, “namun pada akhir pembelajaran pendidik hanya menggunakan tes objektif, tipe pilihan ganda (*multiple choice*) dan tipe menjodohkan (*matching*).

Keterbatasan tes esai dalam menilai sikap dan keterampilan peserta didik tidak berbeda dari kelemahan tes objektif dalam hal yang sama. Kelemahan-kelemahan tersebut mendorong para ahli untuk mencari alternatif lain; menyempurnakan dan melengkapi dengan teknik-teknik baru yang dapat digunakan untuk menilai aspek kognitif secara utuh (termasuk kreatif), afektif dan psikomotor. Untuk itu, penilaian alternatif adalah jawabannya. Menapa demikian?

Penilaian alternatif, (menurut Yusuf, 2015: 270), merupakan satu jenis penilaian dalam pembelajaran dengan menggunakan berbagai teknik penilaian, di mana peserta didik menciptakan suatu respons atau alternatif jawaban sendiri terhadap suatu pertanyaan atau tugas, atau melakukan tindakan, perbuatan, unjuk kerja dalam dunia nyata. Singkatnya, penilaian alternatif mencoba mengurangi kelemahan-kelemahan tes tradisional dalam menilai peserta didik. Penilaian alternatif bisa masalah tersebut apabila peserta didik diminta untuk mengaplikasikan kemampuan dan keterampilannya untuk memecahkan masalah dalam situasi nyata (*as a real-life*), termasuk di dalamnya musik, seni, tari, pendidikan jasmani/olahraga, menulis paper, portofolio, percobaan/eksperimen. Oleh karna itu, penilaian alternatif disebut juga penilaian autentik (*authentic assesment*). Kalau penilaian dilakukan dengan mengamati peserta didik dalam melakukan tugas (*tasks*), seperti rancangan gambar, pemecahan masalah, presentasi,

demonstrasi, mengoperasikan mesin, maka penilaian alternatif disebut juga penilaian unjuk kerja (*performance assessment*).

Dilihat dari segi waktu pemanfaatannya, penilaian alternatif dilakukan secara berkelanjutan, sebagai bagian dari penilaian kelas untuk perbaikan pembelajaran dan cara peserta didik dalam belajar (*assessment for learning*), namun dapat juga dilakukan pada saat tertentu untuk menentukan hasil belajar (*assessment of learning*) yang dilakukan pada akhir semester/caturwulan atau unit pembelajaran.

5.3 LANDASAN RASIONAL

Ada beberapa landasan pemikiran yang dijadikan pertimbangan mengapa penilaian alternatif dikembangkan, antara lain terdapat lima landasan sebagai berikut:

1. Kelemahan Testing

Sebagaimana telah diuraikan di atas, tes yang telah digunakan sebagai instrumen penilaian, baik yang disusun oleh pendidik, maupun tes yang sudah distandarisasikan mempunyai tujuan tertentu pada saat tes itu dirancang; kepada siapa tes itu akan diberikan, serta dalam budaya yang bagaimana tes itu dibakukan. Dengan kata lain, suatu tes yang valid untuk maksud tertentu; tidak berarti valid untuk semuanya.

Tes buatan guru (*teacher made test*) pada umumnya belum dibakukan. Ini berarti, tes yang diberikan belum tentu cocok dengan tujuan belajar yang ingin diujikan. Reliabilitas, objektivitas, dan kepraktisan masing-masing tes buatan guru masih diragukan.

Oleh karena itu, untuk mengetahui apakah peserta didik memahami sesuatu, tes objektif tidak diragukan lagi; namun untuk mengetahui kemampuan dan keterampilan peserta didik dalam “berbicara dengan menggunakan Bahasa Inggris” misalnya, maka ia harus diuji dengan tugas melaksanakan *speaking* yang dimaksud. Kalau yang diharapkan adalah pengetahuannya tentang suatu *grammar* dalam Bahasa Inggris, maka penggunaan tes objektif tepat digunakan. Namun seandainya tujuan pembelajaran yang guru merumuskan adalah pengetahuan dan kemampuan melakukan suatu keterampilan (*speaking* misalnya), maka penilaian yang

digunakan hendaknya mencakup tes objektif dan juga mendemonstrasikan atau melakukan tugas unjuk kerja (*performance tasks*). Dewasa ini pendekatan seperti di atas dikategorikan ke dalam penilaian alternatif.

2. Akurat, Teliti, dan Tepat

Dengan menggunakan berbagai teknik (mungkin juga hanya satu teknik) sesuai dengan target belajar, maka hasil yang didapat betul-betul akan menggambarkan yang sesungguhnya. Peserta didik tidak bisa dikatakan dapat, mampu dan terampil menendang bola, kalau ia memang belum mampu menendang bola dengan baik dan benar. Demikian juga dalam aspek-aspek yang lain. Sebagai ilustrasi, mari kita amati bersama: Seorang peserta didik tidak akan dinyatakan lulus dengan nilai A, kalau dari 100 butir tes objektif, ia hanya betul 40 butir; walaupun ia adalah peserta didik yang mendapatkan skor tertinggi di kelasnya. Lebih baik pendidik mempertimbangkan kembali: "Layakkah saya memberikan nilai A, kalau penguasaannya hanya dua per lima dari yang sesungguhnya?" Kemudian melontarkan pertanyaan kepada diri sendiri: "Di manakah salah saya dalam membelajarkan siswa saya?". Nilai itu akan menjadi lebih akurat, teliti dan tepat, manakala guru menggunakan instrumen lain untuk mengetahui yang sesungguhnya, sesuai dengan target belajar yang dirumuskan sebelumnya.

3. Akuntabilas

Setiap pendidik dalam menilai peserta didiknya seyogyanya memiliki rasa tanggung jawab atas ketepatan dan kebenaran penilaiannya. Pendidik sesungguhnya mempunyai kewajiban mempertanggungjawabkan karyanya kepada sekolah, orangtua, dan masyarakat. Dengan menggunakan berbagai teknik yang sesuai dan tepat, pendidik mengungkapkan target belajar yang sesungguhnya, dan selanjutnya melaksanakan akuntabilitas kepada semua *stakeholder*, orangtua dan masyarakat; dengan baik dan benar.

4. Autentik

Penilaian alternatif menggunakan berbagai cara dalam menilai peserta didik sesuai dengan target belajar. Di samping itu, peserta didik diminta untuk melakukan atau mendemonstrasikan kemampuannya dalam kehidupannya nyata, dalam situasi yang nar-benar relevan dan itulah dia yang sesungguhnya (autentik); bukan permainan angka dalam aspek yang dinilai. Dari sisi aspek

yang dinilai, ia betul-betul dihadapkan pada situasi nyata dalam masyarakat (autentik), bukan diatur *setting* tertentu sebagai penilaian.

5. Menyediakan Informasi komprehensif bagi guru

Mengingat penilaian alternatif menggunakan teknik dalam memahami dan menilai peserta didik, maka akan tersedia informasi komprehensif tentang kekuatan dan kelemahan peserta didik dalam belajar. Catatan anekdot (*anecdotal record*) tentang sikap dan perilaku (positif/negatif) akan terekam dengan baik, kalau pendidik sejak dini membiasakan diri membuat catatan perkembangan sikap dan perilaku menonjol di antara peserta didik yang diasuhnya. Hubungan sosial di antara peserta didik akan dapat diungkap apabila pendidik menggunakan sosiogram. Meminta peserta didik membuat biodata atau bertanya kepada mereka; atau di akhir jam pelajaran mereka diminta untuk bercerita tentang latar belakang dirinya, maka akan tersedia informasi tentang latar belakang dan perbedaan budaya di antara sesama peserta didik.

6. Menyediakan Kesempatan Untuk Menilai Proses dan Produk

Dengan memanfaatkan penilaian alternatif, peserta didik dinilai bukan hanya memilih dari alternatif jawaban yang telah disediakan, tetapi juga mengkonstruksi jawaban sendiri, mendemostrasikan dan melakukan sendiri dalam situasi nyata. Yusuf (2015) memberikan sebuah ilustrasi. Mari kita perhatikan ilustrasi berikut ini: Seorang siswa pada Sekolah Menengah Seni Rupa sedang belajar menyusun *collage* (susunan benda-benda atau potongan kertas dan sebagainya, yang ditempelkan pada bidang datar dan merupakan satu kesatuan karya seni). Selama perbuatan guru dapat menilai bagaimana siswa berpikir dan memikirkan bentuk potongan benda yang tersedia, memikirkan bagaimana dan di mana potongan benda itu seharusnya ditempatkan serta bagaimana proses pembuatan selanjutnya. Pada akhirnya baru menilai produknya.

Dengan kata lain, penilaian alternatif benar-benar menyediakan beragam aspek yang mungkin dinilai dari peserta didik, bukan aspek-aspek kognitif tingkat rendah saja, sehingga penilaian semacam ini memberikan informasi yang lebih komprehensif.

5.4 KARAKTERISTIK ASESMEN ALTERNATIF

Penilaian alternatif sebagai salah upaya untuk mengatasi kelemahan-kelemahan penilaian tradisional memiliki beberapa karakteristik sebagai berikut.

1. Terdapat Saling Keterkaitan Antara Penilaian dan Pembelajaran

Sebagaimana telah dijelaskan pada uraian terdahulu, antara penilaian dan pembelajaran terdapat saling keterkaitan. Penilaian disusun berdasarkan tujuan pembelajaran yang telah dirumuskan, sedangkan tujuan pembelajaran dirumuskan berdasarkan standar yang telah ditetapkan. Pendidik selama membelajarkan peserta didik, tetap mengacu pada rencana pelaksanaan pembelajaran secara aktif, kreatif, inovatif dan menyenangkan, namun tetap dalam koridor yang telah disusun sebelumnya. Dengan kata lain, target belajar dari sisi peserta didik adalah sasaran utama yang tidak dapat diabaikan. Sebaliknya pola pelaksanaan dilakukan dengan mempertimbangkan semua faktor pendukung dan penghambat kegiatan pembelajaran secara kontekstual. Pada saat tertentu (di awal, di tengah dan diakhir kegiatan), penilaian alternatif perlu dilakukan untuk mengetahui dan menentukan tingkat pemahaman serta pencapaian belajar peserta didik. Penilaian itu akan memberikan informasi dan menyediakan umpan balik bagi peserta didik yang sedang belajar dan bagi pendidik yang membelajarkannya.

2. Peserta Didik Dipandang Sebagai Klien Utama

Kecendrungan dewasa ini menempatkan peserta didik sebagai titik sentral dalam pembelajaran (*student centered*) dan pendidik sebagai motivator, fasilitator dan assessor dalam pendidikan. Pendekatan ini berbeda dengan pendekatan tradisional yang menekankan pada pendidik sebagai titik sentral dalam pembelajaran, dan peserta didik sebagai objek yang dibelajarkan. Oleh karena itu, pada pendekatan baru ini peserta didik, dengan semua karakteristiknya yang berbeda antara satu individu dan individu lainnya, perlu mendapat perhatian khusus dalam kegiatan pembelajaran.

Penilaian alternatif sebagai bagian integral dalam proses pembelajaran serta penyedia informasi secara utuh dan berkesinambungan, menempatkan peserta didik sebagai klien utama. Mengapa demikian? Bukan karena ia

“sakit” namun karena ia sedang tumbuh dan berkembang. Peserta didik adalah sumber informasi tentang dirinya. Seandainya pendidik tidak mengungkapkan informasi secara utuh melalui penilaian alternatif tentang dirinya, kegiatan belajarnya, dan informasi lainnya maka peserta didik yang kurang memahami, kurang menguasai, kurang mengerti, kurang mampu melakukan, tidak mau melakukan, kurang siap, dan lain sebagainya; tetap tidak akan diketahui keadaan yang sesungguhnya, sehingga pola perkembangannya menjadi terganggu; pola belajar dan belajarnya tidak optimal. Untuk itu jadikanlah peserta didik sebagai individu yang beruntung dengan mengumpulkan informasi tentang dirinya, kekuatan dan kelemahannya dalam belajar melalui penilaian alternatif, sehingga dapat dilakukan perbaikan dalam kegiatan dan cara belajar oleh peserta didik sendiri dan perbaikan pembelajaran oleh pendidik.

3. Penilaian Alternatif Dilakukan Secara Berkelanjutan

Seandainya informasi dikumpulkan pada saat kegiatan berlangsung dan berkelanjutan, maka perbaikan dan penyempurnaan proses dan hasil belajar akan tercipta dengan baik. Mari kita perhatikan contoh berikut ini:

- a. Dengan hanya mengamati peserta saat diskusi akan dapat diketahui aktivitas peserta didik, dan kemampuannya untuk menyatakan pendapat.
- b. Dengan meminta peserta didik mendemostrasikan sesuatu, sesuai materi yang di berikan akan dapat diketahui penguasaan hasil belajarnya.
- c. Dengan meminta peserta didik membuat tugas di rumah, kemudian memeriksa dan mengembalikan tugas itu setelah diperiksa, akan dapat diketahui hasil belajarnya; namun jangan lupa untuk menyikapi dengan hati-hati (bukan penuh kecurigaan) bahwa tugas itu memang dilakukan peserta didik sendiri.
- d. Dengan meminta peserta didik membuat rubrik tentang suatu topik yang berhubungan dengan materi yang sudah dibelajarkan, pendidik akan mengetahui bagaimana kemampuan peserta didik. Dan seterusnya.

Berdasarkan ilustrasi di atas, jangan tunggu hari esok, ungkapkan hasil belajar peserta didik sebaik dan setepat mungkin, serta lakukan pelaksanaannya secara berkelanjutan. Antara penilaian alternatif yang pertama, kedua, ketiga, dan seterusnya akan dapat dilihat dinamika kemajuan belajar peserta didik. Di samping itu, penilaian alternatif akan

mampu mengungkap multidimensi pencapaian peserta didik dalam belajar dengan menggunakan teknik penilaian alternatif yang bervariasi, namun tetap dalam kerangka acuan target belajar yang telah ditetapkan. Jangan berikan “lambang” yang tidak sesuai dengan isi yang sesungguhnya.

