

A. Supratiknya



PENILAIAN HASIL BELAJAR DENGAN TEKNIK NONTES



iiiiiiiiiiiiiiii2iiiiiiiiiii3iiiiiiiiiii4iiiiiiiiiii5iiiiiiiiiii6iiiiiiiiiii7iiiiiiiiiii8iiiiiiiiiii9iiii

A DAY IN A LIFE - MIKROASESMEN



iiiiiiiiiiiiiiii2iiiiiiiiiii3iiiiiiiiiii4iiiiiiiiiii5iiiiiiiiiii6iiiiiiiiiii7iiiiiiiiiii8iiiiiiiiiii9iiii

A DAY IN A LIFE - MIKROASESMEN



Penilaian Hasil Belajar dengan Teknik Nontes

A. Supratiknya



Penerbit
Universitas Sanata Dharma

Penilaian Hasil Belajar dengan Teknik Nontes

Copyright © 2012

PENERBIT UNIVERSITAS SANATA DHARMA

Jl. STM Pembangunan (Mrican) 1A, Gejayan Yogyakarta 55281

Telp. (0274) 513301, 515253 Ext.1527/1513

Fax (0274) 562383

e-mail: publisher@usd.ac.id

Diterbitkan oleh:

Penerbit Universitas Sanata Dharma
Jl. STM Pembangunan (Mrican) 1A,
Gejayan Yogyakarta 55281
Telp. (0274) 513301, 515253;
Ext.1527/1513
Fax (0274) 562383
e-mail: publisher@usd.ac.id



Penerbit USD

Universitas Sanata Dharma berlambangkan daun teratai coklat bersudut lima dengan sebuah obor hitam yang menyala merah, sebuah buku terbuka dengan tulisan "*Ad Maiorem Dei Gloriam*" dan tulisan "Universitas Sanata Dharma Yogyakarta" berwarna hitam di dalamnya. Adapun artinya sebagai berikut. Teratai: kemuliaan dan sudut lima: Pancasila; Obor: hidup dengan semangat yang menyala-nyala; Buku yang terbuka: ilmu pengetahuan yang selalu berkembang; Teratai warna coklat: sikap dewasa yang matang; "*Ad Maiorem Dei Gloriam*": demi kemuliaan Allah yang lebih besar.

A. Supratiknya

Desain Sampul:

Pius Sigit Kuncara

Tata Letak:

Thoms

Cetakan Pertama

xi, 170 hlm.; 148 x 210 mm.

ISBN: 978-602-9187-19-9

EAN: 9-786029-187199

Hak Cipta Dilindungi Undang-Undang.

Dilarang memperbanyak karya tulis ini dalam bentuk dan dengan cara apa pun, termasuk fotokopi, tanpa izin tertulis dari penerbit.

Sekapur Sirih

Asesmen atau penilaian merupakan salah satu kegiatan utama dalam dunia pendidikan sekolah, mulai jenjang pendidikan dasar sampai pendidikan tinggi. Secara umum dan dalam konteks pendidikan sekolah, asesmen atau penilaian adalah kegiatan pengumpulan informasi tentang murid atau tentang program kegiatan terkait dimensi atau atribut tertentu, seperti hasil belajar murid dalam pengajaran suatu bidang studi atau keefektivan pembelajaran suatu bidang studi yang diselenggarakan oleh guru, untuk selanjutnya dibandingkan dengan kriteria tertentu agar bisa disimpulkan kualitas dimensi atau atribut terkait murid atau program kegiatan yang bersangkutan. Buku ini berfokus pada yang pertama, yaitu penilaian hasil belajar murid.

Ada empat dimensi penting dalam penyelenggaraan asesmen atau penilaian terhadap hasil belajar murid di sekolah, meliputi *purpose* yaitu tujuan dilakukannya asesmen, *mode* atau cara asesmen dilakukan, *content* atau jenis kemampuan yang dijadikan sasaran asesmen, dan *organization* atau tata organisasi penyelenggaraannya. Keempat dimensi asesmen tersebut akan berlainan tergantung antara

lain dari suasana atau semangat zaman yang melingkunginya. Sejak berkembang sekitar awal abad ke-20 hingga memasuki awal abad ke-21 ini, setidaknya telah terjadi sebuah pergeseran yang cukup radikal dalam praktek asesmen terkait keempat dimensinya tersebut (Broadfoot, 2011).

Abad ke-20 lazim dipandang sebagai zaman modern yang ditandai oleh semangat mengutamakan hak dan tanggung jawab pribadi masing-masing orang, mengagungkan rasionalitas atau kemampuan berpikir secara rasional, serta pendewaan ilmu dan segala sesuatu yang ilmiah sebagai kunci tunggal ke arah kemajuan pribadi dan masyarakat. Dalam suasana zaman seperti itu, asesmen lahir dan berkembang sebagai sarana untuk mewujudkan *tujuan* menyediakan ajang persaingan yang adil untuk memperoleh calon-calon terbaik untuk menduduki berbagai jabatan publik. *Mode* atau cara asesmen yang diunggulkan adalah teknik interogasi baik lisan maupun khususnya tertulis melalui *paper and pencil tests* alias tes tulis. Cara tersebut dipandang efektif, sebab *jenis kemampuan* yang hendak diungkap dan dijadikan barometer tunggal keunggulan adalah kemampuan kognitif berupa penguasaan atas materi berbagai disiplin ilmu. Mengingat peran penting asesmen dalam rangka menemukan calon-calon terbaik dan paling sesuai untuk berbagai jabatan dengan kriteria tunggal tersebut, maka *tata organisasi* penyelenggaraannya pun lazimnya dikelola oleh sistem birokrasi resmi baik pemerintah maupun swasta seperti badan ujian nasional, panitia nasional seleksi masuk perguruan tinggi, dan sebagainya, yang memiliki otoritas hampir mutlak. Alhasil, asesmen seolah-olah berkembang menjadi sebuah pranata mandiri yang terpisah bahkan terasing dari kegiatan belajar (dan mengajar) yang menjadi prasyarat atau bahkan dasar keberadaannya.

Memasuki abad ke-21 manusia dibuat sadar tentang kesesatan pikir yang berlangsung dalam zaman modern. Selain menyadari

kemustahilan ilmu pengetahuan memberikan solusi tunggal yang ampuh terhadap aneka problem yang kita hadapi, orang juga dibuat sadar tentang pentingnya keanekaragaman dalam berbagai bidang. Penguasaan bentuk-bentuk pengetahuan ilmiah tradisional sebagaimana lazim diukur dengan tes tradisional tidak lagi memadai digunakan sebagai ukuran tunggal keunggulan seseorang. Aneka jenis kemampuan tradisional seperti yang diunggulkan di zaman modern tentu saja tetap penting, namun kehidupan di zaman baru pascamodern juga menuntut aneka bentuk kemampuan baru seperti kreativitas, kemampuan menganalisis, kemampuan bekerja dalam tim, kemampuan memecahkan masalah, di samping kemampuan membuat perencanaan dan mengelola aktivitas belajarnya sendiri.

Menanggapi tuntutan baru semacam itu, kini asesmen atau penilaian hasil belajar dalam arti dan konteks yang luas lebih ditekankan pada tujuan *sertifikasi*, yaitu memberikan *feedback* baik kepada yang bersangkutan maupun kepada pihak lain terkait taraf penguasaan berbagai kemampuan pada aneka konteks kehidupan dan aneka tingkat perkembangan. Sesuai tujuannya untuk memberikan apa yang oleh Broadfoot (2009) disebut *accreditation of real-life practice* alias pengakuan atas kemampuan melakukan praktek dalam kehidupan nyata, maka modus atau cara asesmen yang ditempuh pun tidak lagi hanya mengandalkan teknik tes tulis tradisional melainkan lebih mengakomodasi teknik-teknik baik tes maupun lebih-lebih nontes yang mampu mengungkap beraneka ragam kemampuan-ketrampilan yang dituntut oleh aneka konteks kehidupan nyata secara lebih otentik. Maka tata organisasi penyelenggaraannya pun tidak lagi mengandalkan usaha formal yang dilakukan oleh sebuah lembaga super, melainkan lebih mengutamakan apa yang oleh Broadfoot (2009) disebut *on-demand, personalized 'micro-assessments'* alias asesmen kecil-kecilan yang lebih dimaksudkan untuk keperluan pribadi dalam

berbagai situasi yang lebih spesifik. Hasil-hasil asesmen yang kendati cenderung informal namun tetap serius dan diperoleh di berbagai konteks belajar semacam ini, secara kumulatif akan menjadi sejenis *personal portfolio*, yaitu rekaman prestasi yang menorehkan narasi pribadi tentang pengalaman belajar seseorang sepanjang hayatnya (Broadfoot, 2009). Selain itu, asesmen menjadi bagian yang tak terpisahkan dari aktivitas belajar-mengajar itu sendiri.

Teks kecil berjudul *Penilaian Hasil Belajar dengan Teknik Nontes* ini ditulis dan dimaksudkan untuk menebarkan spirit dan paradigma baru tentang penilaian hasil belajar sebagaimana diuraikan di atas, khususnya di lingkungan Sekolah Dasar. Teks ini aslinya dimaksudkan sebagai bahan ajar mata kuliah *Evaluasi Pembelajaran 2* dalam kurikulum Program Studi S1 Pendidikan Guru Sekolah Dasar, Fakultas Keguruan dan Ilmu Pendidikan, Universitas Sanata Dharma, yang berfokus pada teknik penilaian nontes. Saat pertama kali menerima tugas mengampu mata kuliah tersebut dalam semester gasal 2011/2012, penulis mengalami kesulitan mempersiapkan materinya. Sejumlah teks berbahasa Indonesia yang tersedia kurang tegas dalam arti kurang sistematis bahkan cenderung rancu di sana-sini dalam menjelaskan dan membedakan teknik nontes dari teknik tes yang sudah lebih dulu dikenal secara luas, maupun dalam menjelaskan persamaan dan perbedaan di antara aneka teknik nontes itu sendiri. Penulis merasa beruntung menemukan buku yang tidak terlalu baru karangan Chatterji, M. (2003) berjudul *“Designing and using tools for educational assessment”* (Boston: Allyn and Bacon). Menurut hemat penulis, Chatterji cukup berhasil menguraikan makna, tujuan, dan penggolongan aneka teknik nontes dalam spirit baru seperti sudah disinggung di muka.

Pada kesempatan ini penulis ingin mengucapkan terima kasih kepada Ketua Program Studi S1 PGSD, FKIP, Universitas Sanata

Dharma saat itu yang memberikan kepercayaan kepada penulis mengampu mata kuliah *Evaluasi Pembelajaran 2* pada salah satu kelas. Penulis juga ingin mengucapkan terima kasih kepada Saudara Agnes Wiwik Wigati, S.Pd., M.Pd., mitra *team teaching* yang bersedia membaca *draft* teks ini dan memberikan masukan perbaikan yang sangat berguna. Akhirnya, penulis juga ingin mengucapkan terima kasih kepada Penerbit Universitas Sanata Dharma yang bersedia menerbitkan teks ini. Sudah barang tentu, tanggung jawab atas semua kekurangan-kesalahan yang mungkin masih terdapat dalam teks ini sepenuhnya berada di pundak penulis.

Yogyakarta, Januari 2012

A. Supratiknya

Daftar Isi

Sekapur sirih	iii
Daftar isi	viii
Daftar Tabel	ix
Bab 1. Pendahuluan	1
Bab 2. Hasil Belajar sebagai Objek Penilaian	5
Bab 3. Pengembangan Alat Penilaian Hasil Belajar: Teknik Tes	25
Bab 4. Pengembangan Alat Penilaian Hasil Belajar: Teknik Nontes	39
Bab 5. Penyusunan Alat Penilaian Nontes	65
Bab 6. Penyusunan Rubrik Penskoran	99
Bab 7. Cara Mengolah Hasil Penilaian Nontes	111
Bab 8. Penentuan dan Pelaporan Nilai Akhir	147
Daftar Pustaka	167
Tentang Penulis	169

Daftar Tabel

Tabel 2.2: Taksonomi Bloom Ranah Kognitif Hasil Revisi Anderson dan Krathwohl	10
Tabel 2.3: Jenis Kemampuan Menurut Dimensi Pengetahuan	11
Tabel 2.4: Taksonomi Tujuan Pengajaran Ranah Afektif	13
Tabel 2.5: Taksonomi Tujuan Pengajaran Ranah Psikomotor	15
Tabel 2.6: Taksonomi Tujuan Pengajaran Ranah Psikomotor ala Dave	18
Tabel 2.7: Taksonomi Tujuan Pengajaran Ranah Psikomotor ala Harrow	19
Tabel 2.6: Taksonomi Fungsional Ranah Pengetahuan, Ketrampilan, dan Perilaku	20
Tabel 4.1: Contoh Daftar Cek: Tingkah Laku di Kelas.	44
Tabel 4.2: Contoh Skala Penilaian Kualitas.	45
Tabel 5.1: Taksonomi Bloom Ranah Kognitif Hasil Revisi Anderson dan Krathwohl	71
Tabel 5.2: Kisi-kisi Hasil Belajar Ranah Pengetahuan sebagai Tujuan Pengajaran Mata Pelajaran	72
Tabel 5.3: Taksonomi Tujuan Pengajaran Ranah Afektif	73

Tabel 5.4: Kisi-kisi Hasil Belajar Ranah Perasaan sebagai Tujuan Pengajaran PKn.	74
Tabel 5.5: Taksonomi Tujuan Pengajaran Ranah Psikomotor	76
Tabel 5.6: Kisi-kisi Hasil Belajar Ranah Ketrampilan Spesifik sebagai Tujuan Pengajaran.	78
Tabel 5.7: Standar Kompetensi Guru Kelas SD/MI.	80
Tabel 6.1: Contoh Penerapan Skala Dua Titik pada Penskoran Analitik.	103
Tabel 6.2: Contoh Penerapan Skala Lebih dari Dua Titik pada Penskoran Analitik.	103
Tabel 6.3: Contoh Penerapan Rubrik Semi Holistik.	105
Tabel 6.4: Contoh Penerapan Rubrik Holistik Murni.	106
Tabel 6.5: Contoh Penyusunan Alat Penilaian Nontes Berbasis Perilaku.	108
Tabel 7.1: Format Pengodean dan Analisis Data Hasil Wawancara atau Observasi.	114
Tabel 7.2: Aneka Taraf Pengukuran Berdasarkan Ciri Bilangan.	120
Tabel 7.3: Aneka Taraf Pengukuran dan Penerapan Teknik Statistik yang Sesuai.	124
Tabel 7.4: Prinsip Pembulatan Bilangan Desimal.	127
Tabel 7.5: Simbol Sejumlah Parameter dan Statistik Deskriptif.	133
Tabel 7.6: Skor Hasil Tes Matematika ($N = 40$).	135
Tabel 7.7: Distribusi Frekuensi Bergolong Skor Hasil Tes Matematika.	135
Tabel 7.8: Distribusi Persentil Berbentuk Empat Persegi Panjang.	138
Tabel 7.9: Contoh Enam Skor dari Data di Tabel 7.6.	139
Tabel 7.10: Data Enam Skor Beserta Deviasi Masing-masing	

Daftar Tabel

dari Mean dan Kuadratnya.....	140
Tabel 7.11: Rekapitulasi Skor Ulangan Matematika Tiga Kelas IV SD Paralel.....	142
Tabel 8.1: Aneka Komponen Penilaian dalam Penentuan Nilai Akhir Mata Pelajaran.	151
Tabel 8.2: Pembobotan Aneka Komponen Penilaian dalam Rangka Penentuan Nilai Akhir.	154
Tabel 8.3: Norma Penilaian Skala-C Berdasarkan Nilai Jenjang Persentil.	157
Tabel 8.4: Norma Penilaian Berdasarkan Pembagian Daerah dalam Kurve Normal.....	158
Tabel 8.5: Contoh Rapor Model Narasi.....	164

Bab 1

Pendahuluan

Proses belajar-mengajar mengandung tiga unsur (Sudjana, 2010): (1) perumusan tujuan pengajaran, (2) perencanaan dan pelaksanaan kegiatan belajar-mengajar, dan (3) penilaian hasil belajar. Dalam perumusan tujuan pengajaran, guru merumuskan bentuk-bentuk perubahan tingkah laku yang diinginkan terjadi dalam diri murid. Dalam kurikulum berbasis kompetensi, bentuk-bentuk perubahan tingkah laku sebagai hasil kegiatan belajar-mengajar yang dimaksud dirumuskan dalam bentuk stándar kompetensi dan kompetensi dasar beserta indikator-indikator pencapaiannya. Perencanaan dan pelaksanaan kegiatan belajar-mengajar mencakup pemilihan strategi meliputi yang utama pemilihan materi serta metode dan alat kegiatan yang hendak ditempuh oleh guru dan murid dalam aktivitas pembelajaran di kelas dalam rangka menghasilkan bentuk-bentuk perubahan tingkah laku dalam diri murid yang diinginkan, serta pelaksanaannya (Suwandi, 2011). Sedangkan penilaian hasil belajar adalah kegiatan untuk melihat sejauh mana tujuan pengajaran telah dicapai atau dikuasai oleh murid dalam bentuk hasil belajar yang bisa mereka tunjukkan setelah menjalani kegiatan belajar-mengajar (Sudjana, 2010).

Penilaian Hasil Belajar

Inti kegiatan penilaian adalah menentukan nilai dari suatu objek dengan cara membandingkannya dengan kriteria tertentu. Dalam penilaian hasil belajar, guru menentukan nilai dari hasil-hasil belajar yang dicapai murid melalui kegiatan belajar-mengajar yang berlangsung di kelas, dengan cara membandingkannya dengan kriteria tertentu (Sudjana, 2010; Suwandi, 2011). Karena penilaian ini dilakukan oleh guru dalam rangka proses pembelajaran, maka ada yang menyebutnya *penilaian berbasis kelas* (KBK, 2004) atau *penilaian kelas* (KTSP, 2006).

Ada tiga istilah yang merujuk pada aktivitas-aktivitas utama dalam kegiatan penilaian kelas, yaitu: (1) asesmen, (2) pengukuran, dan (3) evaluasi. Aktivitas pertama dalam penilaian kelas adalah *asesmen*, yaitu "any systematic method of obtaining information to draw inferences about characteristics of people, objects, or programs" (Friedenberg, 1995; Chatterji, 2003). Artinya, asesmen adalah setiap metode atau prosedur sistematis untuk mendapatkan informasi sebagai dasar membuat inferensi atau kesimpulan tentang aneka karakteristik orang, objek, atau program. Dalam konteks penilaian kelas, asesmen adalah setiap metode atau prosedur mendapatkan data sebagai dasar untuk membuat kesimpulan tentang hasil belajar murid.

Prosedur yang dimaksud bisa berupa *pengukuran*, yaitu teknik pengumpulan data tentang orang yang melibatkan perubahan data menjadi bilangan atau skor. Salah satu teknik khas pengukuran yang identik dengan pengertian pengukuran adalah *tes*, yaitu jenis asesmen yang menggunakan sejumlah prosedur spesifik untuk memperoleh informasi dan mengonversikan atau mengubah informasi tersebut menjadi bilangan atau skor (Friedenberg, 1995). Teknik pengukuran lain adalah *skala*, meliputi *skala kepribadian* maupun *skala penilaian*.

Skala kepribadian merupakan sejenis tes yang bertujuan mengukur aspek-aspek kepribadian yang bersifat abstrak, seperti sikap dan minat, dan yang melibatkan pengonversian informasi yang diperoleh menjadi bilangan atau skor. Skala penilaian adalah teknik untuk menerjemahkan atau mengubah hasil pengamatan menjadi bilangan atau skor. Prosedur pengumpulan data juga bisa berupa teknik-teknik asesmen lain yang lebih bersifat subjektif dan khususnya yang tidak melibatkan pengukuran atau pengonversian data menjadi bilangan. Teknik-teknik ini lazim dikelompokkan sebagai *nontes* (Sudjana, 2010; Suwandi, 2011).

Data atau informasi yang diperoleh baik dengan teknik pengukuran maupun dengan teknik asesmen lain selanjutnya perlu diolah. Pada data yang diperoleh lewat pengukuran (tes dan skala), pengolahan lazimnya dilakukan secara kuantitatif dengan teknik statistik. Melalui teknik statistik deskriptif dapat diketahui setidaknya tiga hal penting, yaitu: (1) distribusi atau persebaran data atau skor; (2) gejala *tendensi sentral*, yaitu kecenderungan data atau skor memusat pada nilai tengah tertentu, serta (3) gejala *variabilitas*, yaitu kecenderungan data atau skor menyebar di sekitar nilai tengah tertentu.

Pada data yang diperoleh lewat teknik-teknik nontes dan lebih khusus lagi lewat teknik-teknik yang bersifat subjektif seperti wawancara, pengolahan data lazimnya dilakukan dengan menggunakan analisis kualitatif. Secara garis besar, analisis ini akan didahului dengan pengkodean atau klasifikasi jawaban atau gejala, sehingga diperoleh kategori-kategori jawaban atau gejala tertentu. Selanjutnya dapat diterapkan pengolahan kuantitatif sederhana meliputi antara lain penghitungan frekuensi atau persentase masing-masing kategori jawaban atau gejala, termasuk penyajiannya secara visual dalam grafik.

Aktivitas terakhir dalam rangkaian kegiatan penilaian kelas adalah *evaluasi*, yaitu "a process that comes after measurement is completed. It involves making a value judgment or interpretation of the resulting data in a decision making context" (Chatterji, 2003). Maksudnya, evaluasi merupakan proses sesudah pengumpulan data atau informasi baik dengan teknik pengukuran (tes dan skala) maupun dengan teknik asesmen lain selesai dilakukan, bahkan sesudah data atau informasi tersebut selesai diolah. Dalam evaluasi, hasil olahan data tersebut dinilai atau diinterpretasikan dalam rangka membuat keputusan atau kesimpulan. Seperti sudah disinggung, penilaian dalam arti evaluasi atau interpretasi ini lazimnya dilakukan dengan cara membandingkan skor atau data dengan sebuah kriteria. Tujuan akhir evaluasi lazimnya memang pembuatan keputusan atau kesimpulan bisa berupa: (a) penentuan posisi setiap murid dibandingkan kelompoknya terkait hasil belajar dalam mata pelajaran tertentu, atau (b) penentuan lulus-tidaknya setiap murid dalam mata pelajaran tertentu (Sudjana, 2010).

Sistematika Buku Ini

Fokus pembahasan buku ini adalah penilaian hasil belajar murid di kelas dengan teknik nontes. Dalam bab-bab berikut sejumlah konsep dan langkah penting sekitar penilaian hasil belajar murid di kelas dengan teknik nontes akan dibahas secara lebih mendalam, meliputi pokok-pokok: hasil belajar sebagai objek penilaian (Bab 2), pengembangan alat penilaian hasil belajar: teknik tes (Bab 3), pengembangan alat penilaian hasil belajar: teknik nontes (Bab 4), penyusunan alat penilaian nontes (Bab 5), penyusunan rubrik penskoran (Bab 6), cara mengolah hasil penilaian nontes (Bab 7), serta penentuan dan pelaporan nilai akhir (Bab 8).

Bab 2

Hasil Belajar sebagai Objek Penilaian

Hasil belajar yang menjadi objek penilaian kelas berupa kemampuan-kemampuan baru yang diperoleh murid sesudah mereka mengikuti proses belajar-mengajar tentang mata pelajaran tertentu. Pemerolehan kemampuan baru tersebut akan terwujud dalam perubahan tingkah laku tertentu, seperti dari tidak tahu menjadi tahu tentang seluk-beluk gejala tertentu, dari acuh-tak-acuh menjadi menyukai objek atau aktivitas tertentu, serta dari tidak bisa menjadi cakap melakukan ketrampilan tertentu seperti membaca tabel, membuat peta, mendayung, mengukir, dan sebagainya.

Tujuan pengajaran dan hasil belajar di sekolah lazimnya dibedakan mengikuti *taksonomi* tertentu. Taksonomi adalah klasifikasi atau penggolongan tentang objek atau gejala berdasarkan satu atau lebih prinsip tertentu. Dalam taksonomi, penggolongan objek atau gejala tersebut lazimnya sekaligus disertai pengurutannya secara hirarkis alias berjenjang, mulai dari yang paling sederhana dan yang ditempatkan pada jenjang atau urutan paling bawah sampai yang paling kompleks dan yang ditempatkan pada jenjang atau urutan puncak dalam hirarki. Sebagai suatu hirarki, diasumsikan bahwa

objek atau gejala yang lebih kompleks dan yang menempati jenjang lebih tinggi dengan sendirinya mencakup objek atau gejala yang lebih sederhana dan yang menempati jenjang di bawahnya. Sebaliknya, objek atau gejala yang menempati jenjang lebih rendah diasumsikan mendasari atau menjadi prasyarat bagi objek atau gejala yang berada di jenjang di atasnya. Dalam taksonomi tujuan pengajaran, objek atau gejala yang menjadi sasaran penggolongan sekaligus penjenjangan adalah jenis-jenis kemampuan yang diharapkan dikuasai oleh murid sebagai hasil kegiatan belajar-mengajar.

Salah satu taksonomi tujuan pengajaran yang dikenal dan dipakai secara luas adalah taksonomi tujuan pengajaran yang digagas oleh Benjamin S. Bloom dan kawan-kawan. Usaha mereka dimulai dengan mengembangkan taksonomi tujuan pengajaran ranah kognitif yang diterbitkan sebagai buku pertama dan yang disunting oleh Benjamin S. Bloom sendiri (1956). Sebenarnya sebelum Bloom, Ragsdale pada 1950 sudah menerbitkan klasifikasi untuk apa yang dia sebut "*motor types of activities learned by children*" atau aneka jenis aktivitas motor yang dipelajari oleh anak-anak. Namun demam taksonomi baru benar-benar berkobar sesudah tersulut oleh karya Bloom dan kawan-kawan. Penerbitan taksonomi-taksonomi yang lain pun segera bermunculan.

Yang pertama tentu saja adalah penerbitan taksonomi tujuan pengajaran ranah afektif sebagai buku kedua yang dikerjakan oleh Krathwohl, Bloom, dan Masia (1964). Pada 1967 Guilford menerbitkan taksonomi tentang kemampuan-kemampuan intelektual yang dia tuangkan dalam sejenis teori yang dia sebut *Structure of Intellect Model*. Antara 1964 dan 1971 Gagne dan Merrill menerbitkan taksonomi yang dipandang sebagai taksonomi pertama yang mengintegrasikan ranah afektif, psikomotor, dan kognitif sekaligus. Pada 1967 Sullivan dan Gerlach mengembangkan taksonomi bukan

berupa deskripsi proses-proses mental seperti yang dilakukan para pengarang lain, melainkan berupa apa yang mereka sebut *observable learner behaviors* yang bisa dipakai sebagai dasar untuk menyusun deskripsi dan analisis tugas. Simpson pada 1966 dan Dave pada 1969 mengembangkan taksonomi dalam ranah psikomotor. Pada 1970 Kibbler mengembangkan taksonomi untuk ranah psikomotor mengacu pada psikologi perkembangan anak mulai masa bayi sampai masa saat anak mengembangkan kemampuan berbahasa. Pada 1972 Harrow mengembangkan taksonomi ranah psikomotor. Pada 1975 De Block mengembangkan taksonomi yang dia namai *model of teaching objectives*. Boleh jadi yang terbaru, pada 2003 Chatterji mengembangkan sebuah taksonomi komprehensif yang dia sebut *a functional taxonomy of knowledge, skills, and behaviors* yang didasarkan pada teori belajar yang dikembangkannya dan yang dia sebut *Dimensions of Learning Model* atau teori tentang dimensi-dimensi belajar.

Ringkas kata, sejak dipopulerkan oleh Bloom dan kawan-kawan pada 1956 hingga kini taksonomi tujuan pengajaran yang meliputi *ranah kognitif* atau intelektual, *ranah afektif*, dan *ranah psikomotor* menjadi dasar penting dalam merumuskan tujuan pengajaran dan dalam penilaian hasil belajar. Dari sekalian taksonomi tersebut, taksonomi Bloom dalam ranah kognitif, taksonomi Krathwohl dalam ranah afektif, dan taksonomi Simpson dalam ranah psikomotor merupakan yang paling luas digunakan hingga kini. Kendati mungkin telah berulang kali dilakukan, namun marilah kita tinjau kembali dulu secara sedikit lebih cermat masing-masing taksonomi yang dimaksud.

Taksonomi Tujuan Pengajaran Ranah Kognitif

Dalam tahun 1990-an Anderson, mantan murid Bloom merevisi taksonomi Bloom dalam ranah kognitif (Tarlinton, 2003). Dua perubahan penting yang dilakukannya meliputi: (1) menghilangkan

kemampuan taraf 5 *Sintesis*, menggeser kemampuan taraf 6 *Evaluasi* menjadi kemampuan taraf 5, dan memasukkan jenis kemampuan baru pada taraf 6 yakni *Creating* atau *Menciptakan*; serta (2) mengubah nama keenam kategori, dari *kata benda* menjadi *kata kerja*. Hasil belajar dalam ranah kognitif tercermin atau terwujud dalam aneka kemampuan intelektual murid. Taksonomi Bloom dalam ranah kognitif hasil revisi Lorin Anderson yang dimaksud menjadi seperti disajikan pada Tabel 2.1. (Tarlinton, 2003).

**Tabel 2.1: Taksonomi Bloom Ranah Kognitif
Hasil Revisi Lorin Anderson**

<p>Menciptakan (<i>Creating</i>)</p>	<p>Menelurkan aneka gagasan, produk, atau cara melihat persoalan yang baru. Kata kunci: Merancang, mengonstruksi, merencanakan, memproduksi, membuat penemuan baru.</p>
<p>Mengevaluasi (<i>Evaluating</i>)</p>	<p>Memberikan membenaran terhadap sebuah keputusan atau rangkaian tindakan tertentu, membuat penilaian tentang nilai dari sebuah gagasan atau benda/hal. Kata kunci: Menguji (<i>checking</i>), menghipotesiskan, memberikan kritik, bereksperimentasi, memberikan penilaian (<i>judging</i>), memberikan apresiasi, membandingkan, menarik kesimpulan, mempertahankan atau memberikan argumentasi pro, mendeskripsikan, memilah, mengevaluasi, menjelaskan, menafsirkan, memberikan membenaran, mengaitkan, memberikan ringkasan.</p>
<p>Menganalisis (<i>Analysing</i>)</p>	<p>Mengurai informasi ke dalam bagian-bagian atau unsur-unsur untuk menjajaki atau menemukan pemahaman dan hubungan-hubungan, memilah materi atau konsep ke dalam bagian-bagian sehingga struktur organisasinya dapat dipahami, membedakan antara fakta dan pendapat. Kata kunci: Membandingkan, mengorganisasikan atau menata, mendekonstruksi, menginterogasi, menemukan, menemukan perbedaan, mengatribusikan atau menemukan penyebab atau kaitan antara dua hal, menganalisis, mengurai, membandingkan, menjelaskan dengan diagram, memberikan ilustrasi, memberikan kerangka garis besar, membedakan.</p>

Hasil Belajar sebagai Obyek Penelitian

<p>Menerapkan (<i>Applying</i>)</p>	<p>Menggunakan informasi dalam situasi lain dalam kehidupan sehari-hari (<i>familiar situation</i>). Menerapkan hasil belajar di kelas dalam situasi baru di luar kelas. Kata kunci: Menerapkan, melaksanakan (<i>carrying out</i>), menggunakan, mengeksekusi, menghitung, menyusun, mendemonstrasikan, memprediksikan.</p>
<p>Memahami (<i>Understanding</i>)</p>	<p>Menjelaskan aneka gagasan atau konsep. Memahami makna, terjemahan, perluasan atau penjabaran, dan penafsiran dari aneka perintah atau masalah. Merumuskan sebuah masalah dengan kata-kata sendiri. Kata kunci: Menginterpretasikan, memberikan contoh, memberikan ringkasan, memparafrasekan atau menjelaskan dengan kata-kata sendiri, mengklasifikasikan, menjelaskan, menginferensikan atau menyimpulkan, membandingkan, membuat estimasi atau perkiraan, membuat generalisasi, memberikan contoh, memprediksikan, menerjemahkan, membedakan, menguraikan lebih lanjut.</p>
<p>Mengingat (<i>Remembering</i>)</p>	<p>Mengingat kembali data atau informasi. Kata kunci: membuat daftar, mendeskripsikan, menabulasikan, menggunakan secara semestinya, mendefinisikan, mengidentifikasi, mengetahui, menemukan kesamaan, membuat kerangka (<i>outline</i>), mengingat kembali, mereproduksi, memilih, menyatakan, mengenali (<i>recognising</i>), melacak, menamai, menemukan.</p>

Pada tahun 2001 Anderson dan Krathwohl mengembangkan lebih lanjut taksonomi Bloom ranah kognitif yang sudah direvisi oleh Anderson tersebut dengan cara menambahkan dimensi pengetahuan pada masing-masing dimensi proses kognitif atau kemampuan berpikir. Taksonomi Bloom ranah kognitif hasil revisi Anderson dan Krathwohl itu kini mencakup *dimensi pengetahuan* yakni jenis pengetahuan yang harus dipelajari serta *dimensi proses kognitif* yakni proses berpikir yang harus digunakan untuk mempelajari masing-masing jenis pengetahuan. Taksonomi Bloom ranah kognitif hasil revisi terbaru tersebut menjadi seperti disajikan dalam Tabel 2.2. (Tarlinton, 2003).