4. Memuat Tugas-tugas Bermakna

Selain apa yang telah dikemukakan di atas, karakteristik berikutnya dari penilaian alternatif ialah tugas-tugas yang diberikan harus bermakna. Hal itu dimaksudkan supaya penilaian tersebut bukan untuk membebani peserta didik, melainkan untuk membantu meningkatkan wawasan, kemampuan dan keterampilan peserta didik, serta berguna dan bermanfaat bagi diri dan kehidupan peserta didik.

5. Kriteria Penilaian Didistribusikan Terlebih Dahulu

Yang tidak kalah pentingnya dengan apa yang sudah diuraikan terdahulu, karakteristik lain yang perlu menjadi perhatian dalam melakukan penilaian alternatif adalah kriteria penilaian didistribusikan terlebih dahulu kepada peserta didik. Sehingga, sebelum mengerjakan suatu tugas, peserta didik sudah mengetahui patokan penilaiannya (pemberian angka dalam menilai).

Penilaian alternatif dimaksudkan untuk mendapatkan informasi yang sesungguhnya tentang kondisi peserta didik yang sudah mengalami pembelajaran secara menyeluruh; tentang hal yang dikuasai dan sekaligus akan menyediakan pula informasi yang belum dikuasai tentang materi yang sudah dibelajarkan sesuai dengan target belajar. Dengan demikian, hasil belajar yang diungkapkan melalui Penilaian Alternatif betul-betul menggambarkan diri siswa dalam aspek-aspek yang sudah dibelajarkan.

5.5 PRO DAN KONTRA TERHADAP PENILAIAN ALTERNATIF

Setiap isu atau konsep baru selalu ada pro dan kontra, begitu juga dengan penilaian alternatif. Banyak yang setuju terhadap penilaian alternatif, namun banyak pula yang tidak setuju. Kalau dikaitkan dengan fungsi penilaian sebagai bagian integral dalam proses pembelajaran, maka penilaian alternatif memiliki nilai positif yang lebih besar ketimbang nilai negatif.

5.5.1 Dasar Pemikiran bagi yang pro

Berdasarkan kelompok yang setuju, secara umum, dasar-dasar pemikiran pentingnya suatu penilaian alternatif dikembangkan sebagai berikut:

1. Holistik

Aneka macam teknik dapat digunakan untuk mengetahui, dan memahami kemampuan menyatakan pendapat maupun melakukan sendiri aspek yang ingin dinilai. Dengan cara demikian gambaran informasi yang diperoleh akan menyeluruh dari berbagai sisi sesuai dengan target belajar. Di samping itu, penilaian alternatif memberikan gambaran yang bersifat analitik tentang peserta didik, baik dilihat dari tingkatan kemampuan maupun ranah kemampuan (kognitif, afektif, dan psikomotor).

2. Berpusat pada Peserta Didik

Kelebihan lain dari penilaian alternatif menempatkan peserta didik sebagai subjek belajar, karena peserta didik merupakan individu yang belajar, sedang tumbuh dan berkembang menuju kemandiriannya. Oleh karena itu, dalam melakukan penilaian alternatif, peserta didik adalah titik sentralnya; bukan pada pendidik yang membelajarkan. Segala teknik perlu digunakan untuk dapat mengungkapkan informasi yang sesungguhnya tentang peserta didik yang merupakan bagian integral dari proses pembelajaran.

3. Menekankan Proses

Dasar pemikiran lain tentang penerapan penilaian alternatif adalah bahwa ia bukanlah penilaian yang hanya memperhatikan produk, melainkan lebih menekankan pada proses. Bagaimana peserta didik memahami dan menghayati sesuatu, bagaimana mereka melakukan dan memberi makna tentang sesuatu yang merupakan fokus penilaian. Berdasarkan informasi penilaian, dengan proses yang berkelanjutan dan perbaikan pembelajaran yang dilakukan secara terus-menerus, maka peningkatan kualitas belajar peserta didik akan terlaksana secara optimal pula.

4. Menuntut Keterampilan Pemecahan Masalah

Tidak sama seperti pendekatan tradisional dalam penilaian hasil belajar, terutama apabila menggunakan tes objektif, target belajar yang dinilai

hanya kemampuan kognitif tingkat rendah; sedangkan dalam penilaian alternatif, peserta didik selalu diminta dan diarahkan supaya mereka menguasai kemampuan dalam pemecahan masalah. Walaupun yang digunakan pertanyaan terbuka sekalipun, namun yang diharapkan bukan peserta didik menyebutkan sesuatu yang sudah dihafalnya, melainkan menggunakan materi yang dipelajari dan dipahaminya untuk memecahkan hal baru yang disampaikan dalam bentuk pertanyaan terbuka.

5. Dunia Nyata Kehidupan Sehari-hari
Berbeda dengan penggunaan tes objektif yang hanya diarahkan untuk pemahaman kognitif, terutama kognitif tingkat rendah, penilaian alternatif (terutama penilaian autentik dan penilaian unjuk kerja) lebih berorientasi kepada dunia nyata. Peserta didik yang dinilai dibawa ke dunia nyata, baik dalam wujud melakukan kegiatan, orientasi pertanyaan, demonstrasi, maupun pertunjukan/pameran peserta didik. Dengan demikian, penilaian alternatif akan sangat membantu dalam menilai peserta didik yang sesungguhnya sesuai dengan target belajar, memahami kelemahannya serta dapat memberikan umpan balik bagi perbaikan pendidikan yang tepat kepada peserta didik.
6. Perspektif Luas
Bila dibandingkan dengan tes objektif dalam menilai peserta didik, maka yang terungkap baru pengetahuan, pemahaman serta aplikasi tertentu (kognitif tingkat rendah) dan belum menyentuh kognitif tingkat tinggi - afektif dan psikomotor. Sebaliknya penggunaan penilaian alternatif akan membawa perspektif lebih luas terhadap diri peserta didik. Dengan menggunakan berbagai teknik dalam pengumpulan informasi sesuai dengan target belajar yang telah ditetapkan; perspektif pengetahuan, pemahaman, kemampuan, sikap, dan keterampilan (*skill*) menjadi jauh lebih luas. Dengan diskusi dalam kelompok kecil, menulis jurnal/refleksi, portofolio, dan rubrik, peserta didik dituntut untuk lebih banyak membaca dan melatih diri dalam menyatakan pendapat, baik secara lisan maupun secara tertulis. Hal itu akan sangat bermanfaat dalam proses pembelajaran dan kemajuan belajar peserta didik.

7. Aktif dan Inovatif

Sejalan dengan kecenderungan dan upaya peningkatan proses pembelajaran dewasa ini, peserta didik merupakan fokus dari kegiatan pendidikan. Cara belajar siswa aktif (*student active learning*) dan pendidikan berpusat pada siswa (*student centered learning*) merupakan dua konsep dalam kegiatan pembelajaran, namun tujuannya sama, yaitu: "Dalam kegiatan pembelajaran, peserta didiklah yang diutamakan sehingga ia mau, mampu dan berhasil dalam belajar." Dalam kaitan itu, penilaian alternatif memberi peluang bagi peserta didik untuk aktif berbuat, menyelesaikan tugas atau melakukan karya inovatif. Melalui penilaian alternatif peserta didik menyusun portofolio, rubrik, jurnal, peta konsep; melakukan kerja kelompok, penilaian diri dan sebagainya. Pilihan teknik yang digunakan sangat terkait dengan target belajar yang telah ditetapkan.

8. Memberanikan peserta didik untuk memberikan refleksi

Pertimbangan lain yang lebih penting, yang tidak tersedia dalam tes tradisional, yaitu penilaian alternatif memberanikan peserta didik untuk memberikan refleksi. Peserta didik diajak dan diikutsertakan dalam penilaian alternatif, walaupun posisi mereka sedang dinilai. Kriteria rubrik penilaian diberikan terlebih dahulu sebelum peserta didik menyusun rubrik yang ditugaskan kepadanya. Nilai yang diberikan pendidik terhadap suatu rubrik yang ia susun dapat dikomentari sebelum menerimanya. Buku catatan harian peserta didik mungkin berisi refleksi terhadap pola pembelajaran yang dilaksanakan pendidik sebelumnya; dan diharapkan ada tanggapan dari guru. Dalam diskusi kelas, peserta didik diberanikan untuk memberikan refleksi, bantahhan, dan persetujuan terhadap pendapat teman-temannya. Oleh karena itu, jika teknik penilaian alternatif digunakan secara tepat dan benar sesuai dengan target belajar, maka penggunaan teknik itu akan mampu memupuk semangat dan keberanian memberikan refleksi, baik dalam proses pembelajaran maupun dalam kehidupan nantinya.

9. Menantang Peserta Didik untuk Menarik Kesimpulan Sendiri

Penilaian alternatif menggunakan berbagai teknik dalam menilai belajar peserta didik. Di samping itu, penilaian alternatif berpusat pada peserta didik dan holistik. Jenis-jenis instrumen yang tersedia dalam penilaian

alternatif memungkinkan peserta didik untuk menarik kesimpulan terhadap tugas yang diberikan kepadanya. Melalui pemanfaatan penilaian alternatif, pendidik mendorong peserta didik untuk menarik kesimpulan sendiri tentang materi yang menjadi fokus penilaian, seperti portofolio, demonstrasi, jurnal peserta didik, fokus group, penyajian kelas maupun interviu.

10. Kencendrungan Dewasa Ini

Pada akhir-akhir ini, penilaian alternatif tumbuh dan berkembang dengan pesat. Kecendrungan itu muncul karena kelemahan tes objektif pilihan ganda, yang selama ini dianggap satu-satunya yang terbaik untuk mengungkap semua hal.

Walaupun banyak yang menyetujui penilaian alternatif sebagai salah satu upaya inovatif dalam bidang pengukuran, penilaian dan evaluasi, namun masih ada sebagian yang kurang setuju dengan pelaksanaannya.

5.5.2 Argumentasi bagi yang Kontra

Ada beberapa argumentasi yang dikemukakan oleh kelompok yang tidak setuju dengan penilaian alternatif. Yusuf (2015) menemukan tujuh buah argumentasi mengapa kelompok yang kontra kurang menyukai penilaian alternatif sebagai mana diuraikan berikut ini:

a. Banyak Waktu yang Digunakan

Tidak dapat dipungkiri bahwa penggunaan teknik penilaian alternatif yang bervariasi sesuai dengan target belajar membutuhkan waktu yang lebih banyak, baik dilihat dari segi persiapan maupun pelaksanaannya.

b. Tidak Semua Orang Suka Kelompok

Sesuai dengan prinsip individualitas, tiap individu mempunyai persamaan di samping banyak pula perbedaanya. Ada individu yang suka kelompok, tetapi banyak pula yang tidak suka kelompok. Kondisi ini kalau diterapkan dalam penilaian alternatif, berarti individu yang tidak suka kelompok akan dirugikan dalam penilaian pembelajaran.

c. Subjektif

Pihak yang kurang setuju dengan penilaian alternatif menyatakan bahwa teknik alternatif itu bersifat subjektif. Walaupun pada beberapa teknik yang digunakan telah memiliki pedoman penilaian, namun pada

sebagian instrumen belum mungkin dibuat secara spesifik. Interview, observasi, dan *checklist*, walaupun memiliki pedoman yang disiapkan; namun dalam pelaksanaannya, subjektivitas penilai dalam menilai sangat menentukan. Demikian juga pada saat menilai *student jurnal*, *student logs*, refleksi, portofolio, tugas rumah, peta konsep, dan peta pikiran yang disusun siswa. Karena jawabannya bukan benar atau salah, atau hanya satu yang benar; maka subjektivitas pendidik dalam menilai peserta didik sangat kuat.

d. Biaya Lebih Besar

Argumentasi lain mengapa banyak orang yang tidak setuju dengan penggunaan penilaian alternatif adalah karena biaya yang digunakan jauh lebih besar daripada penggunaan tes objektif dan tes esai.

e. Kurang Terstruktur

Orang-orang yang sudah terbiasa menggunakan tes objektif dalam menilai pembelajaran peserta didik merasa bahwa penggunaan penilaian alternatif kurang terstruktur; tidak dapat ditata dengan tepat. Misalnya pada hari pertama minggu pertama bulan Oktober 2016, akan diadakan penilaian alternatif; bentuk instrumen yang digunakan adalah *self assessment* atau *collage*. Rencana itu belum tentu dapat dilaksanakan tepat waktu, karena penyelesaian materi dan materi apa yang akan digunakan pada waktu tertentu, lebih banyak dipengaruhi dan terkait dengan kondisi belajar peserta didik dan target belajar selama proses pembelajaran.

f. Kurang Menekankan Produk

Sebagaimana telah dikemukakan di atas, penilaian alternatif lebih mengutamakan proses dan kurang menekankan produk. Menurut kelompok yang kontra pada penilaian alternatif, betapapun bagus proses, jika tidak terlihat produknya, maka nilai kegiatan itu menjadi tidak berarti. Kelompok ini tidak peduli dengan proses pembuatan atau pelaksanaannya.

g. Sulit Dilaksanakan

Alasan lain menyatakan bahwa sesuatu yang baru memang sulit dilaksanakan, apalagi dengan banyaknya instrumen yang mungkin digunakan; sedangkan peserta didik belum semuanya mampu dan

terbiasa dengan instrumen itu. Terlebih lagi, untuk mengetahui satu target belajar yang telah ditetapkan harus didekati dengan banyak instrumen dari sisi yang berbeda sehingga dapat diberikan makna yang tepat dan menyeluruh. “sulit dilaksanakan”, itulah komentar kelompok yang tidak setuju dengan penilaian alternatif.

5.6 TEKNIK-TEKNIK YANG DIGUNAKAN DALAM PENILAIAN ALTERNATIF

Banyak teknik yang dapat digunakan dalam penilaian alternatif sesuai dengan bentuk proses pembelajaran yang dipilih guru, sehingga informasi yang dikumpulkan *valid* dan *reliable* serta benar-benar menggambarkan kondisi peserta didik yang sesungguhnya. Hal itu sangat dimungkinkan, karena suatu aspek yang dinilai, akan ditinjau dan dinilai secara holistik dan berkesinambungan, sesuai dengan target belajar, di mana penilaian melekat di dalamnya (penilaian merupakan bagian integral dalam proses pembelajaran) sebagai berikut:

1. Portofolio
2. Rubrik (*student written rubrics*)
3. Jurnal siswa
4. Buku catatan harian tentang peserta didik (*student logs*)
5. Penilaian diri (*self assessment*)
6. Peta pikiran (*mind concept*) dan peta konsep (*concept maps*)
7. Diskusi kelompok terfokus (*focussed group discussion*)
8. Kerja kelompok
9. Diskusi kelompok kecil (*small group discussion*)
10. Tugas-tugas unjuk kerja (*performance tasks*)
11. Tugas-tugas tertulis tentang suatu kejadian (*event tasks*)
12. Proyek peserta didik (*student project*)
13. Potongan-potongan benda dan/atau kertas yang ditempel pada bidang datar yang merupakan kesatuan karya seni (*collages*)
14. Poster tempelan (*poster sessions*) karya peserta didik
15. Interview/wawancara
16. Observasi
17. Checklist

18. Skala bertingkat (*rating scales*)
19. Kuesioner
18. Catatan anekdot (*anecdotal record*)
19. Simulasi, demonstrasi dan pameran
20. Penyajian lisan/presentasi kelas
21. *Video/audio taping*
22. Pertunjukan peserta didik (*student displays*)
23. Lakon pendek dan lucu (*short plays and game play*),
24. Gambar yang dibuat peserta didik (*student drawing*),
25. Pertanyaan terbuka (*open ended question*),
26. Jawaban pendek, dan tes essay.

Pencapaian penilaian yang sesungguhnya, sebagai bagian integral dalam proses pembelajaran, akan tergantung pada pemilihan teknik penilaian yang tepat. Oleh karena itu, teknik-teknik yang dikemukakan di atas perlu dipilih dengan baik dan benar berdasarkan standar kompetensi dasar; atau target belajar dan indikator keberhasilan. Di samping itu, karakteristik mata pelajaran jangan diabaikan. Misalnya: *collages* sangat tepat digunakan dalam mata pelajaran seni rupa dan kurang berarti untuk mata pelajaran sejarah. Diskusi kelompok akan sangat bermanfaat dan sesuai dengan mata pelajaran bahasa Inggris. Adapun penilaian diri, dapat digunakan untuk semua mata pelajaran.

Pada uraian selanjutnya hanya akan dikemukakan secara mendalam tentang beberapa teknik saja, sedangkan teknik-teknik wawancara, *checklist*, skala bertingkat, observasi dan sesiometri tidak akan dibahas pada bab ini.

5.6.1 Portofolio

a. Pengertian

Secara etimologis, portofolio merupakan perpaduan dua kata *port* dan *folio*. *Port* dapat diartikan sebagai kumpulan dari suatu kegiatan atau *course*, sedangkan *folio* adalah kertas. Portofolio banyak digunakan belakangan ini. Minat dan perhatian penggunaan portofolio sebagai penilaian dalam pendidikan dimulai pada 1980-an.

Penekanan pendidikan yang baik bukan hanya pengetahuan, pemahaman dan wawasan saja, melainkan juga berupa karya lainnya.

Oleh karena itu, karya-karya peserta didik harus dinilai. Penilaian portfolio jauh lebih kompleks dari penilaian dengan menggunakan tes tradisional, yaitu pengumpulan bukti-bukti karya tulis dan sampel kerja peserta didik dari waktu ke waktu sesuai dengan lama waktu setiap materi pendidikan diberikan. Pemilihan karya dilakukan dengan hati-hati serta mewakili karya peserta didik secara keseluruhan dalam materi yang dinilai. Karya-karya itu merupakan gambaran kemajuan belajar peserta didik, berupa kemajuan akademik, belajar, keterampilan, maupun sikap. Jadi, portofolio itu dapat berupa sampel kerja (*work samples*) dan laporan tertulis berkenaan dengan sampel kerja itu, atau berbagai karya tulis peserta didik dari waktu ke waktu selama kegiatan materi pembelajaran. Pendapat yang hampir senada dengan kesimpulan di atas, dikemukakan oleh Jhonson & Jhonson (2002:103) sebagai berikut:

A portfolio is an organized collection of evidence accumulated over time on a student's or group's academic progress, achievements, skill, and attitudes. It consists of work samples and a written rationale connecting the separate items into a more complete and holistic view of the students (or group's) achievement or progress toward learning goals.

Sedangkan Johnsen, 2008, menyatakan: *A portfolio represents a purposeful and systematic collection of student work to showcase skill and accomplishment in one or more areas.*

Portofolio bukan sekedar pemberian tugas, bukan sekedar membuat karangan tentang sesuatu saja, atau presentasi keterampilan dalam menggambar, tari dan lain sebagainya. Portofolio jauh lebih kompleks dan mencakup aspek-aspek lain. Portofolio bisa mencakup kemajuan akademik peserta didik/kelompok, prestasi belajar, keterampilan maupun sikap. Materi yang diportofoliokan dapat berlangsung dalam satu tahun, satu semester, dapat juga satu catur wulan tergantung pada bobot dan beban materi pelajaran.

Pendapat lain yang hampir senada dengan pendapat-pendapat di atas dikemukakan Butter 2006: *"A portfolio is collection of evidence that gathered together to show a person's learning journey over time and to demonstrates their abilities."*

Tujuan pengumpulan karya peserta didik untuk dapat mengetahui usaha, kemajuan dan hasil belajar dalam satu atau lebih materi pembelajaran dari waktu ke waktu. Sebelum pengumpulan karya peserta didik, mereka perlu diikutsertakan dan berpartisipasi aktif dalam pemilihan materi, kriteria pemilihan dan bukti refleksi diri. Sementara itu, isi portofolio disesuaikan dengan bidang yang diportofoliokan, antara lain berupa: laporan tugas lengkap atau tugas dalam kelas, karangan atau cerita, rekaman, laporan investigasi, laporan observasi proyek, *self-refleksi* tentang sesuatu, unjuk kerja, dan kebiasaan kerja dan sikap peserta didik. Oleh karena itu, materi portofolio perlu dibicarakan dan ditetapkan secara bersama oleh peserta didik, dan pendidik.

Apabila ingin menggunakan format portofolio sebagai instrumen penilaian dalam menilai kemampuan belajar, keterampilan atau sikap, hendaklah dirancang secara khusus, sehingga peserta didik tidak dirugikan pada akhir penilaian. Peserta didik hendaknya dibuat betul-betul tahu bidang/materi apa yang akan diportofoliokan, dan bagaimana pula kriteria penilaian yang sesungguhnya. Peserta didik harus mengetahui pola atau matriks penilaian sekaligus batasan, rambu-rambu penilaian, termasuk di dalamnya rentang kategori nilai berdasarkan patokannya.

Secara sederhana portofolio melewati tiga langkah: *Pertama* yaitu persiapan untuk menggunakan portofolio. Pada langkah pertama ini, sekolah atau lembaga harus mengembangkan program portofolio, dengan menentukan tipe portofolio yang digunakan, mengidentifikasi tujuan portofolio, dan memilih kategori sampel kerja yang akan diportofoliokan. Sementara peserta didik dapat memilih pekerjaan terbaik, yang akan mewakili mereka, dan juga memahami bagaimana kriteria penilaian terhadap kerja mereka.

Kedua menata portofolio sesuai dengan waktu yang terdapat dalam kurikulum. Pada tahap ini, tenaga, sekolah atau guru yang bertanggung jawab, hendaknya betul-betul menjelaskan proses portofolio, mulai dari kategori sampel kerja yang termasuk portofolio, mengembangkan pedoman penilaian/penilaian portofolio, dan assignment.

Ketiga menata proses penilaian portofolio pada periode pemberian angka. Pada tahap ini, perlu dipahami bahwa pemberian angka atau skor,

berdasarkan pertimbangan kualitas karya dan dasar-dasar rasional yang dituliskan dalam berbagai laporan tertulis yang menyertai sampel kerja tersebut. Matriks penilaian atau pedoman penilaian yang akan digunakan diberitahukan pada saat awal portfolio akan dilaksanakan.

Pada akhir penilaian dilakukan “konferensi” antara pendidik dan peserta didik, untuk mengetahui tanggapan dan pendapat dari mereka tentang nilai tersebut.

b. Langkah-langkah Pengembangan Portfolio

Sejalan dengan perkembangan konsep pembelajaran yang menekankan peserta didik sebagai fokus (*student centered learning*) dalam kegiatan pembelajaran; dan pendidik berfungsi sebagai fasilitator dan motivator, maka peserta didik harus diajak dan dimotivasi untuk menciptakan sesuatu yang baru atau mampu mengonstruksi sesuatu yang baru. Terkait dengan hal tersebut, portfolio merupakan salah satu bentuk yang dapat digunakan.

Berikut ini adalah beberapa langkah yang perlu dipertimbangkan dalam pengembangan portfolio sebagai berikut:

- a) Tentukan bersama peserta didik, sampel portofolio apa saja yang akan dibuat. Hal ini sangat diperlukan, sebab tujuan utama pembelajaran yaitu membelajarkan peserta didik sehingga mereka belajar dan terus belajar. Penilaian portfolio merupakan bagian integral dari proses pembelajaran; apapun bentuk sampel portofolio yang akan dikerjakan selalu diarahkan supaya mereka belajar. Pengikutsertaan peserta didik, di satu pihak adalah penghargaan pendidik terhadap keberadaan peserta didik; sedangkan di pihak lain, dalam memberikan tugas perlu diperhatikan karakteristik peserta didik dengan semua aspeknya, sehingga tugas tersebut menarik, menyenangkan dan menantang untuk dikerjakan peserta didik.
- b) Kumpulkan portofolio di rumah atau loker pribadi masing-masing di sekolah. Kegiatan ini dimasudkan agar portofolio yang sudah dikerjakan jelas keberadaannya, tertata dengan baik, dan ketika dibutuhkan dapat diambil dengan cepat.
- c) Tentukan aspek-aspek yang akan dinilai dari portofolio beserta pembobotannya. Keterbukaan dan akuntabilitas dalam penilaian perlu

dikembangkan. Sudah bukan waktunya lagi “Hanya guru yang serba tahu apa yang dinilai dari karya peserta didiknya.” Keterbukaan tentang aspek-aspek yang dinilai secara tidak langsung akan membantu peserta didik dalam menyelesaikan karya portofolio ke arah yang lebih baik, karena mereka tahu aspek-aspek apa yang akan dinilai. Di samping itu, guru lebih hati-hati, lebih tepat dan lebih sungguh-sungguh dalam menilai karya peserta didik, sehingga validitas penilaian lebih dipercaya. Akuntabilitas penilaian guru tentang karya peserta didik kepada peserta didik dan orang tua dapat diterima secara bersama.

- d) Di samping itu peserta didik perlu diminta untuk menilai karyanya secara berkesinambungan (*continously self assessment*). Pendidik dapat membimbing peserta didik tentang cara menilai. Kalau belum mencapai standar kompetensi, peserta didik diberi kesempatan memperbaiki, namun buatlah kontrak batas akhir; kapan diserahkan portofolio tersebut.
- e) Bila dianggap perlu dalam menilai karya peserta didik, adakan pertemuan dengan mengundang orang tua dan warga masyarakat.

Suatu hal yang perlu diperhatikan oleh pendidik adalah penyusunan portofolio yang diserahkan kepada peserta didik betul betul karya yang bersangkutan, autentik dan sesungguhnya. Andai kata ada keraguan, gunakan teknik penilaian yang lain, yaitu simulasi dan demonstrasi atau penilaian unjuk kerja (*performance assessment*).

c. Format Penilaian Portfolio

Agar peserta didik selalu mendapat perhatian dan menjadi fokus dalam penembangan dirinya, maka kriteria penilaian disusun dengan mengikutsertakan peserta didik sesuai dengan kapasitasnya. Aspek-aspek yang dinilai dan tuntutan masing-masing aspek, keutamaan masing-masing aspek dalam suatu karya (karya tulis) hendaklah diperhatikan. Salah satu contoh format penilaian sebagai berikut:

**Format Penilaian Portfolio
(untuk satu portfolio)**

Nama :

Judul portofolio :

Skala *scoring* : Rendah 1 2 3 4 5 Tinggi

No.	Aspek-aspek dinilai	Bobot	Skor
1.	Judul jelas, dan teliti	2	
2.	Dimulai dari pernyataan yang terfokus pada rubric	3	
3.	Konsep utama dirumuskan	5	
4.	Dijelaskan mengapa topik menarik	5	
5.	Analisis dan berpikir kritis	15	
6.	Kesimpulan dirumuskan dengan benar	5	
7.	Menggunakan informasi dari berbagai sumber lima tahun terakhir	5	
8.	Tiap paragraf dimulai dari pokok kalimaat	5	
9.	Penggunaan huruf besar, ejaan, tanda baca dan penampilan	5	

Andai kata portofolio dijadikan kegiatan semesteran atau caturwulan, maka format penilaian dapat pula ditata sesuai dengan kondisi tersebut. Salah satu contoh alternatif:

**Format penilaian portfolio
(individual & semester)**

Nama :

Mata pelajaran : Bahasa Indonesia

No.	Standar kompetensi/ kompetensi dasar	Periode (sesuai dengan jadwal pelajaran)	Tata bahasa	Kosa-kata	Kelengkapan gagasan dan analisis	Sistematika penulisan	Keterangan
1.	Menulis karangan deskriptif	10-9 20-9					
2.	Membuat risensi buku	1-10 20-10					
3.					
4.					

5.6.2 Rubrik

a. Pengertian

Rubrik merupakan salah satu teknik dalam penilaian alternatif. Namun apabila dikaitkan dengan kegiatan yang dilakukan peserta didik dalam memecahkan masalah dalam dunia nyata atau dalam kehidupan sehari-hari, maka rubrik dapat juga disebut sebagai salah satu teknik dalam penilaian autentik. Dalam konteks penilaian alternatif sebagai bagian dari kegiatan pembelajaran seutuhnya, rubrik, dalam arti sangat sederhana, dapat diartikan sebagai pedoman penskora (*scoring guide*) yang disusun guru sebelum tugas penyusunan rubrik diberikan. Sehingga memungkinkan guru membuat keputusan yang dapat dipercaya tentang karya peserta didik dan membolehkan peserta didik menilai dirinya sendiri.

Oleh karena itu, rubrik merupakan perangkat deskripsi suatu tugas atau suatu proses dan mungkin juga suatu kontinum kualitas dari yang tidak baik sampai yang terbaik, yang menjadi dasar keseluruhan skor suatu tugas, pekerjaan, karya atau unjuk kerja/performa atau belajar.