Tabel 2.2: Taksonomi Bloom Ranah Kognitif Hasil Revisi Anderson dan Krathwohl

		Dimensi Proses Kognitif				
Dimensi Pengetahuan	Mengingat	Memahami	Mengaplikasikan	Menganalisis	Mengevaluasi	Menciptakan
Pengetahuan Faktual	Membuat daftar (<i>List</i>)	Membuat ringkasan (<i>Summarize</i>)	Mengklasifikasikan (<i>Classify</i>)	Mengurutkan (<i>Order</i>)	Menentukan <i>ranking</i> (<i>Rank</i>)	Mengombinasikan (<i>Combine</i>)
Pengetahuan Konseptual	Mendeskrripsikan (<i>Describe</i>)	Menafsirkan (<i>Interpret</i>)	Bereksperimentasi (<i>Experiment</i>)	Menjelaskan (<i>Explain</i>)	Melakukan penilaian (<i>Assess</i>)	Merencanakan (<i>Plan</i>)
Pengetahuan Prosedural	Menabuliskan (<i>Tabulate</i>)	Memprediksikan (<i>Predict</i>)	Menghitung (<i>Calculate</i>)	Membedakan (<i>Differentiate</i>)	Membuat kesimpulan (<i>Conclude</i>)	Membuat komposisi (<i>Compose</i>)
Pengetahuan Meta-kognitif	Menggunakan secara tepat/ semestinya (<i>Appropriate use</i>)	Mengeksekusi (<i>Execute</i>)	Mengonstruksi (<i>Construct</i>)	Mencapai atau meraih pencapaian (<i>Achieve</i>)	Melakukan tindakan (<i>Action</i>)	Mengaktualisasikan (<i>Actualize</i>)

Uraian yang sedikit lebih rinci tentang masing-masing jenis dimensi pengetahuan adalah seperti disajikan dalam Tabel 2.3. (Tarlinton, 2003).

Tabel 2.3: Jenis Kemampuan Menurut Dimensi Pengetahuan

Dimensi Pengetahuan	Uraian
Pengetahuan Faktual	<p>Pengetahuan tentang unsur-unsur dasar aneka disiplin ilmu/teknologi/seni agar murid memiliki pemahaman dan/atau mampu memecahkan aneka masalah dalam aneka disiplin ilmu/teknologi/seni yang bersangkutan, meliputi:</p> <p><i>Pengetahuan tentang istilah:</i> kosa kata teknis, aneka simbol teknis.</p> <p><i>Pengetahuan tentang aneka detil dan unsur spesifik:</i> aneka sumber daya alam, aneka jenis alat musik, aneka jenis tumbuhan.</p>
Pengetahuan Konseptual	<p>Pengetahuan tentang saling hubungan antar unsur-unsur dasar dalam sebuah struktur yang lebih besar atau lebih kompleks dalam disiplin ilmu/teknologi/seni yang memungkinkan unsur-unsur tersebut berfungsi bersama, meliputi:</p> <p><i>Pengetahuan tentang aneka klasifikasi dan kategori:</i> aneka periode dari zaman atau masa geologis, aneka jenis binatang.</p> <p><i>Pengetahuan tentang aneka prinsip dan generalisasi:</i> prinsip ekuilibrium, hukum penawaran dan permintaan.</p> <p><i>Pengetahuan tentang aneka teori, model, atau struktur:</i> teori evolusi, struktur pemerintahan demokratis.</p>
Pengetahuan Prosedural	<p>Informasi atau pengetahuan yang akan membantu murid melakukan sesuatu yang bersifat spesifik terkait disiplin ilmu/teknologi/seni tertentu, meliputi:</p> <p><i>Pengetahuan tentang ketrampilan spesifik dalam bidang keahlian tertentu:</i> melukis dengan cat air, pembagian bilangan.</p> <p><i>Pengetahuan tentang aneka metode penelitian, aneka teknik atau metodologi dalam aneka disiplin ilmu:</i> teknik wawancara, metode ilmiah.</p> <p><i>Pengetahuan tentang kriteria untuk menentukan kapan menggunakan aneka prosedur yang tepat atau semestinya:</i> kapan atau syarat menerapkan hukum kedua Newton, kapan atau syarat menggunakan uji statistik parametrik.</p>

Dimensi Pengetahuan	Uraian
Pengetahuan Meta-kognitif	<p>Pengetahuan tentang aneka proses berpikir khas tertentu maupun kesadaran tentang pemahamannya sendiri, meliputi tiga komponen:</p> <p><i>Pengetahuan metakognitif</i>, yaitu pemahaman murid tentang dirinya sendiri maupun orang lain sebagai makhluk yang berpikir.</p> <p><i>Regulasi metakognitif</i>, yaitu kemampuan mengatur proses berpikir dan aneka pengalaman belajar lewat serangkaian aktivitas yang membantu murid mengarahkan kegiatan belajarnya.</p> <p><i>Pengalaman metakognitif</i>, yaitu aneka pengalaman terkait kegiatan berpikir yang sedang dilakukannya saat ini.</p>

Taksonomi Tujuan Pengajaran Ranah Afektif

Taksonomi ranah afektif yang paling luas dipakai adalah hasil rumusan Krathwohl, Bloom, dan Masia (1964). Taksonomi ini mengklasifikasikan emosi atau perasaan siswa terhadap aneka pengalaman belajar yang diperolehnya di dalam maupun di luar kelas, atau cara siswa menanggapi orang, benda, atau situasi dengan menggunakan perasaannya. Emosi atau perasaan yang dimaksud meliputi sikap, minat, perhatian, kesadaran, nilai, apresiasi, entusiasme, dan motivasi yang juga diasumsikan tersusun secara hirarkis mulai dari yang paling sederhana sampai ke yang paling kompleks, dan yang tercermin dalam aneka bentuk perilaku siswa. Sepanjang pengetahuan penulis, sejak diterbitkan untuk pertama kali, taksonomi ini tidak pernah direvisi hingga sekarang. Secara ringkas taksonomi ranah afektif menurut Krathwohl, Bloom, dan Masia adalah seperti disajikan dalam Tabel 2.4. (Tarlington, 2003).

**Tabel 2.4: Taksonomi Tujuan Pengajaran Ranah Afektif
(Krathwohl, Bloom, & Masia, 1964)**

Taraf Kemampuan	Uraian
<p>Mau menginternalisasikan nilai-nilai (karakterisasi). Memiliki suatu sistem nilai yang dijadikan pedoman berperilaku, sehingga perilaku menjadi konsisten, bisa diprediksikan, dan yang terpenting menjadi ciri atau karakteristik pribadi yang bersangkutan. (<i>Internalizing Values</i>)</p>	<p>Mau menunjukkan perilaku yang dikendalikan oleh suatu sistem nilai. Kata kunci: bertindak, menunjukkan, mempraktekkan, memodifikasikan, mendengarkan, mengusulkan, mengajukan pertanyaan, memverifikasikan, memberikan layanan. Contoh: Menunjukkan kemandirian saat mengerjakan sesuatu secara mandiri. Mampu bekerjasama dalam aktivitas kelompok. Menerapkan pendekatan sasaran (<i>objective approach</i>) dalam memecahkan masalah. Menunjukkan komitmen terhadap etika dalam praktek sehari-hari. Mau mengubah pendapat dan perilaku menyesuaikan diri dengan bukti-bukti baru. Menghargai orang lain apa adanya, bukan berdasarkan penampilan mereka.</p>
<p>Mau mengorganisasikan nilai-nilai. Mengorganisasikan nilai ke dalam skala prioritas (mengurutkan dari yang paling penting/ bernilai sampai yang paling kurang penting/kurang bernilai) dengan cara membandingkan berbagai nilai yang berbeda, mengatasi konflik-konflik yang terjadi antar nilai-nilai yang berbeda tersebut, dan akhirnya mampu menciptakan suatu sistem nilai yang khas bagi dirinya. (<i>Organization</i>)</p>	<p>Mau mengorganisasikan nilai-nilai mengikuti urutan prioritas tertentu. Kata kunci: menggabungkan, membandingkan, mempertahankan, menjelaskan, merumuskan, menggeneralisasikan, mengintegrasikan, memodifikasikan, mengorganisasikan, menyintesiskan. Contoh: Menyadari pentingnya menyeimbangkan antara kebebasan dan tanggung jawab. Mau bertanggung jawab atas tindakannya. Menjelaskan fungsi perencanaan sistematis dalam pemecahan masalah. Mau menerima dan mengikuti aneka standar etika profesi. Mampu menyusun rencana masa depan selaras dengan kemampuan, minat, dan keyakinan pribadi. Mampu mengatur waktu secara efektif untuk memenuhi kebutuhan belajar dan bermain dengan teman.</p>

Taraf Kemampuan	Uraian
<p>Mau memberikan nilai/ mau memandang bernilai, mulai dari sekadar menerima sesuatu sebagai bernilai sampai menunjukkan komitmen yang lebih kompleks. Kemampuan ini didasari oleh internalisasi terhadap serangkaian nilai-nilai spesifik tertentu. (<i>Valuing</i>)</p>	<p>Mau memberikan nilai pada sesuatu. Kata kunci: menunjukkan, menjelaskan, mengikuti, mempersilakan, memberikan pembenaran, mengusulkan, memilih, mempelajari. Contoh: Menunjukkan keyakinan tentang keunggulan proses yang demokratis. Peka terhadap keberagaman individu maupun budaya. Menunjukkan kemampuan memecahkan aneka masalah. Mau mengusulkan suatu rencana perbaikan kehidupan bersama dan mengikutinya dengan penuh komitmen.</p>
<p>Mau memberikan respon terhadap fenomena tertentu, meliputi mau berpartisipasi aktif, mau memberikan perhatian dan reaksi terhadap fenomena tertentu. Hasil belajar yang ditekankan: mau menjawab dan merasakan kepuasan dengan memberikan respon. (<i>Responding to phenomena</i>)</p>	<p>Mau berperan aktif dalam kegiatan belajar; berpartisipasi. Kata kunci: mau menjawab, memberikan bantuan, mau mengikuti perintah, memberi salam, mau membantu, mau melakukan, memilih. Contoh: Mau berpartisipasi dalam dikusi kelas. Mau memberikan presentasi di depan kelas. Mau mengajukan pertanyaan tentang aneka gagasan, konsep, model yang baru didengar untuk lebih memahaminya. Mengetahui aturan tentang kebersihan dan mau mematuhi.</p>
<p>Mau menerima fenomena tertentu, yaitu mau menyadari, mau mendengarkan, atau mau memberikan perhatian. (<i>Receiving phenomena</i>)</p>	<p>Mau menyadari; menunjukkan kemauan untuk mendengarkan. Kata kunci: bertanya, memilih, mendeskripsikan, mengikuti, memberikan, menyebut nama, menunjuk, duduk, menjawab pertanyaan. Contoh: Mendengarkan guru atau teman dengan rasa hormat. Mendengarkan dan mengingat nama orang atau teman yang baru diperkenalkan.</p>

Taksonomi Tujuan Pengajaran Ranah Psikomotor

Ranah psikomotor mencakup kemampuan menggunakan aneka ketrampilan motor, koordinasi, dan gerakan fisik. Pengembangan aneka ketrampilan ini menuntut praktek atau latihan, dan kemajuan atau keberhasilannya dapat diukur dari peningkatan kecepatan, ketepatan, jarak, prosedur, atau teknik dalam melaksanakan tugas atau aktivitas motor tertentu. Tim peneliti pimpinan Bloom tidak mengembangkan taksonomi ranah psikomotor ini secara lengkap karena merasa kurang memiliki pengalaman mengajarkan ranah ketrampilan ini, namun seorang pakar lain yakni Simpson (1972) mengembangkan sebuah taksonomi ranah psikomotor yang mencakup tujuh kategori kemampuan untuk menuntaskan gagasan Bloom dan kawan-kawan (Tarlinton, 2003). Secara ringkas, taksonomi ranah psikomotor hasil karya Simpson (1972) adalah sebagaimana disajikan dalam Tabel 2.5.

Tabel 2.5: Taksonomi Tujuan Pengajaran Ranah Psikomotor (Simpson, 1972)

Taraf Kemampuan	Uraian
<p>Mampu menghasilkan gerakan baru Mampu menciptakan aneka pola gerakan baru sesuai tuntutan suatu situasi atau problem khusus tertentu. Hasil belajar yang ditekankan berupa kreativitas yang dilandasi aneka ketrampilan taraf tinggi. (<i>Origination</i>)</p>	<p>Mampu menciptakan pola gerakan baru. Kata kunci: menata, membangun, mengombinasikan, mengarang (<i>compose</i>), mengonstruksi, menciptakan, merancang, memulai (<i>initiate</i>), mengawali (<i>originate</i>), membuat. Contoh: Merumuskan suatu teori baru tentang cara mencapai kebugaran. Mengembangkan program pelatihan kebugaran baru yang komprehensif. Menciptakan pola senam kesegaran jasmani baru.</p>

Taraf Kemampuan	Uraian
<p>Mampu beradaptasi Aneka ketrampilan sudah dikuasai dengan baik, sehingga siswa mampu memodifikasikan pola gerakan agar sesuai dengan tuntutan situasi khusus tertentu. (<i>Adaptation</i>)</p>	<p>Mampu memodifikasi aneka ketrampilan motor disesuaikan dengan tuntutan situasi baru. Kata kunci: beradaptasi, mengubah, menata kembali, mengorganisasikan kembali, merevisi, menciptakan variasi. Contoh: Mampu memberikan respon yang efektif terhadap aneka pengalaman yang tidak diharapkan. Mampu memodifikasikan pengajaran olah raga dan kesehatan agar lebih sesuai dengan kebutuhan murid. Melaksanakan tugas dengan sebuah mesin yang aslinya tidak dimaksudkan untuk itu (mesin tidak menjadi rusak dan tindakan tersebut tidak membahayakan).</p>
<p>Mampu melakukan respon kompleks secara lancar Mampu melakukan tindakan motor secara trampil yang melibatkan pola gerakan yang kompleks. Ketrampilan atau ketangkasan ditunjukkan oleh gerakan yang cepat, akurat, dan sangat terkoordinasi, yang dilakukan dengan energi atau upaya minimum. Kategori ini mencakup mengerjakan tugas tanpa ragu-ragu dan melakukan gerakan secara otomatis. (<i>Complex Overt Response</i>)</p>	<p>Menunjukkan tahap agak lanjut menguasai suatu ketrampilan kompleks. Kata kunci: merakit, membangun, mengonstruksi, memperbaiki, memanipulasikan, mengukur, mengorganisasikan, membuat sketsa. Contoh: Memarkir mobil di ruang parkir yang sempit. Mengetik komputer dengan cepat dan tepat. Menunjukkan kepiawaian saat bermain piano. <i>Catatan:</i> Kata-kata kunci kategori ini sama seperti kategori “Mampu melakukan secara mekanik”, namun dengan tambahan kata sifat yang menunjukkan bahwa aktivitas atau gerakan tersebut dilakukan secara lebih cepat, lebih baik, lebih tepat.</p>
<p>Mampu melakukan respon secara mekanik Merupakan tahap piawai dalam mempelajari suatu ketrampilan kompleks. Hasil belajar sudah menyatu dengan kebiasaan, sehingga gerakan-gerakan bisa dilakukan</p>	<p>Mampu melakukan suatu ketrampilan motor yang kompleks. Kata kunci: merakit, membangun, mengonstruksi, memperbaiki, memanipulasikan, mengukur, mengorganisasikan, membuat sketsa.</p>

Taraf Kemampuan	Uraian
dengan percaya diri dan lancar. (<i>Mechanism</i>)	Contoh: Mengetik komputer dengan lancar. Memperbaiki kran yang bocor. Mengemudikan mobil.
<p>Mampu melakukan respon tertentu dengan bimbingan guru Merupakan tahap awal dalam mempelajari suatu ketrampilan kompleks, mencakup kemampuan mencontoh atau coba-salah. Ketrampilan yang memadai akan dicapai lewat latihan. (<i>Guided Response</i>)</p>	<p>Menunjukkan tahap awal menguasai suatu ketrampilan kompleks, meliputi kemampuan mengikuti contoh atau mencontoh. Kata kunci: mencontoh, melacak, mengikuti, mereproduksi, menanggapi, mereaksi. Contoh: Menulis huruf latin tegak bersambung seperti diberi contoh guru di papan tulis. Mengikuti petunjuk dalam rangka membangun sebuah model pesawat. Mengikuti isyarat tangan instruktur saat belajar mengoperasikan <i>forklift</i>.</p>
<p>Memiliki kesiapan untuk bertindak Kesiapan untuk bertindak, meliputi kesiapan mental, fisik, dan emosi. Ketiganya merupakan disposisi yang mendasari respon seseorang terhadap berbagai situasi yang dihadapi (kadang-kadang juga disebut <i>mindset</i>). (<i>Set</i>)</p>	<p>Menunjukkan kesiapan untuk bertindak. Kata kunci: mulai, mempertontonkan, menjelaskan, bereaksi, menyatakan, menunjukkan. Contoh: Tahu dan mampu melakukan serangkaian langkah dalam proses menghasilkan produk tertentu. Menunjukkan hasrat untuk mempelajari sebuah proses baru. <i>Catatan:</i> Sub-kemampuan psikomotor ini memiliki kaitan erat dengan sub-kemampuan “Merespon fenomena” dalam ranah afektif.</p>
<p>Mampu mempersepsikan Mampu menggunakan petunjuk-isyarat sensoris untuk membimbing aktivitas motor, meliputi kepekaan menangkap stimulasi sensoris kemampuan memilih petunjuk-isyarat sensoris, dan kemampuan menerjemahkannya ke dalam tindakan. (<i>Perception</i>)</p>	<p>Mampu menggunakan tanda-tanda sensoris untuk membimbing aktivitas fisik tertentu. Kata kunci: Memilih, mendeskripsikan, mendeteksi, membedakan, mengidentifikasi, mengisolasi, menghubungkan, dan menyeleksi.</p>

Taraf Kemampuan	Uraian
	<p>Contoh: Mampu mendeteksi petunjuk-isyarat komunikasi nonverbal. Mampu memperkirakan arah jatuh dari bola yang dilempar dan selanjutnya menuju lokasi untuk menangkap bola itu. Menyetel panas oven untuk mengurangi temperatur berdasarkan bau makanan yang sedang dipanggang.</p>

Sekadar sebagai perbandingan dan sebagaimana juga sudah disinggung, taksonomi tujuan pengajaran ranah psikomotor yang dikembangkan oleh dua pakar lain yang sesungguhnya juga cukup populer adalah hasil pemikiran Dave (1975) dan Harrow (1972) seperti disajikan pada Tabel 2.6. dan Tabel 2.7.

Tabel 2.6: Taksonomi Tujuan Pengajaran Ranah Psikomotor ala Dave (1975)

Taraf kemampuan	Uraian
Naturalisasi	Mampu melakukan tindakan/kinerja berkualitas tinggi secara wajar-alami, tanpa perlu melibatkan banyak pemikiran. Contoh: Michael Jordan bermain bola basket, Andik Vermansyah bermain sepak bola.
Artikulasi	Mampu mengoordinasikan serangkaian tindakan, mencapai keharmonisan dan konsistensi internal dalam melakukan sesuatu. Contoh: Membuat sebuah video yang melibatkan musik, drama, warna, bunyi, dsb.
Presisi	Mampu memperhalus, menjadi semakin sempurna, menjadi semakin eksak atau akurat, ditandai dengan semakin sedikit membuat kesalahan. Contoh: Terus mencoba melakukan atau mengerjakan sesuatu sampai akhirnya “memuaskan.”

Taraf kemampuan	Uraian
Manipulasi	Mampu melakukan sejumlah tindakan tertentu dengan mengikuti petunjuk dan berlatih. Contoh: Menciptakan sebuah karya sendiri sesudah mengikuti pelajaran/latihan atau sesudah membaca petunjuk.
Imitasi	Mengamati untuk selanjutnya meniru perilaku orang lain. Hasil tiruannya mungkin masih berkualitas rendah. Contoh: Mencontoh sebuah karya seni sederhana.

Tabel 2.7: Taksonomi Tujuan Pengajaran Ranah Psikomotor ala Harrow (1972)

Taraf kemampuan	Uraian
Komunikasi nondiskursif/nonverbal	Mampu melakukan bahasa tubuh secara efektif, seperti <i>gestures</i> alias gerak isyarat dan ekspresi wajah.
Gerak trampil	Mampu melakukan aneka gerakan berkualitas tinggi hasil berlatih seperti yang terjadi dalam olah raga atau seni peran.
Kemampuan fisik	Memiliki stamina yang harus dikembangkan lebih lanjut menjadi kekuatan dan ketangkasan.
Persepsi	Mampu memberikan respon terhadap aneka stimulus seperti melakukan diskriminasi penglihatan, pendengaran, kinestetik, atau perabaan.
Gerak fundamental	Mampu melakukan aneka gerakan dasar seperti berjalan atau memegang.
Gerak refleks	Mampu melakukan aneka reaksi tanpa mempelajarinya.

Taksonomi Fungsional Ranah Pengetahuan, Ketrampilan, dan Perilaku (Chatterji, 2003)

Taksonomi tujuan pengajaran yang dikembangkan oleh Bloom dan kawan-kawan serta para pengarang awal lain sesudah mereka memperlakukan ketiga ranah kemampuan secara terpisah,

seakan-akan masing-masing merupakan wilayah kemampuan tersendiri. Dalam kenyataan sehari-hari fungsi atau peran ketiga jenis kemampuan manusiawi tersebut memang saling terkait secara tak terpisahkan. Masing-masing lebih tepat disebut sebagai dimensi dan bukan sebagai wilayah dari satu kemampuan budi manusiawi. Kurang lebih bertolak dari asumsi semacam itu, Chatterji (2003), seorang ahli psikologi pendidikan Amerika mengembangkan sebuah taksonomi komprehensif yang dia sebut *a functional taxonomy of knowledge, skills, and behavior*. Taksonomi integratif ini dia kembangkan berdasarkan teori belajar yang juga dikembangkannya dan yang dia sebut *Dimensions of Learning Model* atau teori tentang dimensi-dimensi belajar. Taksonomi komprehensif yang mencakup ketiga ranah tradisional kognitif-afektif-psikomotor ini meliputi 7 jenis kemampuan, sebagaimana disajikan dalam Tabel 2.6.

Tabel 2.6: Taksonomi Fungsional Ranah Pengetahuan, Keterampilan, dan Perilaku (Chatterji, 2003)

1. ***Factual Knowledge*** atau ***Pengetahuan Faktual***
Kemampuan mereproduksi informasi meliputi fakta, istilah, definisi, rumus, dan konsep seperti apa adanya dalam arti sebagaimana informasi itu pernah dipelajari, atau dengan menggunakan kata-kata sendiri.
Bentuk perilaku: menyebutkan, menjelaskan, memberikan ilustrasi atau contoh, menerjemahkan, menyatakan kembali dengan kata-kata sendiri, mengingat.
Contoh hasil belajar:
Mampu menjelaskan bagian-bagian dari sepeda.
Mampu menyebutkan hukum gerak ketiga Newton dan contoh penerapannya.

2. ***Application*** atau ***Penerapan***
Mampu menerapkan atau menggunakan prinsip, aturan, rumus, atau prosedur dalam rangka memecahkan sebuah masalah.
Bentuk perilaku: menggunakan, menerapkan, memecahkan, menuliskan, menghitung, membuat, menunjukkan atau mendemonstrasikan.

Contoh hasil belajar:

Menghitung luas sebuah ruangan jika diketahui panjang dan lebarnya.

Menuliskan kalimat dengan tata bahasa, ejaan, dan tanda baca yang benar.

Menerapkan hukum Boyle untuk memecahkan soal hitungan.

3. ***Complex Procedural Skills*** atau ***Ketrampilan Prosedural yang Kompleks***

Mampu menggunakan atau menerapkan prosedur-prosedur yang terdiri dari sejumlah langkah (*multistep procedure*), seringkali juga menuntut kemampuan merumuskan hipotesis, berpikir atau bernalar, menerapkan teknik atau piranti tertentu, atau membuat penilaian dan keputusan, dalam rangka memecahkan sebuah problem atau melaksanakan tugas yang kompleks. Tekanannya adalah pada kemampuan mengimplementasikan proses dan bukan pada hasilnya. Langkah-langkah dalam prosedur yang kompleks tersebut lazimnya mengikuti standar yang berlaku dalam bidang ilmu atau bidang keahlian tertentu.

Bentuk perilaku: menggunakan atau menerapkan sebuah prosedur, mendemonstrasikan sebuah prosedur, menuliskan penerapan prosedur kerja baku.

Contoh hasil belajar:

Menerapkan sebuah metode ilmiah dalam rangka menguji hipotesis tentang gejala fisik atau sosial tertentu.

Menerapkan prosedur baku dalam menyusun rencana tindakan pengerjaan suatu proyek atau menyusun anggaran.

Menerapkan strategi "plan, solve, check, and revise" (susun rencana, laksanakan, evaluasi, dan perbaiki) dalam memecahkan sebuah masalah.

4. ***Higher Order Thinking and Problem Solving*** atau ***Berpikir Tingkat Tinggi dan Memecahkan Masalah***

Mampu mengidentifikasi atau merumuskan masalah dan menemukan aneka kemungkinan strategi atau prinsip untuk memecahkan masalah tersebut. Kemampuan ini bisa mencakup gabungan antara kemampuan menganalisis (menguraikan masalah ke dalam bagian-bagian atau unsur-unsurnya), melakukan sintesis (menyatukan atau menggabungkan bagian-bagian atau unsur-unsur) dan/atau melakukan evaluasi (menilai manfaat atau keuntungan-kerugian).Tergantung dari jenis tugas yang dihadapi, bisa menuntut penerapan aneka bentuk perilaku dan ketrampilan berpikir, termasuk mengembangkan atau menghasilkan produk tertentu.

Contoh hasil belajar:

Membandingkan dua teori tentang penyebab kepunahan dinosaurus.
Membandingkan media sastra yang dipakai oleh dua penyair abad ke-20 dan menilai gaya sastra mereka.

Membuat rancangan tata ruang dan fasilitas yang diperlukan untuk sebuah proyek.

5. ***Attitudes and Values*** atau ***Sikap dan Nilai***

Kemampuan mengungkapkan diri dan atau mengambil tindakan berdasarkan perasaan, pendapat, dan keyakinan pribadi menyangkut orang, objek, atau peristiwa.

Contoh hasil belajar:

Menunjukkan sikap hormat terhadap milik, hak, dan pandangan orang lain.

Menerima keberagaman antar budaya.

6. ***Social Behavior*** atau ***Perilaku Sosial***

Mampu menunjukkan bentuk-bentuk perilaku yang sesuai dengan norma masyarakat dalam situasi sosial baik yang bersifat resmi maupun tidak resmi.

Contoh hasil belajar:

Melaksanakan peran dalam mengerjakan tugas kerja sama kelompok di sekolah.

Mengungkapkan berbagai gagasan yang bertujuan melancarkan berbagai fungsi kelompok.

7. ***Motor Skills*** atau ***Ketrampilan Motor***

Menunjukkan koordinasi, kekuatan, penguasaan, dan ketrampilan dalam melaksanakan tugas-kegiatan fisik dan olah raga.

Contoh hasil belajar:

Lari sejauh satu kilometer sekeliling lapangan dalam waktu kurang dari 10 menit.

Menunjukkan ketrampilan mengetik dengan 10 jari.

Dalam kenyataan, model taksonomi yang membedakan kemampuan budi manusiawi ke dalam tiga ranah atau wilayah sebagaimana dikembangkan oleh Bloom dan kawan-kawan masih luas dipakai di dunia pendidikan sekolah hingga kini, apalagi khususnya sesudah taksonomi tujuan pengajaran ranah kognitifnya direvisi. Dengan adanya model taksonomi baru yang lebih integratif

seperti dikemukakan oleh Chatterji (2003), sudah barang tentu pilihan guru menjadi semakin luas. Mana pun model yang dipilih, taksonomi tujuan pengajaran akan sangat membantu guru dalam merumuskan aneka tujuan pengajaran berbagai mata pelajaran yang diampunya. Selain itu, kehadiran taksonomi tujuan pengajaran juga akan membantu guru dalam mengembangkan aneka alat penilaian hasil belajar para murid untuk memastikan sejauh mana aneka tujuan pengajaran tersebut telah dicapai oleh para murid. Dalam pembicaraan selanjutnya, kita masih akan bertolak dari taksonomi tujuan pengajaran sebagaimana dikembangkan oleh Bloom dan kawan-kawan.

Bab 3

Pengembangan Alat Penilaian Hasil Belajar: Teknik Tes

Seperti sudah disinggung, secara garis besar ada dua kategori alat penilaian hasil belajar, yaitu *tes* dan *nontes*. Kendati fokus pembahasan kita adalah teknik nontes, namun kiranya perlu kita ketahui garis besar seluk-beluk tes, sebab sejumlah prinsip dalam penyusunan tes juga berlaku dalam penyusunan alat nontes. Selain itu, garis besar seluk-beluk teknik tes perlu kita simak untuk menjernihkan pemahaman kita tentang persamaan dan perbedaannya dengan teknik nontes.

Tes adalah "*a type of assessment that uses specific procedures to obtain information and convert that information to numbers or scores*" (Friedenberg, 1995). Artinya, tes merupakan salah satu jenis asesmen yang menggunakan aneka prosedur spesifik untuk memperoleh informasi dan mengonversikan atau mengubah informasi tersebut ke dalam skor atau bilangan.

Ada tiga ciri utama tes yang terkandung dalam definisi di atas (Friedenberg, 1995). Pertama, aneka prosedur spesifik yang dimaksud meliputi: (1) memilih serangkaian pertanyaan atau tugas untuk dijadikan item atau butir tes, (2) menentukan berbagai kondisi atau

persyaratan terkait pengadministrasian atau penyelenggaraan tes yang dimaksud, meliputi antara lain alokasi waktu pengerjaannya, dan (3) mengembangkan sistem penskoran dan penginterpretasian jawaban *testi* (dari kata Inggris *testee*) alias subjek atau orang yang dites.

Kedua, tes bersifat mengkuantifikasikan jawaban dengan cara mengonversikan atau mengubah jawaban *testi* menjadi bilangan atau skor. Tes bertujuan mengukur berbagai *atribut* atau *aspek kemampuan* atau *aspek kepribadian* dari *testi* dengan cara menera masing-masing atribut dengan nilai numerik atau bilangan.

Ketiga, setiap tes harus dipandang sebagai *a sample of behavior* atau sebuah sampel atau contoh perilaku. Jawaban atas pertanyaan-pertanyaan dan hasil pelaksanaan tugas-tugas dalam tes merupakan sampel atau contoh dari pengetahuan, ketrampilan, atau sifat kepribadian *testi*. Maka, kualitas sebuah tes akan sangat ditentukan oleh pemilihan pertanyaan-pertanyaan atau tugas-tugas yang menjadi item-item atau butir-butirnya. Tes yang baik harus mampu berperan sebagai sampel perilaku yang representatif, yaitu mengandung item-item yang benar-benar mencerminkan pengetahuan, ketrampilan, atau sifat kepribadian sebagai hasil belajar yang hendak diukur atau diungkap dari *testi*.

Dari segi ranah kemampuan yang diukur, tes bisa dibedakan ke dalam dua golongan besar: (1) *maximal performance tests* atau tes yang bertujuan mengukur kinerja maksimal, dan (2) *typical performance tests* atau tes yang bertujuan mengukur kinerja khas. Dalam konteks penilaian hasil belajar di sekolah, tes golongan pertama (*maximal performance test*) kiranya tepat dipakai untuk melakukan asesmen tentang semua kemampuan dalam dimensi proses kognitif mulai dari pengetahuan faktual sampai ke pengetahuan metakognitif dalam ranah kognitif, serta beberapa kemampuan dalam ranah

psikomotor, khususnya *kemampuan melakukan respon secara mekanik* dan *kemampuan melakukan respon kompleks secara lancar*. Pengukuran semua jenis kemampuan tersebut sebagai hasil belajar (*achievement*) dilakukan dengan cara menentukan *the upper limits of test-taker's knowledge or skill* alias batas atas pengetahuan atau ketrampilan testi (Friedenberg, 1995). Batas atas kemampuan atau ketrampilan tersebut lazimnya terungkap dari keberhasilan murid menjawab pertanyaan atau melaksanakan tugas yang paling sulit. Item-item tes hasil belajar golongan *maximal performance* memang lazim memiliki *taraf kesulitan* yang berlainan. Selain itu, jawaban terhadap item tes *maximal performance* harus bisa dinyatakan *benar* atau *berhasil* sehingga diberi skor, atau sebaliknya bisa dinyatakan *salah* atau *gagal* sehingga tidak diberi skor. Dengan begitu, makin tinggi skor yang dicapai seorang testi dalam tes *maximal performance* akan mencerminkan batas atas kemampuannya yang makin tinggi pula.

Dalam konteks penilaian hasil belajar di sekolah, tes golongan kedua (*typical performance test*) kiranya tepat dipakai untuk melakukan asesmen tentang cara atau kebiasaan dalam berpikir, merasa, dan berperilaku yang bersifat khas pada masing-masing murid atau jenis-jenis kemampuan taraf tinggi dalam ranah afektif sebagai hasil belajar. Yang menjadi fokus golongan tes ini bukan batas atas kinerja, melainkan cara berperilaku yang khas (*typical performance*) pada masing-masing murid. Sekalipun merupakan hasil belajar, namun jenis-jenis kemampuan afektif yang diungkap dengan golongan tes *typical performance* lazim berupa apa yang disebut *konstruk* yakni aspek atau atribut kepribadian yang bersifat abstrak seperti aneka sifat kepribadian, sikap terhadap aneka objek atau situasi, minat, atau nilai. Tidak seperti hasil belajar berupa pengetahuan atau ketrampilan dalam mata pelajaran tertentu yang diungkap dengan tes *maximal performance*, hasil belajar berupa konstruk tidak memiliki

ranah isi yang spesifik atau jelas sehingga keberadaannya hanya bisa disimpulkan berdasarkan indikator-indikator tingkah lakunya baik yang bersifat *favorable* alias mendukung kehadirannya maupun yang bersifat *unfavorable* alias menyangkal kehadirannya. Akibatnya juga, pada golongan tes *typical performance* tidak dikenal jawaban yang bisa dinyatakan benar atau salah, dalam arti bahwa semua jawaban adalah benar karena dipandang mencerminkan kekhasan berperilaku masing-masing murid. Dengan kata lain, item-item dirancang sedemikian rupa sehingga jawaban testi dapat dipakai sebagai dasar untuk memasukkannya ke dalam kategori tertentu terkait sifat kepribadian, sikap, minat, atau nilai tertentu, misal sifat ekstravert atau introvert, sikap positif atau negatif terhadap mata pelajaran tertentu, minat yang tinggi atau rendah terhadap kegiatan seni atau olah raga, dan sebagainya.