Rubrik dapat dibedakan atas (1) rubrik analisis (*analytic rubric*) dan rubrik holistik (*wholistic rubric*). Rubrik analitik hanya menggambarkan dan menilai bagian dari suatu produk yang telah selesai, sedangkan rubrik holistik memandang suatu produk pekerjaan secara menyeluruh.

Suatu rubrik dikatakan baik dan efektif apabila: (1) terfokus pada satu aktibut; (2) menunjukkan gradiasi yang jelas tentang skala, baik secara kualitatif maupun kuantitatif; dan (3) mengomunikasikan standar kepada peserta didik atau yang lain, sehingga dapat menggunakannya dengan standar kepada peserta didik atau yang lain, sehingga dapat menggunakannya dengan standar yang sama.

Pedoman *scoring* rubrik hendaknya selesai sebelum penilaian dilakukan. Pedoman tersebut kemudian diedarkan dan/atau disampaikan kepada peserta didik, sehingga peserta didik mengetahui secara benar dan tepat apa yang diharapkan dari mereka dan bagaimana mereka dapat mencapai nilai terbaik. Di samping itu perlu didasari bahwa pedoman *skoring* rubrik adalah tipe penilaian formatif, karena penilaian ini sebagai bagian dari proses belajar dan pembelajaran, namun dapat juga digunakan sebagai penilaian sumatif atau dalam pemberian angka.

b. Suatu hal yang perlu diingat

Seandainya guru/pendidik akan menggunakan rubrik sebagai salah satu instruments penilaian, maka guru/pendidik harus menyusun dan mengomunikasikan suatu daftar rubrik sebagai kegiatan pembelajaran yang dilakukannya, kemudian menentukan tipe rubrik apa yang akan disusun oleh peserta didik.

c. Kriteria penskoran rubrik

Dalam menyusun pedoman penskoran perlu kehati-hatian, sehingga pedoman penskoran yang disusun betul-betul menggambarkan komposisi yang sesungguhnya. Apakah akan digunakan sistem pembobotan atau tidak. Mana yang dipilih tergantung pada kemampuan, ketajaman dan profesionalisme guru/pendidik yang membina mata pelajaran tersebut.

Secara umum, penyusunan rubrik analitik sebagai berikut:

- (1) Tentukan jumlah kolom dan baris yang dibutuhkan. Jumlah baris sangat tergantung pada seberapa banyak tujuan atau karakter yang diinginkan.
- (2) Kolom bagian kiri diisi dengan No. Urut, dan pada kolom berikutnya diisi dengan pernyataan tujuan/kriteria aspek-aspek yang dinilai. Satu pernyataan untuk tiap baris.
- (3) Adapun pada kolom berikutnya (kalau ada) diisi dengan bobot.
- (4) Kolom berikutnya adalah kolom untuk menentukan hasil akhir seseorang.
- (5) Pada kolom berikutnya adalah kolom untuk menentukan hasil akhir seseorang, yaitu dengan mengalikan nilai masing-masing aspek yang dinilai dengan bobot yang ditentukan sebelumnya.

**Format penskoran rubrik yang ditulis peserta didik/siswa
(student written journal)**

Nama : _____ tanggal

Judul rubrik :

Skala skoring : rendah 1 2 3 4 5 tinggi

No.	Aspek-aspek dinilai	bobot	skor	hasil
1.	Titel jelas, dan teliti	2		
2.	Dimulai dari pernyataan yang terfokus pada rubrik	3		
3.	Konsep utama dirumuskan	5		
4.	Dijelaskan mengapa topik menarik	5		
5.	Analisis dan berpikir kritis	15		
6.	Kesimpulan dirumuskan dengan benar	5		
7.	Menggunakan informasi dari berbagai sumber dan lima tahun terakhir	5		
8.	Tiap paragraf dimulai dari pokok kalimat	5		
9.	Penggunaan huruf besar, ejaan, tanda baca dan penampilan	5		

Bentuk lainya, merupakan skor angka yang berbeda antar kriteria, di mana pebobotan secara eksplisit sudah ada didalam skor, sebagai berikut:

No.	Kriteria	Skala Scoring					Skor
		2	4	6	8	10	
1.	Kejelasan judul	2	4	6	8	10	
2.	Layout	2	4	6	8	10	
3.	Kejelasan perumusan konsep utama	3	4	9	12	15	
4.	Analisis, keakuratan, dan kemutahiran sumber	10	20	30	40	50	
5.	Penggunaan huruf,tanda baca, dan style	1	2	3	4	5	
6.	Kesimpulan	2	4	6	8	10	

Andai kata karya rubrik disajikan dalam diskusi kelompok, maka format penilaiannya dapat dilakukan dalam bentuk, antara lain:

Format penilaian rubrik yang disajikan dalam diskusi kelompok

Nama :

Tanggal :

No.	Kriteria	Bobot	Skor			Hasil
			1	2	3	
1	Sikap, prilaku, dan penampilan	10				
2	Penguasaan materi	1				
3	Penggunaan bahasa, tata bahasa, kosakata, dan bahasa tubuh	10				
4	Kesadaran dan pemahaman audiens	5				
5	Cara menjawab pertanyaan	5				
6	Menyimpulkan pendapat	5				

Setiap pendidik perlu meyakini bahwa tidak ada suatu jenis rubrikpun yang paling baik, yang bisa berlaku untuk semua karakter atau aspek yang dinilai. Dalam kaitan itu, guru hendaknya betul-betul memahami karakteristik materi ajar sesuai dengan rencana pelaksanaan pembelajaran dan tujuan pembelajaran, dan merumuskan kriteria aspek yang dinilai dengan benar. Suatu rubrik dikatakan baik jika: (1) membantu perkembangan kesadaran peserta didik dalam belajar, (2) menyediakan iklim yang adil, jujur dan fleksibel; dan (3) memberanikan peserta didik untuk bertanggung jawab terhadap cara belajar mereka masing-masing.

5.6.3 Jurnal belajar (*learning journal*) atau jurnal siswa (*student journal*)

Peserta didik adalah orang yang paling tahu tentang kondisi dirinya yang sesungguhnya. Karena itu, dalam proses penilaian dia pula yang paling tepat menyatakan kemajuan itu apabila ditinjau dari sudut pandang peserta didik sebagai fokus pembelajaran, berkenaan dengan pemahaman, penguasaan, sikap dan keterampilannya tentang sesuatu yang sudah dibelajarkan. Untuk itu, peserta didik menyusun suatu jurnal belajar (*learning journal*), yaitu suatu

catatan yang dibuat peserta didik sendiri (*self-report*) secara naratif; catatan pribadi yang terpusat pada aspek-aspek materi pelajaran yang dipelajari, yang mempunyai nilai personal dan relevan. Catatan tersebut dapat berupa observasi pribadi, perasaan, opini dalam merespons bacaan-bacaan, kejadian dan pengalaman.

Catatan siswa itu dapat berupa sikap, perasaan, pikiran, persepsi maupun kesan tentang suatu materi pelajaran. Pada saat penyekoran, jangan memandang catatan itu dari segi baik atau buruknya, namun lihatlah dari kemampuannya dalam menganalisis, menerangkan atau menjelaskan sesuatu.

5.6.4 Catatan ringkas tentang belajar (*learning log*)

Learning log merupakan catatan ringkas yang disusun peserta didik sendiri (*self-report*) tentang materi pelajaran yang telah dipelajarinya. Catatan ringkas tersebut dapat berupa: (1) perubahan unjuk kerja dalam belajar, (2) dokumen kemajuan yang telah dicapai peserta didik; dan (3) dokumentasi partisipasi peserta didik dalam belajar.

5.6.5 Penilaian diri (*self assessment*)

Penilaian diri adalah suatu teknik penilaian, di mana peserta didik diminta untuk menilai dirinya sendiri berkaitan dengan status dan tingkat pencapaiannya kompetensi yang dipelajarinya dalam mata pelajaran tertentu didasarkan atas kriteria atau acuan yang telah disiapkan. Oleh karena itu, penilaian diri merupakan suatu teknik penilaian di mana seseorang belajar lebih tentang dirinya. "lebih" dalam konteks ini, berarti lebih banyak, lebih mendalam, lebih seru dan lebih hati-hati tentang dirinya (sendiri), sehubungan dengan materi pelajaran/tugas-tugas yang disampaikan.

Dalam pelaksanaannya, penilaian diri mengikuti langkah-langkah yang sistematis, sebagai berikut:

- a. Menjelaskan kepada peserta didik tentang penilaian diri.
- b. Menentukan kompetensi yang akan dinilai.
- c. Menentukan kriteria penilaian.
- d. Merumuskan dan menyusun format penilaian.

- e. Meminta peserta didik untuk menilai diri sendiri.
- f. Pendidik mengkaji penilaian peserta didik.
- g. Melakukan tindak lanjut; pendidik dan peserta didik membahas penilaian secara bersama-sama.

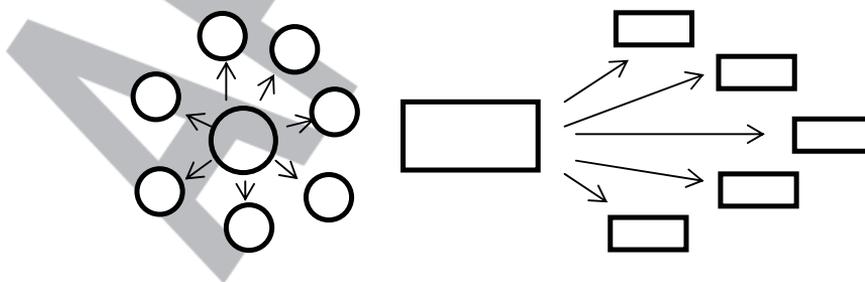
Dengan mengikuti langkah-langkah di atas, subjektivitas peserta didik dalam memberi angka dapat diminimalkan.

5.6.6 Peta Konsep (*Concept of Map*)

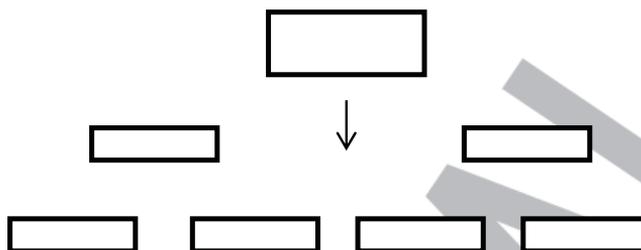
Dilihat dari segi fungsinya, peta konsep merupakan suatu teknik penilaian yang dapat digunakan dalam proses pembelajaran maupun untuk menentukan tingkat pencapaian peserta didik dalam belajar. Peta konsep merupakan suatu grafik keterwakilan tentang suatu konsep dan hubungannya dengan sub-aspek dalam konsep tersebut, yang merefleksikan suatu pemahaman tentang bagaimana suatu konsep berhubungan antara satu dan yang lain dan kepada suatu skema yang sudah ada. Oleh karena itu, ada konsep inti (*core concept*), yang biasanya diletakkan dalam suatu lingkaran atau kotak. Relasi dinyatakan dengan garis dan panah yang menghubungkan konsep-konsep, dan dengan proposisi atau pernyataan garis yang menyatakan hakikat hubungan.

Peta konsep digunakan untuk mengomunikasikan idea-idea dan relasi-relasi yang kompleks dan membuat struktur berpikir peserta didik menjadi lebih sederhana, sehingga mendorong belajar menjadi lebih bermakna. Beberapa bentuk peta konsep yang dapat disusun sebagai berikut:

- a. Jaringan laba-laba, di mana konsep utama diletakkan pada bagian tengah dan sub-sub konsep lain mengitarinya.



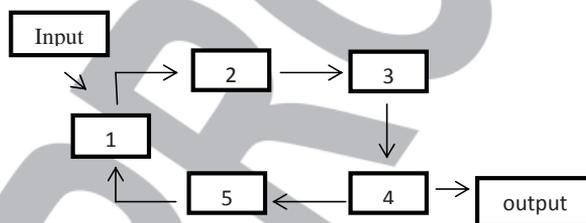
b. Bertingkat (Hierarchy)



c. Tata alir (Flowchart): dengan informasi berurutan (order) dalam format linerar



d. Sistem (system), dengan arah hubungan seperti dalam tata alir; informasi bertingkat dengan sistem input dan output.



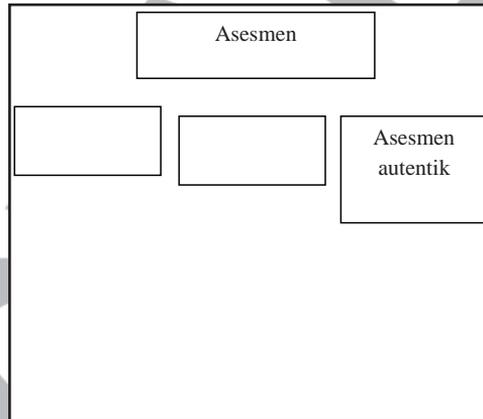
Peta konsep dapat digunakan sebelum proses pembelajaran dimulai; dengan menugaskan peserta didik membaca materi yang akan disampaikan, kemudian peserta didik disarankan untuk membuat peta konsep berkenaan dengan materi pelajaran yang sudah dibacanya. Dapat juga dilakukan di akhir pembelajaran, dengan menugaskan peserta didik membuat intisari materi yang telah disampaikan dalam bentuk peta konsep dan kemudian menyerahkannya kepada pendidik.

Disamping itu, sebagai penilaian sumatif, peta konsep sangat bermanfaat diaplikasikan apabila pendidik ingin memahami kerangka pemikiran, pemahaman konsep dan luasnya cakupan penguasaan peserta didik tentang materi yang telah diberikan. Hal itu dapat dilakukan dengan menggunakan peta konsep berupa kotak-kotak yang masih dikosongkan. Selanjutnya, kotak-kotak tersebut diisi dengan materi konsep-konsep

yang sudah dipelajari pada posisi/hirerarkhi/level yang sesuai dengan seharusnya.

Isilah: peta konsep dengan daftar yang telah disediakan pada posisi yang tepat

1. Asesmen
2. Asesmen kelas
3. Asesmen alternatif
4. Asesmen autentik
5. Rubrik
6. Portofolio
7. Meringkas dalam satu kalimat
8. One minute paper
9. Debat
10. Demonstrasi



5.6.7 Diskusi Kelompok Terfokus (*Focused Grup Discussion*)

Walaupun diskusi kelompok terfokus banyak digunakan untuk keperluan pemantapan suatu konsep melalui masukan dari peserta diskusi atau menghimpun sejumlah pendapat, namun tidak kalah pentingnya digunakan dalam menilai kemampuan berargumentasi peserta didik dalam suatu diskusi. Fokus diskusi adalah materi pelajaran yang ditugaskan untuk dipelajari. Strategi dan metode pembelajaran yang digunakan adalah diskusi dan peserta diskusi adalah peserta didik yang ambil bagian dalam materi tersebut.



Bab 6

PENILAIAN AUTENTIK (*AUTHENTIC ASSESSMENT*)

Pada prinsipnya, penilaian autentik adalah penilaian alternatif, namun tidak semua penilaian alternatif adalah penilaian autentik. Dikatakan penilaian autentik jika guru meminta peserta didik mengaplikasikan keterampilan dan kemampuannya untuk memecahkan masalah seperti dalam situasi di dalam kehidupan nyata.

6.1 PENGERTIAN PENILAIAN AUTENTIK

Munculnya konsep penilaian autentik di awal 1990-an, sebagai wujud ketidakpuasan para ahli terhadap kelemahan-kelemahan tes objektif, terutama untuk menilai kemampuan kognitif tingkat tinggi dalam melakukan sesuatu di kehidupan yang sesungguhnya atau *real world setting* (Yusuf, 2015). Menurut Mueller (2008): *Assessment Authentic: A form of assessment in which students are asked to perform real-world tasks that demonstrate meaningful application of essential knowledge and skills*. Adapun Linn dan Gronlund (1995) menyatakan bahwa: *Authentic assessment is any type of alternative assessment done in a "realworld setting,"* yang berarti bahwa penilaian autentik itu adalah sejenis penilaian alternatif yang dikerjakan di dalam setting dunia nyata. Oleh karena itu penilaian autentik juga merupakan penilaian alternatif, apabila peserta didik melakukan, menerapkan atau melaksanakan suatu tugas di dalam kehidupan nyata. Tetapi, tidak berarti semua teknik penilaian alternatif dapat digunakan dalam penilaian autentik.

6.2 CIRI-CIRI PENILAIAN AUTENTIK

Beberapa ciri asesmen autentik, sehingga tampak bedanya dari asesmen dengan pendekatan tradisional, atau dengan teknik asesmen alternatif yang lain, adalah sebagai berikut.

1. *Real-Life* dan *On Going*

Peserta didik terlibat dan melibatkan diri dalam tugas-tugas nyata dalam kehidupan nyata, menampilkan unjuk kerja/kinerja atau melakukan pemecahan masalah di lapangan. Oleh karena itu, penilaian autentik mengharuskan peserta didik melakukan, mendemonstrasikan, menciptakan dan/atau mengerjakan suatu tugas dalam kehidupan nyata, sehingga mereka memperoleh pemahaman yang lebih kaya mengenai apa yang dipikirkannya dan bagaimana mengontruksi makna (*meaning*).

2. Kriteria yang Akan Digunakan Dimenegerti Sejak Awal

Kisi-kisi, kriteria dan format penilaian yang akan digunakan sudah peserta didik telah pahami sejak awal. Kriteria dan Format tersebut telah dikomunikasikan kepada peserta didik sejak awal kegiatan pembelajaran dimulai, bahkan mereka boleh bertanya tentang teknik dan format yang akan digunakan. Dengan kata lain, peserta didik sejak awal betul-betul tidak ragu lagi tentang aspek-aspek yang akan dinilai, bagaimana cara menilainya dan bobot skor setiap aspek yang dinilai.

3. Valid dan Reliabel

Instrumen yang digunakan betul-betul dirancang berdasarkan target belajar, tujuan dan kompetensi; serta sesuai dengan karakteristik materi pelajaran yang diberikan dan dengan pengalaman belajar yang telah berlangsung. Penilaian buku sekali jadi, dan bukan pula hanya dengan satu jenis instrumen. Kesahihan dan keterandalan instrumen, tidak diragukan lagi. Instrumen-instrumen yang digunakan telah divalidasi oleh teman sebaya, dan mudah dibaca oleh peserta didik karena telah disiapkan jauh sebelumnya. Kritik dan saran tentang instrumen penilaian telah ditampung sebelum kegiatan pembelajaran berlangsung, kemudian disempurnakan lagi berdasarkan saran-saran tersebut.

4. Peserta Didik Merancang Sendiri Tugasnya
Menggunakan tes objektif dalam menilai kemajuan belajar peserta didik, berarti “menggiring” peserta didik memilih jawaban dari alternatif jawaban yang telah disediakan; sedangkan menggunakan berbagai teknik penilaian autentik dalam menilai kemajuan belajar peserta didik, berarti pendidik mengembangkan kemampuan peserta didik dalam merancang sendiri tugasnya. Dengan cara demikian, penilaian autentik ikut membantu dalam mengembangkan nalar dan struktur berpikir sehingga sadar bagaimana cara menarik kesimpulan yang benar; bukan hanya memilih dari alternatif yang sudah disediakan.
5. Mengembangkan dan Mengutamakan Kemampuan Berpikir Tingkat Tinggi
Penilaian autentik mewajibkan peserta didik untuk merancang sendiri, melakukan atau menerapkannya dalam kehidupan nyata. Oleh karena itu, dengan menggunakan berbagai teknik-teknik dalam penilaian autentik, berarti peserta didik minimal mengaplikasikan pengetahuan, pemahaman baru dan kemampuan berpikirnya; bukan menyebutkan atau mengulang sesuatu yang sudah dipelajari.
6. Autentik/Dalam Situasi Nyata
Karakteristik lain adalah peserta didik dinilai pada saat ia menerapkan atau melakukan sesuatu dalam kehidupan nyata. Apa yang ia tampilkan dan apa yang ia kerjakan itulah kemampuan ia yang sesungguhnya.
7. Komprehensif dan Terintegritas
Sisi lain yang perlu mendapatkan perhatian dalam penilaian autentik adalah aspek-aspek yang diamati maupun yang dinilai bukan hanya dari satu sisi, seperti fakta atau kognitif saja, melainkan terpadu secara utuh dan menyeluruh. Hal ini sangat ditentukan oleh bentuk tugas dan seni mengelaborasi tugas dalam kehidupan nyata. Bukan faktanya yang ingin diungkap, melainkan kemampuan mengaplikasikan fakta dalam situasi baru, kemampuan mengorganisasikan, serta kejelasan dan ketepatan sikap dan prilakunya dalam bertindak atau menyatakan sesuatu.
8. Menekankan Proses dan Produk
Perhatian dan pertimbangan penilaian autentik tertuju bukan saja pada proses pelaksanaan suatu tugas akan tetapi pada produk yang

dihasilkan. Pengetahuan dan kemampuan seseorang dalam membuat suatu tugas, belum bisa mewakili secara utuh. Karena itu, produk yang dikerjakannya perlu pula dinilai. Keterpaduan kedua fokus - proses dan produk - penilaian tersebut menggambarkan peserta didik yang sesungguhnya dalam mata pelajaran yang dinilai. Hasil penilaian akan baik kalau prosesnya bernilai baik dan produknya juga bernilai baik.

9. Mengutamakan Fakta dan Bukti Langsung
Fakta dan bukti-bukti langsung adalah bagian yang tak terpisahkan dari dari penilaian autentik. Penilai tidak boleh cepat percaya tanpa melihat bukti-bukti tugas atau kegiatan yang telah dilaksanakan oleh peserta didik, dalam kaitannya dengan proses pembelajaran secara utuh.

6.3 TEKNIK-TEKNIK YANG DIGUNAKAN DALAM PENILAIAN AUTENTIK

Sebagaimana telah dibahas pada uraian terdahulu, penilaian autentik adalah penilaian alternatif manakala pelaksanaannya dilakukan dalam kehidupan nyata (*real world/real life setting*). Karena itu semua teknik penilaian alternatif dapat digunakan dalam penilaian autentik, kalau dapat dikaitkan dengan kehidupan nyata.

Secara umum teknik-teknik yang dapat digunakan dalam penilaian autentik adalah:

1. Observasi,
2. Pertanyaan Lisan/Pertanyaan Terbuka,
3. Presentasi Kelas,
4. Proyek,
5. Tugas-tugas,
6. Jurnal,
7. Kerja Kelompok,
8. Portofolio,
9. Rubrik,
10. Interviu,
11. Kelompok Terfokus,
12. Tes Unjuk Kerja,

13. Percobaan/Demonstrasi
14. Debat/Diskusi
15. Peta Konsep
16. Ekshibisi, dan
17. Poster.

-oo0oo-

APPROVAL

Bab 7

PENENTUAN NILAI ATAU GRADE

Tujuan Bab Ini: Setelah mempelajari bab ini mahasiswa diharapkan:

1. Bisa mendefinisikan *grade* atau nilai.
2. Bisa menjelaskan makna *grade* hasil belajar.
3. Bisa mengaplikasikan dan mempertimbangkan perbedaan individual dalam menentukan *grade*.
4. Bisa mengidentifikasi dan menjelaskan macam-macam sistem *grade*.
5. Bisa menganalisis perbedaan antara *grade* tunggal, *grade* ganda dan *grade* kategorik.
6. Bisa mengaplikasikan cara menentukan sistem penilaian terakhir.
7. Bisa menerapkan penyelenggaraan evaluasi hasil belajar di kelas.

Secara periodik, pembelajaran siswa dan program instruksional yang telah dilakukan oleh seorang guru dapat dinilai dengan beberapa cara. Misalnya, diberi pertanyaan untuk diberikan nilai akhir yang dapat dicapai oleh siswa. Diuji dengan tes tertulis untuk mengetahui kemampuan siswa dalam menerima materi yang telah diberikan, atau dapat juga dijumlah total skor hasil belajar dalam satu semester dan dibagi dengan jumlah siswa yang mengikuti ujian.

Guru dapat mengajar para siswa tanpa memberikan *grade* atau skor hasil belajar mereka. Pada sisi lain, siswa juga bisa belajar tanpa diberi skor atau *grade*. Skor atau *grade* juga tidak menghasilkan peningkatan pengajaran dan juga tidak meningkatkan pembelajaran pada siswa. Hampir semua

guru diminta atasannya, dalam hal ini kepala sekolah, kepala ranting, kepala cabang dan kepala wilayah pendidikan nasional untuk memberikan keputusan penilaian yang di dalamnya mencakup *grade* dalam interval yang resmi dan baku.

Grade hasil belajar siswa dapat didefinisikan sebagai berikut: *Grade is major division of the instructional program on schools, representing the working of one school year* (Good: 2013). *Grade* dapat juga diartikan derajat atau angka yang merupakan bagian program instruksional di sekolah, dan menggambarkan kinerja siswa dalam periode satu tahun. Pengertian tentang *grade* juga dinyatakan oleh Johnson dan Johnson, *grade* sebagai simbol yang mungkin berupa huruf, angka, atau kata-kata yang menggambarkan pertimbangan nilai relatif pencapaian hasil belajar selama waktu tertentu. Waktu tertentu ini bisa satu tahun, satu semester, atau satu kuartal, tergantung sistem yang berlaku di suatu lembaga sekolah. Dalam keadaan nyata, *grade* kadang rancu dengan skor atau nilai yang dicapai siswa dalam mengikuti satu ujian. *Grade* yang hendak dibahas pada bab ini adalah penentuan hasil akhir setelah siswa mengikuti proses pembelajaran dalam satu semester atau satu kuartal. *Grade* atau nilai akhir ini sesuai dengan Kurikulum Tingkat Satuan Pendidikan tahun 2006, terdiri atas tiga macam penilaian, yaitu

- a. Penilaian pendidik atau para guru yang lebih memfokuskan pada perubahan perilaku dan penguasaan pengetahuan,
- b. Penilaian tingkat satuan yang merupakan penilaian sekolah yang juga memerhatikan penilaian guru dan penguasaan keterampilan yang sesuai dengan tingkat satuan pendidikan, dan
- c. Penilaian pemerintah, yaitu penilaian yang dipresentasikan pada penguasaan beberapa mata pelajaran yang telah ditentukan pemerintah.

Isu yang berkaitan dengan sistem penentuan *grade* hasil belajar, secara keseluruhan sudah ada sejak lama di dunia pendidikan. *Grade* berkembang sesuai dengan tuntutan masyarakat pengusaha, industri, atau lembaga lain yang memerlukan tenaga para lulusan sekolah. Pada tahun 1950-an sampai 1960-an, berlaku nilai 6 dan 7. Ada dua kali latihan ujian sekolah untuk menghadapi ujian yang sebenarnya. Hal ini merupakan hal yang wajar. Mengingat bahwa ketika beberapa siswa tidak lulus atau tidak naik kelas disebabkan tidak mencapai hasil belajar atau nilai rapor lebih kecil dari 6, maka tidak ada protes, apalagi membenci para guru kelas.

Pada tahun 1970-an sampai tahun 1990-an, semua sekolah harus meluluskan semua siswanya dengan alasan menjaga “kestabilan nasional”, anak yang tidak lulus atau tidak naik, bisa protes atau mengancam gurunya. Pada waktu itu, siswa pandai maupun bodoh dapat naik kelas. Pada tahun 2000-an, pemerintah menetapkan nilai 4,25 sebagai batas minimal kelulusan. Tidak ada latihan ujian serta tidak ada ulangan sebagai persiapan ujian nasional. Hal itu menimbulkan, banyak protes dari orang tua, lembaga sosial masyarakat, atau masyarakat yang peduli pendidikan. Itulah dinamika perkembangan sistem *grade* atau nilai pada pendidikan nasional di Indonesia.