Dari segi penyusunnya, tes bisa digolongkan ke dalam: (1) tes baku (*standardized tests*) yang disusun oleh pakar pengukuran umum maupun pakar pengukuran berbagai bidang studi, meliputi antara lain tes inteligensi, tes bakat, tes minat, tes kepribadian, dan tes baku hasil belajar berbagai bidang studi seperti Matematika, IPA, bahasa, dan (2) tes buatan guru (*teacher made tests*), meliputi tes hasil belajar berbagai bidang pelajaran yang disusun sendiri oleh guru untuk keperluan pengajaran. Kita hanya akan berfokus pada jenis tes hasil belajar mata pelajaran buatan guru.

Jadi, baik berdasarkan pemahaman kita tentang taksonomi tujuan pengajaran versi baru sebagaimana diuraikan di bagian sebelumnya maupun berdasarkan pemahaman kita tentang makna tes yang sesungguhnya, kiranya tidak bisa serta merta kita katakan bahwa *tes* merupakan alat yang cocok untuk melakukan asesmen aneka kemampuan dalam ranah kognitif sedangkan *nontes* merupakan alat paling cocok untuk melakukan asesmen aneka kemampuan

dalam ranah afektif, sebagaimana beredar di kalangan yang cukup luas hingga kini.

Syarat Utama Tes: Valid

Salah satu syarat mutlak yang harus dipenuhi oleh tes hasil belajar baik yang baku maupun yang dibuat sendiri oleh guru adalah validitas, yaitu sejauh mana tes benar-benar mengukur pengetahuan atau sifat yang tepat seperti yang dimaksud oleh tujuan tes itu (Friedenberg, 1995). Ada tiga jenis pemeriksaan validitas: (1) *validitas isi*, memeriksa sejauh mana tes mengukur ranah isi pengetahuan atau ketrampilan di bidang tertentu yang tepat sebagaimana dimaksud, (2) *validitas konstruk*, memeriksa sejauh mana tes mengukur suatu konstruk tertentu entah berupa kemampuan yang lebih didominasi oleh fungsi kognitif seperti inteligensi maupun kecenderungan kepribadian tertentu yang lebih didominasi oleh fungsi afektif seperti sifat, dan (3) *validitas kriteria*, memeriksa sejauh mana tes mampu memprediksikan kinerja testi dalam bidang tertentu di masa mendatang atau memprediksikan variabel-variabel tertentu sebagaimana diharapkan berdasarkan penalaran teoretis tertentu. Seperti akan kita bahas, tidak semua jenis validitas cocok diterapkan pada proses validasi atau pemeriksaan semua jenis tes.

Validasi isi secara khusus cocok diterapkan pada jenis tes yang memiliki *content domain* atau ranah isi yang jelas, khususnya tes hasil belajar berbagai mata pelajaran di sekolah baik yang bersifat baku maupun susunan guru. Setiap mata pelajaran memiliki ranah atau cakupan isi pengetahuan yang jelas, lazimnya dituangkan dalam silabus atau bahkan disajikan sebagai materi dalam buku teks atau buku pelajaran. Mengikuti taksonomi tujuan pengajaran ranah kognitif Bloom yang sudah direvisi oleh Anderson dan Krathwohl, pengetahuan terkait mata pelajaran tertentu tersebut mencakup

pengetahuan tentang aneka fakta (pengetahuan faktual), pengetahuan tentang aneka konsep (pengetahuan konseptual), pengetahuan tentang aneka prosedur (pengetahuan prosedural), dan sejenis ketrampilan intelektual untuk mengolah dan mengembangkan aneka pengetahuan yang sudah dimiliki tersebut untuk menghasilkan tindakan, produk, atau pengetahuan baru tertentu (pengetahuan meta-kognitif). Sebuah tes hasil belajar dinyatakan memiliki validitas isi yang baik jika item-item dalam tes tersebut merepresentasikan secara memadai aspek-aspek yang penting terkait dengan baik materi pelajaran alias dimensi pengetahuan maupun jenis kemampuan alias dimensi proses kognitifnya.

Validasi konstruk secara khusus cocok diterapkan pada proses validasi atau pemeriksaan validitas tes yang mengukur konstruk psikologis. Konstruk psikologis merupakan suatu atribut psikologis atau aspek kepribadian orang. Namun berbeda dengan atribut atau aspek fisik yang bersifat konkret seperti warna kulit, warna rambut, bentuk hidung, tinggi badan dan sebagainya, konstruk psikologis bersifat abstrak. Selain abstrak, konstruk psikologis juga merupakan sesuatu yang diciptakan oleh ilmuwan psikologi atau pendidikan berdasarkan kemampuannya berimajinasi, bukan merupakan aspek fisik kepribadian yang bisa diisolasi dan bisa diamati secara langsung.

Konstruk psikologis bisa berupa kemampuan yang didominasi oleh fungsi kognitif. Jenis konstruk psikologis ini lazim dibedakan menjadi dua kategori besar, yaitu inteligensi atau kecerdasan umum dan berbagai *aptitude* alias bakat atau kemampuan khusus seperti bakat atau kemampuan di bidang bahasa, matematika, dan sebagainya. Dalam teori *multiple intelligences* keberadaan dan fungsi kecerdasan umum kurang dikedepankan, sebaliknya lebih ditonjolkan keberadaan dan fungsi aneka kecerdasan atau kemampuan khusus meliputi kecerdasan linguistik, kecerdasan matematis, kecerdasan

spasial, kecerdasan musikal, kecerdasan kinestetik, kecerdasan interpersonal, kecerdasan intrapersonal, dan kecerdasan naturalistik (Gardner, 1999, dalam Shermis & Di Viesta, 2011). Selain itu, pembatasan penggunaan istilah kecerdasan semata-mata untuk jenis kemampuan yang didominasi fungsi kognitif juga ditinggalkan, terbukti dari pengenalan konsep kecerdasan emosi. Namun dari sudut kepentingan penyusunan alat asesmen hasil belajar, pemilahan antara konstruk psikologis yang didominasi fungsi kognitif dan konstruk psikologis yang didominasi fungsi afektif kiranya tetap bermanfaat. Dari sisi ini kecerdasan emosi lebih tepat ditempatkan ke dalam kategori konstruk psikologis yang didominasi fungsi afektif.

Konstruk psikologis tentu saja juga bisa berupa kemampuan yang didominasi fungsi afektif. Jenis-jenis kemampuan atau lebih tepat kecenderungan kepribadian yang masuk dalam golongan konstruk psikologis yang didominasi fungsi afektif ini lazim disebut *personality traits* atau sifat kepribadian. Sifat kepribadian lazim dibedakan ke dalam empat kategori (Nunnally, Jr., 1974), meliputi: (1) *social traits* atau sifat sosial, yaitu cara khas bertingkah laku seseorang dalam setiap situasi sosial, atau dalam bahasa sehari-hari disebut “sifat” yang ragam dan jumlahnya sebanyak kata sifat yang terdapat dalam kamus bahasa ibu, seperti sifat jujur, setia, dan sebagainya; (2) *motif*, yaitu aneka “dorongan” atau “kebutuhan” yang bersifat nonbiologis atau sosial, seperti dorongan atau kebutuhan untuk berprestasi, untuk berafiliasi atau berteman, dan sebagainya; (3) *personal conceptions* atau konsepsi pribadi, yaitu cara seseorang berpikir tentang dirinya sendiri dan dunia sekelilingnya, seperti konsep diri, minat, sikap, dan nilai-nilai hidup; dan (4) *penyesuaian-diri* (dan lawannya, *salah-suai*), yaitu keadaan atau kecenderungan kepribadian terkait dengan sejauh mana seseorang relatif terbebas dari gangguan emosi atau gangguan tingkah laku, seperti kecemasan, stres, fobia, dan sebagainya.

Konstruk psikologis seperti diuraikan di atas, lazimnya tidak memiliki *content domain* atau ranah isi yang jelas. Kendati abstrak konstruk psikologis semacam itu dipostulasikan atau diyakini ada, namun hanya dapat diamati melalui indikator perilaku (*behavioral indicators*) khususnya melalui kinerja (*performance*) atau hasil seseorang dalam mengerjakan suatu tugas atau tes. Sebuah tes dinyatakan memiliki validitas konstruk yang baik antara lain jika tes itu terbukti memiliki struktur internal yang jelas dan hasilnya sejalan atau sesuai dengan yang diimajinasikan oleh inventor atau pengembangnya, dan tentu saja sesuai dengan imajinasi masyarakat luas alias akal sehat.

Validasi kriteria atau pemeriksaan validitas kriteria cocok diterapkan untuk semua tes. Salah satu fungsi utama tes adalah dipakai sebagai dasar untuk membuat prediksi tentang tingkah laku nyata tertentu. Tingkah laku yang diprediksikan tersebut bisa bersifat spesifik, misal *performance* atau hasil kinerja dalam sebuah tes lain, atau bisa bersifat luas dan kompleks, misal keberhasilan menyelesaikan program pendidikan dalam suatu program studi di perguruan tinggi atau keberhasilan menjalankan tugas jabatan tertentu misal sebagai guru atau pilot.

Bentuk tingkah laku baik yang bersifat spesifik atau kompleks yang diprediksikan disebut *criterion* atau kriteria, sedangkan tes yang dipakai sebagai dasar prediksi disebut *prediktor*. Dalam validasi kriteria, pengukuran atau pengumpulan data tingkah laku yang dijadikan sebagai kriteria dapat dilakukan *bersamaan waktu* atau *berselang waktu* dengan pengukuran atau pengumpulan data prediktornya. Selang waktu tersebut bisa singkat misal dalam hitungan hari, atau bisa juga panjang misal dalam hitungan tahun. Pada kasus pertama, jenis validitas kriteria yang diperoleh disebut *concurrent validity*. Pada kasus kedua, jenis validitas kriteria yang

diperoleh disebut *predictive validity*. Validitas kriteria secara khas ditunjukkan oleh koefisien korelasi yang signifikan antara prediktor dan kriterionnya. Koefisien korelasi signifikan ini bisa positif atau negatif, tergantung landasan konseptual atau teoretisnya. Sebagai contoh, hasil tes kecerdasan yang tinggi sebagai prediktor harus berkorelasi positif dan signifikan dengan rerata nilai rapor yang tinggi. Sebaliknya, hasil tes kecemasan yang tinggi harus berkorelasi secara negatif dan signifikan dengan hasil ujian lisan, misalnya. Validitas kriteria, baik validitas konkuren maupun validitas prediktif, bisa dipakai sebagai salah satu evidensi atau bukti baik validitas isi maupun validitas konstruk suatu tes.

Pemeriksaan Validitas Isi, Validitas Konstruk, dan Validitas Kriteria

Pemeriksaan validitas isi lazim dilaksanakan pada awal penyusunan tes melalui penilaian rasional yang didasarkan keahlian (*rational expert judgment*). Pemeriksaan validitas konstruk suatu tes juga dilakukan pada awal penyusunan tes dengan sejenis penalaran rasional yang didasarkan keahlian juga, namun masih harus dilengkapi dengan pemeriksaan lain secara empiris sesudah tes itu selesai disusun dan dengan menggunakan teknik statistik tertentu. Pemeriksaan validitas kriteria dilakukan secara empiris sesudah tes selesai disusun, prinsipnya dengan memeriksa korelasi atau hubungan antara tes itu dengan ukuran kinerja atau variabel lain tertentu yang ditetapkan sebagai kriteria.

Sarana atau instrumen yang diperlukan dalam rangka validasi yang dilakukan pada awal penyusunan tes ada tiga macam, tergantung dari jenis validasi yang dikerjakan. Sebagaimana sudah kita lihat, jenis validasi ini sendiri juga ditentukan oleh *test domain* atau ranah tes. Ranah tes adalah *fokus tes* (Friedenberg, 1995). Sebagaimana

sebagian sudah kita lihat, secara garis besar ranah tes bisa dibedakan menjadi tiga kategori, yaitu *ranah pengetahuan* atau *keampilan*, *ranah konstruk*, dan *ranah tingkah laku*. Dalam penyusunan tes, ranah tes ini selanjutnya akan menentukan materi tes, yaitu jenis tugas atau item yang harus dikerjakan oleh orang yang dites dalam rangka mengukur suatu atribut psikologis tertentu di dalam dirinya.

Sesuai ranah atau fokus tes, ada tiga macam sarana atau instrumen validasi tes, yaitu (Friedenberg, 1995): (1) *test plan* atau kisi-kisi jika ranah atau fokus tes berupa pengetahuan atau ketrampilan; (2) *eksplikasi konstruk* jika ranah atau fokus tes berupa konstruk, yaitu dimensi kepribadian tertentu yang bersifat hipotetis dan yang dipandang membedakan orang yang satu dari yang lain; dan (3) *analisis tugas*, jika ranah atau fokus tes berupa perilaku atau pola perilaku. Penyusunan masing-masing sarana validasi tersebut akan dibahas dengan ringkas dan secara berturut-turut dalam bagian di bawah ini.

Menyusun Kisi-kisi

Jika tes yang disusun bertujuan mengukur pengetahuan dan atau ketrampilan dalam bidang studi atau pelajaran tertentu, validitas isi tes diusahakan melalui penyusunan *test plan* atau kisi-kisi untuk memastikan bahwa isi tes merepresentasikan cakupan isi pengetahuan dan atau ketrampilan dalam bidang studi atau pelajaran yang bersangkutan. Kisi-kisi adalah “*a written list of the information to be covered by the test items and the behaviors required to answer the questions correctly*” (Friedenberg, 1995). Artinya, kisi-kisi merupakan daftar tertulis tentang informasi, yaitu pengetahuan dan atau ketrampilan dalam bidang studi atau pelajaran tertentu, yang harus dicakup oleh item-item tes serta jenis-jenis perilaku yang dituntut untuk menjawab item-item atau pertanyaan-pertanyaan secara tepat.

Penyusunan kisi-kisi lazim didasarkan pada taksonomi tujuan pengajaran tertentu. Seperti sudah kita lihat, salah satu taksonomi tujuan pengajaran yang paling luas dijadikan pedoman dalam pengajaran di sekolah adalah taksonomi tujuan pengajaran ranah kognitif yang disusun oleh Bloom dan kawan-kawan. Jika mengikuti taksonomi Bloom ranah kognitif sebelum mengalami revisi, maka secara garis besar kisi-kisi tes hasil belajar suatu bidang studi atau pelajaran tertentu akan memiliki dua komponen: (1) *content objectives* atau sasaran materi, dan (2) *behavioral objectives* atau sasaran perilaku.

Sasaran materi menentukan jenis informasi/pengetahuan dan/atau ketrampilan yang harus dicakup tes berupa daftar topik dan subtopik tentang materi bidang studi atau pelajaran yang bersangkutan. Sasaran perilaku menentukan jenis respon atau proses kognitif yang harus ditunjukkan oleh testi mengacu pada taksonomi tujuan pengajaran tertentu, dalam contoh ini taksonomi tujuan pengajaran ranah kognitif Bloom dan kawan-kawan sebelum direvisi. Seperti kita ketahui, jenis respon atau proses kognitif yang dimaksud meliputi pengetahuan, pemahaman, penerapan, analisis, sintesis, dan evaluasi.

Jika kita mengacu pada taksonomi ranah kognitif dari Bloom dan kawan-kawan yang sudah direvisi, sasaran materinya perlu dipilah-pilah ke dalam empat kategori pengetahuan, yaitu pengetahuan faktual, pengetahuan konseptual, pengetahuan prosedural, dan pengetahuan meta-kognitif. Sedangkan *dimensi proses kognitif*-nya perlu disesuaikan dengan klasifikasi yang baru, yaitu mengetahui, memahami, menerapkan, menganalisis, mengevaluasi, dan menciptakan.

Menyusun Eksplikasi Konstruk

Ada tes yang bertujuan mengukur konstruk psikologis, yaitu dimensi atau atribut kepribadian tertentu bisa berupa kemampuan maupun berupa sifat atau kecenderungan kepribadian. Konstruk psikologis semacam ini adalah hasil pemikiran ahli sehingga disebut konstruk hipotetis atau konstruk teoretis. Terkait pemilikan dimensi atau atribut kepribadian tertentu yang merupakan konstruk psikologis, diasumsikan setiap orang memiliki *individual differences* atau kekhususan individual. Artinya, terkait pemilikan dimensi atau atribut kepribadian tertentu, setiap orang akan berbeda dari orang lain.

Konstruk psikologis bersifat abstrak dan tidak bisa diukur secara langsung. Ada atau tidak adanya atribut kepribadian tertentu yang merupakan konstruk atau lebih tepat, seberapa besar kadar keberadaan atribut kepribadian tertentu yang merupakan konstruk dalam diri seseorang harus diinferensikan atau disimpulkan dari hasil pengukuran bentuk-bentuk perilaku spesifik tertentu yang dilakukan oleh orang yang bersangkutan. Bentuk-bentuk tingkah laku spesifik tertentu ini dipandang sebagai indikator atau pencerminan keberadaan konstruk yang bersangkutan. Cara mengidentifikasi bentuk-bentuk atau jenis-jenis tingkah laku yang dipandang mencerminkan keberadaan atau sebaliknya ketidakhadiran atribut kepribadian tertentu yang merupakan konstruk disebut *eksplikasi konstruk*.

Dalam eksplikasi konstruk, “*the test developer compiles a list of specific behaviors, beliefs, and attitudes that demonstrate the presence of the construct and specific behaviors, beliefs, and attitudes inconsistent with its presence*” (Friedenberg, 1995). Artinya, dalam eksplikasi konstruk penyusun tes mengumpulkan daftar perilaku, keyakinan, dan sikap yang mencerminkan keberadaan atau kehadiran konstruk

yang diukur (*favorable*), serta perilaku, keyakinan, dan sikap yang bertentangan atau mencerminkan ketiadaan konstruk yang diukur (*unfavorable*).

Daftar kategori tingkah laku spesifik baik yang *favorable* maupun yang *unfavorable* terhadap konstruk psikologis yang menjadi objek atau sasaran pengukuran ini selanjutnya dipakai sebagai dasar atau pedoman untuk menyusun item-item tes. Maka tes yang mengukur konstruk psikologis semacam ini lazim akan mencakup item-item dengan isi yang *favorable* atau sejalan, serta item-item dengan isi yang *unfavorable* atau bertentangan, dengan konstruk psikologis seperti sikap-minat-nilai atau aspek kepribadian lain yang sedang diukur.

Menyusun Analisis Tugas

Jika tes dimaksudkan untuk mengukur pola perilaku yang kompleks, maka penentuan isi tesnya dilakukan lewat *task analysis* atau *job analysis* alias analisis tugas (Friedenberg, 1995). Dalam analisis tersebut, tugas atau pekerjaan yang menjadi sasaran pengukuran diuraikan ke dalam komponen-komponennya yang spesifik. Komponen-komponen inilah yang akan menjadi isi tes. Masing-masing komponen selanjutnya bisa diuraikan lebih lanjut ke dalam rangkaian perilaku sehingga masing-masing perilaku bisa dijadikan item-item tes. Analisis tugas perlu dilakukan untuk menyusun tes hasil belajar tentang kemampuan tertentu yang terdiri atas pola perilaku kompleks melibatkan baik ranah kognitif, afektif, maupun psikomotor, misalnya berpidato di depan kelas.

Fokus kita tentu saja adalah teknik atau pendekatan nontes. Gagasan tentang pentingnya memastikan validitas tes sebelum menggunakannya dalam penilaian hasil belajar tetap relevan untuk kita pegang sebagai pedoman. Namun seperti akan kita lihat, paling tidak langkah-langkah dalam memastikan validitas alat teknik

nontes akan sedikit berbeda dari langkah-langkah dalam memastikan validitas alat teknik tes sebagaimana sudah diuraikan, kendati prinsip-prinsipnya sama.

Bab 4

Pengembangan Alat Penilaian Hasil Belajar: Teknik Nontes

Sebenarnya, perbedaan alat asesmen atau penilaian hasil belajar secara hitam-putih antara tes dan nontes sama menyatakannya dengan perbedaan membabi-butakan bahwa tes lebih cocok untuk mengumpulkan data hasil belajar ranah kognitif sedangkan nontes lebih cocok untuk mengumpulkan data hasil belajar ranah afektif, sebagaimana lazim berlangsung di kalangan guru di semua jenjang pendidikan sekolah bahkan juga di kalangan dosen di perguruan tinggi. Dalam kenyataan terjadi tumpang-tindih yang wajar antara tes dan nontes, sesuai sifat dan kompleksitas ranah hasil belajar yang hendak dinilai, baik kognitif, afektif, psikomotorik secara terpisah maupun sebagai kombinasi antara ketiganya.

Untuk membahas alat asesmen nontes, kita akan meminjam sebuah sistematika cukup logis yang dikembangkan oleh Chatterji (2003). Pakar penilaian ini mengklasifikasikan jenis-jenis alat asesmen berdasarkan *mode of response* atau *jenis respon* yang dituntut dari murid. Pembagian ini dengan sendirinya tidak secara eksklusif hanya berlaku bagi jenis alat nontes, namun kita memang hanya akan berfokus pada jenis alat itu sesuai tema pembahasan kita.

Chatterji (2003) membedakan lima *modes of response* atau *cara memberikan respon* yang lazim diterapkan dalam pengerjaan tugas-tugas dalam rangka penilaian hasil belajar, yaitu: (1) respon *tertulis*, (2) respon berupa *tingkah laku* atau proses, (3) respon berupa *produk*, (4) respon *wawancara*, dan (5) respon berupa *portofolio*. Berdasarkan lima jenis cara respon tersebut, dia membedakan lima jenis penilaian hasil belajar nontes sebagaimana diuraikan di bawah ini.

Penilaian Tertulis

Di sini murid harus mengerjakan tugas-tugas secara tertulis, baik dalam format *paper and pencil* atau manual, maupun dengan menggunakan komputer. Ciri utama jenis penilaian ini adalah sifat tertulisnya, jadi harus mengandalkan kemampuan verbal atau berbahasa, khususnya membaca dan menulis.

Jenis penilaian tertulis nontes yang sangat lazim adalah *kuesioner* atau daftar pertanyaan. Dari segi isi informasi yang dikumpulkan, *kuesioner* dibedakan menjadi dua, yaitu *kuesioner* yang bertujuan mengungkap *fakta* dan *kuesioner* yang bertujuan mengungkap *pendapat* atau jenis-jenis tanggapan pribadi lainnya, seperti pikiran, penalaran, tanggapan perasaan dan sikap.

Dari segi format atau *sifat* atau *kadar keterstrukturannya* jawaban yang diminta, secara garis besar *kuesioner* dibedakan menjadi *kuesioner* (dengan jawaban) *tertutup* dan *kuesioner* (dengan jawaban) *terbuka*. *Kuesioner* *tertutup* lazimnya dipakai untuk menjangkau informasi berupa fakta. Jawaban yang berisi kemungkinan fakta yang diharapkan itu bisa sudah disediakan atau disajikan sehingga murid tinggal memilih dengan cara memberi tanda centang, tanda silang, melingkari, atau menghitamkan salah satu jawaban yang paling sesuai dengan keadaan dirinya, atau harus mengisikannya sendiri dengan cara menuliskan jawabannya pada ruang yang disediakan.

Pertanyaan dalam kuesioner tertutup seringkali juga disajikan dalam bentuk tabel. Jawaban yang diminta dituliskan dalam kolom-kolom yang disediakan.

Kuesioner terbuka lazimnya dipakai untuk menjangkau informasi berupa pendapat, pikiran, penalaran, tanggapan perasaan, sikap, kesan atau jenis-jenis ungkapan pribadi lainnya. Murid diberi kebebasan untuk mengungkapkan dirinya dengan cara menuliskannya pada ruang cukup leluasa yang disediakan, bahkan jika perlu bisa menggunakan lembar tambahan baik yang disediakan oleh guru maupun yang boleh diusahakan sendiri oleh murid. Contoh-contohnya adalah seperti disajikan di bawah ini (Chatterji, 2003):

1. Kuesioner Tertutup: Mengungkap Informasi Faktual

a. *Keterangan diri*

- Nama :
- Tempat lahir :
- Tanggal lahir :
- Jenis kelamin : () Perempuan () Laki-laki
- Golongan darah :
- Tinggi badan : cm
- Berat badan : kg
- Warna kulit : () Kuning () Sawo matang
() Hitam
- Jenis rambut : () Lurus () Ikal () Keriting
- Ukuran baju : S M L XL
- Ukuran celana :
- Ukuran sepatu :
- Dsb.

b. *Riwayat pendidikan*

No.	Jenjang Pendidikan	Jenis Sekolah	Nama Sekolah	Tempat	Tahun Lulus	Gelar
1	Dasar					NA
						NA
2	Menengah					NA
3	Tinggi	D-I				
		D-II				
		D-III				
		D-IV/S1				
		S-2				
		S-3				

2. **Kuesioner terbuka**

- a. Tuliskanlah pendapat Anda terhadap gagasan untuk menghapuskan pemberian remisi hukuman terhadap terpidana kejahatan korupsi.

.....

.....

.....

.....

- b. Tuliskanlah sikap Anda terhadap gagasan untuk mensahkan aborsi. Berilah alasan-alasan untuk mendukung sikap Anda.

.....

.....

.....

.....

- c. Uraikanlah gagasan Anda, mengapa kita perlu menjaga kebersihan lingkungan.

.....

.....

.....

.....

Penilaian Berbasis Perilaku

Di sini murid dituntut mendemonstrasikan aneka tingkah laku atau aneka proses sebagai hasil belajarnya, yang bisa diamati secara langsung. Ada tiga ciri utama jenis penilaian ini (Chatterji, 2003), yaitu: (a) bersifat *live* atau *on the spot*, dalam arti bahwa penilaian dilakukan saat aneka tingkah laku, kinerja, atau demonstrasi itu sedang berlangsung, (b) menggunakan format jawaban terbuka, dalam arti murid diberikan kebebasan penuh dalam memberikan jawaban atau respon, dan (c) mengandalkan metode observasi atau pengamatan. Ada satu kualifikasi lain yang lazim diterapkan, yaitu apakah observasi atau penilaian tersebut *timed* atau diberi pembatasan durasi atau waktu, atau *untimed* alias tanpa pembatasan durasi atau waktu.

Selain itu, karena jenis penilaian ini pada dasarnya mengandalkan metode observasi atau pengamatan, maka lazimnya dibedakan menjadi (a) pengamatan terstruktur, dan (b) pengamatan tidak terstruktur atau lazim disebut *pengamatan naturalistik* (Chatterji, 2003). Ada teknik-teknik khusus yang lazim diterapkan untuk melaksanakan masing-masing jenis observasi atau pengamatan. Pengamatan terstruktur lazim dilaksanakan dengan bantuan teknik *check-list* atau daftar cek dan *rating scales* atau skala penilaian (Sudjana, 2010). Pengamatan naturalistik lazimnya dilaksanakan dengan bantuan teknik *anecdotal records* atau catatan anekdot yang bersifat insidental.

1. Pengamatan terstruktur

Dengan bantuan Daftar Cek. Daftar cek pada dasarnya merupakan daftar tingkah laku sebagai sasaran pengamatan, untuk mengecek apakah masing-masing tingkah laku yang tercantum dalam daftar muncul atau ditemukan (*Ada* atau *Ya*) atau tidak muncul alias tidak ditemukan (*Tidak Ada* atau *Tidak*) selama pengamatan berlangsung. Hasil pengamatan tersebut dinyatakan dengan

memberikan *tanda cek* (V) pada kolom yang sesuai di belakang masing-masing tingkah laku. Maka, lazimnya format Daftar Cek akan berupa matriks yang terdiri atas minimal 4 kolom (berisi nomor, bentuk tingkah laku, *Ada/Ya*, dan *Tidak Ada/Tidak*) dan baris-baris sebanyak jenis atau bentuk tingkah laku yang diharapkan muncul.

Tabel 4.1: Contoh Daftar Cek: Tingkah Laku di Kelas

Kelas :
 Nama Murid :
 Mata Pelajaran :
 Nama Guru :

No.	Bentuk Tingkah Laku	Ya	Tidak
1	Memperhatikan penjelasan guru.	V	
2	Membuat catatan.	V	
3	Menyimak buku pelajaran.	V	
4	Mengajukan pertanyaan.		V
5	Dsb.		

Yang perlu diperhatikan, *arah* atau *isi* masing-masing bentuk tingkah laku sebaiknya sama, sehingga item-item dalam Daftar Cek itu menjadi *homogen* atau *tidak heterogen*. Jika item-item *homogen*, jika perlu setiap tanda cek di belakang masing-masing item atau bentuk tingkah laku pada kolom *Ya* bisa diberi skor 1, sedangkan setiap tanda cek pada kolom *Tidak* bisa diberi skor 0, sehingga jumlah skor pada kolom *Ya* bisa mencerminkan "kualitas" tingkah laku di kelas masing-masing siswa yang diamati. Sehingga, misalnya, bisa ditarik kesimpulan bahwa makin tinggi skor total seorang murid maka makin "baik" pula tingkah lakunya di kelas. Namun, kesimpulan semacam ini tetap harus didasarkan pada pemeriksaan yang jeli terhadap isi masing-masing item atau bentuk tingkah lakunya, apakah semuanya memang bermakna positif alias baik.

Dengan bantuan Rating Scales atau Skala Penilaian. Format skala penilaian pada dasarnya mirip daftar cek. Bedanya, yang dinilai atau diamati bukan hanya ada atau tidaknya bentuk tingkah laku tertentu, melainkan jika ada seberapa tinggi kadarnya. Tinggi-rendah kadar tingkah laku itu bisa berkenaan dengan kualitas baik-buruknya (*quality scale*) atau bisa berkenaan dengan frekuensi atau jumlah kali kemunculannya (*frequency scale*).

Tabel 4.2: Contoh Skala Penilaian Kualitas

Kelas :
 Nama Murid :
 Mata Pelajaran : Bahasa Indonesia
 Nama Guru :

No	Bentuk Tingkah Laku	Baik	Sedang	Kurang
1	Ucapan dalam membaca.			
2	Kepekaan pada tanda baca dalam membaca.			
3	Kelancaran dalam membaca.			
4	Kejelasan pembedaan huruf besar dan kecil dalam menulis.			
5	Dsb.			

Tabel 4.3:

Contoh Skala Penilaian Frekuensi

Kelas :
 Nama Murid :
 Mata Pelajaran : Bahasa Indonesia
 Nama Guru :

No	Bentuk Tingkah Laku	Sering	Jarang	Tidak Pernah
1	Keliru mengucapkan bunyi huruf dalam membaca.			
2	Keliru menafsirkan tanda baca dalam membaca.			
3	Mengulang-ulang kata atau kumpulan kata dalam membaca.			
4	Mengacaukan cara penulisan huruf besar dan kecil.			
5	Dsb.			

Seperti dalam *daftar cek, arah* atau *isi* masing-masing bentuk tingkah laku sebaiknya sama, sehingga item-item dalam Skala Penilaian baik penilaian kualitas maupun penilaian frekuensi menjadi *homogen* atau *tidak heterogen*. Jika item-item *homogen*, jika perlu setiap tanda cek di belakang masing-masing item atau bentuk tingkah laku pada masing-masing kolom bisa diberikan skor yang berlainan, misal untuk kedua contoh di atas sebagai berikut:

Baik	= 2	Sering	= 2
Sedang	= 1	Jarang	= 1
Kurang	= 0	Tidak Pernah	= 0

Dengan diberi skor, jumlah total skor pada ketiga kolom bisa ditafsirkan mencerminkan "kualitas" kemampuan dalam pelajaran Bahasa Indonesia masing-masing siswa yang diamati. Sehingga, misalnya, bisa ditarik kesimpulan bahwa makin tinggi skor total seorang murid maka makin "baik" pula kemampuannya dalam pelajaran Bahasa Indonesia (pada contoh Skala Penilaian Kualitas), atau sebaliknya (pada contoh Skala Penilaian Frekuensi). Namun, kesimpulan semacam ini tetap harus didasarkan pada pemeriksaan yang jeli terhadap isi masing-masing item atau bentuk tingkah lakunya, apakah semuanya memang bermakna positif alias baik, atau sebaliknya bermakna negatif atau buruk.

2. Pengamatan naturalistik

Dalam pengamatan naturalistik atau tidak terstruktur, guru sebagai pengamat mengamati tingkah laku murid secara *live* atau *on the spot*, yaitu dalam situasi aktual atau nyata dengan hanya sedikit atau bahkan sama sekali tanpa campur tangan pengamat serta bebas dari berbagai faktor atau kendala eksternal. Tanpa campur tangan mencakup pengertian bahwa pengamat tidak secara eksplisit

mempersiapkan daftar jenis tingkah laku yang akan menjadi sasaran pengamatannya. Pengamat tetap memiliki tujuan tertentu dalam melakukan pengamatan, namun tujuan itu tidak dijabarkannya ke dalam daftar terstruktur yang akan dipakainya sebagai pedoman seperti yang dilakukannya dalam pengamatan terstruktur dengan daftar cek maupun skala penilaian. Salah satu teknik khas pengamatan naturalistik atau tidak terstruktur adalah *anecdotal records* atau catatan anekdot (Chatterji, 2003).

Catatan anekdot adalah deskripsi atau catatan rekaman tentang episode-episode atau peristiwa-peristiwa yang berlangsung dalam situasi natural alias wajar atau alamiah. Lazimnya pencatatan peristiwa ini difokuskan pada seseorang murid yang sedang menjadi perhatian guru, sehingga himpunan dari catatan-catatan anekdot semacam ini akan memberikan deskripsi atau gambaran tentang pola tingkah laku murid yang bersangkutan (Chatterji, 2003).