7.1 MAKNA *GRADE* HASIL BELAJAR

Keberadaan dan penggunaan nilai akhir atau *grade* masih relevan sampai sekarang, walaupun secara faktual *grade* itu tidak mencerminkan adanya peningkatan pembelajaran dan pengajaran seperti tersebut di atas. *Grade* atau skor diberikan sebagai simbol yang mempresentasikan hasil belajar seorang siswa. Di samping itu, *grade* juga merupakan simbol yang merefleksikan komunikasi evaluasi sumatif yang diberikan guru sebagai media komunikasi dan laporan kepada orang tua, kepala sekolah, dan para *stakeholders* yang berkepentingan. Apakah *grade* atau nilai itu? Menurut Good (2013), “*grades are symbols that represent a value judgment concerning the relative quality of students achievements during specific periods of interaction*” atau *grade* adalah simbol (angka, huruf, atau kata) yang menggambarkan nilai pertimbangan yang berkaitan dengan kualitas siswa dalam berprestasi selama periode pengajaran.

Pada prinsipnya, penentuan *grade* dengan penilaian skor dari suatu hasil evaluasi hampir sama. Penentuan *grade* biasanya dilakukan setelah beberapa kali evaluasi, sedangkan skor penilaian merupakan hasil yang dicapai siswa hanya untuk satu kali evaluasi saja. Kemudian, skor penilaian didokumentasikan dan digabungkan dengan skor evaluasi lain sehingga dapat digunakan untuk menentukan *grade* pada akhir semester atau kuartal. *Grade* tersebut diambil dari mata pelajaran yang ada pada setiap semester atau kuartal, lalu *grade* yang ada digunakan sebagai materi utama laporan hasil belajar atau rapor kepada orang tua dan kepada kepala sekolah. Jadi *grade*, dengan skor penilaian itu berbeda, yaitu bahwa skor penilaian

menggambarakan satu kali hasil evaluasi, sedangkan *grade* merupakan hasil rerata atau gabungan skor yang dicapai oleh setiap siswa dalam mengikuti proses evaluasi pada setiap unit.

Untuk menentukan *grade*, guru meringkas dan menggabungkan beberapa aspek pencapaian hasil belajar, misalnya kehadiran; partisipasi di kelas; ketaatan dalam mengikuti ujian awal, tengah, dan akhir masa belajar.

Grade, atau nilai akhir memiliki arti yang sangat penting karena nilai akhir tersebut dapat menentukan apakah siswa dikatakan pandai atau tidak, bisa melanjutkan ke jenjang yang lebih tinggi atau tidak, bahkan *grade* selalu menjadi bagian integral yang dipertimbangkan ketika mereka akan bekerja. *Grade* yang baik menentukan karier siswa atau prospek jenjang sekolah atau bahkan beberapa kemungkinan jabatan terbuka bagi siswa yang bersangkutan.

Kita bisa mengamati pengaruh *grade* terhadap perilaku seseorang dalam kehidupan sehari-hari. Misalnya, sering kita saksikan orang tua bisa marah besar ketika melihat *grade* anaknya sebagian besar merah. Sebaliknya, orang tua akan tersenyum puas memuji putra-putrinya ketika hasil pencapaian belajar mereka mendapatkan *grade* yang semuanya bagus. *Grade* siswa juga menentukan apakah ia menerima gelar penghargaan akademik dari lembaga sekolah, misalnya *cumlaude*, beasiswa, atau promosi. Dengan kata lain, kekuatan *grade* menentukan prospek hidup siswa di masyarakat. Salah satu tujuan penentuan *grade* adalah mengomunikasikan nilai pada *stakeholders* dalam evaluasi pendidikan, di antaranya adalah orang, lembaga, atau masyarakat yang di dalamnya termasuk orang tua, siswa, kepala sekolah, dan pejabat di kantor Diknas.

Manfaat dan arti *grade* bervariasi sesuai dengan fungsi dan peranannya terhadap para pelaku yang berkepentingan. Bagi siswa, nilai menunjukkan pencapaian hasil belajar siswa. Oleh karena itu, para siswa perlu mengetahui sistem *grade* dengan baik agar mereka tetap termotivasi untuk belajar secara kontinu. Sedangkan bagi para guru, *grade* mempunyai makna yang bervariasi dengan melihat skor pencapaian hasil belajar. Seorang guru akan dapat menebak dan mengatakan kamu tidak belajar, ya, dalam ulangan yang lalu. Sebaliknya, seorang guru akan tersenyum dan memuji siswa untuk belajar terus karena melihat skor hasil belajar yang menunjukkan keberhasilan dalam ulangan.

Grade hasil belajar akhir yang didasarkan atas tingkah laku dan penampilan yang terarah dalam tes yang terorganisasi dengan baik, memiliki derajat yang lebih tinggi dibanding dengan *grade* yang hanya didasarkan pada tes kertas dan pena saja. Pada lingkup yang lebih luas, termasuk lingkup sekolah atau lembaga pendidikan, *grade* sebagai symbol yang menunjukkan keberhasilan siswa. Sebagaimana besar orang tua akan cepat memahami jika para siswa menunjukkan *grade* yang tinggi, misalnya A dan B. Mereka merasa bangga dan mendorong anaknya untuk lebih menekuni lagi apa yang telah bisa dicapai sehingga menjadi lebih baik lagi.

7.2 MEMPERTIMBANGKAN PERBEDAAN INDIVIDUAL

Banyak dimensi yang berkaitan dengan *grade* yang harus guru dan evaluator perhatikan. Salah satunya yang harus selalu diperhatikan dalam penentuan *grade* adalah bagaimana penilaian yang diberikan mampu memberikan peluang diketahuinya perbedaan potensi di antara individual siswa. Untuk mencapai hal itu, penilaian kertas dan pena saja belum cukup. Hal ini dikarenakan penilaian menggunakan kertas dan pena lebih menekankan pada aspek kognitif dari pada aspek lainnya.

Apalagi jika pendidikan menekankan pada aspek lain, misalnya aspek keterampilan dan kualitas pribadi yang pada umumnya terdapat dalam ranah afektif siswa. Untuk itu, jenis butir tes alternatif lain yang lebih mengakomodasikan keperluan tersebut dapat dipilih untuk mengungkap kemampuan masing-masing siswa.

Siswa yang pandai secara akademik memiliki kecenderungan dapat mengerjakan dengan baik tes kertas dan pena, tes interpretif, pertanyaan *problem solving*, maupun tes dalam bentuk pertanyaan esai. Pada sisi lain, siswa yang lambat pada umumnya mengalami kesulitan dengan bentuk-bentuk tes tersebut. Biasanya, siswa yang lambat lebih berhasil dalam tes alternative. Oleh karna itu, guru perlu memahami fenomena tersebut dengan cara bijak. Para siswa yang tergolong lambat mungkin memerlukan bentuk tes lain, selain kertas dan pena.

Masih banyak guru atau evaluator mengabaikan aspek kualitas pribadi yang secara umum dapat diungkap dari ranah afektif dan ranah keterampilan. Untuk dapat mengungkap potensi siswa yang kurang berhasil

dengan tes dapat diungkap melalui kegiatan evaluasi, seperti laboratorium dan pekerjaan rumah (*project work*). Pada umumnya, evaluasi kualitas pribadi tersebut sangat memperhatikan aspek-aspek penilaian yang kurang kurang diakomodasikan oleh tes kertas dan pena. Beberapa aspek lain yang mempengaruhinya antara lain aspek penghargaan, nilai, kepercayaan, kerja sama, hubungan antar pribadi, dan sikap siswa.

Penentuan *grade* yang mampu mengungkap potensi siswa pandai, rerata, dan lambat, pada kenyataannya adalah sangat sulit dan menantang untuk direalisasikan. Oleh karena itu, sistem penilaian yang lebih baik masih perlu dikembangkan. Tujuan utama penentuan *grade* bukan saja menunjukkan kemampuan kognitif atau aspek akademik saja, tetapi juga mampu mengungkapkan pengetahuan yang berasal dari aspek keterampilan dan kualitas personal secara proposional. Semua aspek itu memberikan kontribusi nyata bagi seorang siswa dalam upaya menjadikannya sebagai warga masyarakat yang produktif dan berhasil.

7.3 MACAM-MACAM SISTEM GRADE

Secara garis besar, sistem *grading* dalam evaluasi pendidikan termasuk pendidikan bahasa dapat dibedakan menjadi tiga macam, yaitu *grade* tunggal, *grade* ganda, dan *grade* kategorik.

1. *Grade* Tunggal

Grade tunggal bisa didefinisikan sebagai sistem penentuan *grade* yang bentuknya paling sederhana dan paling banyak digunakan.

Keunggulan *grade* tunggal ini meliputi, antara lain:

- a. Memberikan pesan yang ringkas tentang pencapaian hasil belajar.
- b. Lebih mudah dipahami.
- c. Memberikan hasil prediksi keberhasilan siswa dalam belajar, dan
- d. Memberikan motivasi untuk belajar lebih baik.

Namun di samping keunggulan tersebut, *grade* tunggal juga memiliki kelemahan, yang meliputi antara lain:

- a. Tidak memberikan gambaran hasil yang jelas.
- b. Acuan penilaiannya masih terbatas.
- c. Bisa menimbulkan keraguan pada siswa yang bersangkutan, dan

- d. Bisa membuat benji, karena adanya perbedaan antara usaha dengan hasil yang dicapai.

Kita bisa mengenali *grade* tunggal dengan mudah. Misalnya, pada sistem *grade* tunggal ini, siswa menerima hasil belajar mungkin dalam bentuk angka, seperti 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 untuk rentang 1 sampai 10. Di samping itu, ada pula *grade* dengan rentang 10 sampai 100. *Grade* tunggal juga dapat menggunakan huruf atau sistem abjad, yaitu A, B, C, D, dan E. Sistem *grade* tunggal dengan huruf mempunyai makna sebagai berikut.

Table 7.1 Makna *Grade* Hasil Belajar Siswa

<i>Grade</i> yang Dicapai	Makna Huruf	Poin
A	Sangat bagus	4
B	Bagus	3
C	Cukup atau rerata	2
D	Kurang berhasil	1
E	Gagal	0

2. *Grade ganda (Multigrade)*

Definisi *grade* ganda atau *multigrade* adalah sistem penentuan hasil belajar yang banyak digunakan dalam konteks evaluasi pendidikan. Istilah *multigrade* dapat diartikan sebagai penentuan skor yang terdiri atas ketentuan nilai hasil belajar yang memiliki makna berbeda untuk sistem instruksional yang berbeda (Yusuf, 2015). Misalnya, dalam konteks evaluasi program, nilai hasil belajar untuk penilaian produk, berbeda dengan nilai hasil belajar untuk penilaian proses. Kedua nilai tersebut juga berbeda maknanya dengan nilai hasil belajar pada kriteria progres. Untuk mendapatkan keputusan ketentuan nilai akhir, seorang guru dapat mengkombinasikan aspek-aspek pendidikan yang dimaksud sehingga memiliki hasil akhir dari hasil belajar.

Guru yang hendak menentukan nilai akhir atau *grade* dengan cara mengkombinasikan nilai perlu memperhatikan oleh seorang bahwa semakin banyak aspek kombinasi yang digunakan, semakin banyak unsur subjektif berpengaruh. Untuk mengatasi hal tersebut, guru itu perlu menggunakan formulasi baku sehingga dapat digunakan dengan lebih objektif.

Berikut adalah contoh formal penilaian tesis yang menggunakan evaluasi model *grade* ganda.

Penilaian Seminar Proposal Tesis

Nama :

Nomor Registrasi :

Judul Proposal :

No	Aspek yang Dinilai	Bobot	Skor	Nilai
1.	Signifikasi dan kontribusi topic terhadap teori dan praktik kependidikan			
2.	Logika alur permasalahan: latar belakang, identifikasi, pembatasan, dan rumusan masalah			
3.	Kedalaman dan relevansi kajian literature, kerangka berpikir, pertanyaan penelitian, dan atau hipotesis penelitian			
4.	Ketepatan pendekatan penelitian dengan permasalahan yang hendak dipecahkan			
5.	Kelengkapan aspek metodologi penelitian dengan permasalahan yang dipilih			
6.	Kelengkapan administrasi (persetujuan pembimbing ketepatan waktu, tata tulis)			
Jumlah Nilai:				
Catatan khusus:				

3. *Grade* Kategorik

Sistem lain yang juga banyak digunakan di sekolah menengah ataupun perguruan tinggi adalah system *grade* dengan dua kategori, yaitu lulus-tidak lulus. Sistem ini juga dapat dinyatakan dengan cara lain, di antaranya memuaskan-tidak memuaskan atau lulus-gagal. Sistem penilaian dengan dua kategori ini tidak memiliki pilihan, misalnya rerata.

Sistem *grade* kategorik, pada umumnya, digunakan untuk memberikan kesempatan kepada para siswa atau mahasiswa yang ingin mengeksplorasi cakupan pengetahuan baru, dengan tetap di bawah bimbingan dosen atau guru pengampu. Cara ini lebih baik dari pada cara belajar dengan menggunakan model otodidak, karena para siswa masih tetap mendapatkan bimbingan, petunjuk, dan pemberian materi baru oleh dosen atau guru pengampu. Di samping itu, sistem *grade* kategorik ini juga digunakan untuk mengakomodasi mahasiswa atau siswa yang wajib mengambil mata kuliah atau pelajaran yang memerlukan jangka waktu lebih dibandingkan waktu yang umum digunakan, misalnya semester atau kuartal. Mata kuliah yang dimaksud, yaitu mata kuliah penulisan tugas akhir, seperti skripsi, tesis, atau disertasi.

7.4 MENENTUKAN SISTEM PENILAIAN TERAKHIR

Seorang guru sering kali hanya menerima ketentuan *grade* yang ada dari pimpinan sekolah. Dalam hal ini, guru hanya melaksanakan tugas sesuai dengan ketentuan yang ditentukan. Mengingat pentingnya *grade* hasil belajar terhadap prospek siswa yang bersangkutan, maka sebaiknya guru juga memahami bagaimana mekanisme *grade* akhir dilakukan agar tetap bijak dalam menentukan *grade* hasil belajar. Berikut ini diuraikan beberapa langkah penting yang perlu diperhatikan.