Mengutip pendapat Mehrens dan Lehman (1984), Chatterji (2003) menyatakan bahwa catatan anekdot yang baik harus memiliki ciri-ciri sebagai berikut:

1. Berupa deskripsi singkat peristiwa faktual, misal "Murid berbisik kepada teman yang duduk sebangku dengannya."
2. Catatan tersebut tidak boleh mengandung inferensi atau kesimpulan, pendapat, atau penilaian dari pihak pengamat. Misal, contoh catatan di atas tidak boleh berbunyi sebagai berikut, "Murid berbisik kepada teman yang duduk sebangku, seperti bertanya tentang jawaban soal yang sedang ditugaskan oleh guru," atau "Murid berbisik kepada teman yang duduk sebangku, menguatkan kesan bahwa dia memang tidak bisa diam."
3. Catatan tersebut harus berisi rekaman tentang *critical incident* atau kejadian penting terkait si murid. Penentuan nilai

penting atau kurang pentingnya suatu peristiwa ditentukan oleh tujuan pengamatan. Kembali pada contoh kita, misal guru ingin mengamati tingkah laku baik di dalam maupun di luar kelas dari seorang murid yang oleh orang tuanya dikeluhkan sangat pendiam dan sulit bergaul. Dengan tujuan untuk menemukan fakta-fakta untuk membuktikan keluhan orang tua murid tersebut, maka peristiwa "murid berbisik kepada temannya sebangku" menjadi penting untuk direkam oleh guru dalam catatan anekdotnya, sebab peristiwa ini merupakan sebuah fakta yang tidak sejalan dengan keluhan orang tua. Tetapi kemudian, misalnya, guru melanjutkan pengamatannya saat istirahat. Tatkala hampir semua murid lain bergembira ria bermain bersama selama lima belas menit waktu istirahat, murid yang menjadi sasaran pengamatan itu ternyata "memilih duduk menyendiri sambil membaca." Terkait tujuan seperti sudah dikemukakan, peristiwa ini tentu juga memiliki makna penting sebagai fakta yang mendukung keluhan orang tua.

4. Hanya sesudah memperoleh rekaman peristiwa dalam jumlah yang dipandang memadai, pengamat boleh membuat kesimpulan tentang adanya pola perilaku tertentu pada subjek yang menjadi sasaran pengamatan. Pada contoh kita di atas, sesudah guru memiliki rekaman tentang berbagai peristiwa terkait subjek sasaran yang dikumpulkan selama sehari-hari dan dalam berbagai situasi yang berlainan seperti di dalam kelas saat mengikuti pelajaran biasa, di dalam kelas saat berlangsung kerja kelompok, di dalam kelas saat mengerjakan tugas presentasi di depan kelas, di luar kelas saat pelajaran olah raga, di luar kelas saat mengerjakan tugas kerja bakti membersihkan halaman sekolah, di luar kelas

saat mengikuti kegiatan upacara bendera, dan sebagainya, guru sebagai pengamat sampai pada kesimpulan, misalnya, sebagai berikut: "Murid itu bisa berkomunikasi dengan baik dengan temannya saat menghadapi situasi yang jelas, seperti ketika sedang mengerjakan tugas kelompok, namun lebih suka menghabiskan waktu sendiri mengerjakan pekerjaan yang disukainya dalam situasi bebas, seperti saat istirahat. Dengan demikian bisa disimpulkan, mungkin tepat untuk mengatakan murid itu cenderung pendiam, namun kiranya kurang tepat menyebutnya kurang bisa bergaul, bahkan murid itu terkesan lebih matang dibandingkan teman-teman seusianya."

Contoh yang dikemukakan pada awal uraian bagian ini merupakan contoh peristiwa yang sederhana, yaitu "Murid berbisik kepada teman yang duduk sebangku dengannya." Dalam praktek, catatan anekdot bisa berupa rekaman peristiwa yang cukup kompleks, sekalipun syarat singkat tetap harus terpenuhi. Berikut ini adalah contoh catatan anekdot yang dikemukakan oleh Chatterji (2003).

Contoh Catatan Anekdote

Kelas :
Nama murid : Bambang
Catatan ke : 3
Tanggal : 18 Mei 1992
Waktu/pukul : 09.00
Tempat : Di dalam kelas

Bambang terlambat masuk kelas selama 15 menit. Dia menjelaskan alasan keterlambatannya dengan mengatakan bahwa dia harus menunggu lama giliran mendapatkan pesanan makanan untuk sarapannya di kantin sekolah. Sesudah duduk, dia segera mengeluarkan

buku dan alat tulisnya dan segera mengerjakan tugas yang diberikan oleh guru. Sesudah menyelesaikan 5 dari 8 soal matematika yang ditugaskan oleh guru, dia meletakkan kepalanya di atas meja. Dalam beberapa menit kemudian, dia tertidur.

Jelas kiranya, untuk bisa membuat sesuatu kesimpulan tentang diri Bambang guru membutuhkan tambahan hasil pengamatan yang jauh lebih banyak dan yang diperoleh dari berbagai situasi yang berbeda-beda pula.

Penilaian Berbasis Produk

Metode penilaian ini menuntut murid menciptakan atau mengonstruksikan produk atau hasil karya tertentu. Produk atau hasil karya ini selanjutnya dipakai sebagai dasar pengukuran dan penilaian. Beberapa contoh produk atau hasil karya misalnya adalah buku harian, makalah atau karangan, laporan eksperimen di laboratorium, laporan pengerjaan tugas IPA, karya seni atau kerajinan, dan sebagainya. Penilaian berbasis produk lazimnya memiliki ciri-ciri sebagai berikut (Chatterji, 2003):

1. Bersifat *open-ended* atau terbuka. Murid diberi kebebasan menciptakan produk atau hasil karyanya dengan diberi sedikit parameter atau kriteria sebagai persyaratan atau tuntutan yang sekaligus berfungsi sebagai pedoman bagi murid dalam melaksanakan tugas menciptakan produk atau hasil karya tertentu. Namun, seperti juga dinyatakan oleh Chatterji (2003), adakalanya tugas menciptakan produk ini bersifat tertutup dalam rupa satu tugas tertentu yang berlaku bagi semua murid. Misal, untuk mengukur kemampuan murid mengetik di atas komputer dengan sepuluh jari, guru memilih sebuah naskah pendek sebagai materi. Semua murid diminta

mengetik naskah tersebut dengan diberi pembatasan waktu, hasilnya dicetak untuk selanjutnya dinilai oleh guru terkait aspek ketepatan, kecepatan, dan kerapian, misalnya. Dalam lebih banyak kasus lain, tugas menciptakan produk tersebut bersifat terbuka. Maka, sifat berikutnya adalah:

2. Disertai pemberian waktu yang longgar untuk mengerjakan tugas menciptakan produk atau hasil karya ini, seringkali bahkan berupa tugas yang harus dikerjakan di rumah (*take-home exercises*).
3. Produk atau hasil karya yang berhasil diciptakan oleh murid dipandang sebagai evidensi atau bukti terhadap kepemilikan murid atas suatu jenis kemampuan tertentu yang lazimnya bersifat kompleks, sebagai hasil belajar.

Contoh: laporan tugas IPA kelas V sekolah dasar (Chatterji, 2003)

Perintah:

Tulislah sebuah laporan berisi perbandingan antara dua dari antara sembilan planet di dalam Tata Surya, yang boleh kamu pilih sendiri. Perbandingan tersebut harus mencakup hal-hal sebagai berikut: nama masing-masing planet, waktu yang dibutuhkan oleh masing-masing planet untuk berotasi pada porosnya, waktu yang dibutuhkan oleh masing-masing planet untuk mengelilingi matahari, dua fakta tentang bulan atau bulan-bulan dari masing-masing planet, rerata temperatur masing-masing planet, ukuran masing-masing planet, rerata jarak masing-masing planet dari matahari, nama julukan masing-masing planet, dan apakah ada kehidupan di masing-masing planet sejauh yang kita ketahui hingga kini. Dalam membuat laporan, tunjukkanlah perbedaan antara fakta dan pendapat pribadimu. Berikanlah alasan atau alasan-alasan untuk mendukung setiap pendapatmu. Isi laporan harus akurat dan jelas, serta panjang laporan tidak boleh lebih dari dua halaman kertas buku tulis atau satu lembar bolak-balik.

Waktumu 3 jam pelajaran (atau 3 x 45 menit) untuk mengumpulkan informasi dan menulis laporan. Kamu boleh menggunakan sumber-sumber buku pelajaran atau buku catatanmu, buku dan sumber lain yang ada di kelas, buku dan sumber lain di perpustakaan, atau sumber-sumber di Laboratorium IPA. Seluruh tugas harus kamu selesaikan di sekolah dalam waktu yang disediakan. Hasil pekerjaanmu harus dikumpulkan pada akhir jam pelajaran yang ketiga.

Laporan ini memiliki bobot 50% untuk menentukan nilai rapor mata pelajaran IPA. Saya akan menggunakan kriteria yang tercantum dalam daftar cek yang akan dibagikan untuk memberikan skor laporanmu. Ikutilah daftar cek tersebut sebagai pedoman dalam mengerjakan tugasmu.

Daftar Cek Kriteria Penskoran

No.	Unsur	Penyebutan dan Ketepatan	Beri tanda centang Ya atau Tidak (Masing-masing centang pada kolom Ya diberi skor 1)	
			Ya	Tidak
1.	Nama masing-masing planet	Disebutkan		
		Secara tepat		
2.	Julukan masing-masing planet	Disebutkan		
		Secara tepat		
3.	Ukuran masing-masing planet	Disebutkan		
		Secara tepat		
4.	Temperatur masing-masing planet	Disebutkan		
		Secara tepat		
5.	Waktu rotasi masing-masing planet	Disebutkan		
		Secara tepat		
6.	Jarak dari matahari masing-masing planet	Disebutkan		
		Secara tepat		
7.	Waktu mengelilingi matahari masing-masing planet	Disebutkan		
		Secara tepat		
8.	Kehidupan di masing-masing planet	Disebutkan		
		Secara tepat		

No.	Unsur	Penyebutan dan Ketepatan	Beri tanda centang <i>Ya</i> atau <i>Tidak</i> (Masing-masing centang pada kolom <i>Ya</i> diberi skor 1)	
			<i>Ya</i>	<i>Tidak</i>
9.	Bulan atau bulan-bulan dari masing-masing planet	Disebutkan		
		Secara tepat		
10.	Alasan atau alasan-alasan untuk pendapat yang dikemukakan	Disebutkan		
		Secara tepat		

Penilaian Berbasis Wawancara

Dalam *interview* atau wawancara, murid diminta memberikan jawaban secara lisan terhadap pertanyaan atau pertanyaan yang diajukan oleh guru atau pewawancara. Wawancara memiliki kemiripan dengan kuesioner dan observasi. Kemiripannya dengan kuesioner terletak pada format jawaban, yaitu bisa bersifat *open-ended* alias terbuka berupa jawaban bebas, atau bersifat terstruktur atau tertutup berupa jawaban *Ya* atau *Tidak*, atau memberikan jawaban lisan singkat. Keunggulan wawancara sebenarnya terletak pada pemberian kesempatan sebebaskan-bebasnya kepada murid untuk memberikan jawaban bisa berupa ungkapan pendapat, sikap, atau perasaan, masing-masing bisa disertai penjelasan maupun alasan-alasan. Kemiripannya dengan observasi, wawancara juga bisa dilakukan dengan atau tanpa pembatasan waktu (*timed* versus *untimed*).

Contoh:

Wawancara untuk menilai "locus of control"

Pertanyaan	"Look fors" atau Unsur-unsur yang dicari dalam Jawaban	Penskoran terhadap Unsur-unsur yang dicarai dalam Jawaban
<p><i>Pikirkanlah atau ingatlah sebuah PR atau tugas rumah yang baru kamu selesaikan belum lama ini.</i></p> <p>Pertanyaan: <i>Bagaimana cara kamu menilai bahwa kamu telah mengerjakan tugas itu secara berhasil?</i></p> <p>Pertanyaan pendalaman: <i>Apakah kamu merasa telah mencapai prestasi tertentu?</i> <i>Jelaskan alasan-alasan mengapa kamu merasa puas dengan hasil pekerjaanmu.</i></p>	<p>Hasil pekerjaan itu memenuhi standar mutu saya; memenuhi target waktu saya; memenuhi tujuan-sasaran saya.</p> <p>Orang-orang di rumah (ayah, ibu, kakak, dsb.) memuji hasil pekerjaan saya.</p> <p>Mereka puas dengan hasil pekerjaan saya.</p>	<p><i>Internal locus of control</i> (Kode 1)</p> <p><i>External locus of control</i> (Kode 2)</p>

Locus of control secara harfiah berarti "letak kendali" atau dalam terjemahan bebas berarti "kecenderungan menjelaskan peristiwa yang dialami" merupakan sebuah ciri kepribadian yang berdampak menentukan cara atau arah seseorang bertindak laku, khususnya sesudah mengalami pengalaman yang menggembirakan atau menyedihkan, yang digagas oleh seorang ahli psikologi bernama Julian B. Rotter (1966). Menurut teori ini, setiap orang berlainan dalam hal cara menjelaskan peristiwa-peristiwa yang dialaminya, yaitu cenderung menunjuk faktor-faktor di dalam (*internal locus of control*) atau faktor-faktor di luar (*external locus of control*) dirinya sebagai sumber atau penyebab. Kecenderungan tersebut tampak kentara saat seseorang mengalami peristiwa yang menggembirakan seperti keberhasilan atau sebaliknya mengalami peristiwa yang

menyedihkan seperti kegagalan. Orang yang *internal* cenderung menjelaskan sumber atau penyebab keberhasilan atau kegagalannya pada faktor-faktor yang terdapat di dalam dirinya sendiri, seperti kemampuan, motivasi, atau usahanya. Sebaliknya, orang yang *eksternal* cenderung menjelaskan sumber atau penyebab keberhasilan atau kegagalannya pada faktor-faktor yang terdapat di luar dirinya, seperti tugas yang dianggap mudah atau sukar, guru yang pemurah atau pelit memberikan nilai, atau faktor nasib atau keberuntungan belaka. Sebagaimana tersirat, masing-masing kecenderungan akan menentukan arah yang berbeda bagi yang bersangkutan dalam merespon keberhasilan atau kegagalan. Mengalami kegagalan, orang yang *internal* akan berusaha meningkatkan motivasi dan usaha belajarnya, sedangkan orang yang *eksternal* mungkin hanya akan bersedih dan justru kehilangan motivasi belajar karena merasa bahwa segala usahanya tidak membuahkan hasil seperti diharapkan. Pada kolom 2 dalam contoh di atas, diperoleh unsur-unsur dalam jawaban murid yang menjadi indikator tentang kecenderungannya untuk menjelaskan sumber pengalamannya terletak di dalam dirinya (*internal*) atau di luar dirinya (*eksternal*), seperti diidentifikasi dalam kolom 3.

Tidak jarang wawancara dikombinasikan dengan jenis penilaian lain, lazimnya dipakai sebagai pelengkap untuk mengungkap jenis-jenis informasi lain yang tidak atau kurang bisa diungkap dengan jenis penilaian kombinasinya itu. Salah satu contoh yang lazim adalah *viva voce examination* atau *oral examination* (Chatterji, 2003), yaitu ujian lisan secara individual atau kelompok kecil di hadapan satu atau lebih penguji untuk melengkapi ujian tertulis atau ujian praktik laboratorium.

Penilaian Berbasis Portofolio

Jenis penilaian ini dilaksanakan dengan cara “mengumpulkan secara terencana berbagai hasil karya atau rekaman tingkah laku yang bersama-sama memberikan gambaran yang komprehensif tentang pencapaian seseorang dalam sebuah bidang keahlian yang cukup luas” (Chatterji, 2003). Sebagai metode penilaian hasil belajar di sekolah, definisinya yang lebih lengkap sebagaimana dikutip oleh Chatterji (2010) dari ahli lain adalah sebagai berikut:

“A portfolio is a purposeful collection of student work that tells the story of a students’ efforts, progress, or achievement in (a) given area(s). This collection must include student participation in selection of portfolio contents; the guidelines for selections; the criteria for judging merit; and evidence of student self-reflection” (Arter & Spandel, 1992, dalam Chatterji, 2003, h. 94).

Artinya, portofolio merupakan sebuah kumpulan terencana dari hasil karya murid yang akan mampu berkisah tentang usaha, kemajuan, atau prestasi murid dalam suatu bidang pelajaran atau keahlian tertentu. Kumpulan ini harus mencakup keterlibatan murid dalam menyeleksi atau memilih materi portofolionya; pedoman untuk menyeleksi materi portofolio; kriteria untuk menentukan kualitas; dan bukti refleksi-diri yang dilakukan oleh murid yang bersangkutan.

Metode ini kini dipandang sebagai metode pendokumentasian yang komprehensif tentang perkembangan ketrampilan dan pengetahuan seseorang dalam suatu bidang keahlian tertentu, misal kompetensi sebagai seorang guru kelas Sekolah Dasar. Sejumlah kata kunci yang tercantum dalam definisi yang dikemukakan baik oleh Chatterji (2003) maupun oleh Arter dan Spandel (1992, dalam Chatterji, 2003) di atas adalah sebagai berikut.

Pertama, koleksi atau pengumpulan hasil karya tersebut dilakukan *secara terencana (purposeful)*, dalam arti bukti-bukti karya dalam bidang keahlian tertentu yang dihasilkan murid dan yang akan dimasukkan ke dalam koleksi dalam rangka menyusun portofolio harus benar-benar dipilih *secara selektif*. Kedua, agar seleksi atau pemilihan materi yang dimasukkan dalam portofolio tersebut benar-benar tepat, maka guru perlu membuatkan *pedoman* yang jelas. Ketiga, pemilihan materi tersebut tidak boleh dilakukan sepihak oleh guru, melainkan harus melibatkan murid yang bersangkutan. Keberadaan pedoman yang jelas akan sangat membantu murid dalam ikut memilih apa saja yang merupakan bukti usaha, kemajuan atau prestasinya dalam mempelajari keahlian dalam suatu bidang tertentu. Keempat, guru juga perlu membuatkan kriteria untuk menentukan kualitas masing-masing bukti karya yang dipilih sebagai materi dalam portofolio. Kelima, sebagai bentuk lain keterlibatan murid dalam menyusun portofolionya sendiri, maka dia juga harus menuliskan refleksi diri khususnya tentang proses dia sampai pada apa yang berhasil dicapainya, pelajaran yang diperoleh selama menjalani proses untuk mencapai prestasinya itu, serta mana bagian dari proses mencapai prestasinya itu dia rasakan paling sulit atau sebaliknya, yang paling menantang dan memberikan kepuasan batin baginya.

Contoh:

Portofolio untuk murid SD dalam mata pelajaran bahasa Indonesia

Petunjuk:

Kumpulkanlah satu contoh puisi hasil karyamu sendiri untuk masing-masing kategori puisi berikut ini:

1. pantun
2. sajak
3. gurindam
4. syair
5. soneta

Puisi-puisi hasil karyamu harus mengandung sarana-sarana puitik (poetic devices) berikut ini:

- a. *Personifikasi*
- b. *Onomatopoea*
- c. *Simili*
- d. *Metaphor*
- e. *Irama*

Uraikanlah proses yang kamu lakukan sejak menuliskan draft, mengedit, merevisi, dan akhirnya mencipta puisi dalam bentuk yang layak diterbitkan.

Kumpulkanlah semua hasil pekerjaanmu dalam portofolionmu.

Untuk melengkapi portofolionmu, tulislah sebuah karangan reflektif (kira-kira dua halaman kuarto, diketik dengan spasi rangkap), yang mengandung unsur-unsur:

- a. *Penjelasan tentang pada bagian mana saja kamu menggunakan sarana-sarana puitik dan alasan kamu menggunakan sarana-sarana tersebut.*
- b. *Penjelasan tentang pelajaran apa saja yang kamu petik dari tugas atau latihan ini.*
- c. *Bagian mana dari keseluruhan tugas ini yang kamu rasakan paling sulit maupun yang kamu rasakan paling menantang-menyenangkan.*

Jenis-jenis Portofolio

Portofolio pada dasarnya bertujuan memberikan representasi atau gambaran yang lengkap tentang kemampuan murid (Layton & Lock, 2008). Namun demikian, ada ahli yang membedakan portofolio menjadi dua kategori besar berdasarkan tujuannya, yaitu (1) *working portfolio* atau portofolio kerja, dan (2) *show portfolio* atau portofolio kinerja (Taylor, 2003, dalam Layton & Lock, 2008). Seperti tersirat dari namanya, portofolio kerja bertujuan mengumpulkan seluruh contoh hasil karya atau hasil pekerjaan harian murid selama mengerjakan tugas tertentu sejak awal hingga akhir. Portofolio kerja menjadi

sejenis dokumentasi lengkap hasil pekerjaan murid untuk setiap mata pelajaran selama satu semester, misalnya. Hasil-hasil pekerjaan yang memuaskan maupun yang tidak memuaskan dimasukkan atau dihimpun dalam portofolio, sehingga diperoleh rekaman lengkap tentang dinamika kemajuan belajar murid dalam masing-masing mata pelajaran, misalnya.

Portofolio kinerja bertujuan hanya mengumpulkan hasil-hasil karya atau hasil-hasil pekerjaan pilihan, terutama hasil-hasil pekerjaan yang memuaskan sehingga dipandang merepresentasikan prestasi murid dan bisa dipakai sebagai dasar untuk melakukan penilaian hasil belajar secara formal, misalnya.

Secara lebih rinci dan masih berdasarkan tujuan namun sekaligus merujuk jenis-jenis hasil karya atau hasil pekerjaan spesifik yang harus dijadikan isi masing-masing, Swicegood (1994, dalam Layton & Lock, 2008) membedakan portofolio menjadi lima jenis:

1. Untuk membuat semacam *exhibit* atau media unjuk kinerja tentang hasil karya terbaik murid dan perkembangannya.
2. Untuk memperoleh kumpulan hasil karya murid yang bersifat multidimensional atau mencakup seluruh kualitas yang pernah dicapainya, dalam jangka waktu tertentu.
3. Untuk mendokumentasikan gambaran nyata tentang rentang atau jenis-jenis kemampuan murid.
4. Untuk menciptakan sarana guna memudahkan dan meningkatkan kualitas refleksi guru dan murid tentang kemajuan belajar dan tingkat penguasaan murid atas materi atau kemampuan tertentu yang dipelajari.
5. Untuk meningkatkan dialog dan komunikasi antara murid dan guru serta pihak-pihak lain.

Menyangkut isi portofolio, Chang (2001, dalam Layton & Lock, 2008) membedakan lima jenis item penting yang perlu dimasukkan ke dalam portofolio, baik secara tradisional dalam bentuk *hardcopy* maupun secara elektronik dalam bentuk *softcopy*, yaitu:

1. Hasil-hasil belajar yang dipilih murid sendiri.
2. Hasil-hasil refleksi murid.
3. Tujuan-tujuan yang dirumuskan dengan jelas.
4. Sampel atau contoh hasil akhir sebuah karya (*completed work*).
5. Contoh-contoh nyata bukti terjadinya perkembangan murid.

Berdasarkan sifat hasil karya yang dikumpulkan atau yang dijadikan isi portofolio, Nitko (2004, dalam Layton & Lock, 2008) membedakan portofolio menjadi lima jenis:

1. ***Hasil karya terbaik***. Jenis ini berisi kumpulan contoh hasil pekerjaan atau aneka produk lain yang oleh murid dan guru dipandang merupakan indikator paling jelas yang menunjukkan tingkat penguasaan murid atas materi atau kemampuan yang sedang dipelajari. Sebagian unsur yang dimasukkan ke dalam portofolio bisa jadi sudah ditentukan sebelumnya sesuai kriteria penilaian. Selanjutnya murid diminta memilih contoh-contoh yang dia pandang paling baik mencerminkan kemajuan atau prestasi belajarnya. Melalui proses evaluasi-diri semacam ini, murid akan semakin memahami seperti apa indikator-indikator penguasaan materi yang diharapkan sehingga juga akan semakin trampil mengevaluasi aspek-aspek lain dari kinerja atau hasil belajarnya (Nitko, dalam Layton & Lock, 2008).
2. ***Perkembangan pendidikan murid***. Jenis portofolio ini berfokus pada berbagai proses belajar dan berpikir murid, termasuk aneka kesulitan yang mungkin dialaminya, dan bertujuan mengarahkan murid untuk menemukan berbagai strategi belajar yang baru. Jenis portofolio ini

berisi contoh-contoh hasil karya murid yang dikumpulkan pada periode waktu yang berlainan selama seluruh masa berlangsungnya proses belajar, meliputi contoh-contoh hasil pekerjaan pada masa *beginning* atau awal, *middle* atau tengah, dan *end* atau akhir, masing-masing mencakup hasil-hasil karya yang bersifat *beginning*, *intermediate*, dan *final* (Nitko, dalam Layton & Lock, 2008).

3. ***Hasil karya yang paling favorit.*** Jenis portofolio ini bertujuan mendokumentasikan kemampuan murid untuk melakukan evaluasi-diri, yaitu mengevaluasi kemajuan belajarnya sekaligus mendokumentasikan kemajuan kebiasaan kerjanya (*work habits*) selama terlibat dalam tugas menciptakan aneka produk yang mencerminkan bidang-bidang minatnya. Jenis portofolio ini berisi contoh-contoh hasil karya atau aneka produk lain, refleksi murid tentang masing-masing hasil karyanya, dan sebuah rubrik penilaian atau penskoran. Rubrik penskoran semacam ini perlu diberikan sejak awal pemberian tugas, agar murid bisa memilih item-item yang sungguh-sungguh mencerminkan tujuan pengajaran. Selain itu, keberadaan rubrik penskoran yang jelas juga akan membantu murid mengevaluasi setiap produk dengan cara yang benar (Layton & Lock, 2008).
4. ***Proses dalam pengerjaan proyek berjangka panjang.*** Jenis portofolio ini efektif untuk mengamati aneka ketrampilan yang dijadikan tujuan pengajaran dan perkembangannya dari waktu ke waktu selama proses pembelajaran berlangsung. Akan sangat bermanfaat jika guru bisa memberikan petunjuk baik berupa nasehat atau penilaian lisan agar murid mampu mencapai kemajuan belajar ke arah *mastery* atau penguasaan atas materi atau kemampuan yang dipelajari. Jenis portofolio ini bisa mendokumentasikan kemajuan bertahap (*incremental progress*) yang dialami murid, sekaligus memberikan informasi tentang

kemampuan berfikir murid pada taraf yang lebih tinggi (Layton & Lock, 2008).

5. ***Gambaran menyeluruh untuk dibandingkan dengan standar yang disusun sebelumnya.*** Jenis portofolio ini bertujuan membandingkan prestasi murid dengan sebuah kriteria ketrampilan akhir yang menjadi tujuan pembelajaran. Di sini guru dan murid bersama-sama mengumpulkan contoh-contoh hasil pekerjaan murid untuk dibandingkan dengan sebuah kriteria yang sudah dipilih sebelumnya. Murid juga diberi kesempatan untuk menilai sendiri contoh-contoh hasil pekerjaannya. Murid diharapkan memiliki pemahaman yang lebih baik tentang materi atau kemampuan yang dipelajari melalui proses refleksi yang juga harus dibuatnya (Layton & Lock, 2008).

Terkait jenis-jenis sifat hasil karya yang dijadikan isi portofolio ini, Nitko menyarankan tentang pentingnya membuat portofolio berisi aneka jenis hasil karya (*an assortment of portfolio styles*). Selain itu, mengutip pendapat Losardo dan Notari-Syverson (2001, dalam Layton & Lock, 2008) Layton dan Lock (2008) mengingatkan bahwa komponen penilaian portofolio bukan terletak pada materi yang berhasil dikumpulkan itu sendiri, melainkan pada penerapan sebuah *rubrik penskoran* atau *lembar penskoran* yang sudah disiapkan sebelumnya. Penilaian harus dilakukan terhadap seluruh materi yang terdapat di dalam portofolio serta harus dilakukan oleh guru dan murid bersama-sama dalam serangkaian pertemuan penilaian portofolio (*portfolio review conferences*) yang dilaksanakan secara terjadwal dan teratur. Di sini portofolio memberikan kesempatan kepada baik guru maupun murid untuk mengevaluasi kemajuan bersama-sama dengan cara melibatkan murid dalam proses penilaian melalui evaluasi dan pemeriksaan diri (*self-evaluation and review*).

Demikianlah sistematika pembagian teknik-teknik penilaian nontes yang dikemukakan oleh Chatterji (2003) dan yang cukup memudahkan kita memahami esensi dan karakteristik teknik penilaian nontes. Semoga pemahaman ini juga memudahkan kita dalam mengembangkan dan menyusun metode-metode penilaian nontes dalam penilaian hasil belajar murid di kelas.

Bab 5 Penyusunan Alat Penilaian Nontes

Alat penilaian nontes pada dasarnya merupakan anak kandung dari pembaharuan penilaian pendidikan yang dikenal sebagai *authentic assessment* atau penilaian otentik. Penilaian otentik mengandalkan penggunaan jenis-jenis tugas yang otentik (*authentic tasks*) untuk menilai hasil belajar murid. Tugas otentik adalah jenis tugas yang “*replicate the kinds of challenges that adults face in the workplace, in civic affairs, or in their personal lives*” (Wiggins, 1998, h. 22). Artinya, jenis tugas yang mereplikasikan atau memiliki kemiripan dengan jenis-jenis tantangan yang dihadapi oleh orang dewasa pada umumnya di tempat kerja, dalam berbagai urusan kehidupan bermasyarakat, atau dalam kehidupan pribadi mereka. Tugas otentik akan menuntut murid melakukan tugas atau aktivitas nyata (*real performance*) yang memiliki kemiripan sedekat mungkin dengan tugas-tugas yang lazim kita hadapi dalam kehidupan sehari-hari, entah di rumah, di tempat kerja, atau di tempat-tempat lain di mana aktivitas kehidupan sehari-hari berlangsung.

Tugas otentik berbeda dengan jenis-jenis tugas yang lazim dipakai dalam *typical assessment* atau penilaian konvensional dengan menggunakan tes. Dalam penilaian konvensional yang lazim

dilakukan di sekolah-sekolah dengan menggunakan berbagai macam tes itu, sebenarnya yang terjadi adalah menerapkan "*indirect ways of testing performance directly*." Maksudnya, menerapkan cara yang tidak langsung untuk mengetes kinerja murid secara langsung (Wiggins, 1998). Artinya, jumlah item-item tes berapa pun banyaknya tetap bukan merupakan *performance* atau kinerja nyata alias "langsung", kendati diujikan secara langsung. Di pihak lain, penilaian hasil belajar semestinya sungguh-sungguh bersifat edukatif, yaitu berupa *pengetesan langsung (direct testing)* yang mampu membukakan hati dan pikiran murid tentang tujuan pendidikan yang sesungguhnya serta wujud nyata dari pekerjaan orang dewasa.

Menurut Wiggins (1998), sebuah tugas, problem, atau proyek yang diberikan kepada murid untuk dikerjakan dalam rangka penilaian bersifat otentik apabila memenuhi syarat-syarat sebagai berikut:

1. *Bersifat realistik*. Tugas itu mereplikasikan atau memiliki kemiripan dengan cara-cara pengetahuan atau kemampuan seseorang diuji dalam situasi kehidupan nyata.
2. *Menuntut kemampuan memberikan "judgment" atau penilaian dan melakukan inovasi*. Tugas itu bersifat terbuka atau tidak terstruktur dan untuk menyelesaikannya menuntut murid menggunakan pengetahuan dan ketrampilannya secara efektif, tidak sekedar mengikuti sejenis prosedur tertentu yang sudah bersifat rutin. Misal, menuntut murid terlebih dulu membuat rencana.
3. *Menuntut murid sungguh-sungguh bekerja atau mengerjakan sesuatu*. Tugas itu menuntut murid melakukan eksplorasi atau sejenis penelitian serta menerapkan metode kerja sebagaimana lazim dipraktekkan dalam disiplin ilmu tertentu seperti Ilmu Pengetahuan Alam, sejarah, atau disiplin lain.
4. *Mereplikasikan atau mensimulasikan konteks tempat orang menghadapi "ujian" nyata entah di tempat kerja, dalam kehidupan bermasyarakat,*

atau dalam kehidupan pribadi. Konteks adalah situasi spesifik yang memiliki unsur-unsur tujuan, pelaku, serta peluang dan kendala tertentu. Contohnya, tempat kerja. Ada tujuan tertentu, masing-masing orang punya peran tertentu, ada aneka sumber daya namun juga selalu terbuka kemungkinan bagi munculnya kendala, masalah, dan sejenisnya yang bisa menghalangi setiap orang menjalankan perannya secara mulus, sehingga tujuan bersama tidak tercapai secara optimal. Murid perlu mengalami seperti apa rasanya mengerjakan tugas-tugas dalam konteks kehidupan nyata. Situasi semacam itu menuntut kemampuan berpikir secara nyata. Tes yang lazim diterapkan dalam penilaian di sekolah jelas sama sekali tidak memiliki kaitan dengan konteks.

5. *Menguji kemampuan murid menerapkan secara efisien dan efektif bekal pengetahuan dan ketrampilan yang dimilikinya untuk secara nyata mengerjakan sebuah tugas yang kompleks.* Lagi-lagi tes konvensional seperti yang lazim dipakai di sekolah-sekolah tidak memiliki ciri seperti ini. Sesulit apa pun tes konvensional tidak akan menuntut kemampuan berpikir secara menyeluruh dari pihak murid seperti ketika dia dituntut mengerjakan sebuah tugas nyata yang kompleks, misalnya melakukan percobaan tentang prinsip kerja mesin jet dalam pelajaran IPA.
6. *Memberi kesempatan kepada murid untuk berlatih, berpraktik, mencari sumber-sumber, memperoleh feedback, serta menyempurnakan kinerja atau hasil karyanya.* Penilaian otentik bersifat edukatif, dalam arti melalui siklus *bekerja-memperoleh feedback-melakukan revisi-bekerja (performance-feedback-revision-performance)* murid sungguh-sungguh dibantu untuk menghasilkan karya yang berkualitas tinggi melalui bekerja atau mengerjakan tugas nyata dalam konteks tertentu seraya belajar memanfaatkan informasi dan aneka sumber lain.

Langkah Pertama: Pentingnya Kisi-kisi Tujuan Pengajaran

Memperhatikan ciri-cirinya seperti diuraikan pada bagian sebelumnya, maka penilaian otentik akan mengandalkan penggunaan tugas-tugas kompleks yang harus dikerjakan oleh murid sedangkan penilaiannya akan mengandalkan teknik-teknik penilaian nontes untuk mengungkap aneka kemampuan yang bersifat kompleks pula. Pada bagian sebelumnya kita juga sudah melihat lima teknik penilaian nontes, yaitu penilaian tertulis, penilaian berbasis perilaku, penilaian berbasis produk, penilaian berbasis wawancara, dan penilaian berbasis portofolio. Dalam rangka membahas penyusunan rencana penilaian dan rubrik penskorannya, kita akan lebih berfokus pada empat teknik penilaian nontes selain teknik penilaian berbasis portofolio. Sebagaimana akan kita lihat, penilaian berbasis portofolio akan melibatkan keempat teknik penilaian nontes lainnya dalam kadar yang berlainan.

Namun sebelumnya perlu disadari bahwa penilaian kelas lazimnya dibedakan menjadi dua kategori, masing-masing mengusung tujuan yang berbeda, yaitu *penilaian formatif* dan *penilaian sumatif*. Sebagaimana dinyatakan oleh Bloom, Hastings, dan Madaus (1971, dalam Friedenber, 1995), penilaian sumatif dilakukan oleh guru dalam proses memberikan nilai akhir atau nilai rapor untuk suatu mata pelajaran. Dengan kata lain, penilaian sumatif dilakukan dengan tujuan menunjukkan taraf pengetahuan dan/atau ketrampilan murid *kini* (*current level of knowledge and/or skill*) terkait suatu mata pelajaran. Sebaliknya menurut para pengarang yang sama, penilaian formatif dilakukan oleh guru dengan tujuan memberikan umpan balik tentang kemajuan belajar murid baik bagi guru maupun bagi murid yang bersangkutan (dan pihak lain terkait seperti orang tua) atau memberikan umpan balik tentang keefektivan pembelajaran khususnya bagi guru sendiri.

Aneka jenis teknik penilaian nontes tepat digunakan baik untuk penilaian formatif maupun penilaian sumatif. Khusus tentang empat jenis teknik penilaian yang pertama, yaitu penilaian tertulis, penilaian berbasis perilaku, penilaian berbasis produk, dan penilaian berbasis wawancara, aneka teknik ini dapat diterapkan untuk menilai secara mandiri ketercapaian komponen-komponen atau aspek-aspek kemampuan yang diajarkan dalam suatu mata pelajaran. Hasil penilaian secara mandiri dengan keempat teknik penilaian nontes yang dilakukan guru sepanjang semester pembelajaran suatu mata pelajaran dapat dihimpun sebagai bahan untuk melakukan penilaian portofolio. Hasil penilaian portofolio ini selanjutnya dapat dipakai dalam rangka menentukan secara komprehensif dan final taraf penguasaan murid atas pengetahuan-ketrampilan dalam mata pelajaran tertentu yang dinyatakan dalam sebuah nilai akhir. Bertolak dari asumsi kecil tersebut, pembicaraan selanjutnya kurang lebih akan difokuskan pada seluk-beluk penyusunan alat penilaian nontes selain portofolio, mulai dari perencanaan sampai ke penyusunan rubrik penskorannya.

Seperti alat penilaian tes, alat penilaian nontes juga harus valid, yaitu benar-benar mengukur pengetahuan, perasaan, atau ketrampilan yang tepat seperti yang dimaksud oleh tujuan penilaian itu (Friedenberg, 1995). Secara umum, salah satu jenis validitas yang perlu ditegakkan pada awal penyusunan alat penilaian adalah validitas isi. Sebagaimana sudah kita bahas, metode untuk memastikan validitas isi suatu tes adalah dengan menyusun kisi-kisi bertolak dari suatu taksonomi tujuan pengajaran tertentu. Idealnya, kisi-kisi tes hasil belajar pun dibuat pada awal, sebagai bagian dari kegiatan persiapan guru sebelum proses pembelajaran dilaksanakan. Dalam praktek guru sering baru menyusun kisi-kisi tes saat memerlukannya, yaitu di tengah atau menjelang akhir semester kegiatan pembelajaran ketika hendak menyusun tes hasil belajar.

Untuk penerapan teknik penilaian nontes penyusunan kisi-kisi mutlak dilakukan pada awal sebelum kegiatan pembelajaran berlangsung, sebagai bagian dari usaha guru melakukan persiapan mengajar, khususnya dalam bentuk merumuskan aneka hasil belajar sebagai tujuan pengajaran secara komprehensif untuk mata pelajaran tertentu yang akan diasuhnya. Dengan kata lain, langkah pertama dalam perencanaan penyusunan alat nontes adalah menyusun kisi-kisi hasil belajar sebagai tujuan pengajaran untuk mata pelajaran yang bersangkutan.

Dalam praktek berbagai hasil belajar sebagai tujuan pengajaran dalam suatu mata pelajaran tersebut bervariasi kompleksitasnya terkait ranah kemampuan yang dilibatkan, ada yang cukup sederhana didominasi oleh ranah kognitif, ada yang cukup sederhana didominasi oleh ranah afektif, namun mungkin ada yang kompleks berupa gabungan antara ranah kognitif, afektif, dan psikomotorik. Pada bagian ini akan dibahas cara menyusun hasil belajar sebagai tujuan pengajaran untuk masing-masing ranah kemampuan secara terpisah disertai contoh, bertolak dari taksonomi tujuan pengajaran yang lazim dipakai sebagaimana sudah kita bahas sebelumnya.

Hasil Belajar dengan Ranah Isi Pengetahuan dan Perasaan

Untuk mengidentifikasi jenis-jenis hasil belajar dengan ranah isi pengetahuan dan perasaan, kita mengandalkan taksonomi yang pernah dibuat oleh Bloom dan kawan-kawan. Untuk ranah isi pengetahuan kita akan menggunakan taksonomi Bloom hasil revisi kedua yang dilakukan oleh Anderson dan Krathwohl. Untuk ranah isi perasaan, kita akan menggunakan taksonomi yang dikembangkan oleh Krathwohl, Bloom, dan Masia (1964).

Kisi-kisi hasil belajar dengan ranah isi pengetahuan. Sebelum menyusun kisi-kisi hasil belajar dengan ranah isi pengetahuan sebagai tujuan pengajaran, baiklah kita ingat kembali taksonomi tujuan

pengajaran Bloom ranah kognitif hasil revisi kedua yang dilakukan oleh Anderson dan Krathwohl.

Tabel 5.1: Taksonomi Bloom Ranah Kognitif Hasil Revisi Anderson dan Krathwohl

Dimensi Pengetahuan	Dimensi Proses Kognitif					
	Mengingat	Mema-hami	Meng-aplikasikan	Meng-analisis	Meng-evaluasi	Mencip-takan
Pengetahuan Faktual	Membuat daftar (<i>List</i>)	Membuat ringkasan (<i>Summarize</i>)	Mengklasifikasi (<i>Classify</i>)	Mengurutkan (<i>Order</i>)	Menentukan <i>ranking</i> (<i>Rank</i>)	Mengombinasikan (<i>Combine</i>)
Pengetahuan Konseptual	Mendeskripsikan (<i>Describe</i>)	Menafsirkan (<i>Interpret</i>)	Bereksperimentasi (<i>Experiment</i>)	Menjelaskan (<i>Explain</i>)	Melakukan penilaian (<i>Assess</i>)	Merencanakan (<i>Plan</i>)
Pengetahuan Prosedural	Menabuliskan (<i>Tabulate</i>)	Memprediksikan (<i>Predict</i>)	Menghitung (<i>Calculate</i>)	Membedakan (<i>Differentiate</i>)	Membuat kesimpulan (<i>Conclude</i>)	Membuat komposisi (<i>Compose</i>)
Pengetahuan Meta-kognitif	Menggunakan secara tepat/semesta (<i>Appropriate use</i>)	Mengeksekusi (<i>Execute</i>)	Mengonstruksi (<i>Construct</i>)	Mencapai atau meraih pencapaian (<i>Achieve</i>)	Melakukan tindakan (<i>Action</i>)	Mengaktualisasikan (<i>Actualize</i>)

Dari tabel di atas tampak, seluruh kategori hasil belajar mulai dari yang relatif paling sederhana (Pengetahuan Faktual – Mengingat) sampai ke hasil belajar yang relatif paling kompleks (Pengetahuan Metakognitif – Mengaktualisasikan) dapat didayagunakan untuk menilai hasil belajar secara otentik, yaitu hasil belajar berupa kemampuan yang mereplikasikan atau memiliki kemiripan dengan kemampuan nyata yang diperlukan oleh orang dewasa untuk menghadapi aneka tantangan nyata di tempat kerja, dalam berbagai urusan kehidupan bermasyarakat, atau dalam kehidupan pribadi mereka.

Maka, secara umum kisi-kisi hasil belajar ranah kognitif sebagai tujuan pengajaran suatu mata pelajaran akan memiliki struktur dasar sebagai berikut:

Tabel 5.2: Kisi-kisi Hasil Belajar Ranah Pengetahuan sebagai Tujuan Pengajaran Mata Pelajaran

Isi/ Kompe- tensi Dasar	Dimensi Penge- tahuan	Dimensi Proses Kognitif					
		Meng- ingat	Mema- hami	Mengapli- kasikan	Mengana- lisis	Meng- evaluasi	Mencip- takan
	Pengetahuan Faktual						
	Pengetahuan Konseptual						
	Pengetahuan Prosedural						
	Pengetahuan Metakognitif						
	Pengetahuan Faktual						
	Pengetahuan Konseptual						
	Pengetahuan Prosedural						
	Pengetahuan Metakognitif						

Seperti bisa dilihat dalam Tabel 5.2., tujuan pengajaran sebuah mata pelajaran perlu dijabarkan dalam serangkaian isi atau kompetensi dasar atau bahkan rumusan hasil belajar yang lebih rinci. Rumusan isi, kompetensi dasar atau hasil belajar yang lebih rinci ini dicantumkan pada kolom pertama. Kemudian pada kolom ke-3 sampai dengan kolom ke-8 masing-masing isi, kompetensi dasar, atau rumusan hasil belajar yang lebih rinci tersebut perlu dijabarkan lebih lanjut ke dalam indikator-indikator yang lebih spesifik lagi sesudah mempertimbangkan jenis dimensi pengetahuannya sebagaimana dispesifikasikan pada kolom ke-2.

Kisi-kisi hasil belajar dengan ranah isi perasaan. Seperti penyusunan kisi-kisi hasil belajar dengan ranah isi pengetahuan, sebelum menyusun kisi-kisi hasil belajar dengan ranah isi perasaan sebagai tujuan pengajaran, baiklah kita ingat kembali taksonomi

tujuan pengajaran ranah perasaan yang dikembangkan oleh Krathwohl, Bloom, dan Masia (1973).

**Tabel 5.3: Taksonomi Tujuan Pengajaran Ranah Afektif
(Krathwohl, Bloom, & Masia, 1964)**

Taraf Kemampuan	Uraian
Mau menginternalisasikan nilai-nilai (karakterisasi) (<i>Internalizing Values</i>)	<p><i>Mau menunjukkan perilaku yang dikendalikan oleh suatu sistem nilai.</i></p> <p>Memiliki suatu sistem nilai yang dijadikan pedoman berperilaku, sehingga perilaku menjadi konsisten, bisa diprediksikan, dan yang terpenting menjadi ciri atau karakteristik pribadi yang bersangkutan.</p>
Mau mengorganisasikan nilai-nilai (<i>Organization</i>)	<p><i>Mau mengorganisasikan nilai-nilai mengikuti urutan prioritas tertentu.</i></p> <p>Mengorganisasikan nilai ke dalam skala prioritas (mengurutkan dari yang paling penting/bernilai sampai yang paling kurang penting/kurang bernilai) dengan cara membandingkan berbagai nilai yang berbeda, mengatasi konflik-konflik yang terjadi antar nilai-nilai yang berbeda tersebut, dan akhirnya mampu menciptakan suatu sistem nilai yang khas bagi dirinya.</p>
Mau memberikan nilai/mau memandang bernilai (<i>Valuing</i>)	<p><i>Mau memberikan nilai pada sesuatu.</i></p> <p>Mulai dari sekadar menerima sesuatu sebagai bernilai sampai menunjukkan komitmen yang lebih kompleks. Kemampuan ini didasari oleh internalisasi terhadap serangkaian nilai-nilai spesifik tertentu.</p>
Mau memberikan respon terhadap fenomena tertentu. (<i>Responding to phenomena</i>)	<p><i>Mau berperan aktif dalam kegiatan belajar; berpartisipasi.</i></p> <p>Meliputi mau berpartisipasi aktif, mau memberikan perhatian dan reaksi terhadap fenomena tertentu. Hasil belajar yang ditekankan: mau menjawab dan merasakan kepuasan dengan memberikan respon.</p>
Mau menerima fenomena tertentu (<i>Receiving phenomena</i>)	<p>Mau menyadari; menunjukkan kemauan untuk mendengarkan.</p> <p>Mau menyadari, mau mendengarkan, atau mau memberikan perhatian.</p>

Misal, dalam rangka pengajaran mata pelajaran Pendidikan Kewarganegaraan guru ingin mengajarkan nilai cinta dan hormat terhadap keberagaman budaya dalam masyarakat. Berdasarkan taksonomi yang dikembangkan oleh Krathwohl, Bloom, dan Masia (1973) seperti disajikan di atas, maka kisi-kisi hasil belajar ranah afektif atau perasaan sebagai tujuan pengajarannya kurang lebih akan seperti yang disajikan dalam Tabel 5.4.

Tabel 5.4: Kisi-kisi Hasil Belajar Ranah Perasaan sebagai Tujuan Pengajaran PKn

Isi/ Kompetensi Dasar	Taraf Kemampuan/ Indikator	Uraian
Cinta dan hormat terhadap keberagaman budaya dalam masyarakat	Mau menginternalisasikan nilai-nilai (karakterisasi) <i>(Internalizing Values)</i>	Mau menunjukkan perilaku yang dikendalikan oleh suatu sistem nilai yang berbasis cinta dan hormat terhadap keberagaman budaya dalam masyarakat.
	Mau mengorganisasikan nilai-nilai <i>(Organization)</i>	Mau mengorganisasikan nilai-nilai mengikuti urutan prioritas tertentu, dengan menempatkan nilai cinta dan hormat terhadap keberagaman budaya dalam masyarakat sebagai poros.
	Mau memberikan nilai/mau memandang bernilai <i>(Valuing)</i>	Mau memberikan nilai pada bentuk-bentuk perwujudan nilai cinta dan hormat terhadap keberagaman budaya dalam masyarakat.
	Mau memberikan respon terhadap fenomena tertentu. <i>(Responding to phenomena)</i>	Mau berperan serta aktif dalam kegiatan yang mencerminkan nilai cinta dan hormat terhadap keragaman budaya dalam masyarakat.
	Mau menerima fenomena tertentu <i>(Receiving phenomena)</i>	Mau menyadari atau mau memberikan perhatian terhadap kenyataan tentang adanya keberagaman budaya dalam masyarakat.

Hasil Belajar dengan Ranah Isi Ketrampilan dan Pola Perilaku

Untuk mengidentifikasi jenis-jenis hasil belajar ranah isi ketrampilan kita bisa menggunakan taksonomi tujuan pengajaran ranah psikomotor yang pernah dibuat oleh Simpson. Jenis hasil belajar ini lazimnya memiliki cakupan yang agak spesifik, seperti misalnya ketrampilan mengetik pada komputer dengan sepuluh jari, ketrampilan mengoperasikan aplikasi komputer tertentu, dan sebagainya.

Lain halnya dengan hasil belajar dengan ranah isi pola perilaku yang kompleks. Hasil belajar semacam ini lazimnya berupa gugusan atau rangkaian kompetensi kompleks yang tercakup dalam sebuah *job* alias pekerjaan atau *role* alias peran tertentu. Sebagai contoh, pekerjaan sebagai guru kiranya akan mencakup rangkaian kompetensi seperti menyusun perencanaan pembelajaran, menyusun evaluasi pembelajaran, melaksanakan pembelajaran, kemampuan bertanya, kemampuan memberikan *feedback*, kemampuan menampilkan diri sebagai model karakter yang baik, dan sebagainya. Contoh lain, peran sebagai ketua kelas akan mencakup rangkaian kompetensi seperti memiliki disiplin diri, memiliki sifat pemberani, mampu mengungkapkan pikiran secara jelas, memiliki rasa tanggung jawab, mampu menampilkan diri sebagai teladan, mampu mendengarkan ungkapan pendapat atau perasaan teman, dan sebagainya. Maka, untuk mengidentifikasi jenis-jenis hasil belajar dengan ranah isi pola perilaku yang kompleks terkait pekerjaan atau peran tertentu, kita perlu menemukan sumber entah berupa sumber tertulis atau pendapat narasumber yang berwenang untuk kita jadikan pedoman agar bisa melakukan analisis tugas secara cermat.

Kisi-kisi hasil belajar dengan ranah ketrampilan spesifik.
Sebelum menyusun kisi-kisi hasil belajar dengan ranah isi ketrampilan

spesifik sebagai tujuan pengajaran, baiklah kita ingat kembali taksonomi tujuan pengajaran ranah psikomotor yang dikembangkan oleh Simpson.

Tabel 5.5: Taksonomi Tujuan Pengajaran Ranah Psikomotor (Simpson, 1972)

Taraf Kemampuan	Uraian
Mampu menghasilkan gerakan baru (<i>Origination</i>)	<i>Mampu menciptakan pola gerakan baru.</i> Mampu menciptakan aneka pola gerakan baru sesuai tuntutan suatu situasi atau problem khusus tertentu. Hasil belajar yang ditekankan berupa kreativitas yang dilandasi aneka ketrampilan taraf tinggi.
Mampu beradaptasi (<i>Adaptation</i>)	<i>Mampu memodifikasi aneka ketrampilan motor disesuaikan dengan tuntutan situasi baru.</i> Aneka ketrampilan sudah dikuasai dengan baik, sehingga siswa mampu memodifikasikan pola gerakan agar sesuai dengan tuntutan situasi khusus tertentu.
Mampu melakukan respon kompleks secara lancar (<i>Complex Overt Response</i>)	<i>Menunjukkan tahap agak lanjut menguasai suatu ketrampilan kompleks.</i> Mampu melakukan tindakan motor secara trampil yang melibatkan pola gerakan yang kompleks. Ketrampilan atau ketangkasan ditunjukkan oleh gerakan yang cepat, akurat, dan sangat terkoordinasi, yang dilakukan dengan energi atau upaya minimum. Kategori ini mencakup mengerjakan tugas tanpa ragu-ragu dan melakukan gerakan secara otomatis.
Mampu melakukan respon secara mekanik (<i>Mechanism</i>)	<i>Mampu melakukan suatu ketrampilan motor yang kompleks.</i> Merupakan tahap piawai dalam mempelajari suatu ketrampilan kompleks. Hasil belajar sudah menyatu dengan kebiasaan, sehingga gerakan-gerakan bisa dilakukan dengan percaya diri dan lancar.
Mampu melakukan respon tertentu dengan bimbingan guru (<i>Guided Response</i>)	<i>Menunjukkan tahap awal menguasai suatu ketrampilan kompleks, meliputi kemampuan mengikuti contoh atau mencontoh.</i> Merupakan tahap awal dalam mempelajari suatu ketrampilan kompleks, mencakup kemampuan mencontoh atau coba-salah. Ketrampilan yang memadai akan dicapai lewat latihan.

Taraf Kemampuan	Uraian
Memiliki kesiapan untuk bertindak <i>(Set)</i>	<i>Menunjukkan kesiapan untuk bertindak.</i> Kesiapan untuk bertindak, meliputi kesiapan mental, fisik, dan emosi. Ketiganya merupakan disposisi yang mendasari respon seseorang terhadap berbagai situasi yang dihadapi (kadang-kadang juga disebut <i>mindset</i>).
Mampu mempersepsikan <i>(Perception)</i>	<i>Mampu menggunakan tanda-tanda sensoris untuk membimbing aktivitas fisik tertentu.</i> Mampu menggunakan petunjuk-isyarat sensoris untuk membimbing aktivitas motor, meliputi kepekaan menangkap stimulasi sensoris kemampuan memilih petunjuk-isyarat sensoris, dan kemampuan menerjemahkannya ke dalam tindakan.

Sebagai contoh, dalam mata pelajaran tertentu guru ingin mengajarkan ketrampilan spesifik tertentu, misalnya menulis tegak bersambung di papan tulis putih dengan menggunakan spidol papan tulis. Berdasarkan taksonomi yang dikembangkan oleh Simpson (1972) seperti disajikan di atas, maka kisi-kisi hasil belajar sebagai tujuan pengajarannya kurang lebih akan seperti yang disajikan dalam Tabel 5.6.

Tabel 5.6: Kisi-kisi Hasil Belajar Ranah Ketrampilan Spesifik sebagai Tujuan Pengajaran

Isi/Kompetensi Dasar	Taraf Kemampuan/ Indikator	Uraian
Menulis tegak bersambung di papan tulis putih dengan menggunakan spidol papan tulis	Mampu menghasilkan gerakan baru (<i>Origination</i>)	<i>Mampu menciptakan pola gerakan baru.</i> Mampu menuliskan sejumlah kalimat pada papan putih dengan menggunakan spidol dengan benar, sejajar dengan garis panjang papan putih, dengan ciri tulisan tangan khas, dan dengan indah.
	Mampu beradaptasi (<i>Adaptation</i>)	<i>Mampu memodifikasi aneka ketrampilan motor disesuaikan dengan tuntutan situasi baru.</i> Mampu menuliskan sejumlah kalimat pada papan putih dengan menggunakan spidol dengan benar, sejajar dengan garis panjang papan putih, dan dengan menunjukkan ciri tulisan tangan yang mulai menetap.
	Mampu melakukan respon kompleks secara lancar (<i>Complex Overt Response</i>)	<i>Menunjukkan tahap agak lanjut menguasai suatu ketrampilan kompleks.</i> Mampu menuliskan sejumlah kalimat pada papan putih dengan menggunakan spidol dengan benar dan sejajar dengan garis panjang papan putih.
	Mampu melakukan respon secara mekanik (<i>Mechanism</i>)	<i>Mampu melakukan suatu ketrampilan motor yang kompleks.</i> Mampu menuliskan sebuah kalimat pada papan putih dengan menggunakan spidol secara benar.
	Mampu melakukan respon tertentu dengan bimbingan guru (<i>Guided Response</i>)	<i>Menunjukkan tahap awal menguasai suatu ketrampilan kompleks, meliputi kemampuan mengikuti contoh atau mencontoh.</i> Mampu menggerakkan spidol dengan benar pada papan putih untuk menulis sejumlah huruf dengan bimbingan petunjuk dari guru.
	Memiliki kesiapan untuk bertindak (<i>Set</i>)	<i>Menunjukkan kesiapan untuk bertindak.</i> Mampu menempatkan spidol pada posisi yang benar di papan putih.
	Mampu mempersepsikan (<i>Perception</i>)	<i>Mampu menggunakan tanda-tanda sensoris untuk membimbing aktivitas fisik tertentu.</i> Mampu memegang spidol dengan benar.

Kisi-kisi hasil belajar dengan ranah pola perilaku kompleks.

Sudah disinggung bahwa hasil belajar dengan ranah pola perilaku kompleks sebagai tujuan pengajaran lazimnya terkait dengan *job* alias pekerjaan atau *role* alias peran tertentu. Untuk mengidentifikasi bentuk-bentuk perilaku yang merupakan komponen atau unsur pelaksanaan pekerjaan atau peran tertentu, taksonomi tujuan pengajaran seperti yang disusun oleh Simpson (1972) jelas tidak lagi memadai. Sebagai gantinya, identifikasi ini harus didasarkan pada sebuah *job analysis* atau *task analysis*.

Analisis pekerjaan atau analisis tugas merupakan proses untuk mendapatkan informasi tentang seluk-beluk suatu pekerjaan atau tugas (McCormick & Ilgen, 1980). Seperti dinyatakan oleh McCormick dan Ilgen (1980), ada beberapa jenis informasi yang lazim dikumpulkan tentang suatu pekerjaan, yaitu: (1) aneka aktivitas yang merujuk pada pekerjaan atau tugas yang harus dilaksanakan, misal mengembangkan kurikulum, mengembangkan materi pembelajaran; (2) aneka aktivitas yang merujuk pada pekerja yang melaksanakan tugas, misal menyelenggarakan pembelajaran yang mendidik, menyelenggarakan penilaian dan evaluasi proses dan hasil belajar; (3) mesin, perkakas, alat bantu, dan sebagainya yang khas untuk melaksanakan pekerjaan atau tugas, misal *laptop* dan *LCD*, kapur atau spidol, penggaris, peta, dan sebagainya; (4) aneka jenis material yang diproses; (5) jenis pengetahuan yang diperlukan; (6) aneka kondisi kerja; dan (7) aneka persyaratan yang dituntut dari orang yang akan melaksanakan pekerjaan atau tugas yang bersangkutan.

Masih menurut kedua pakar yang sama, analisis tugas atau pekerjaan bermanfaat untuk berbagai keperluan, meliputi: (1) administrasi pegawai, mencakup rekrutmen, seleksi dan penempatan pegawai, pelatihan dan pengembangan pegawai, penilaian kinerja, administrasi gaji dan upah, serta urusan hubungan perburuhan;

(2) perancangan pekerjaan dan peralatan, mencakup perancangan metode kerja, perancangan pekerjaan; (3) pengendalian administrasi, mencakup perencanaan organisasi, perencanaan dan pengendalian sumber daya manusia; dan (4) manfaat lain, mencakup perencanaan kurikulum pendidikan dan pelatihan, konseling vokasional, penyusunan system klasifikasi pekerjaan.

Berikut adalah contoh hasil analisis pekerjaan atau tugas guru kelas Sekolah Dasar/Madrasah Ibtidaiyah yang sudah disahkan sebagai standar kompetensi guru kelas SD/MI dalam Peraturan Menteri Pendidikan Nasional Nomor 16 Tahun 2007.

Tabel 5.7: Standar Kompetensi Guru Kelas SD/MI*

No.	Kompetensi Inti Guru	Kompetensi Guru Kelas SD/MI
	<i>Kompetensi pedagogik</i>	
1	Menguasai karakteristik peserta didik dari aspek fisik, moral, sosial, kultural, emosional, dan intelektual.	1.1. Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek fisik, intelektual, sosial-emosional, moral, spiritual, dan latar belakang sosial-budaya. 1.2. Mengidentifikasi potensi peserta didik usia SD dalam lima mata pelajaran SD. 1.3. Mengidentifikasi kemampuan awal peserta didik usia SD dalam lima mata pelajaran SD. 1.4. Mengidentifikasi kesulitan belajar peserta didik usia SD dalam lima mata pelajaran SD.
2	Menguasai teori belajar dan prinsip-prinsip pembelajaran yang mendidik.	1.1. Memahami berbagai teori belajar dan prinsip-prinsip pembelajaran yang mendidik terkait dengan lima mata pelajaran SD. 1.2. Menerapkan berbagai pendekatan, strategi, metode, dan teknik pembelajaran yang mendidik secara kreatif dalam lima mata pelajaran SD. 1.3. Menerapkan pendekatan pembelajaran tematis, khususnya di kelas-kelas awal SD.
3	Mengembangkan kurikulum yang terkait dengan mata pelajaran/bidang	1.1. Memahami prinsip-prinsip pengembangan kurikulum. 1.2. Menentukan tujuan lima mata pelajaran SD.

Penyusunan Alat Penilaian Nontes

No.	Kompetensi Inti Guru	Kompetensi Guru Kelas SD/MI
	pengembangan yang diampu.	1.3. Menentukan pengalaman belajar yang sesuai untuk mencapai tujuan lima mata pelajaran SD. 1.4. Memilih materi lima mata pelajaran SD yang terkait dengan pengalaman belajar dan tujuan pembelajaran. 1.5. Mengembangkan indikator dan instrumen penilaian
4	Menyelenggarakan pembelajaran yang mendidik.	1.1. Memahami prinsip-prinsip perancangan pembelajaran yang mendidik. 1.2. Mengembangkan komponen-komponen rancangan pembelajaran. 1.3. Menyusun rancangan pembelajaran yang lengkap, baik untuk kegiatan di dalam kelas, laboratorium, maupun lapangan. 1.4. Melaksanakan pembelajaran yang mendidik di kelas, di laboratorium, dan di lapangan. 1.5. Menggunakan media pembelajaran sesuai dengan karakteristik peserta didik dan lima mata pelajaran SD untuk mencapai tujuan pembelajaran secara utuh. 1.6. Mengambil keputusan transaksional dalam lima mata pelajaran SD sesuai dengan situasi yang berkembang.
5	Memanfaatkan teknologi informasi dan komunikasi untuk kepentingan pembelajaran.	1.1. Memanfaatkan teknologi informasi dan komunikasi dalam pembelajaran.
6	Memfasilitasi pengembangan potensi peserta didik untuk mengaktualisasikan berbagai potensi yang dimiliki.	1.1. Menyediakan berbagai kegiatan pembelajaran untuk mendorong peserta didik mencapai prestasi belajar secara optimal. 1.2. Menyediakan berbagai kegiatan pembelajaran untuk mengaktualisasikan potensi peserta didik termasuk kreativitasnya.
7	Berkomunikasi secara efektif, empatik, dan santun dengan peserta didik.	1.1. Memahami berbagai strategi berkomunikasi yang efektif, empatik, dan santun baik secara lisan maupun tulisan.

No.	Kompetensi Inti Guru	Kompetensi Guru Kelas SD/MI
		<p>1.2. Berkomunikasi secara efektif, empatik, dan santun dengan peserta didik dengan bahasa yang khas dalam interaksi pembelajaran yang terbangun secara siklikal dari (a) penyiapan kondisi psikologis peserta didik, (b) memberikan pertanyaan atau tugas sebagai undangan kepada peserta didik untuk merespon, (c) respon peserta didik, (d) reaksi guru terhadap respon peserta didik, dan seterusnya.</p>
8	<p>Menyelenggarakan penilaian dan evaluasi proses dan hasil belajar.</p>	<p>1.1. Memahami prinsip-prinsip penilaian dan evaluasi proses dan hasil belajar sesuai dengan karakteristik lima mata pelajaran SD. 1.2. Menentukan aspek-aspek proses dan hasil belajar yang penting untuk dinilai dan dievaluasi sesuai dengan karakteristik lima mata pelajaran SD. 1.3. Menentukan prosedur penilaian dan evaluasi proses dan hasil belajar. 1.4. Mengembangkan instrumen penilaian dan evaluasi proses dan hasil belajar. 1.5. Mengadministrasikan penilaian proses dan hasil belajar secara berkesinambungan dengan menggunakan berbagai instrumen. 1.6. Menganalisis hasil penilaian proses dan hasil belajar untuk berbagai tujuan. 1.7. Melakukan evaluasi proses dan hasil belajar.</p>
9	<p>Memanfaatkan hasil penilaian dan evaluasi untuk kepentingan pembelajaran.</p>	<p>1.1. Menggunakan informasi hasil penilaian dan evaluasi untuk menentukan ketuntasan belajar. 1.2. Menggunakan informasi hasil penilaian dan evaluasi untuk merancang program remedial dan pengayaan. 1.3. Mengomunikasikan hasil penilaian dan evaluasi kepada pemangku kepentingan. 1.4. Memanfaatkan informasi hasil penilaian dan evaluasi pembelajaran untuk meningkatkan kualitas pembelajaran.</p>
10	<p>Melakukan tindakan reflektif untuk peningkatan kualitas pembelajaran.</p>	<p>1.1. Melakukan refleksi terhadap pembelajaran yang telah dilaksanakan. 1.2. Memanfaatkan hasil refleksi untuk perbaikan dan pengembangan lima mata pelajaran SD. 1.3. Melakukan penelitian tindakan kelas untuk meningkatkan kualitas pembelajaran lima mata pelajaran SD.</p>

Penyusunan Alat Penilaian Nontes

No.	Kompetensi Inti Guru	Kompetensi Guru Kelas SD/MI
<i>Kompetensi kepribadian</i>		
11	Bertindak sesuai dengan norma agama, hukum, sosial, dan kebudayaan nasional Indonesia.	1.1. Menghargai peserta didik tanpa membedakan keyakinan yang dianut, suku, adat istiadat, daerah asal, dan gender. 1.2. Bersikap sesuai dengan norma agama yang dianut, hukum dan norma sosial yang berlaku dalam masyarakat, serta kebudayaan nasional Indonesia yang beragam.
12	Menampilkan diri sebagai pribadi yang jujur, berakhlak mulia, dan teladan bagi peserta didik dan masyarakat.	1.1. Berperilaku jujur, tegas, dan manusiawi. 1.2. Berperilaku yang mencerminkan ketaqwaan dan akhlak mulia. 1.3. Berperilaku yang dapat diteladani oleh peserta didik dan anggota masyarakat di sekitarnya.
13	Menampilkan diri sebagai pribadi yang mantab, stabil, dewasa, arif, dan berwibawa.	1.1. Menampilkan diri sebagai pribadi yang mantab dan stabil. 1.2. Menampilkan diri sebagai pribadi yang dewasa, arif, dan berwibawa.
14	Menunjukkan etos kerja, tanggung jawab yang tinggi, rasa bangga menjadi guru, dan rasa percaya diri.	1.1. Menunjukkan etos kerja dan tanggung jawab yang tinggi. 1.2. Bangga menjadi guru dan percaya pada diri sendiri. 1.3. Bekerja mandiri secara profesional.
15	Menjunjung tinggi kode etik profesi guru.	1.1. Memahami kode etik profesi guru. 1.2. Menerapkan kode etik profesi guru. 1.3. Berperilaku sesuai dengan kode etik guru.
<i>Kompetensi sosial</i>		
16	Bersikap inklusif, bertindak objektif, serta tidak diskriminatif karena pertimbangan jenis kelamin, agama, ras, kondisi fisik, latar belakang keluarga, dan status sosial ekonomi.	1.1. Bersikap inklusif dan objektif terhadap peserta didik, teman sejawat, dan lingkungan sekitar dalam melaksanakan pembelajaran. 1.2. Tidak bersikap diskriminatif terhadap peserta didik, teman sejawat, orang tua peserta didik, dan lingkungan sekolah karena perbedaan agama, suku, jenis kelamin, latar belakang keluarga, dan status sosial ekonomi.