1. Sejak awal pertemuan dengan para siswa, tentukan makna dari *grade* atau skor yang direncanakan. Penentuan *grade* diberitahukan bersamaan dengan pertemuan awal ketika guru mendiskusikan materi dalam satuan pelajaran selama satu semester.
2. Tentukan penampilan (*performance*) apa yang perlu dimasukkan dalam *grade*. Apa yang perlu diperhatikan oleh para siswa dalam memenuhi penampilan yang diperlukan untuk penentuan *grade*.
3. Beritahukan kepada para siswa tentang bagaimana proses penentuan *grade* dilakukan, agar mereka dapat meningkatkan motivasi belajar dengan tetap memperhatikan proses penentuan *grade* hasil belajar.
4. Penilaian akhir hasil belajar siswa, sebaiknya tetap memperhatikan bahwa penentuan *grade* diberikan untuk merepresentasikan hasil belajar secara individual bukan secara bersama. Oleh karena itu, semua siswa mendapatkan keputusan tentang *grade* hasil belajar masing-masing.

5. Buatlah estimasi bagaimana presentase siswa menerima nilai dalam setiap tingkatan *grade* yang ada.
6. Buatlah kriteria atau bobot guna membuat pertimbangan kategori tentang penampilan siswa.
7. Bakukanlah cara *scoring* atau penilaian dan tentukanlah jumlah total yang dicapai untuk setiap siswa.
8. Sertakan catatan yang perlu pada setiap siswa, termasuk catatan seperti berapa kali absen karena sakit, tidak ada berita, atau seizing guru kelas.
9. Hindari tindakan penentuan *grade* yang menghasilkan interpretasi yang keliru atau meragukan. Misalnya, rerata siswa adalah pada batas 6,6; padahal ujian akhirnya mencapai nilai 9. Untuk itu, guru bisa mengecek kembali daftar penilaian dalam setiap ulangan, dan keaktifannya dalam mengerjakan tugas atau pekerjaan rumah dari guru. Pilih tingkatan *grade* hasil belajar siswa yang lebih atas, sehingga memperoleh nilai 7 atau memperoleh *grade* B, daripada ia mendapatkan nilai 6,6 atau C+. Berikut ini diberikan contoh format penentuan *grade* para siswa.

Tabel 4.2 Model Penilaian Akhir

Pencapaian/ Penampilan	Persentase Butir	Nilai		total
		Bobot	Skor	
Ujian MID				
Ujian Akhir				
Presentasi				
Eksperimen di Laboratorium				
Tugas Pekerjaan Rumah				
Partisipasi dalam Diskusi				
Total				

7.5 MENYELENGARAKAN EVALUASI HASIL BELAJAR DI KELAS

Guru atau dosen sering ditanya “Bagaimana menyelenggarakan evaluasi hasil belajar di kelas yang baik?”. Pertanyaan tersebut memang tidak terlalu sukar untuk dijawab, terutama oleh para guru senior yang telah memiliki pengalaman luas, tetapi menjadi hal yang mungkin menyulitkan jika pertanyaan tersebut ditujukan kepada guru yang masih junior. Hal ini disebabkan beberapa alasan. *Pertama*, guru junior pada umumnya masih

didampingi oleh guru pembimbing atau guru yang lebih senior. Jadi, mereka hanya sebagai pelaksana, ketika ada anjuran untuk membuat soal sesuai dengan jadwal yang telah ditentukan, maka ia membuat soal tersebut. *Kedua*, sebagian guru atau dosen hanya peduli dengan evaluasi yang berkaitan dengan tugas mengajar mereka dan tidak mau tahu kapan guru lain melakukan evaluasi. Kecuali guru tersebut juga menjabat posisi lain, seperti sekretaris jurusan atau prodi yang harus membuat jadwal evaluasi secara keseluruhan, maka ia akan peduli dan memerhatikan semua kepentingan guru-guru yang satu jurusan atau prodi yang juga ikut menggunakan ruang evaluasi.

Bagi guru yang mungkin memperoleh tugas untuk menyelenggarakan evaluasi pembelajaran, pada umumnya mereka melakukan beberapa tahapan sebagai berikut: a) persiapan, b) penyusunan instrument evaluasi, c) pelaksanaan evaluasi, d) pengolahan hasil penilaian, dan e) pemberitahuan hasil akhir nilai atau *grade*. Masing-masing akan dibahas pada paragraf berikut ini.

1. Langkah Persiapan

Pada langkah ini termasuk di dalamnya adalah kegiatan perencanaan; memberikan informasi kepada seluruh guru yang bersangkutan; pemberian jadwal yang berisikan mengajar apa guru tersebut, kapan, dan juga ruangan tempat ujian.

2. Langkah Penyusunan Instrumen

Pada tahap penyusunan instrument tes evaluasi, guru dianjurkan untuk membuat soal dan secepatnya diserahkan kepada panitia penyelenggara evaluasi, misalnya tanggal penyerahan tiga hari sebelum waktu ujian yang diselenggarakan.

3. Pelaksanaan Evaluasi

Pelaksanaan evaluasi merupakan proses di mana seorang guru mengevaluasi kepada para siswanya. Waktu pelaksanaan ini perlu diatur agar tidak bersamaan dengan guru lain atau siswa tidak sedang melakukan evaluasi materi pembelajaran dari guru lain. Pada pelaksanaan evaluasi ini, di samping guru yang bersangkutan, kadang-kadang untuk sekolah yang mempunyai guru banyak, panitia penyelenggara mengundang guru lain untuk membantu mengawasi pelaksanaan evaluasi.

4. Pengolahan Hasil Evaluasi

Pada tahap ini guru mengumpulkan hasil jawaban dari siswa untuk kemudian dikoreksi dan mendapat nilai akhir. Pada umumnya, dalam tahap pengolahan ini guru diberi peluang antara satu sampai dua minggu untuk dapat memberikan hasil akhir.

5. Pemberitahuan Hasil Evaluasi

Pemberitahuan hasil evaluasi merupakan tahapan akhir, di mana siswa dapat mengetahui hasil belajar mereka. Nilai akhir tersebut biasanya diadministrasi untuk keperluan kenaikan kelas atau kenaikan jenjang para siswa.

-oo0oo-

APPROVED



Bab 8

PENILAIAN PERFORMA (*PERFORMANCE ASSESSMENT*)

Tes objektif dan tes esai lebih banyak mengukur penguasaan hasil belajar dalam domain/kawasan kognitif dan afektif; dan kadang-kadang kurang terkait dengan keterampilan seseorang dalam melakukan sesuatu. Kendala yang demikian membawa dampak bahwa penguasaan hasil belajar kurang didukung oleh kemampuan dan keterampilan melaksanakan sesuatu yang sudah dipelajari. Mereka lebih mampu berteori, tetapi kurang terampil melakukan teori tersebut.

Sehubungan dengan itu, di dalam lembaga pendidikan perlu dikembangkan lebih banyak penilaian performa baik berupa tes performa (performance test) untuk bidang studi tertentu, maupun teknik-teknik inventori lainnya, sehingga guru atau pendidik dapat mengungkapkan kemampuan, keterampilan dan sikap yang dimiliki peserta didik secara menyeluruh.

8.1 PENGERTIAN PENILAIAN PERFORMA (PERFORMA)

Penilaian performa atau performa merupakan penilaian yang dilakukan guru dengan mengamati kegiatan peserta didik dalam melakukan suatu tugas. Oleh karena itu, penilaian performa lebih diarahkan untuk menilai ketercapaian kompetensi yang menuntut peserta didik melakukan suatu tugas.

Ada tiga karakteristik utama penilaian performa sebagai berikut:

1. Peserta didik mendesain penilaian sendiri; lebih dari memilih atau merespons.
2. Format penilaian mengikuti format guru, yang mengamati tingkah laku peserta didik tentang kemampuan merefleksikan penguasaannya dalam dunia nyata.
3. *Scoring* diarahkan pada pola berpikir dan belajar peserta didik yang tampak.

Oleh karena itu, penilaian performa dapat digunakan dalam berbagai mata pelajaran, seperti: Bahasa Inggris, Bahasa Indonesia, olah raga, seni rupa, pekerjaan tangan, dan drama.

8.2 KARAKTERISTIK PENILAIAN PERFORMA

Selain karakteristik yang telah disebutkan di atas, ada karakteristik lain yang lebih spesifik. Karakteristik penilaian performa tidak begitu berbeda dengan ciri-ciri penilaian autentik, yakni:

1. Menyusun Respon Sendiri
Penilaian performa tidak seperti penilaian hasil belajar dengan pendekatan tradisional. Penilaian performa mempunyai ciri yang sama dengan penilaian autentik. Pada penilaian performa (*performance assessment*), peserta didik bukan memilih dari jawaban yang telah disediakan; melainkan menyusun respons sendiri. Bentuk respons tersebut berupa tindakan melakukan tugas-tugas yang diberikan. Dengan melakukan observasi, guru dapat menilai kemampuan siswa melakukan tugas yang diberikan kepadanya dalam situasi nyata. Dalam menilai, pendidik dapat menggunakan berbagai cara sesuai dengan tugas yang diberikan.
2. Berpikir pada Tingkat Lebih Tinggi
Tidak seperti dalam penilaian dengan pendekatan tradisional, di mana guru lebih sering menggunakan tes objektif. Hal itu disebabkan karena keterbatasan guru dalam menyusun butir soal untuk kawasan kognitif tinggi dan/atau karena tujuan pembelajaran adalah “peserta didik mengetahui.” Namun di dalam penilaian performa penekanannya terletak pada kemampuan melakukan suatu tugas dalam kehidupan nyata, soalnya pun dirancang tidak dalam bentuk pilihan ganda.

3. Keautentikan Tugas-tugas

Keautentikan Tugas-tugas merupakan ciri utama penilaian performa ini. Jangan memberi tugas yang tidak terdapat dalam kehidupan nyata kepada peserta didik. Tugas-tugas tersebut hendaknya merupakan pemantapan materi yang telah diberikan melalui kegiatan pembelajaran di dalam kelas. Selain itu, pemberian tugas yang bermakna dalam kehidupan merupakan kelanjutan, penguatan dan pemantapan penguasaan konsep beserta aplikasinya dalam kehidupan nyata di dalam kehidupan masyarakat.

Konsekuensinya, tugas-tugas yang diberikan harus berhubungan erat dengan tugas nyata dalam masyarakat. Seandainya tujuan belajar adalah: "Peserta didik mampu membuat patung", maka setelah melakukan proses pembelajaran di dalam kelas, cara-cara membuat patung (konsep dan latihan di kelas), perlu dilanjutkan dengan pemberian tugas: "membuat patung manusia".

4. Proses dan Produk

Salah satu ciri utama dalam penilaian performa/kinerja ini adalah proses pembuatan dan hasil kerja selalu menjadi target penilaian. Oleh karena itu, penilaian harus dilakukan melalui: Bagaimana hasil kerja dibuat, kemudian baru menilai hasil kerjanya. Dengan cara bagaimana hasil kerja dibuat, kemudian baru menilai hasil kerjanya. Dengan cara demikian, gambaran penguasaan peserta didik tentang suatu materi, benar-benar demikian. Itulah gambaran penguasaan peserta didik yang sebenarnya.

5. Mengutamakan Kedalaman Bukan Keluasan

Fokus utama penilaian performa adalah kedalaman penguasaan peserta didik tentang suatu materi/keterampilan, bukan pada luasannya. Penilaian performa dapat dilakukan dengan menggunakan tes dan dapat pula dengan nontes.

8.3 TES PERFORMA

1. Pengertian Tes Performa

Tes performa (*performance test*) merupakan suatu jenis tes, di mana peserta didik diminta untuk melakukan sesuatu tugas, dengan menggunakan bahasa

sedikit mungkin; bahkan kalau memungkinkan dapat dihilangkan sama sekali, kecuali untuk menguji keterampilan bahasa, misalnya keterampilan berbicara. Tes ini pada prinsipnya berada antara "*paper and pencil test*" yang menguji kemampuan kognitif dan afektif dengan performa dalam situasi alamiah yang sebenarnya, di mana proses belajar diaplikasikan.

Hakikat tes performa dapat dikemukakan sebagai berikut:

Menurut Blythe C. Michael (Tuckman, 1975),

Performance test is a test usually involves some motor or manual response on the test taker's part, generally a manipulation of concrete equipment or materials as contrasted to paper and pencil test. The term "performance" is also used to denote a test that is actually a work sample; in this sense, it may include paper and pencil test, as for example a test in bookkeeping, in short hand, or in proof reading where no materials other than paper and pencil maybe required, and the test response is identical with the behavior about which informations is desired.

Lindquist (1951) menyatakan bahwa: *The term performance test has been used in connection with both the measurement of aptitude and the measurement of achievement.* Pada bagian lain ia menyatakan *performance test ... demonstrates his skill through any actual performance by constructing something. Performing a physical feat, or producing an artistic composition.*

Dengan demikian, bisa kita pahami dengan jelas bahwa tes performa bukan semata-mata tes performa fisik dalam situasi yang sebenarnya (dalam realita kehidupan), melainkan juga mencakup tes keterampilan mental seperti tes performa yang digunakan untuk mengukur *aptitude*, yakni kecerdasan khusus dan penguasaan peserta didik dalam belajar menyusun dan mengembangkan sesuatu, seperti komposisi lagu.

Hal tersebut bisa terjadi mengingat banyak cara yang dipakai para ahli di bidang pendidikan dan psikologi untuk mengungkapkan kemampuan umum (*general intellectual*) seseorang. Ada yang menggunakan bahasa dan ada pula yang tidak memakai bahasa, yakni bagi seseorang yang mengalami gangguan atau kesukaran dalam berbahasa. Maka kemampuan umumnya dapat diungkapkan melalui tes performa. Ini berarti, tes yang diberikan bersifat *non verbal*, seperti *knox cube test*. Selain itu, tes performa dapat juga digunakan untuk memperkirakan kemampuan seseorang apakah bisa berhasil dalam latihan tertentu yang membutuhkan keterampilan khusus,

misalnya dalam menentukan *neoromuscules coordination*, persepsi spesial dan lain-lain. Dalam pengukuran penguasaan belajar, pemanfaatan tes performa akan sangat menentukan dalam berbagai bidang studi, seperti dalam keterampilan musik, keterampilan berkomunikasi, keterampilan membuat pita rekaman, kursus-kursus PKK (*home economics*), dan pendidikan jasmani.

Dalam penilaian penguasaan belajar di sekolah, tes performa dapat di arahkan pada penilaian *product* dan penilaian proses/prosedur. Misalnya, dalam menilai keterampilan mengetik, penilaian harus dilakukan pada proses mengetik dan hasil karya produknya.