No.	Kompetensi Inti Guru	Kompetensi Guru Kelas SD/MI
17	Berkomunikasi secara efektif, empatik, dan santun dengan sesama pendidik, tenaga kependidikan, orang tua, dan masyarakat.	1.1. Berkomunikasi dengan teman sejawat dan komunitas ilmiah lainnya secara santun, empatik, dan efektif. 1.2. Berkomunikasi dengan orang tua peserta didik dan masyarakat secara santun, empatik, dan efektif tentang program pembelajaran dan kemajuan peserta didik. 1.3. Mengikutsertakan orang tua peserta didik dan masyarakat dalam program pembelajaran dan dalam mengatasi kesulitan belajar peserta didik.
18	Beradaptasi di tempat bertugas di seluruh wilayah Republik Indonesia yang memiliki keragaman sosial budaya.	1.1. Beradaptasi dengan lingkungan tempat bekerja dalam rangka meningkatkan efektivitas sebagai pendidik, termasuk memahami bahasa daerah setempat. 1.2. Melaksanakan berbagai program dalam lingkungan kerja untuk mengembangkan dan meningkatkan kualitas pendidikan di daerah yang bersangkutan.
19	Berkomunikasi dengan komunitas profesi sendiri dan profesi lain secara lisan dan tulisan atau bentuk lain.	1.1. Berkomunikasi dengan teman sejawat, profesi ilmiah, dan komunitas ilmiah lainnya melalui berbagai media dalam rangka meningkatkan kualitas pendidikan. 1.2. Mengomunikasikan hasil-hasil inovasi pembelajaran kepada komunitas profesi sendiri secara lisan dan tulisan atau bentuk lain.
Kompetensi profesional		
20	Menguasai materi, struktur, konsep, dan pola pikir keilmuan yang mendukung mata pelajaran yang diampu.	Bahasa Indonesia 1.1. Memahami hakikat bahasa dan pemerolehan bahasa. 1.2. Memahami kedudukan, fungsi, dan ragam bahasa Indonesia. 1.3. Menguasai dasar-dasar dan kaidah bahasa Indonesia yang baik dan benar. 1.4. Memiliki ketrampilan berbahasa Indonesia (menyimak, berbicara, membaca, dan menulis). 1.5. Memahami teori dan genre sastra Indonesia. 1.6. Mampu mengapresiasi karya sastra Indonesia secara reseptif dan produktif.

No.	Kompetensi Inti Guru	Kompetensi Guru Kelas SD/MI
		<p><i>Matematika</i></p> <p>1.7. Menguasai pengetahuan konseptual dan prosedural serta keterkaitan keduanya dalam konteks materi aritmatika, aljabar, geometri, trigonometri, pengukuran, statistika, dan logika matematika.</p> <p>1.8. Mampu menggunakan matematisasi horizontal dan vertikal untuk menyelesaikan masalah matematika dan masalah dalam dunia nyata.</p> <p>1.9. Mampu menggunakan pengetahuan konseptual, prosedural, dan keterkaitan keduanya dalam pemecahan masalah matematika, serta penerapannya dalam kehidupan sehari-hari.</p> <p>1.10. Mampu menggunakan alat peraga, alat ukur, alat hitung, dan piranti lunak komputer.</p> <p><i>Ilmu Pengetahuan Alam (IPA)</i></p> <p>1.11. Mampu melakukan observasi gejala alam baik secara langsung maupun tidak langsung.</p> <p>1.12. Memanfaatkan konsep-konsep dan hukum-hukum ilmu pengetahuan alam dalam berbagai situasi kehidupan sehari-hari.</p> <p>1.13. Memahami struktur ilmu pengetahuan alam, termasuk hubungan fungsional antar konsep yang berhubungan dengan mata pelajaran IPA.</p> <p><i>Ilmu Pengetahuan Sosial (IPS)</i></p> <p>1.14. Menguasai materi keilmuan yang meliputi dimensi pengetahuan, nilai, dan ketrampilan IPS.</p> <p>1.15. Mengembangkan materi, struktur, dan konsep keilmuan IPS.</p> <p>1.16. Memahami cita-cita, nilai, konsep, dan prinsip-prinsip pokok ilmu-ilmu sosial dalam konteks kebhinekaan masyarakat Indonesia dan dinamika kehidupan global.</p> <p>1.17. Memahami fenomena interaksi perkembangan ilmu pengetahuan, teknologi, seni, kehidupan agama, dan perkembangan masyarakat serta saling ketergantungan global.</p>

No.	Kompetensi Inti Guru	Kompetensi Guru Kelas SD/MI
		<p><i>Pendidikan Kewarganegaraan (PKn)</i></p> <p>1.18. Menguasai materi keilmuan yang meliputi dimensi pengetahuan, sikap, nilai, dan perilaku yang mendukung kegiatan pembelajaran PKn.</p> <p>1.19. Menguasai konsep dan prinsip kepribadian nasional dan demokrasi konstitusional Indonesia, semangat kebangsaan, dan cinta tanah air serta bela Negara.</p> <p>1.20. Menguasai konsep dan prinsip perlindungan, kemajuan HAM, serta penegakan hukum secara adil dan benar.</p> <p>1.21. Menguasai konsep, prinsip, nilai, moral, dan norma kewarganegaraan Indonesia yang demokratis dalam konteks kewargaan Negara dunia.</p>
21	Menguasai standar kompetensi dan kompetensi dasar mata pelajaran/bidang pengembangan yang diampu.	<p>1.1. Memahami standar kompetensi lima mata pelajaran SD.</p> <p>1.2. Memahami kompetensi dasar lima mata pelajaran SD.</p> <p>1.3. Memahami tujuan pembelajaran lima mata pelajaran SD.</p>
22	Mengembangkan materi pembelajaran yang diampu secara kreatif.	<p>1.1. Memilih materi lima mata pelajaran SD yang sesuai dengan tingkat perkembangan peserta didik.</p> <p>1.2. Mengolah materi lima mata pelajaran SD secara integratif dan kreatif sesuai dengan tingkat perkembangan peserta didik.</p>
23	Mengembangkan keprofesionalan secara berkelanjutan dengan melakukan tindakan reflektif.	<p>1.1. Melakukan refleksi terhadap kinerja sendiri secara terus-menerus.</p> <p>1.2. Memanfaatkan hasil refleksi dalam rangka peningkatan keprofesionalan.</p> <p>1.3. Melakukan penelitian tindakan kelas untuk peningkatan keprofesionalan.</p> <p>1.4. Mengikuti kemajuan zaman dengan belajar dari berbagai sumber.</p>
24	Memanfaatkan teknologi informasi dan komunikasi untuk berkomunikasi dan mengembangkan diri.	<p>1.1. Memanfaatkan teknologi informasi dan komunikasi dalam berkomunikasi.</p> <p>1.2. Memanfaatkan teknologi informasi dan komunikasi untuk pengembangan diri.</p>

* Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 16 Tahun 2007 tentang Standar Kualifikasi Akademik dan Kompetensi Guru.

Dalam praktek, Kurikulum Tingkat Satuan Pendidikan untuk masing-masing jenjang sekolah mulai Sekolah Dasar sampai dengan Sekolah Menengah Atas sudah menyediakan rumusan tujuan pengajaran secara hirarkis minimal mulai dari standar kompetensi sampai ke kompetensi dasarnya.

Langkah Kedua: Menyusun Alat Nontes

Secara umum, alat penilaian nontes cocok untuk melakukan *performance assessment* alias penilaian kinerja. Secara garis besar, teknik penilaian ini akan sangat sesuai untuk menilai jenis-jenis hasil belajar yang kompleks dalam arti baik berupa jenis kemampuan yang berada pada jenjang hirarki tinggi dalam masing-masing taksonomi tujuan pengajaran khususnya ranah kognitif dan afektif, maupun berupa kemampuan kompleks yang merupakan kombinasi dari aspek-aspek kemampuan tertentu dalam kedua ranah tujuan pengajaran tersebut. Selain itu, karena sifat kompleksnya di satu sisi dan karena sifatnya yang berbasis kinerja di sisi lain, beberapa dari antara teknik ini hanya cocok diterapkan pada kelompok murid yang sudah menguasai minimal kemampuan membaca dan menulis dengan baik seperti teknik penilaian tertulis, atau sebaliknya justru sangat sesuai untuk diterapkan pada kelompok murid usia dini yang belum menguasai kemampuan baca-tulis seperti teknik penilaian berbasis wawancara. Berikut akan disajikan beberapa pedoman dalam menyusun masing-masing teknik penilaian berbasis kinerja yang dimaksud.

Sesudah memiliki kisi-kisi hasil belajar sebagai tujuan pengajaran, yang dalam Kurikulum Tingkat Satuan Pendidikan dirumuskan secara hirarkis mulai dari standar kompetensi sampai kompetensi dasar, selanjutnya guru bisa mulai memilih hasil belajar yaitu kompetensi dasar mana yang mau dia nilai dengan alat nontes serta memilih jenis alat nontes yang sesuai. Salah satu format umum

berupa tahap-tahap dan langkah-langkah penyusunan alat penilaian nontes beserta contoh aplikasinya adalah seperti dikemukakan oleh Chatterji (2003), sebagai berikut:

Tahap-tahap Umum Menyusun Alat Penilaian Nontes (Chatterji, 2003)

Chatterji, pakar pendidikan berkebangsaan Amerika kelahiran India yang sudah kita sebut berulang kali, menyarankan beberapa tahap masing-masing dengan beberapa langkah (kecuali tahap terakhir) dalam menyusun alat penilaian nontes. Berikut disajikan adaptasinya, sesudah disesuaikan dengan terminologi dalam Kurikulum Tingkat Satuan Pendidikan.

Tahap 1. Menyusun alat penilaian nontes untuk kompetensi dasar tertentu

1. Tentukanlah populasi sasaran penilaian, yaitu “siapa” yang akan dikenai penilaian.
Misal, murid kelas 5 atau kelas 6 Sekolah Dasar.
2. Tentukanlah tujuan penilaian, yaitu “mengapa” atau “untuk apa” penilaian dilakukan, yaitu dalam rangka penilaian formatif atau penilaian sumatif.
3. Tentukan konstruk atau kompetensi dasarnya. Misal, mampu atau trampil menyusun dan membaca grafik.
4. Tentukan ranah atau cakupan hasil belajarnya, berupa indikator-indikator pencapaian kompetensi dasarnya. Chatterji (2003) menyarankan agar guru memilah ke dalam indikator umum dan indikator khusus. Dalam praktek guru bisa langsung merumuskan indikator khusus, dan itu pun bisa mengambil dari sumber yang sudah banyak tersedia atau guru merumuskannya sendiri.
5. Tentukan metode atau alat penilaian nontesnya.

6. Tentukan aneka kondisi penilaiannya, meliputi antara lain langkah-langkah penyusunan sampai pelaksanaannya, tempat pelaksanaannya.
7. Tentukan prosedur pengadministrasiannya.
8. Tentukan prosedur penskorannya (rubrik penskoran).
9. Tentukan pembobotan aneka indikator pencapaiannya.
10. Susunlah semua komponen penilaian di atas menjadi sebuah instrumen yang padu.

Tahap 2. Mengevaluasi kualitas alat penilaian nontes secara keseluruhan

1. Lakukanlah validasi isi alat penilaian nontes tersebut dengan bantuan *reviewers* eksternal, dalam arti minta bantuan teman atau ahli lain dan jangan dilakukan sendiri.
2. Lakukanlah revisi atau perbaiki seperlunya, sesuai masukan dari *reviewer* eksternal.
3. Lakukanlah uji coba dan pemeriksaan validitas secara empiris, sesuai dengan tujuan penilaian.
4. Lakukanlah revisi atau perbaiki seperlunya baik terhadap masing-masing komponen maupun terhadap instrumen penilaian secara keseluruhan.

Tahap 3. Alat penilaian nontes siap diterapkan

Menyusun Masing-masing Alat Penilaian Nontes

Seperti sudah disinggung di muka teknik penilaian nontes tepat digunakan baik untuk tujuan penilaian formatif maupun penilaian sumatif. Dalam kerangka penilaian sumatif keempat teknik penilaian nontes selain portofolio, yaitu penilaian tertulis, penilaian berbasis perilaku, penilaian berbasis produk, dan penilaian berbasis wawancara

tepat dipakai untuk menilai secara mandiri taraf penguasaan murid atas pengetahuan-ketrampilan tertentu yang merupakan komponen-komponen atau aspek-aspek kemampuan yang hendak diajarkan dalam suatu mata pelajaran. Hasil penilaian dengan keempat teknik penilaian nontes selain portofolio yang dilakukan guru sepanjang semester pembelajaran suatu mata pelajaran selanjutnya dapat dihimpun sebagai bahan untuk melakukan penilaian portofolio. Dengan kata lain, teknik penilaian portofolio dapat diterapkan dalam rangka penilaian sumatif dengan tujuan menentukan taraf penguasaan murid atas pengetahuan-ketrampilan dalam mata pelajaran tertentu secara komprehensif dan final yang dinyatakan dalam sebuah nilai akhir. Bertolak dari asumsi tersebut, pembicaraan kurang lebih akan difokuskan pada jenis-jenis teknik penilaian nontes selain portofolio.

Dari sudut penyusunan rencana penilaian dan rubrik penskorannya keempat teknik penilaian nontes di luar penilaian portofolio sendiri pada dasarnya dapat dibagi menjadi dua golongan, yaitu: (1) penilaian tertulis, penilaian berbasis wawancara, dan penilaian berbasis produk di satu sisi, dan (2) penilaian berbasis perilaku di sisi lain. Golongan pertama merupakan penilaian yang sesuai untuk melakukan asesmen terhadap hasil belajar dengan *content* berupa pengetahuan dan perasaan mulai dari yang sederhana sampai yang kompleks. Golongan kedua merupakan penilaian yang sesuai untuk melakukan asesmen terhadap hasil belajar dengan *content* berupa ketrampilan dan pola perilaku yang kompleks. Berikut akan dibahas pedoman umum penyusunan masing-masing alat penilaian nontes, termasuk penilaian portofolio.

Menyusun Alat Penilaian Nontes Tertulis

Sejenis kuesioner tertulis dengan pertanyaan terbuka sebenarnya juga sangat sesuai untuk menilai hasil belajar yang kompleks, bukan sekadar mengungkap pendapat atau perasaan murid tentang sesuatu. Misalnya, lewat penilaian tertulis ini murid bisa diminta mengemukakan pendapatnya tentang pentingnya menjaga kebersihan lingkungan dalam mata pelajaran IPA. Lewat pemberian tugas semacam itu, guru bisa menilai sejumlah aspek dari kemampuan murid, seperti keruntutan berpikir, penguasaan bahasa tulis, kerapihan tulisan, dan sebagainya. Namun, karena tugas ini menuntut kemampuan baca dan tulis dari pihak murid, maka teknik ini dipandang kurang sesuai bagi kelompok murid berusia dini atau pelajar dewasa yang buta aksara (Chatterji, 2003). Beberapa pedoman penting dalam menyusun alat penilaian tertulis adalah sebagai berikut (Chatterji, 2003):

1. Tugas tertulis dengan pertanyaan terbuka ini sebaiknya hanya digunakan untuk menilai jenis-jenis hasil belajar tertentu yang tidak akan memadai jika dinilai dengan tes tulis tradisional atau dengan jenis teknik penilaian kinerja lainnya. Jenis-jenis hasil belajar tertentu yang dimaksud meliputi:
 - a. Hasil belajar proses (*process outcomes*) yang sering disebut *complex procedural skills* alias pengetahuan atau ketrampilan prosedural kompleks, seperti menguraikan secara lengkap langkah-langkah membuat karya seni kerajinan batik.
 - b. Hasil belajar luas (*broad outcomes*) yang mencakup sejumlah indikator namun yang harus dinilai bersama-sama secara keseluruhan sekaligus, seperti menulis surat atau menyajikan data dalam sebuah grafik dan menafsirkannya.
 - c. Hasil belajar yang melibatkan kemampuan bernalar, memberikan alasan, memberikan bukti, atau memberikan

- penjelasan tentang sebuah masalah atau isu, seperti menyusun esai tentang pentingnya menjaga kebersihan lingkungan.
- d. Hasil belajar lain yang melibatkan berbagai kemampuan tingkat tinggi khususnya dalam ranah kognitif dan afektif, baik secara sendiri-sendiri maupun sebagai kombinasi.
 2. Petunjuk dan perintah lain harus menggambarkan tugas yang diminta secara jelas. Misal, jika tugasnya adalah membuat sebuah esai, maka petunjuk atau perintahnya harus secara jelas memberikan informasi tentang *panjang* tulisan yang diharapkan, *kriteria penskoran*-nya, *contents* atau isi tulisan yang diharapkan, dan persyaratan lain (misal, batas waktu pengerjaan, sifat tugas yaitu buku terbuka atau buku tertutup, dan sebagainya).
 3. Untuk menjamin reliabilitas penilaian, perlu dikumpulkan bukan hanya satu melainkan beberapa hasil pengerjaan tugas dalam rangka penilaian tertulis.
 4. Guru perlu melengkapinya dengan *rubrik penskoran* dan *anchor* alias jangkar atau sering juga disebut *benchmark*. Jangkar dalam arti *anchor* atau *benchmark* merupakan contoh jawaban yang mewakili berbagai taraf kualitas pengerjaan tugas yang dikaitkan dengan skor tertentu. Jangkar ini akan berfungsi sebagai pedoman dalam penskoran demi menjamin objektivitas dan keajegan dalam pemberian skor (Chatterji, 2003).

Menyusun Alat Penilaian Nontes Berbasis Perilaku

Teknik penilaian ini mengandalkan observasi langsung oleh penilai terhadap perilaku murid *in situ* alias dalam konteks atau situasi nyata-wajar-alamiah atau dalam konteks atau situasi buatan (*simulated contexts*). Berapa pedoman yang perlu diperhatikan meliputi (Chatterji, 2003):

1. Perlu dipastikan bahwa teknik ini sungguh-sungguh sesuai untuk menilai isi/kompetensi dasar, perilaku, atau aspek lain yang menjadi sasaran penilaian.
2. Pemilihan situasi, pemberian perintah dan petunjuk harus benar-benar mampu memunculkan jenis-jenis perilaku yang diharapkan serta memungkinkan untuk diobservasi secara langsung.
3. Untuk menjamin reliabilitas penilaian, observasi harus dilakukan beberapa kali baik secara langsung atau menggunakan alat-alat perekam.
4. Perlu dilengkapi dengan rubrik penilaian yang disusun dengan baik.
5. Jika penilaian dilakukan dengan menggunakan materi atau peralatan khusus, maka perlu disiapkan perintah dan petunjuk lengkap yang jelas sehingga mudah diikuti.

Menyusun Alat Penilaian Nontes Berbasis Produk

Dalam penilaian nontes berbasis produk, murid diminta menciptakan sebuah *work sample* alias contoh hasil pekerjaan atau sebuah produk yang pengerjaannya menuntut penerapan berbagai pengetahuan-ketrampilan yang tercakup dalam hasil belajar dari suatu mata pelajaran tertentu. Lazimnya murid diberi waktu yang leluasa untuk mengerjakan tugas, bisa berupa *take-home projects* alias tugas untuk dikerjakan di rumah, makalah, atau laporan. Beberapa pedoman dalam menyusun alat penilaian nontes berbasis produk ini adalah sebagai berikut (Chatterji, 2003):

1. Perlu dipastikan bahwa teknik ini sungguh-sungguh sesuai untuk menilai isi/kompetensi dasar, perilaku, atau aspek lain yang menjadi sasaran penilaian. Teknik penilaian ini paling sesuai untuk menilai hasil belajar yang menuntut penerapan pengetahuan-ketrampilan prosedural yang kompleks serta penguasaan

teknik-teknik untuk menciptakan produk tertentu, namun untuk mendapatkan evidensi atau bukti atas pencapaian hasil belajar murid tersebut guru tidak perlu melakukan observasi langsung terhadap prosesnya, melainkan cukup menilai produk yang dihasilkan oleh murid. Contoh-contoh produk yang dimaksud meliputi antara lain laporan praktek laboratorium, naskah cerita pendek, naskah puisi, hasil karya seni atau kerajinan.

2. Persyaratan, perintah dan petunjuk pengerjaan tugas harus memberikan batasan yang jelas tentang produk yang diharapkan bisa dihasilkan oleh murid. Petunjuk yang dimaksud meliputi antara lain bentuk produk yang diharapkan, bahan dasar yang boleh digunakan, batas waktu pengerjaan. Karena lazimnya tugas ini dikerjakan di luar kelas bahkan di rumah, guru harus waspada terhadap kemungkinan murid mendapatkan bantuan dari orang lain dalam mengerjakan tugas atau produknya.
3. Untuk menjamin reliabilitas penilaian, murid perlu mengumpulkan beberapa contoh hasil pekerjaan atau produk.
4. Perlu dilengkapi dengan rubrik penilaian yang disusun dengan baik.

Menyusun Alat Penilaian Nontes Berbasis Wawancara

Dalam teknik penilaian ini murid diminta menjawab secara lisan pertanyaan-pertanyaan yang diajukan oleh guru dalam situasi pertemuan tatap muka. Murid dituntut mendemonstrasikan atau menunjukkan penguasaannya atas materi yang dipelajari lewat jawaban yang mereka berikan terhadap pertanyaan wawancara yang diajukan oleh guru. Selain harus mempersiapkan pertanyaan utama, guru juga perlu menyiapkan *probes*, yaitu pertanyaan-pertanyaan lanjutan terkait dengan pertanyaan utama yang mungkin perlu diajukan kepada murid untuk membantunya kembali ke fokus

manakala ada gejala murid mulai menyimpang dari pertanyaan utama (Chatterji, 2003).

Selain itu, karena lazimnya pertanyaan-pertanyaan yang diajukan bersifat terbuka sehingga memberi keleluasaan kepada murid untuk menjawab, maka dalam rangka penskoran guru perlu terlebih dulu melakukan *coding* atau pengodean, yaitu mengidentifikasi konsep-konsep atau tema-tema kunci dalam jawaban murid untuk diberi skor. Beberapa pedoman dalam menyusun alat penilaian nontes berbasis wawancara adalah sebagai berikut (Chatterji, 2003):

1. Perlu dipastikan bahwa teknik ini sungguh-sungguh sesuai untuk menilai isi/kompetensi dasar, perilaku, atau aspek lain yang menjadi sasaran penilaian. Teknik ini sangat sesuai untuk menilai hasil belajar di bidang bahasa, seperti Bahasa Indonesia atau Bahasa Inggris, menilai pemahaman mendalam dan proses penalaran terkait pokok persoalan tertentu sebagaimana lazim berlangsung dalam ujian pendadaran skripsi, tesis, atau disertasi, dan menilai hasil belajar pada umumnya pada subjek anak-anak usia dini. Wawancara informal sering juga efektif digunakan untuk melakukan sejenis penilaian formatif dalam berbagai mata pelajaran (Chatterji, 2003).
2. Situasi wawancara perlu dipersiapkan dan ditata sedemikian rupa sesuai dengan tujuan penilaian. Secara khusus, yang perlu dipersiapkan adalah apa yang disebut "*looks for*", yaitu daftar kata-kata atau tema-tema yang harus dicari dalam jawaban murid. Jika kata-kata atau tema-tema yang diharapkan semacam ini belum muncul dalam jawaban murid, maka guru perlu melakukan *probes*, yaitu pertanyaan-pernyataan yang semakin spesifik mengarah ke kata-kata atau tema-tema yang dicari.
3. Untuk menjamin reliabilitas hasil penilaian, perlu ditanyakan sejumlah pertanyaan terkait dengan hasil belajar tertentu.

4. Perlu dilengkapi dengan daftar pertanyaan, daftar *probes*, pedoman pengodean, dan rubrik penskoran yang dipersiapkan dengan baik.

Menyusun Alat Penilaian Nontes Berbasis Portofolio

Keunggulan teknik penilaian portofolio yang tidak ditemukan pada teknik-teknik penilaian kinerja lainnya antara lain adalah (1) mampu memberikan evidensi atau bukti tentang hasil karya terbaik murid, (2) mampu mendokumentasikan perubahan atau perkembangan dari waktu ke waktu dalam bidang-bidang khusus tertentu, (3) mampu menilai proses dan produk belajar sekaligus, dan (4) mampu melibatkan murid yang bersangkutan dalam proses penilaian lewat refleksi-diri, pemilihan materi portofolio, penilaian-diri, atau tugas-kegiatan metakognitif lainnya (Chatterji, 2003). Beberapa pedoman dalam menyusun alat penilaian nontes berbasis portofolio adalah sebagai berikut (Chatterji, 2003):

1. Tujuan menerapkan teknik penilaian nontes berbasis portofolio harus jelas. Di Bab 4 kita sudah melihat jenis-jenis portofolio, seperti hasil karya terbaik, perkembangan pendidikan murid, hasil karya yang paling favorit, proses dalam pengerjaan proyek berjangka panjang, dan gambaran menyeluruh untuk dibandingkan dengan standar yang disusun sebelumnya. Masing-masing efektif diterapkan untuk mencapai tujuan yang berlainan.
2. Perlu dipastikan bahwa teknik ini sungguh-sungguh sesuai untuk menilai isi, perilaku, atau aspek lain yang menjadi sasaran penilaian.
3. Harus ada pedoman yang jelas terkait hal-hal berikut ini:
 - a. Siapa yang akan merancang penilaian ini?
 - b. Siapa yang akan memilih contoh-contoh hasil karya atau contoh-contoh perilaku?

- c. Siapa yang akan menskor contoh-contoh hasil karya atau contoh-contoh perilaku?
 - d. Apa saja syarat yang harus dipenuhi murid dalam menghasilkan contoh-contoh hasil karya? Di mana tugas itu dikerjakan? Bentuk-bentuk dukungan atau bantuan seperti apa saja yang diperbolehkan dan yang tidak diperbolehkan, dan sebagainya.
4. Untuk menjamin reliabilitas penilaian, perlu dikumpulkan contoh-contoh hasil karya dalam jumlah yang memadai.
 5. Setiap jenis hasil karya berbeda yang dikumpulkan dalam portofolio harus dilengkapi dengan rubrik penskoran masing-masing.

Kendati sudah disinggung di sana-sini, topik tentang rubrik belum dibahas secara spesifik dalam bab ini. Topik itu memang baru akan dibahas dalam bab berikut. Agar pembaca tentang format umum contoh penyusunan alat penilaian nontes sebagai hasil dari langkah-langkah seperti diuraikan di atas lebih bermakna, maka contoh tersebut baru akan disajikan dalam bab berikut sesudah dibahas tentang rubrik penilaian.

Bab 6

Penyusunan Rubrik Penskoran

Salah satu masalah kunci dalam penilaian hasil belajar dengan teknik nontes adalah sistem penskorannya, yaitu cara mentransformasikan atau mengubah informasi-informasi yang berhasil dikumpulkan ke dalam skor atau bilangan yang mencerminkan perbedaan kualitas agar mudah mengolah dan melaporkannya kepada pihak-pihak terkait. Memang perlu kita perhatikan, dari segi informasi yang dihasilkannya seluruh jenis teknik penilaian nontes dapat kita bedakan menjadi dua golongan besar, yaitu (1) teknik nontes yang menghasilkan informasi murni kualitatif berupa rekaman ungkapan murid baik lisan maupun tertulis yang bersifat terbuka, hasil pengamatan perilaku murid yang juga bersifat terbuka, dan produk atau hasil karya sebagai bentuk ungkapan diri atau kemampuan lain yang bukan berupa tulisan dan yang juga bersifat “terbuka”, serta (2) teknik nontes yang menghasilkan informasi yang sudah berupa bilangan, baik merupakan skor maupun sekadar frekuensi atau persentase. Fokus kita pada bagian ini adalah penskoran terhadap informasi hasil penerapan teknik nontes dalam penilaian hasil belajar murid yang masih bersifat murni kualitatif berupa rekaman tulisan, rekaman gambar, atau hasil karya lain yang berupa bukan tulisan.

Rubrik atau Kriteria Penskoran

Sistem penskoran terhadap hasil belajar murid berupa informasi terbuka yang dikumpulkan dengan teknik penilaian nontes disebut *rubrik* alias *kriteria penskoran* (Chatterji, 2003). Rubrik atau kriteria penskoran adalah “*guidelines that we use to ensure that our judgments of open-ended responses are accurate, consistent, and fair*” (Chatterji, 2003, h.234). Artinya, rubrik atau kriteria penskoran merupakan pedoman yang kita ikuti untuk memastikan atau menjamin bahwa penilaian kita terhadap respon atau jawaban murid yang bersifat terbuka kita lakukan secara akurat, konsisten, dan adil.

Penilaian disebut akurat dalam arti bahwa yang kita beri skor sungguh-sungguh unsur-unsur hasil belajar penting yang menjadi tujuan pembelajaran, dan bukan hal-hal lain. Penilaian disebut konsisten dalam arti bahwa kriteria yang kita pakai untuk menilai hasil belajar dari murid ke murid tetap atau sama, tidak berubah-ubah. Dengan begitu penilaian kita juga akan bersifat adil, dalam arti hasil belajar seluruh murid dalam satu mata pelajaran yang sama dinilai pada unsur-unsur yang sama dan dengan menggunakan kriteria penilaian yang sama pula.

Penggunaan rubrik sebagai pedoman penilaian juga menjamin keakuratan dalam arti memungkinkan guru melakukan “*finer discriminations in performances that vary in degrees of quality*” (Chatterji, 2003, h. 234). Artinya, penggunaan rubrik menjamin guru mampu melakukan pembedaan secara lebih cermat atau lebih teliti terhadap variasi kualitas kinerja atau hasil kerja murid.

Cara Menyusun Rubrik

Rubrik lazimnya berupa serangkaian kriteria yang sengaja disusun oleh guru untuk menilai jawaban, kinerja, atau hasil karya murid. Dalam rubrik, guru mendeskripsikan taraf kualitas jawaban,

kinerja, atau hasil karya yang diharapkan mampu dicapai oleh murid. Maka rubrik akan berupa deskripsi atau uraian tentang aneka ciri atau tanda yang harus terdapat dalam jawaban, kinerja atau hasil karya murid serta penempatan masing-masing ciri tersebut dalam sebuah skala. Rubrik perlu disiapkan oleh guru pada awal bahkan sebelum proses penilaian berlangsung, serta harus dikomunikasikan kepada murid. Maka bagi guru rubrik akan berfungsi sebagai kriteria nyata untuk menilai hasil belajar murid, sedangkan bagi murid akan berfungsi sebagai *scaffolding* yaitu semacam pedoman kerja yang akan membantu murid mencapai hasil belajar yang optimal.

Ada dua kategori besar langkah dalam menyusun rubrik, yaitu identifikasi ciri atau kualitas serta penentuan sistem atau cara penskorannya. Secara lebih rinci, kedua kategori langkah yang dimaksud adalah sebagai berikut:

Identifikasi Ciri atau Kualitas

1. Identifikasikanlah atau daftarkanlah semua ciri atau kualitas yang harus terdapat dalam jawaban, kinerja, atau hasil karya murid sebagai hasil belajarnya. Dalam praktek, identifikasi ciri ini dapat kita temukan dalam rumusan indikator-indikator pencapaian tujuan pengajaran. Indikator-indikator ini lazim dijabarkan atau diturunkan secara langsung dari kompetensi dasar, atau melalui rumusan tentang hasil-hasil belajar yang lebih spesifik.
2. Urutkanlah semua ciri atau kualitas tersebut mulai dari yang paling penting sampai dengan yang paling kurang penting.

Penentuan Cara Penskoran

Ada dua kategori cara penskoran terhadap ciri atau kualitas yang dijadikan kriteria penilaian dalam rubrik, yaitu *penskoran analitik* dan *penskoran holistik*, sehingga juga dikenal dua macam rubrik yaitu *rubrik*

analitik dan *rubrik holistik*. Dalam *penskoran analitik* atau *rubrik analitik*, masing-masing ciri atau kualitas yang berhasil diidentifikasi diberi skor sendiri-sendiri, dengan asumsi bahwa masing-masing ciri atau kualitas tersebut memiliki bobot yang sama dalam mencerminkan atau membentuk suatu kemampuan tertentu sebagai hasil belajar. Hasilnya adalah sebuah profil yang menunjukkan, dari antara ciri atau kualitas yang berhasil diidentifikasi mana yang muncul dan mana yang tidak muncul, atau pada ciri atau kualitas mana saja seorang murid unggul dan pada ciri atau kualitas mana saja dia terpuruk. Kita ambil sebagai contoh, jika terhadap kompetensi guru SD sebagaimana disajikan di Bab 5 kita terapkan rubrik analitik, berarti skor harus kita berikan pada masing-masing dari 91 komponen kompetensi yang ada. Jika kemudian rubrik tersebut kita terapkan untuk menilai kompetensi seorang guru SD, maka akan kita peroleh profil kompetensinya yang menunjukkan pada komponen mana saja guru yang bersangkutan menunjukkan kinerja yang unggul, yang sedang, dan yang lemah atau kurang. Inilah kelebihan atau kekuatan rubrik analitik, yaitu mampu memberikan sebuah *profil* dan bukan hanya informasi global tentang hasil belajar murid.

Seperti tersirat dalam uraian di atas, pada cara penskoran analitik ini ada dua cara penentuan skala, yaitu *skala dua titik* (*two-point scale*) seperti pada *check-list* alias daftar cek, dan *skala lebih dari dua titik* (bisa berupa *3-point scale*, *4-point scale*, *5-point scale*, *7-point scale*, atau lebih) seperti pada *rating scale* atau skala penilaian. Contoh penerapannya dalam penilaian kompetensi seorang guru SD melanjutkan contoh yang sudah kita singgung di atas, akan seperti disajikan pada Tabel 6.1. dan Tabel 6.2.