Dengan memperhatikan sistem yang dipakai, instruktur menilai proses pengetikan mulai dari penempatan jari pada *keyboard*, proses pengetikan dan hasil ketikan. Dalam ketiga fokus tersebut selalu diperhatikan ketepatan, kecepatan, dan ketelitian sehingga dapat menggambarkan kemampuan dan keterampilan yang sebenarnya.

2. Jenis-jenis Tes Performa

Sebagaimana tes objektif dan tes esai, tes performa sebenarnya telah lama berkembang tetapi belum banyak dimanfaatkan dalam proses pembelajaran menurut keadaan yang sebenarnya. Banyak jenis tes performa yang telah disusun para ahli, beberapa di antaranya, sebagai berikut:

- a. Paper and pencil performance,
- b. Recognition test ,
- c. Simulated performance,
- d. Work sample test.

a. Paper and Pencil Test Performance

Paper and pencil performance test merupakan tes performa yang menggunakan kertas dan pensil dalam konstruksi tes, kemudian peserta didik melakukan tugas berdasarkan petunjuk tersebut.

Misalnya:

- 1) Susunlah satu set soal untuk menilai kemampuan peserta didik.
Atau
- 2) Susunlah satu set rencana eksperimen pelaksanaan pengalaman lapangan.

Atau

- 3) Susunlah satu sketsa wajah duka seorang ibu yang kehilangan anak tunggalnya akibat gempa tektonik dan tsunami.

Dengan kata kerja perintah *susunlah*, berarti peserta didik melakukan sesuatu yang ditugaskan kepadanya dengan menggunakan kertas dan pensil. Tes performa seperti ini merupakan langkah awal ke arah performa yang sebenarnya. Umpamanya: Sebelum peserta didik dapat melakukan penggunaan laboratorium bahasa dengan baik, ia terlebih dahulu difasilitasi, dimotivasi serta diarahkan agar membaca cara penggunaan laboratorium bahasa dan menguasai langkah-langkah penggunaannya dengan benar.

b. *Recognition Test*

Jenis tes ini disusun dengan maksud untuk mengukur kemampuan seseorang mengenal hal-hal esensial dalam suatu unjuk kerja/penampilan atau dalam mengidentifikasi suatu objek.

Misalnya: dua orang peserta didik disuruh berlari 100 meter. Kemudian peserta didik yang diuji:

- 1) Disuruh menunjukkan di mana kekurangan kedua orang peserta didik tersebut.
- 2) Disuruh menunjukkan kesalahan yang dilakukan kedua peserta didik tersebut.

Contoh lain:

Kepada peserta didik disajikan penggunaan alat tiup trompet. Peserta didik yang ikut ujian disuruh mengidentifikasi kekurangan-kekurangan maupun kelebihan penggunaan alat tersebut.

Jenis tes performa *recognition test* ini mudah dipersiapkan dan disesuaikan dengan situasi yang sering berubah. Walaupun tes tipe ini mengukur aspek yang penting dari performa, tetapi tidak mengukur secara langsung keterampilan penguasaan dan prosedur-prosedurnya sekaligus, karena orang yang diuji tidak melakukan performa yang sebenarnya.

c. *Simulated Performance*

Kondisi lingkungan yang digunakan dalam tipe ini adalah miniatur dari keadaan yang sebenarnya. Oleh karena itu, tes jenis ini dapat juga dikatakan

sebagai *miniature test*.

Contohnya: Apabila pendidik ingin menguji kemampuan peserta didik SMK jurusan mesin dalam mengoperasikan dan menguji reaksinya sebagai sopir, maka pendidik dapat menguji keterampilan itu dengan menciptakan situasi mini dari keadaan yang sebenarnya. Perlengkapan-perlengkapan diatur sedemikian rupa, kemudian peserta didik yang diuji ditempatkan dalam kondisi tersebut untuk melakukan kegiatan sesuai petunjuk.

Ilustrasi lain: Untuk memperkirakan kemampuan seseorang dalam praktik mengajar yang sebenarnya, maka sebelum diturunkan ke lapangan (untuk praktik) mereka mengikuti kegiatan praktik mengajar dengan mengambil sejumlah murid sekolah, kemudian mencoba praktik mengajar (*micro teaching*). Sehingga dapat menguji kemampuan seluruh peserta praktik. Dalam hal ini, teman sebaya bisa juga dijadikan sebagai murid (*peer-teaching*).

d. Work Sampel Test

Peserta didik, dalam tes tipe ini, dihadapkan pada situasi yang sebenarnya. Mereka diberi sampel kerja yang harus dilakukannya dalam penampilan yang aktual. Apabila sampel kerja diambil secara tepat, diadministrasikan dan diskor secara benar, maka keterampilan dan tingkah laku yang diketahui tidak akan jauh menyimpang.

Misalnya:

- a. Keterampilan membuat wayang golek.
Yang akan dinilai adalah kerja peserta didik. Dalam pelaksanaannya, sampel kerja yang diambil hendaklah representatif/mewakili wayang golek yang dibuat. Pendidik harus mampu menyeleksi secara benar dan menentukan sampel kerja yang akan dinilai, sehingga betul-betul mewakili berbagai jenis wayang golek yang dibuat oleh peserta didik. Namun kalau yang akan dinilai adalah keterampilan dalam membuatnya dalam setiap langkah, dan kesesuaiannya dengan pola dan langkah-langkah yang sesungguhnya, maka penilaiannya dilakukan dengan mengikuti proses penyusunannya.
- b. Keterampilan untuk menyatakan pendapat dalam bahasa verbal (bahasa Inggris).

Untuk mengetahui dan menilai keterampilan itu, pendidik harus membentuk dan mengorganisasi suatu kelompok diskusi, kemudian menetapkan topik untuk dibicarakan. Selanjutnya peserta didik berdiskusi dan mencoba menyatakan pendapatnya dalam bahasa Inggris.

3. Manfaat Tes Performa

Manfaat tes performa sama seperti manfaat tes objektif dan tes esai dalam proses pembelajaran. Tes performa dimaksudkan untuk mengukur perbedaan individual dalam kecakapan, keahlian atau kepandaian mengenai performa tertentu. Pada umumnya, tes performa ini dimanfaatkan:

- a. Untuk meramalkan kesuksesan dalam melaksanakan keterampilan tertentu.
- b. Untuk mendiagnosis kesulitan-kesulitan dalam penampilan.
- c. Sebagai alat bantu dalam pendidikan peserta didik.
- d. Sebagai patokan atau kriteria untuk menilai kemajuan seseorang, dalam penelitian atau dalam promosi dalam latihan.
- e. Untuk menilai keterampilan dan sekaligus kualitas produk yang dihasilkan.

Gambaran keterampilan peserta didik tentang suatu objek akan lebih baik, apabila pendidik melakukan penilaian produk atau penilaian proses dan produk.

4. Langkah-langkah Penyusunan Tes Performa

Bentuk tes yang sering digunakan berupa *non verbal test*, *rating scale*, *check list*, *paired comparison*, *method of equal appearances intervals*.

Adapun langkah-langkah dalam menyusun tes performa adalah sebagai berikut:

- a. Susun secara cermat keterampilan dan kemampuan khusus yang terlibat dalam kegiatan tertentu yang akan diukur.
- b. Pilih keterampilan yang akan diuji di antara keterampilan yang telah ditetapkan. Keterampilan yang dipilih harus mewakili yang sesungguhnya.
- c. Menetapkan karakteristik aspek yang akan diukur; apakah proses, produk atau keduanya, kemudian kembangkan atau pilih bentuk tes yang tepat.

- d. Selidiki keterbatasan yang mungkin ada:
 - 1) Apakah jumlah waktunya cukup.
 - 2) Apakah *equipment* lengkap dan sesuai dengan jumlah peserta didik.
 - 3) Apakah SDM/personel cukup.
- e. Formulasikan dan kembangkan rencana pelaksanaan yang bersifat sensitif.
- f. Uji coba dan revisi tes.
- g. Persiapkan petunjuk pengadministrasian.

Dalam hal ini keterbatasan tersebut harus dianalisis sehingga mencakup persiapan pengadministrasian, pelaksanaan tes dan cara menskor test.

5. Macam-macam Model Tes Performa

Proses pemilihan model yang akan digunakan oleh pendidik di sekolah, berkaitan erat dengan tujuan penilaian dan kadar validitas serta reabilitas performa yang diinginkan: Apakah tujuannya memahami proses/tindakan, atau menggunakan. Atau mungkin juga memahami dan mampu menggunakan.

Beberapa contoh tes yang dapat digunakan disajikan berikut ini :

- a. *Vigilance test* yang dikembangkan oleh De Silva dan Channel.
Tes ini dimaksudkan untuk mengukur operasi dan reaksi supir mobil dalam menggunakan rem, stir, gas atau alat yang lain.
- b. *Miniatur punch press test*

8.4 INVENTORI PERFORMA

Banyak teknik nontes dapat digunakan untuk menilai performa peserta didik dalam melakukan suatu kegiatan, antara lain: Portofolio, Rubrik, Daftar cek (*check list*), skala bertingkat (*rating scale*), eksibisi (*exhibition*), demonstrasi (*demonstration*), investigasi (*investigation*).



Bab 9

ITEM ANALYSIS

**UNTUK BAB 9 MASIH DLM BAHASA INGGRIS, MOHON
DITERJEMAHKAN KE DLM BAHASA INDONESIA ATAU
DIHILANGKAN SAJA**

APPROVAL



DAFTAR PUSTAKA

- Anderson, L.W., & Krathwohl, D.R. (Eds.). 2001. *A Taxonomy of Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- ASC. 1989-2006. *User's Manual for the ITEMAN™ Conventional Item Analysis Program*. St. Paul, Minnesota: Assessment Systems Corporation
- Bott, P.A. 1996. *Testing and assessment in occupational and technical education*. Boston: Allyn and Bacon.
- Cross, A. 2013. *Home economic evaluation*. Columbus, Ohio: A Bell & Howel Company.
- Freeman, Frank, S. 2012. *Theory and practice of psychological testing*. New York: Holt, Rine Hart, and Winston.
- Gronlund, N.E. & Waugh, C.K. 2009. *Assessment of student assessment*. Upper Saddle River, New Jersey: Pearson Education.
- Gronlund, N.E. 2010. *Measurement and evaluation in teaching*. New York: Macmillan Publishing Company.

- Jhonson, D.W. & Jhonson, R.T. 2002. *Meaningful assessment: A managable and cooperative process*. Boston: Allyn and Bacon.
- Johnson, D.W. 2012. *Meaningful assessment: A managable and coopertive process*. Boston: Allyn and Bacon Publisher.
- Kusuma, M. 2016. *Evaluasi Pendidikan: Pengantar, Kompetensi dan Implementasi*. Yogyakarta: Parama Ilmu.
- Lindquist, E.F. 1951. *Educational measurement*. Washington, D.C.: George Banta Publishing Company.
- Linn, R.L. & Gronlund, N.E. 1995. *Measurement and assessment in teaching. Seventh Edition*. Englewood Cliffs, NJ: Prentice Hall
- Linn, R.L. & Gronlund, N.E. 2000. *Measurement and assessment in teaching*. Upper Saddle River, New Jersey: Merrill/Prentice-Hall.
- Linn, R.L. & Gronlund, N.E. 2013. *Measurement and assessment in teaching*. Illinois: Pearson.
- Matlock-Hetzel, S. 1997. *Basic Concepts in Item and Test Analysis*. Texas: A&M University
- McMillan, J.H. 2004. *Classroom assessment: principles and practice for effective instruction*. Boston: Allyn and Bacon.
- Miriam, K.L. 1996. *An introduction to psychological tests and scales*. London: UCL Press.
- Mueller, J. 2008. *Authentic assessment of core concepts*. <http://Jonathan.mueller.faculty.noctrl.edu>. Diakses 10 Oktober 2016.
- Ngadimun, H. 2004. *Analisis butir soal dengan komputer dan menafsirkannya*. Makalah disampaikan pada sosialisasi KBK bagi guru SMP Kabupaten Tanggamus di Pulau Panggung, tanggal 22-24 Juli 2004. Bandar Lampung: HEPI.
- Nitko, A.J. 2001. *Educational assessment of students*. Upper Saddle River, New Jersey: Merrill/Prentice-Hall.
- Oosterhoff, A.C. 1999. *Developing and using classroom assessments*. Upper Saddle River, New Jersey: Prentice-Hall.

- Power, T. 2012. *Methods of assessment*. <http://www.tedpower.co.uk/esl0706.html> (Accessed: 09 May 2013)
- Roid, G.H. & Haladyna. 2012. *A technology for test-item writing*. Orlando: Academic Press, Inc.
- Smith, J.K. Smith, L.F. & DeLisi, R. 2001. *Natural classroom assessment*. Thousand Oaks, C.A.: Corwin Press
- Stiggins, R.J. 2001. *Student-involved classroom assessment*. Upper Saddle River, New Jersey: Merrill/Prentice-Hall.
- Suparman, U. 2011. The implementation of Iteman to improve the quality of English test items as a foreign language. *Aksara*, Vol. XII, No. 1 pp. 85-96. Bandar Lampung, April 2011.
- Suparman, U. 2013. *How to improve the quality of English test items as an FI using iteman: (assessment and evaluation in language teaching)*. Makalah disajikan dalam Seminar Nasional ke 6 di Universitas Pamulang, Rabu, 20 November, 2013.
- Suparman, U. 2013. *Improving the quality of English test items as a foreign language by means of Iteman: (an assessment analysis)*. A paper presented on the First International Teacher Education Conference (ITEC) organized by Lampung Univeristy in Bukit Randu Hotel, June 30th - July 3rd, 2013.
- Tuchman, B.W. 1975. *Measuring educational outcomes: Fundamental of testing*. New York: Harcourt Brace Jovanovich.
- Wiggins, G.P. 1998. *Educative assessment: Designing assessments to inform and improve student performance*. San Fransisco, C.A.: Jossey-Bass.
- Yusuf, A.M. 2015. *Asesmen dan evaluasi pendidikan: Pilar penyedia informasi dan kegiatan pengendalian mutu pendidikan*. Rawamangun, Jakarta: Prenadamedia Group.