Tabel 6.1: Contoh Penerapan Skala Dua Titik pada Penskoran Analitik

No	Kompetensi guru kelas SD/MI	Ya	Tidak
1	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek fisik.		
2	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek intelektual.		
3	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek sosial-emosional.		
4	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek moral.		
5	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek spiritual.		
6	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek sosial-budaya.		

Keterangan:

- Tanda cek pada kolom “Ya” untuk setiap komponen kompetensi diberi skor = 1.
- Tanda cek pada kolom “Tidak” untuk setiap komponen kompetensi diberi skor = 0.

Tabel 6.2: Contoh Penerapan Skala Lebih dari Dua Titik pada Penskoran Analitik

No	Kompetensi guru kelas SD/MI	0	1	2	3
1	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek fisik.				
2	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek intelektual.				
3	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek sosial-emosional.				
4	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek moral.				
5	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek spiritual.				
6	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek sosial-budaya.				

- Keterangan:**
- 0 = Tidak teramati.
 - 1 = Kadang-kadang teramati.
 - 2 = Sering teramati.
 - 3 = Selalu teramati.

Penskoran holistik juga disebut penskoran global, fokusnya pada jawaban, kinerja, atau hasil karya murid secara keseluruhan bukan pada bagian-bagian atau komponen-komponennya (Chatterji, 2003). Ada dua model penskoran holistik, yaitu *model semi holistik* dan *model holistik murni*. *Penskoran semi holistik* masih memperhatikan aneka ciri atau kualitas yang merupakan unsur pembentuk suatu kemampuan kompleks tertentu. Masing-masing ciri atau kualitas tersebut juga diberi skor sendiri-sendiri, namun dengan asumsi bahwa masing-masing memiliki bobot yang berlainan dalam menentukan kualitas jawaban, kinerja, atau hasil karya murid secara keseluruhan. Maka, sesudah semua ciri atau kualitas yang dipakai sebagai kriteria penilaian berhasil diidentifikasi dan diurutkan menurut penting-tidaknya, langkah penentuan skornya adalah sebagai berikut:

1. Tentukan rentang nilai untuk masing-masing ciri atau kualitas, misal 1-3 atau 1-4 atau rentang lain, dan gunakan rentang yang sama untuk seluruh ciri atau kualitas.
2. Tentukan bobot masing-masing ciri atau kualitas dengan persentase tertentu, sehingga jumlahnya = 100%.
3. Skor untuk masing-masing ciri atau kualitas adalah hasil perkalian antara bobot dan skor yang diberikan untuk ciri atau kualitas yang bersangkutan.
4. Skor final murid adalah jumlah dari skor seluruh ciri atau kualitas. Skor final tunggal ini akan memberikan gambaran global tentang hasil belajar murid sebagaimana tercermin dari hasil penilaian guru terhadap kualitas jawaban, kinerja, atau hasil karyanya.

Contoh penerapan penskoran semi holistik dalam penilaian kompetensi seorang guru SD melanjutkan contoh yang sudah kita singgung di atas, akan seperti disajikan pada Tabel 6.3.

Tabel 6.3: Contoh Penerapan Rubrik Semi Holistik

No	Kompetensi guru kelas SD/MI	Bobot	Nilai (1-4)	Skor
1	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek fisik.	5		
2	Memahami karakteristik peserta didik usia SD yang berkaitan dengan aspek intelektual.	15		
3	Mengidentifikasi potensi peserta didik usia SD dalam	20		
4	Mengidentifikasi kemampuan awal peserta didik dalam lima mata pelajaran SD.	30		
5	Mengidentifikasi kesulitan belajar peserta didik usia SD dalam lima mata pelajaran SD.	30		
		100%		

Seperti tampak pada contoh di Tabel 6.3., pada penskoran semi holistik asesor perlu menentukan skor masing-masing komponen. Caranya, nilai masing-masing komponen yang ditentukan oleh asesor antara 1 sampai 4 dikalikan dengan bobot masing-masing komponen yang bersangkutan. Selanjutnya skor masing-masing komponen tersebut dijumlahkan, sehingga diperoleh *skor total* tunggal yang dipandang mewakili kesan umum tentang kompetensi guru yang bersangkutan terkait empat komponen kompetensi secara keseluruhan. Jadi, yang diutamakan di sini adalah gambaran umum tentang kompetensi guru yang dinilai berdasarkan skor total yang diperolehnya.

Dalam *penskoran holistik murni*, yang menjadi fokus perhatian semata-mata gambaran umum kualitas jawaban, kinerja, atau hasil karya murid secara keseluruhan, praktis tanpa memberi perhatian pada unsur-unsurnya. Gambaran umum kualitas tersebut dibedakan ke dalam beberapa tingkatan, masing-masing tingkatan dikenai skor yang berlainan. Kembali pada contoh penilaian kompetensi guru,

maka salah satu contoh penerapan penskoran holistik murni adalah seperti disajikan dalam Tabel 6.4.

Tabel 6.4: Contoh Penerapan Rubrik Holistik Murni

Deskripsi Kompetensi	Skor
Menunjukkan seluruh unsur kompetensi pedagogik, kompetensi kepribadian, kompetensi sosial, dan kompetensi profesional secara sangat baik.	4
Menunjukkan seluruh unsur kompetensi pedagogik, kompetensi kepribadian, kompetensi sosial, dan kompetensi profesional secara baik.	3
Menunjukkan seluruh unsur kompetensi pedagogik, kompetensi kepribadian, kompetensi sosial, dan kompetensi profesional secara sedang atau cukup.	2
Menunjukkan seluruh unsur kompetensi pedagogik, kompetensi kepribadian, kompetensi sosial, dan kompetensi profesional secara kurang.	1

Dalam rubrik yang menerapkan penskoran dengan lebih dari dua titik dan teristimewa dalam penerapan rubrik holistik yang hasilnya berupa skor global tunggal yang mencerminkan kualitas umum taraf prestasi murid, misalnya “baik”, “sedang”, atau “kurang”, lazimnya masing-masing taraf prestasi disertai dengan sebuah *benchmark* atau baku mutu. *Benchmark* atau baku mutu adalah contoh jawaban, kinerja, atau hasil karya yang berfungsi sebagai standar kongkret untuk menilai jawaban, kinerja, atau hasil karya serupa yang dihasilkan oleh murid. Guru bisa memilih dari antara jawaban, kinerja, atau hasil karya murid untuk dijadikan *benchmark* masing-masing taraf prestasi. *Benchmark* ini selanjutnya dipakai sebagai standar untuk memberikan penilaian terhadap jawaban, kinerja, atau hasil karya murid-murid lainnya.

Kapan Menggunakan Rubrik Analitik dan Rubrik Holistik?

Melihat kelebihan dan kekurangan masing-masing rubrik atau sistem penskoran, muncul pertanyaan apa dasar yang bisa kita pakai untuk memilih menggunakan yang satu atau lainnya? Menurut Chatterji (2003), salah satu dasar pertimbangan penting dalam memilih jenis rubrik yang hendak kita pakai adalah tujuan kita melakukan penilaian. Jika tujuan kita adalah menentukan nilai keseluruhan pada akhir proses pembelajaran dalam rangka penilaian sumatif, kiranya lebih tepat menggunakan rubrik atau sistem penskoran holistik atau global. Sebaliknya, jika tujuan kita adalah melakukan diagnosis tentang kekuatan dan kelemahan masing-masing murid dalam rangka penilaian formatif selama proses pembelajaran masih berlangsung, kiranya lebih tepat menggunakan rubrik atau sistem penskoran analitik.

Contoh format umum penyusunan alat penilaian nontes

Berikut ini disajikan contoh hasil penyusunan alat penilaian nontes berbasis perilaku untuk menilai pencapaian kompetensi dasar dalam ranah afektif dalam bidang budi pekerti, yang diadaptasikan dari Chatterji (2003). Seperti sudah disinggung, langkah-langkah umum penyusunan penilaian nontes ini sudah disajikan dalam Bab 5, namun contoh implementasinya baru disajikan pada bab ini menunggu uraian tentang rubrik penskoran.

Tabel 6.5: Contoh Penyusunan Alat Penilaian Nontes Berbasis Perilaku

Populasi	: Siswa kelas 3 – 5 Sekolah Dasar.
Tujuan	: Sumatif, untuk nilai rapor semesteran bidang budi pekerti.
Kompetensi dasar	: Perilaku bertanggung jawab di sekolah.
Indikator	:

Murid mampu:

Indikator umum:

Menunjukkan sikap bertanggung jawab dalam menyelesaikan tugas-tugas sekolah (Ranah afektif).

Indikator khusus:

- Mampu merumuskan tujuan untuk berbagai tugas sekolah.
- Mampu memusatkan perhatian pada aneka tugas yang diberikan guru di sekolah.
- Mampu mengubah perilaku berdasarkan masukan dari guru atau teman.
- Mampu menggunakan aneka strategi dalam menyelesaikan aneka tugas secara tepat waktu (misal, membuat jadwal kerja, daftar cek tugas, dsb.)
- Mampu menanggapi pujian atau kritikan dengan cara yang bisa diterima oleh guru.
- Memenuhi aneka kewajiban secara tepat waktu.

Metode penilaian: Penilaian berbasis perilaku.

- Observasi terstruktur.
- Catatan anekdot atas aneka peristiwa mencolok.

Kondisi penilaian dan pengadministrasiannya:

- Pada awal semester, guru menjelaskan apa yang diharapkan dari murid terkait budi pekerti.
- Guru menyusun format observasi terstruktur berdasarkan aneka indikator perilaku yang telah ditentukan.
- Guru melakukan observasi antara 2-4 kali per siswa, pada awal, pertengahan, dan menjelang akhir semester.
- Observasi dilakukan di kelas, perpustakaan, di dalam atau luar kelas saat mengerjakan tugas perorangan maupun kelompok.
- Di antara kegiatan observasi pada awal, tengah, dan menjelang akhir semester guru membuat 1-3 catatan anekdot untuk masing-masing murid.

Prosedur penskoran: Setiap indikator akan dinilai dengan 3 kategori penilaian:

- Skor 1: Tidak pernah atau jarang dilakukan.
- Skor 2: Sering dilakukan.
- Skor 3: Selalu dilakukan.

Format observasi :

Nama murid:		Kelas:			
Skala penilaian: 1 : Tidak pernah atau jarang dilakukan. 2 : Sering dilakukan. 3 : Selalu dilakukan. TB : Tidak Berlaku, murid tidak memiliki kesempatan untuk menunjukkan perilaku yang bersangkutan					
		Skala penilaian			
		Minggu 1-4	Minggu 5-8	Minggu 9-12	Minggu 13-16
1.	Mampu merumuskan tujuan untuk berbagai tugas sekolah. Jenis/bentuk tugas sekolah: a. b. c. dst.				
2.	Mampu memusatkan perhatian pada aneka tugas yang diberikan guru di sekolah, meliputi: a. Di kelas. b. Di laboratorium. c. Di tempat bermain. d. dst.				
3.	Mampu mengubah perilaku berdasarkan masukan dari: a. Guru b. Teman				
4.	Mampu menggunakan aneka strategi dalam menyelesaikan aneka tugas secara tepat waktu, meliputi: a. Daftar cek tugas. b. Kalender/jadwal tugas. c. Lain, yaitu:				

5. Mampu menanggapi pujian atau kritikan dengan cara yang bisa diterima oleh guru. a. Pujian dari guru. b. Pujian dari teman.				
6. Memenuhi aneka kewajiban secara tepat waktu.				
7. Catatan anekdot:				
Yogyakarta, Guru kelas				

Catatan: Diadaptasikan dari Chatterji (2003), *op. cit.*, h. 204, 220.

Dari contoh penyusunan alat penilaian nontes berbasis perilaku di atas dapat kita simpulkan minimal dua hal krusial sebagai berikut. Pertama, satuan hasil belajar yang dijadikan sasaran penilaian adalah *kompetensi dasar*. Kedua, butir-butir karakteristik yang dijadikan kriteria dalam penyusunan rubrik penilaian adalah *indikator-indikator* pencapaian tujuan pengajaran yang diturunkan dari kompetensi dasar. Maka, hal paling krusial dalam penyusunan alat penilaian nontes kiranya adalah perumusan kompetensi dasar atau turunannya sebagai satuan tujuan pengajaran serta penjabarannya ke dalam indikator-indikator pencapaian tujuan pengajaran sebagai komponen-komponen kompetensi dasar yang akan dijadikan butir-butir kriteria dalam rubrik penilaian.

Bab 7

Cara Mengolah Hasil Penilaian Nontes

Penerapan berbagai teknik penilaian nontes yang sudah kita bahas pada bab-bab sebelumnya terhadap hasil belajar murid menghasilkan *raw data* atau data mentah. Bisa dikatakan, diperolehnya data mentah tentang hasil belajar siswa ini menandai berakhirnya kegiatan asesmen dalam arti sempit, yaitu pengumpulan informasi tentang hasil belajar siswa. Agar data mentah tersebut bisa diproses lebih lanjut baik dalam rangka penilaian untuk menentukan kualitasnya maupun dalam rangka pelaporan kepada pihak-pihak terkait, maka data mentah itu terlebih dulu perlu diolah. Inti pengolahan yang dimaksud adalah mengubah data mentah menjadi satuan-satuan informasi yang lebih mudah diorganisasikan atau ditata sehingga juga lebih mudah disajikan dan dibaca, maupun juga lebih mudah dimanipulasi dalam arti dikenai berbagai operasi pengolahan data baik secara kualitatif maupun khususnya secara kuantitatif, sehingga lebih mudah ditafsirkan dan dikomunikasikan.

Salah satu faktor penentu cara pengolahan data tentu saja adalah jenis datanya. Dari pengenalan kita terhadap lima jenis metode penilaian nontes yang sudah kita bahas, secara garis besar

dapat dibedakan dua jenis data mentah yang akan kita peroleh dari penerapan kelima metode penilaian nontes tersebut. Pertama, jenis data yang sudah berupa *bilangan* hasil penghitungan frekuensi atau skor yang diperoleh melalui penerapan daftar cek atau skala penilaian maupun penerapan rubrik penskoran. Jenis data ini bisa kita sebut sebagai data kuantitatif yang sudah lebih siap untuk diolah lebih lanjut. Kedua, jenis data yang masih berupa narasi, baik berupa rekaman ungkapan verbal murid hasil penerapan teknik penilaian tertulis atau teknik penilaian berbasis wawancara maupun rekaman tingkah laku hasil penerapan teknik penilaian berbasis perilaku. Jenis data ini bisa kita sebut sebagai data kualitatif, pengolahannya masih menuntut minimal satu langkah tambahan agar selanjutnya bisa dikonversikan atau diubah ke dalam bilangan.

Pengolahan Data Kualitatif

Data kualitatif bisa berupa narasi verbal meliputi transkripsi rekaman ungkapan pikiran dan perasaan dalam wawancara, ungkapan pikiran dan perasaan dalam puisi, prosa, atau esai, rencana kegiatan, laporan kegiatan, deskripsi perilaku atau situasi hasil pengamatan yang diperoleh baik lewat pengamatan langsung maupun dengan perekam audio-visual, maupun berupa produk atau hasil karya nonverbal bisa berupa gambar, bagan atau pola, dan benda.

Pembahasan tentang pengolahan data kualitatif pada bagian ini akan difokuskan pada data berupa narasi verbal, bahkan lebih khusus lagi berupa rekaman ungkapan pikiran dan perasaan dalam wawancara dan deskripsi perilaku hasil observasi atau pengamatan. Ada beberapa alasan yang mendasari pembatasan ini. Pertama, dalam penilaian berdasarkan data kualitatif berupa karangan (puisi, prosa, esai), rencana dan laporan kegiatan, serta hasil karya (gambar, bagan, pola, benda), lazimnya sudah disediakan rubrik penskoran yang mantab, sementara dalam penilaian

berdasarkan data kualitatif berupa narasi verbal hasil wawancara atau observasi lazimnya tidak dilengkapi dengan rubrik penskoran. Kedua, prinsip pengolahan data kualitatif berupa narasi verbal hasil wawancara dan observasi ini diharapkan bisa juga diterapkan dalam pengolahan jenis-jenis data kualitatif lain, sudah barang tentu dengan penyesuaian-penyesuaian seperlunya.

Dalam pengolahan data kualitatif berupa narasi verbal seperti yang dimaksud pada bagian ini, berlaku semacam prinsip bahwa *"data don't speak for themselves"* (Kristi Poerwandari, 2005), artinya bahwa data tidak akan mengungkapkan dirinya, melainkan perlu diolah agar data itu berbicara. Sedangkan inti pengolahan dan analisis data kualitatif sebagaimana dimaksud pada bagian ini adalah *"breaking up, separating, or disassembling ... materials into pieces, parts, elements, or units. With facts broken into manageable pieces, ... sorts and sifts them, searching for types, classes, sequences, patterns, or wholes"* (dalam Kristi Poerwandari, 2005). Artinya, materi atau data mentah tersebut dipecah, dipilah, atau diurai menjadi kepingan, bagian, unsur, atau satuan. Sesudah dipecah menjadi kepingan-kepingan yang mudah digarap, selanjutnya data tersebut dipilah dan disaring untuk menemukan aneka tipe, kelas, sekuensi, pola, atau kesatuan gagasan utuh tertentu. Untuk itu, ada tiga tahap penting dalam pengolahan (dan analisis) data kualitatif berupa narasi verbal sebagaimana dimaksudkan dalam bagian ini, yaitu (1) *coding* atau pengodean, (2) analisis tematik, dan (3) interpretasi.

1. Tahap pengodean

Inti kegiatan pada tahap ini adalah membubuhkan kode pada materi atau data mentah yang diperoleh (Kristi Poerwandari, 2005). Untuk itu, pertama-tama transkripsi verbatim atau kata demi kata dari hasil wawancara atau observasi diketik dalam format yang terdiri

dari tiga kolom. Data mentah diketik di kolom kedua yang terletak di tengah, dari atas ke bawah. Perpindahan dari gagasan yang satu ke gagasan lain atau dari fase wawancara/observasi yang satu ke fase lain bisa ditandai dengan garis pemisah horizontal. Kolom ketiga di sebelah kanan disediakan untuk menuliskan kata kunci atau tema yang berhasil ditemukan dari data mentah di kolom kedua atau untuk menuliskan temuan tentang sekuensi, pola, atau kesatuan gagasan utuh tertentu yang ditemukan dari hasil menghubungkan-hubungkan sejumlah kata kunci atau tema sebagai persiapan untuk merumuskan interpretasi. Kolom pertama di sebelah kiri disediakan untuk menuliskan nomor urut baris-baris transkrip data mentah atau catatan-catatan lain yang dipandang perlu, termasuk hasil refleksi pribadi guru atau asesor tentang data yang diperoleh. Bentuk format yang dimaksud akan seperti yang ditampilkan pada Tabel 7.1.

Tabel 7.1: Format Pengodean dan Analisis Data Hasil Wawancara atau Observasi

Nama :
 Kelas :
 Tanggal :
 Waktu :
 Tempat :

Catatan/ Refleksi	Transkrip Verbatim Hasil Wawancara/Observasi	Analisis Awal
1 2 3 4 5 6 7 8 9 10	<p>Contoh catatan anekdot dari Bab 4: <i>Johny terlambat masuk kelas selama 15 menit. Dia menjelaskan alasan keterlambatannya dengan mengatakan bahwa dia harus menunggu lama giliran mendapatkan pesanan makanan untuk sarapannya di kantin sekolah. Setelah duduk, dia segera mengeluarkan buku dan alat tulisnya dan segera mengerjakan tugas yang diberikan oleh guru. Setelah menyelesaikan 5 dari 8 soal matematika yang ditugaskan oleh guru, dia meletakkan kepalanya di atas meja. Dalam beberapa menit kemudian, dia tertidur.</i></p>	

Kedua, memomori masing-masing baris transkrip data mentah secara urut dan kontinyu dari atas ke bawah. Menurut Kristi Poerwandari (2005), selain penomoran berdasarkan satuan baris, ada pula yang menyarankan melakukan penomoran berdasarkan satuan paragraf, artinya berdasarkan satuan gagasan atau satuan fase wawancara/ observasi. Karena pembagian gagasan yang menjadi dasar pembagian paragraf seringkali tidak mudah diidentifikasi, kiranya lebih aman menggunakan penomoran berdasarkan satuan baris. Ketiga, keseluruhan transkrip sebagai satuan berkas data mentah perlu diberi identitas secara jelas, misal dengan membubuhkan identitas murid yang menjadi sasaran penilaian (nama, kelas, dan sebagainya), serta minimal keterangan tentang waktu dan tempat wawancara atau observasi dilaksanakan.

2. Tahap analisis tematik

Inti kegiatan pada tahap ini adalah melakukan apa yang disebut *open coding* atau pengodean terbuka, yaitu menemukan *kata kunci* atau *tema* dalam data mentah berupa transkrip narasi hasil wawancara atau observasi (Strauss & Corbin, 1990, dalam Kristi Poerwandari, 2005). Kata kunci atau tema bisa muncul atau hadir secara *manifest* atau secara *laten*. Kata kunci atau tema yang manifest adalah kata kunci atau tema yang secara langsung eksplisit muncul dalam narasi sehingga asesor atau penilai tinggal memungutnya dan mencatatnya di kolom yang disediakan. Kata kunci atau tema yang laten tidak muncul secara eksplisit, namun hadir mendasari, membayangi, atau mewarnai narasi secara implisit atau tersembunyi. Asesor atau penilai harus memiliki kepekaan dan kejelian untuk menangkap dan memungutnya.

Upaya menemukan kata kunci atau tema baik yang bersifat manifest maupun laten ini dapat dilakukan secara *deduktif* atau *induktif*

(Boyatzis, 1998, dalam Kristi Poerwandari, 2005). Dalam pencarian kata kunci atau tema secara deduktif, asesor atau penilai sudah memiliki kata kunci atau tema yang akan dicarinya dalam narasi berpegang pada teori yang sudah disiapkan sebelumnya. Misal, saat seorang asesor melakukan penilaian tentang kecenderungan seorang murid untuk berpikir eksternal atau internal mengikuti teori *internal versus external control of reinforcement* dari Rotter (1965) dengan teknik penilaian berbasis wawancara, dia sudah memiliki kategori eksternal dan internal dalam benaknya yang dia deduksikan dari teori Rotter tersebut. Selanjutnya dia tinggal menemukan ungkapan-ungkapan yang secara eksplisit maupun implisit menunjukkan kecenderungan berpikir eksternal atau internal dalam transkrip wawancara dari murid yang bersangkutan.

Sebaliknya, dalam pencarian kata kunci atau tema secara induktif, asesor atau penilai “belum” memiliki kata kunci atau tema apa pun dalam benaknya saat membaca transkrip data mentah penilaian yang dikumpulkan dengan teknik berbasis wawancara. Kata “belum” sengaja ditempatkan dalam tanda petik, sebab lazimnya seorang asesor atau penilai minimal memiliki tujuan tertentu dalam melakukan penilaian. Maka pastilah dia sudah memiliki kategori-kategori tertentu yang hendak dicarinya dalam narasi kendati kategori-kategori tersebut tidak diturunkan dari teori formal tertentu seperti dalam pendekatan deduktif. Namun demikian, tanpa berpedoman pada kategori formal tertentu seorang asesor atau penilai yang menerapkan pendekatan induktif bisa menemukan kata kunci atau tema yang lebih kaya dari upayanya menggali data mentah berupa transkrip hasil wawancara atau observasi.

Sesudah menemukan berbagai kata kunci atau tema, langkah berikutnya adalah mengklasifikasikannya dengan cara memberi *label*, yaitu sejenis nama atau konsep untuk mendefinisikan atau

mendeskripsikan tema yang bersangkutan (Boyatzis, 1998, dalam Kristi Poerwandari, 2005). Dalam memberikan label, asesor atau penilai bisa menggunakan *indigenous concepts*, yaitu label berupa konsep yang berasal dari jawaban atau narasi murid sendiri, atau menggunakan *sensitizing concepts* yaitu label berupa konsep yang dipilih sendiri oleh asesor atau penilai dan lazimnya diambilkan dari teori atau dari khazanah disiplin ilmu tertentu yang sudah dikenal secara luas (Patton, 1990, dalam Kristi Poerwandari, 2005).

Kedua kegiatan utama di atas, yaitu menemukan dan memberi label pada kata-kata kunci atau tema-tema, secara praktis dapat dilakukan dengan mengikuti langkah-langkah sebagai berikut (Kristi Poerwandari, 2005):

- a. Membaca transkrip hasil wawancara atau observasi yang sudah disusun dalam tabel berkolom tiga, secara berulang-ulang untuk memperoleh pemahaman yang baik.
- b. Mulai menemukan kata-kata kunci atau tema-tema, dan hasilnya dituliskan pada kolom ketiga di sebelah kanan.
- c. Membuat catatan-catatan lain berisi interpretasi atau kesimpulan sementara, pada kolom pertama di sebelah kiri.
- d. Mengumpulkan kata-kata kunci dan tema-tema yang berhasil ditemukan dalam sebuah daftar.

3. Tahap interpretasi

Inti kegiatan interpretasi adalah memahami data yang sudah diperas ke dalam kata-kata kunci atau tema-tema di atas, secara lebih meluas dan lebih mendalam. Salah satu cara sederhana dalam menginterpretasikan kata-kata kunci atau tema-tema adalah melanjutkan proses pengodean mengikuti langkah-langkah analisis yang disarankan oleh Strauss dan Corbin (1990, dalam Kristi Poerwandari, 2005). Sesudah diperoleh kata-kata kunci atau tema-

tema sebagai hasil penerapan pengodean terbuka, langkah berikutnya adalah melakukan *axial coding* atau pengodean aksial dan *selective coding* atau pengodean selektif.

Dalam pengodean aksial, asesor atau penilai menyusun kembali kata-kata kunci atau tema-tema dengan cara menemukan hubungan-hubungan antar tema, antara tema dan subtema, dan sebagainya, sehingga diperoleh pola atau pola-pola baru yang lebih padat dalam arti lebih ringkas pula dan lebih bermakna. Dalam pengodean selektif sebagai langkah akhir interpretasi, asesor memilih pola yang paling mendasar sebagai bentuk interpretasi atau kesimpulan akhir. Pola paling mendasar yang layak digunakan sebagai interpretasi atau kesimpulan final semacam ini antara lain akan ditandai oleh tidak adanya *disconfirmatory data* alias data penyangkal atau *kasus negatif* yang bisa melemahkan interpretasi atau kesimpulan final tersebut.

Setiap asesor atau penilai bisa menghasilkan interpretasi yang berbeda terhadap data yang sama. Dalam analisis data kualitatif hal semacam ini bisa dan biasa terjadi. Ada tiga macam pendekatan untuk menguji kesahihan sebuah interpretasi data, sesuai apa yang disebut konteks atau tujuan validasi. Pertama, jika validasinya dilakukan dalam konteks atau tujuan memberikan pemahaman diri kepada murid, maka sebuah interpretasi dipandang valid sejauh hal itu dibenarkan atau diterima oleh murid yang bersangkutan. Kedua, jika validasinya dilakukan dalam konteks atau tujuan memberikan pemahaman kritis yang masuk akal kepada masyarakat, maka sebuah interpretasi dipandang valid sejauh hal itu dibenarkan atau diterima oleh masyarakat luas. Ketiga, jika validasinya dilakukan dalam konteks atau tujuan memberikan pemahaman teoretis kepada kelompok ahli, maka sebuah interpretasi dipandang valid sejauh hal itu dibenarkan atau diterima oleh komunitas ahli terkait (Kvale, 1996, dalam Kristi Poerwandari, 2005).

Dalam penilaian hasil belajar dengan teknik nontes, kiranya yang utama adalah upaya menemukan tema-tema atau pola-pola jawaban dalam transkrip narasi murid hasil penerapan teknik penilaian berbasis wawancara atau berbasis perilaku, baik sebagai hasil *open coding* maupun hasil *selective coding*. Tema-tema atau pola-pola ini kemudian diolah lebih lanjut secara kuantitatif dengan menghitung frekuensi atau jumlahnya, persentase atau proporsinya, dan sebagainya. Begitu pula, jika data yang dihasilkan oleh penerapan teknik penilaian nontes sudah berupa bilangan, baik akibat penerapan rubrik penilaian maupun sebagai hasil skala penilaian, langkah selanjutnya adalah mengolah data tersebut secara kuantitatif sederhana dengan teknik statistik deskriptif. Hal ini membawa kita beralih ke pembahasan tentang pengolahan data kuantitatif.

Pengolahan data Kuantitatif

Data hasil penilaian nontes lazimnya juga berupa bilangan sebab penilaian nontes seringkali tetap melibatkan kegiatan pengukuran baik secara langsung melalui penggunaan *skala penilaian*, maupun secara tidak langsung melalui penerapan rubrik penskoran. Pengolahan data kuantitatif semacam ini memerlukan penerapan teknik statistik, khususnya statistik deskriptif. Namun sebelum membahas teknik statistik yang dimaksud, terlebih dulu akan dibahas secara garis besar seluk-beluk pengukuran.

Seperti sudah disinggung, pengukuran atau *measurement* adalah “*the assigning of numbers to individuals in a systematic way as a means of representing properties of individuals*” (Allen & Yen, 1979, h. 6). Artinya, pengukuran merupakan kegiatan menerangkan bilangan pada murid mengikuti sistem tertentu sebagai cara untuk menunjukkan aneka karakteristik murid. Dalam konteks pembicaraan kita kini, karakteristik yang dimaksud adalah pencapaian hasil belajarnya.

Penerangan bilangan dalam kegiatan pengukuran dilakukan dengan memanfaatkan minimal empat ciri dasar yang bisa melekat dalam bilangan, yaitu (1) *distinctiveness* alias sifat khusus atau unik, maksudnya masing-masing bilangan adalah unik alias berbeda satu sama lain, 1 bukan 2 dan sebaliknya; implikasinya, bilangan yang berbeda akan diterangkan pada murid yang memiliki besaran berbeda terkait karakteristik yang sedang diukur; sifat unik ini dimiliki oleh *semua* bilangan; (2) *ordering in magnitude* alias menunjukkan tingkatan atau urutan besar, khususnya makin tinggi bilangan maka makin tinggi pula besaran yang ditunjukkannya; implikasinya, bilangan yang lebih tinggi akan diterangkan pada murid yang memiliki besaran yang lebih besar terkait karakteristik yang sedang diukur; *tidak semua* bilangan memiliki sifat ini; (3) *equal intervals* alias jarak sama antar bilangan; juga *tidak semua* bilangan memiliki sifat ini; dan terakhir, (4) *absolute zero* alias pemilikan nol absolut atau mutlak sebagai bilangan yang menunjukkan besaran terkecil; juga *tidak semua* bilangan memiliki sifat ini.

Berdasarkan jenis atau jenis-jenis sifat yang dimiliki bilangan yang dipakai dalam kegiatan pengukuran, seorang pakar statistik bernama Stevens (Downie & Heath, 1984) membedakan empat taraf pengukuran, yaitu pengukuran nominal, ordinal, interval, dan rasio, sebagaimana disajikan dalam Tabel 7.2.

Tabel 7.2: Aneka Taraf Pengukuran Berdasarkan Ciri Bilangan

Ciri Bilangan	Taraf Pengukuran			
	Nominal	Ordinal	Interval	Rasio
Sifat khusus atau unik	Ya	Ya	Ya	Ya
Menunjukkan urutan	Tidak	Ya	Ya	Ya
Jarak sama antar bilangan	Tidak	Tidak	Ya	Ya
Nol mutlak sebagai besaran terkecil	Tidak	Tidak	Tidak	Ya

Catatan: Diadaptasikan dari Allen & Yen (1979), *op. cit.*, h. 7.

Taraf pengukuran nominal hanya memanfaatkan ciri sifat khusus atau unik alias identitas bilangan. Bilangan hanya dipakai sebagai *nomen* (bahasa Latin) atau nama untuk membedakan misalnya, murid atau kelompok murid yang satu dari yang lain. Contohnya, guru memanggil murid dengan nomor urut masing-masing dalam daftar nama murid untuk kelas yang bersangkutan. Atau, saat meng-*entry* data hasil belajar murid dalam tabel yang akan diolah dengan aplikasi komputer, guru memberi label bilangan “1” untuk murid lelaki dan label bilangan “2” untuk murid perempuan. Data kuantitatif hasil pengukuran nominal semacam ini hanya bisa dikenai operasi statistik sederhana seperti menghitung jumlah, proporsi, atau persentase kasus dalam setiap kelompok atau kategori (Downie & Heath, 1984).

Taraf pengukuran ordinal memanfaatkan dua ciri bilangan, yaitu sifat khusus atau unik serta sifat menunjukkan urutan pada bilangan. Sifat khusus atau unik atau identitas bilangan sudah kita bahas dan cukup jelas dalam arti *self-evident* alias tidak memerlukan penjelasan tambahan, yaitu 1 bukan 2, 2 bukan 3, dan seterusnya. Sifat menunjukkan urutan sebenarnya juga jelas, yaitu bilangan yang lebih besar menunjukkan kuantitas atau jumlah yang lebih banyak terkait kualitas atau karakteristik tertentu yang sedang menjadi sasaran pengukuran dibandingkan bilangan yang lebih kecil. Salah satu jenis pengukuran ordinal yang paling populer adalah *rank order* alias penjenjangan atau pemeringkatan. Murid yang memiliki kualitas atau karakteristik yang sedang menjadi sasaran pengukuran dalam jumlah yang lebih banyak ditera dengan bilangan yang lebih kecil atau lebih rendah. Sebagai contoh, tiga orang murid kita urutkan berdasarkan tinggi badan mereka mulai dari yang terendah sampai yang tertinggi. Kemudian kepada murid yang paling rendah atau paling pendek badannya kita terakan bilangan 3, yang lebih tinggi kita terakan bilangan 2, dan yang paling tinggi kita terakan bilangan

1. Hasilnya merupakan sebuah *rank order* alias penjenjangan atau pemeringkatan murid berdasarkan tinggi badan mereka.

Pengukuran pada taraf ordinal ini memiliki dua kekurangan. Pertama, kita tidak bisa mengandaikan adanya interval alias jarak yang sama antar bilangan. Selisih tinggi badan antara murid yang terpendek dan yang berbadan sedang bisa jadi hanya 5 sentimeter, sedangkan selisih tinggi badan antara murid dengan tinggi badan sedang dan yang paling jangkung bisa jadi 10 sentimeter atau lebih, namun ketiganya tetap ditempatkan dalam peringkat 3, 2, dan 1. Kedua, pengukuran taraf ordinal tidak mengenal bilangan nol absolut alias mutlak. Tidak mungkin menerangkan bilangan 0 yang berarti tidak memiliki tinggi badan sama sekali. Data hasil pengukuran pada taraf ordinal bisa dikenai operasi statistik sederhana seperti dihitung median, nilai-nilai persentil, dan koefisien korelasi jenjangnya, namun tetap harus ditafsirkan secara hati-hati.

Taraf pengukuran interval menghasilkan data berupa bilangan yang memiliki tiga ciri, yaitu sifat khusus atau unik, sifat menunjukkan urutan, dan memiliki interval atau jarak yang sama antar bilangan. Contoh dalam pengukuran fisik adalah pengukuran suhu dengan menggunakan thermometer. Suhu 36° Celsius di daerah tropis lebih panas dibandingkan suhu 26° Celsius di daerah subtropis, sedangkan suhu 26° di daerah subtropis sendiri lebih panas dibandingkan suhu 16° Celsius di daerah dingin, namun selisih panas suhu di tiga daerah tersebut adalah sama yaitu sebesar 10° . Pengukuran pada taraf interval ini memiliki satu kelemahan, yaitu tidak mengenal bilangan nol absolut atau mutlak. Suhu 0° Celsius bukan berarti tanpa panas sama sekali, hanya panas tersebut terlalu rendah sehingga yang terasa adalah dingin. Data hasil pengukuran interval dapat dikenai sejumlah operasi statistik yang bermakna, meliputi dihitung mean aritmetik, deviasi standar, dan koefisien korelasi *product moment*-nya serta bisa diuji signifikansi perbedaannya dengan *t-test* dan *F-test*.

Taraf pengukuran rasio menghasilkan data kuantitatif yang memiliki empat atau seluruh ciri bilangan. Taraf pengukuran ini secara lazim berlaku pada pengukuran atribut atau benda fisik. Sebagai contoh, andaikan kita memiliki beberapa batang tongkat bambu. Pertama, tongkat-tongkat bambu tersebut dapat kita ukur panjangnya, sehingga masing-masing memiliki ukuran panjang yang khas (*distinctiveness*). Kedua, tongkat-tongkat tersebut dapat kita urutkan berdasarkan urutan panjangnya, tongkat yang lebih panjang kita beri ukuran jenjang yang lebih rendah sedangkan yang lebih pendek kita berikan ukuran jenjang yang lebih tinggi (*rank order*). Ketiga, interval atau jarak satuan ukuran panjang untuk semua tongkat itu adalah sama (*interval* yang sama). Keempat, ukuran nol terletak sama pada pangkal masing-masing tongkat yang berarti tidak atau belum memiliki panjang sama sekali (*nol absolut*). Dalam pengukuran rasio, bilangan sungguh-sungguh menunjukkan rasio atau perbandingan antar benda atau objek pengukuran lain yang diwakilinya. Akibatnya, kita bisa melakukan perbandingan antara benda atau objek pengukuran secara sungguh-sungguh bermakna. Data pengukuran rasio dapat dikenai semua operasi statistik maupun matematik sampai ke tingkat tertinggi.

Pengukuran pada taraf rasio sukar diterapkan dalam pengukuran atribut psikologis, kecuali data kuantitatif atau bilangan yang dihasilkan dipakai sekadar untuk menunjukkan *frekuensi* kejadian sehingga bilangan "0" berarti tidak ada kejadian sama sekali. Misal, guru ingin menilai tingkat kemandirian murid-muridnya dengan menghitung frekuensi atau berapa kali masing-masing murid menangis di sekolah dalam satu minggu. Asumsinya, makin rendah frekuensi menangis seorang murid, semakin mandirilah dia. Maka akan ada murid yang frekuensi menangnya tinggi, rendah, atau bahkan nol yang berarti tidak pernah menangis sama sekali dalam

satu minggu. Siswa yang frekuensi menangisnya 6 kali dalam seminggu berarti 2 kejadian lebih sering dibandingkan siswa yang frekuensi menangisnya 4 kali dalam seminggu, dan 2 kali lebih sering dibandingkan siswa yang frekuensi menangisnya 3 kali dalam seminggu.

Pengolahan data kuantitatif memerlukan penerapan teknik statistik. Namun tidak setiap teknik statistik cocok diterapkan pada semua data yang bisa jadi dihasilkan pada taraf pengukuran yang berlainan. Dengan kata lain, harus ada kesesuaian antara jenis pengukuran pada taraf pengukuran mana data dihasilkan dan teknik statistik yang diterapkan. Tabel 7.3. menyajikan ringkasan tentang taraf pengukuran dan penerapan teknik statistik yang sesuai (Downie & Heath, 1984; Blalock, Jr., 1979; Guilford, 1954).

Tabel 7.3: Aneka Taraf Pengukuran dan Penerapan Teknik Statistik yang Sesuai

Jenis Pengukuran	Teknik Statistik												
	<i>f</i>	<i>p</i>	<i>P</i>	<i>Rasio</i>	<i>Mo</i>	<i>Me</i>	<i>Persentil</i>	<i>Mean</i>	<i>SD</i>	<i>R</i>	<i>r</i>	<i>t</i>	<i>F</i>
Pengukuran nominal	Ya	Ya	Ya	Ya	Ya								
Pengukuran ordinal	Ya	Ya	Ya	Ya	Ya	Ya	Ya			Ya			
Pengukuran interval	Ya	Ya	Ya	Ya	Ya	Ya	Ya	Ya	Ya		Ya	Ya	Ya
Pengukuran rasio	Ya	Ya	Ya	Ya	Ya	Ya	Ya	Ya	Ya	Ya	Ya	Ya	Ya

Statistik Deskriptif

Jenis teknik statistik yang banyak digunakan untuk mengolah data hasil belajar di sekolah adalah *statistik deskriptif*, yaitu jenis statistik yang bertujuan mendeskripsikan aneka karakteristik sebuah distribusi skor (Allen & Yen, 1979). Pendeskripsian data yang dimaksud pada dasarnya dilakukan dengan cara diringkas

atau disederhanakan ke dalam sejumlah bilangan ukuran mengikuti sejumlah konsep tertentu (Blalock, Jr., 1979). Sejumlah konsep penting yang dimaksud akan diuraikan di bagian-bagian berikut ini.

Konstan dan Variabel

Sebuah konstan adalah lambang bilangan tertentu yang tidak berubah sebagaimana telah ditetapkan (Allen & Yen, 1979). Lazimnya lambang yang dipakai adalah huruf kecil Latin n . Misal, guru memutuskan seluruh murid di kelasnya yang berjumlah 40 orang dilibatkan sebagai subjek dalam sebuah penelitian tentang penilaian hasil belajar. Maka, $n = 40$, dan jumlah atau bilangan ini tidak akan berubah selama proses penelitian guru tersebut berlangsung.

Sebuah variabel adalah sebuah karakteristik yang mampu menampakkan perbedaan atau variasi dalam hal besar atau kuantitasnya (Downie & Heath, 1984). Sebuah variabel lazim dilambangkan dengan huruf kecil Latin, misal x . Setiap hasil pengukuran atau observasi atas variabel tersebut dilambangkan dengan huruf besar Latin padanannya, yaitu X . Dengan sendirinya X bisa bervariasi, dalam arti berisi aneka nilai numerik atau bilangan (Allen & Yen, 1979).

Sebuah variabel lain bisa dilambangkan dengan y , sehingga setiap hasil pengukuran atau observasi atas variabel tersebut dilambangkan dengan Y . Seperti X , Y dengan sendirinya juga bisa berisi aneka nilai numerik atau bilangan. Misal, x adalah kemampuan berbahasa Indonesia sedangkan y adalah kemampuan menyelesaikan soal-soal Matematika sebagai hasil belajar dalam masing-masing mata pelajaran dari murid-murid dalam penelitian guru di atas. X dan Y masing-masing murid, yaitu hasil belajar mereka sebagaimana ditunjukkan oleh skor yang mereka capai dalam tes akhir semester masing-masing mata pelajaran terkait, akan berlainan alias berbeda-

beda sesuai besar alias kuantitas atau taraf kemampuan yang berhasil mereka kuasai dalam masing-masing mata pelajaran.

Variabel Kontinyu dan Variabel Diskontinyu

Secara garis besar variabel dibedakan ke dalam dua kategori, yaitu variabel diskret dan variabel kontinyu (Downie & Heath, 1984; Allen & Yen, 1979). Variabel kontinyu adalah variabel atau karakteristik yang bisa menampakkan perbedaan dalam variasi bilangan atau nilai numerik yang tak terbatas, sehingga mampu memberikan pengukuran dengan tingkat presisi atau ketepatan yang semakin optimal. Misal, ukuran meter bisa ditingkatkan ketepatannya menjadi desimeter, sentimeter, milimeter, bahkan lebih tepat lagi, sampai akhirnya tampak sebagai garis lurus karena sedemikian rapat jarak antara bilangan yang satu dengan bilangan lainnya. Akibatnya, setiap bilangan dalam variabel kontinyu selalu memiliki *batas atas* dan *batas bawah*, yang menjadikannya seolah-olah kontinyu alias bersambung dengan bilangan di atas maupun di bawahnya (Downie & Heath, 1984).

Sebaliknya, variabel diskontinyu atau diskret adalah variabel atau karakteristik yang hanya bisa menampakkan perbedaan dalam sejumlah bilangan atau nilai numerik tertentu, khususnya dalam satuan-satuan bilangan utuh. Akibatnya, setiap bilangan dalam variabel diskontinyu atau diskret tidak memiliki batas atas maupun batas bawah yang menghubungkannya dengan bilangan di atas atau di bawahnya (Downie & Heath, 1984).

Penggunaan Bilangan

Menghadapi variabel kontinyu kita akan menemukan bilangan pecahan yang lazimnya dituliskan di belakang koma alias desimal. Jika pecahan ini terlalu “besar” dalam arti bahwa jumlah bilangan

di belakang koma terlalu banyak pastilah merepotkan. Untuk menyederhanakannya, lazimnya dilakukan pembulatan. Prinsip umum dalam melakukan pembulatan bilangan desimal adalah seperti disajikan dalam Tabel 7.4. (Downie & Heath, 1984):

Tabel 7.4: Prinsip Pembulatan Bilangan Desimal

No.	Prinsip	Contoh		
		Desimal	Dibulatkan	
1.	Jika bilangan desimal terakhir < 5, dihapuskan.	7.2	7	
		7.11	7.1	
		0.674	0.67	
2.	Jika bilangan desimal terakhir > 5, bilangan yang tepat mendahuluinya dinaikkan menjadi bilangan tepat di atasnya.	7.8	8	
		7.17	7.2	
		1.098	1.10	
3.	Jika bilangan desimal terakhir = 5, maka:			
		a. Jika bilangan yang tepat mendahuluinya adalah ganjil, maka bilangan tersebut dinaikkan menjadi bilangan tepat di atasnya.	8.875	8.88
			66.975	66.98
		b. Jika bilangan yang tepat mendahuluinya adalah genap, maka bilangan desimal terakhir 5 tersebut dihapuskan.	8.05	8.0
			5.25	5.2

Bilangan Bermakna

Masih terkait bilangan pecahan, sering muncul pertanyaan berapa bilangan atau angka di belakang koma yang perlu kita cantumkan dalam jawaban kita saat mengolah atau menganalisis data yang melibatkan bilangan pecahan desimal? Prinsip umumnya adalah: *Kita cantumkan dalam jawaban kita satu bilangan lebih banyak dibandingkan jumlah bilangan yang terdapat dalam data asli kita* (Downie & Heath, 1984). Misal, skor Bahasa Indonesia sampel yang terdiri dari 3 murid dalam contoh penelitian guru kita di atas adalah 46, 33, dan 21. Artinya, data asli kita terdiri dari serangkaian dua bilangan

bermakna. Maka saat harus melaporkan *mean* atau reratanya cukup kita tambahkan satu bilangan bermakna, sehingga kita tuliskan rerata = 33.3, tidak perlu rerata = 33.333.

Proporsi, Persentase, dan Rasio

Data hasil pengukuran pada taraf nominal bisa diolah dan dilaporkan dalam bentuk *proporsi*, *persentase*, dan *rasio*. **Proporsi** adalah bagian dari sebuah kesatuan atau keseluruhan yang utuh. Proporsi lazim dilambangkan dengan huruf kecil Latin *p*. Sebagai contoh, jika di salah satu kelas I SD ada total 40 murid terdiri dari 24 perempuan dan 16 lelaki, maka proporsi murid perempuan di kelas itu adalah $p_{\text{perempuan}} = 0,6$ (24 dari total murid 40), sedangkan proporsi murid lelakinya adalah $p_{\text{lelaki}} = 0,4$ (16 dari total murid 40). Blalock, Jr. (1979) menegaskan, “*the value of a proportion cannot be greater than unity.*” Maksudnya, nilai sebuah proporsi mustahil melebihi kesatuan utuhnya. Maka, jumlah seluruh proporsi selalu kembali utuh alias sama dengan 1.

Jika sebuah proporsi dikalikan dengan 100, proporsi tersebut terkonversikan menjadi sebuah *persentase*. Istilah persentase berasal dari kata *per cent* yang berarti *per seratus*. Persentase lazim dilambangkan dengan huruf besar Latin *P*. Sebuah proporsi $p = .25$ dikonversikan ke dalam persentase, artinya dikalikan dengan 100, menjadi $P = 25\%$. Jumlah seluruh persentase yang merupakan hasil konversi dari proporsi juga selalu kembali ke kesatuan utuhnya alias sama dengan 100%.

Jika jumlah kasus atau data kecil, yaitu kurang dari 100 buah, persentase menjadi tidak stabil. Sebagai contoh, laporan kejahatan dalam tahun tertentu di sebuah wilayah kepolisian menyajikan data sebagai berikut: jumlah pencurian meningkat 15%, jumlah pemerkosaan meningkat 50%, dan jumlah pembunuhan meningkat

130%. Sepintas, peningkatan tertinggi terjadi dalam jenis kejahatan pembunuhan. Ketika dicermati jumlahnya, ternyata data kasarnya sebagai berikut: jumlah pencurian meningkat dari 70 menjadi 81 kasus; jumlah pemerkosaan meningkat dari 2 menjadi 3 kasus; sedangkan jumlah pembunuhan meningkat dari 9 menjadi 12 kasus. Dengan kata lain, semata-mata mengandalkan persentase bisa menghasilkan informasi yang menyesatkan. Maka ada yang mengusulkan, sebaiknya dihindari menggunakan persentase jika jumlah kasus atau datanya kurang dari 100, dan setiap kali menggunakan persentase sebaiknya selalu disertakan informasi tentang jumlah kasus atau data yang mendasari atau diwakilinya (Downie & Heath, 1984).

Rasio adalah pembagian sebuah bilangan dengan sebuah bilangan lain. Rasio bilangan *A terhadap* bilangan *B* adalah bilangan *A* dibagi bilangan *B*. Kuncinya adalah kata *terhadap*. Dalam rumusan rasio, bilangan yang mendahului kata “terhadap” disebut *numerator*, sedangkan yang mengikutinya disebut *denominator*. Maka, rasio murid lelaki *terhadap* murid perempuan di kelas I pada contoh kita di atas adalah = 16 : 24. Rasio murid perempuan *terhadap* murid lelaki di kelas I adalah = 24 : 16. Biasanya rasio dilaporkan dalam bentuknya yang paling sederhana, dengan cara menghapuskan faktor yang sama-sama terdapat dalam numerator maupun denominator. Maka, rasio murid lelaki *terhadap* murid perempuan = 2 : 3, sedangkan rasio murid perempuan *terhadap* murid lelaki = 3 : 2. Kadang-kadang orang lebih suka melaporkan rasio dengan menggunakan satuan utuhnya sebagai denominator, sehingga rasio murid lelaki *terhadap* murid perempuan = 0,66 (terhadap 1), sedangkan rasio murid perempuan *terhadap* murid lelaki = 1,5 (terhadap 1). Perhitungan rasio menjadi kompleks jika dalam satu satuan keseluruhan terdapat lebih dari 2 kategori seperti contoh di atas. Misal, dari antara 40 murid kelas I yang sama di atas, 26 beragama Islam, 8 beragama Katolik, 4 beragama Protestan, dan 2

beragama Hindu. Menyadari situasi ini Blalock, Jr. (1979) berpendapat bahwa kecuali memang diperlukan, kiranya lebih mudah menghitung dan memahami laporan berupa persentase atau proporsi.

Sampel dan Populasi

Dengan menyinggung tentang populasi dan sampel, sesungguhnya kita telah masuk ke wilayah statistik lain yang disebut *statistik induktif* (Blalock, Jr., 1979). Tujuan statistik induktif adalah "*inferring properties of a population on the basis of known sample results.*" Artinya, membuat kesimpulan tentang sifat-sifat sebuah populasi berdasarkan hasil-hasil pengamatan terhadap sampel yang diketahui. Maka, sering juga disebut *statistik inferensial*. Proses pembuatan kesimpulan semacam ini disebut *statistical inference* alias inferensi secara statistis (Blalock, Jr., 1979; h. 5). Pembahasan hanya akan dibatasi tentang konsep populasi dan sampel.

Populasi atau universum adalah keseluruhan subjek yang memiliki kesamaan minimal dalam satu hal dan yang sedang menjadi fokus perhatian kita. Misal, penelitian tentang hasil belajar siswa yang dilakukan guru di atas merupakan bagian dari sebuah penelitian diagnostik tentang hasil belajar murid kelas V SD yang dilakukan oleh Kementerian Pendidikan dan Kebudayaan secara nasional. Maka, populasi subjek penelitian tersebut adalah seluruh murid kelas V di seluruh SD di Tanah Air. Artinya, setiap murid kelas V di seluruh SD di Indonesia merupakan anggota populasi. Satu-satunya kesamaan yang mempersatukan seluruh subjek yang berjumlah jutaan tersebut sebagai satu populasi adalah status mereka sebagai murid kelas V SD. Nilai-nilai yang mendeskripsikan keadaan populasi atau universum, seperti antara lain rerata, *range* atau rentang skor terendah dan tertinggi dalam hasil belajar mata pelajaran tertentu, disebut *parameter*.

Murid kelas V di SD tempat guru itu mengajar yang terpilih untuk berpartisipasi sebagai subjek dalam penelitian tersebut merupakan salah satu sampel atau contoh dari populasi murid kelas V SD di seluruh Indonesia. Sudah barang tentu, peneliti utama dalam penelitian diagnostik tersebut juga mengambil sampel-sampel lain dari SD-SD lain di kota dan kabupaten lain di seantero Nusantara. Nilai-nilai yang mendeskripsikan keadaan masing-masing sampel (misal rerata dan *range*) disebut *statistik*. Jadi, ada dua pengertian statistik. Dalam arti luas, statistik adalah metode pengolahan data kuantitatif dengan menerapkan prinsip-prinsip matematik. Dalam arti sempit, statistik adalah nilai-nilai yang mendeskripsikan keadaan sampel.

Besar suatu populasi tidak harus berskala nasional dengan anggota berjumlah jutaan seperti dalam contoh di atas. Seandainya di sekolah tempat guru tersebut mengajar terdapat tiga kelas V paralel, penelitian yang dilakukan guru tersebut merupakan penelitian mandiri bukan bagian dari penelitian payung mana pun, dan guru tersebut tidak punya niat untuk menerapkan hasil penelitiannya untuk murid-murid kelas V di sekolah lain, maka keseluruhan murid kelas V di tiga kelas paralel di sekolah tempat guru itu mengajar bisa dia perlakukan sebagai populasi subjek bagi penelitiannya.

Populasi seperti contoh terakhir di atas tidak terlampau besar dan terjangkau oleh guru sebagai peneliti. Menghadapi populasi yang terjangkau semacam itu, idealnya sampel diambil dengan cara *random* atau tepatnya *simple random* alias *random* sederhana. Prinsip dasar pengambilan sampel secara (*simple*) *random* adalah memberikan kesempatan yang sama kepada setiap anggota populasi untuk terambil sebagai bagian atau anggota sampel. Caranya, dibuat dulu sebuah daftar yang memuat nama seluruh murid kelas V dan diberi nomor urut mulai dari nomor 1 sampai nomor ke-*n* sesuai

jumlah total murid. Selanjutnya, entah dengan menggunakan teknik bilangan random atau teknik *fishbowl draw* alias teknik undian, diambil anggota sampel secara individual satu per satu sampai mencapai jumlah sampel yang direncanakan. Sampel yang terbentuk dengan sendirinya akan mencakup murid-murid dari ketiga kelas V paralel di sekolah tersebut. Jika guru mengambil sampel berbasis kelas, apalagi memilih kelas yang diajarnya seperti tersirat dalam contoh di atas, maka sampel yang diperoleh jelas tidak bersifat *random*.

Keunggulan utama pengambilan sampel secara random adalah bahwa jika terjadi bias secara individual yang bersumber dari faktor masing-masing anggota sampel, bias tersebut akan saling meniadakan sehingga hasil pengukuran atau pengamatan yang diperoleh dipandang lebih mendekati keadaan yang sebenarnya. Sebaliknya jika sampel diambil secara tidak random seperti pada contoh penggunaan satu kelas murid sebagai sampel seperti terjadi dalam contoh penelitian yang dilakukan oleh guru di atas, jika kebetulan kelompok kelas tersebut bias ke arah tertentu (misal terdiri atas murid-murid yang mendaftar pada hari terakhir menjelang penerimaan murid baru di sekolah tersebut ditutup) maka hasil pengukuran atau pengamatan terkait hasil belajar mereka bisa tidak jernih. Akibatnya, hasil pengukuran atau pengamatan tersebut tidak layak digeneralisasikan ke populasinya atau tepatnya dalam contoh ini ke kedua kelompok kelas lainnya, sebab sampelnya tidak representatif atau mewakili.

Dalam kasus seperti di atas, lebih baik guru tersebut memperlakukan kelompok kelasnya sebagai populasi. Artinya, penelitian guru tersebut menjadi penelitian populasi, hasilnya hanya diberlakukan bagi kelompok kelas yang bersangkutan, dan tidak perlu digeneralisasikan ke kedua kelompok kelas lainnya. Dengan cara ini tidak ada persoalan terkait populasi dan sampel.

Beberapa nilai yang lazim dipakai untuk mendeskripsikan keadaan populasi atau sampel dalam statistik (dalam arti luas) deskriptif adalah jumlah kasus, *Mean* alias rerata skor, deviasi standar atau simpangan baku yaitu rerata jarak antara masing-masing skor dari *Mean*, dan varians yaitu kuadrat dari simpangan baku. Simbol untuk masing-masing nilai baik sebagai parameter maupun sebagai statistik dalam arti sempit disajikan dalam Tabel 7.5.

Tabel 7.5: Simbol Sejumlah Parameter dan Statistik Deskriptif*

Nilai	Parameter	Statistik
Mean atau rerata	μ (dibaca "mu")	\bar{x}
Deviasi standar (<i>SD</i>)	σ (dibaca "sigma")	<i>s</i>
Varians	σ^2	s^2
Jumlah kasus	<i>n</i>	<i>N</i>

Sumber: Downie dan Heath (1984), *op.cit.*, h. 12.

Distribusi Frekuensi

Jika data kita peroleh dari pengukuran pada taraf interval, pengelompokan data tersebut harus kita lakukan lewat penyusunan *distribusi frekuensi*. *Distribusi frekuensi* menunjukkan frekuensi alias jumlah kali kejadian masing-masing nilai atau skor hasil pengukuran dari sebuah variabel yang sedang menjadi sasaran pengukuran. Jika jumlah variasi skor kecil, *range* atau rentang skor sempit, antara lain karena jumlah sampel kecil, misal $n = 15$, cukup dibuat *distribusi frekuensi tunggal*, yaitu distribusi frekuensi yang didasarkan pada rentang urutan skor tunggal yang mengandung masing-masing skor yang muncul sebagai hasil pengukuran.

Namun jika jumlah variasi skor besar, *range* atau rentang skornya lebar atau luas, antara lain karena jumlah sampel besar, misal $n = 40$ atau bahkan lebih, maka lebih efisien dibuat *distribusi frekuensi bergolong*, yaitu distribusi frekuensi yang didasarkan pada rentang

urutan kelompok-kelompok skor yang mengandung masing-masing skor yang muncul sebagai hasil pengukuran. Kelompok-kelompok skor ini disebut *interval kelas* atau sering disingkat *interval*.

Penyusunan distribusi frekuensi bergolong juga memiliki manfaat lebih banyak dibandingkan distribusi frekuensi tunggal, seperti akan menjadi jelas dalam uraian berikut. Langkah-langkah menyusun distribusi frekuensi bergolong adalah sebagai berikut (Downie & Heath, 1984):

1. Tentukan *range* atau rentang skor, yaitu selisih antara skor tertinggi dan skor terendah.
2. *Range* tersebut dibagi 15, untuk memperoleh estimasi besar atau lebar interval yang efisien.
3. Susunlah daftar interval, mulai dari interval paling bawah. Mulailah interval paling bawah tersebut dengan bilangan yang merupakan kelipatan dari besar intervalnya.
4. Hitunglah frekuensi masing-masing interval dengan *tally*.
5. Tuliskanlah jumlah frekuensi masing-masing interval di bawah kolom *f*.
6. Jumlahkanlah kolom *f* dan tuliskanlah jumlah kasus di baris paling bawah. Jumlah *f* harus sama dengan *N*, namun ini bukan jaminan bahwa perhitungan jumlah frekuensi dengan *tally* sudah benar. Periksalah kembali.

Misal, guru dalam contoh penelitian kelas kita memperoleh data skor kasar hasil tes mata pelajaran Matematika seperti disajikan dalam Tabel 7.6.

Tabel 7.6: Skor Hasil Tes Matematika (N = 40)

56	78	62	37	54	39	62	60
28	82	38	72	62	44	54	42
42	55	57	65	68	47	42	56
56	56	55	66	42	52	48	48
47	41	50	52	47	48	53	68

Sumber: Downie & Heath (1984), *op. cit.*, h. 14.

Hasil penyusunan distribusi frekuensi bergolong dari data di atas dengan menerapkan langkah-langkah seperti dikemukakan oleh Downie dan Heath (1984) adalah seperti disajikan dalam Tabel 7.7. Namun distribusi frekuensi tersebut sudah dilengkapi dengan perhitungan *cumulative frequencies* (*cf*) alias frekuensi kumulatif, serta *cumulative proportions* (*cp*) dan *cumulative percentage*-nya (*cP*).

Tabel 7.7: Distribusi Frekuensi Bergolong Skor Hasil Tes Matematika

Interval	Tally	<i>f</i>	<i>cf</i>	<i>cp</i>	<i>cP</i>
80 - 84	/	1	40	1.000	100
75 - 79	/	1	39	.975	97.5
70 - 74	/	1	38	.950	95.0
65 - 69	////	4	37	.925	92.5
60 - 64	////	4	33	.825	82.5
55 - 59	///// //	7	29	.725	72.5
50 - 54	///// /	6	22	.550	55.0
45 - 49	///// //	6	16	.400	40.0
40 - 44	///// //	6	10	.250	25.0
35 - 39	///	3	4	.100	10.0
30 - 34		0	1	.025	2.5
25 - 29	/	1	1	.025	2.5
		$N = \sum f = 40$			

Persentil

Pada kolom cP dalam Tabel 7.7. kita bisa mengetahui persentase murid yang memperoleh skor Matematika sampai interval skor tertentu. Sebagai contoh, ada 25% dari jumlah murid, tepatnya 10 orang, yang memperoleh skor dalam interval 40-44 ke bawah. Namun kita tidak tahu berapa tepatnya skor yang memisahkan 25% murid dari 75% murid sisanya. *Persentil* atau *titik persentil* adalah “*a specific point in a distribution that has a given percent of the cases below it*” (Downie & Heath, 1984). Artinya, persentil atau titik persentil adalah titik berupa skor khusus dalam sebuah distribusi frekuensi yang memiliki persentase tertentu dari keseluruhan jumlah subjek atau kasus yang berada di bawahnya, yaitu sebesar titik persentil yang bersangkutan. Sebagai contoh, persentil 25 (C_{25} alias 25^{th} centile) adalah skor dalam sebuah distribusi frekuensi yang memiliki 25% dari seluruh jumlah kasus yang disajikan dalam distribusi frekuensi berada di bawah skor yang bersangkutan.

Sebagai contoh dan seraya memperhatikan Tabel 7.7. kita akan menghitung Persentil 50 (C_{50}), yaitu skor yang membagi seluruh jumlah kasus menjadi 50% di atas dan 50% di bawahnya. Jumlah murid atau kasus dalam contoh di atas adalah $N = 40$ orang. Maka, 50% dari jumlah murid = $50\% \times N = 20$ orang. Kita simak kolom cf atau frekuensi kumulatif dari bawah ke atas sampai ke cf yang mengandung atau memuat 20 kasus. Ternyata, jumlah tersebut terletak di dalam interval 50-54. Interval tersebut mengandung frekuensi = 6. Padahal di bawah interval tersebut sudah ada 16 kasus, berarti kita hanya memerlukan tambahan 4 kasus dari antara 6 kasus yang terdapat dalam interval 50-54 untuk mendapatkan dua belahan yang sama besar. Bagaimana cara menentukan skor yang memuat tambahan 4 kasus seperti yang kita perlukan? Caranya dengan melakukan interpolasi, yaitu mengambil 4 dari 6 kasus yang terdapat dalam interval 50-54 yang memiliki panjang interval = 5.

$$C_{50} = B_b + \left[\frac{\frac{50}{100} N - cfb}{fd} \right] i$$

$$C_{50} = 49.5 + \left[\frac{20-16}{6} \right] 5$$

$$= 49.5 + \left[\frac{4}{6} \right] 5$$

$$= 49.5 + 3.33$$

$$= 52.83$$

$$= 52.8$$

Jadi, rumus perhitungan titik persentil, jika dihitung dari bawah seperti contoh di atas, adalah seperti disajikan dalam Rumus 7.1. (Downie & Heath, 1984):

Rumus 7.1: Perhitungan titik persentil

$$C_x = B_b + \left[\frac{\frac{x}{100} N - cfb}{fd} \right] i$$

Keterangan rumus:

C_x = titik persentil yang dicari

B_b = batas bawah interval yang mengandung titik persentil.

N = jumlah kasus.

c_{fb} = frekuensi kumulatif kasus di bawah interval tempat dilakukan interpolasi

f_d = frekuensi interval tempat dilakukan interpolasi

i = besar atau lebar interval kelas

C_{50} adalah *Median*, dan titik ini bisa kita pakai jika hendak membagi distribusi frekuensi menjadi dua kategori atau kelompok.

Jika kita hendak membagi distribusi frekuensi ke dalam empat kategori atau kelompok yang sama besar, maka kita bisa menggunakan tiga titik *kuartil* sebagai nilai batasnya, yaitu Q_1 alias C_{25} , Q_2 alias C_{50} atau *Median*, dan Q_3 alias C_{75} . Jika kita hendak membagi distribusi frekuensi ke dalam 10 kategori atau kelompok yang sama besar, kita bisa menggunakan 9 titik *desil* sebagai nilai batasnya, yaitu D_1 alias C_{10} , D_2 alias C_{20} dan seterusnya sampai dengan D_9 alias C_{90} . Distribusi persentil dalam satuan desil akan berbentuk empat persegi panjang seperti disajikan pada Tabel 7.8.

Tabel 7.8: Distribusi Persentil Berbentuk Empat Persegi Panjang

10%	10%	10%	10%	10%	10%	10%	10%	10%	10%
C_{10}	C_{20}	C_{30}	C_{40}	C_{50}	C_{60}	C_{70}	C_{80}	C_{90}	

Seorang guru bisa tergoda untuk menggunakan distribusi persentil ini sebagai norma untuk mengonversikan skor mentah ke dalam rentangan nilai 1 sampai dengan 10. Jika benar hal ini akan dilakukan, guru harus berhati-hati memperhatikan bentuk distribusinya khususnya dengan syarat bentuk distribusi tersebut mendekati bentuk simetris dengan frekuensi tertinggi kira-kira di tengah dan semakin rendah atau kecil di kedua ujungnya. Syarat ini penting sebab persentil merupakan satuan pengukuran yang tidak setara (Downie & Heath, 1984). Sebagaimana akan kita lihat, salah satu jenis norma dalam penilaian beracuan norma memang disusun berdasarkan distribusi persentil semacam ini. Maka, kiranya guru perlu memiliki pemahaman yang baik tentang persentil (termasuk desil, kuartil, dan median) berikut cara perhitungannya.

Tendensi Sentral

Tendensi sentral adalah gejala kecenderungan data atau skor memusat ke titik atau nilai tengah distribusi. Ada yang mendefinisikannya sekadar sebagai *rerata*, yaitu “*a single figure that stands for or represents a group of figures*” (Downie & Heath, 1984; h. 35). Artinya, sebuah bilangan tunggal yang mewakili sekelompok bilangan. Ada tiga jenis rerata yang lazim dipakai untuk mendeskripsikan data, yaitu *mean* (\bar{x}), *median* (*Md*), dan *mode* (*Mo*).

Mean. Salah satu jenis *mean* yang dipakai secara luas adalah *mean aritmatik*, yaitu rerata sebagaimana lazim kita pahami dalam kehidupan sehari-hari. *Mean* memiliki sejumlah ciri khas sebagai berikut:

Andaikan dari data di Tabel 7.6. kita ambil enam skor dengan komposisi seperti disajikan dalam Tabel 7.9.

Tabel 7.9: Contoh Enam Skor dari Data di Tabel 7.6

62	60	58	56	54	52
----	----	----	----	----	----

$$\begin{aligned}
 \text{Mean} &= \frac{62 + 60 + 58 + 56 + 54 + 52}{6} \\
 &= \frac{342}{6} \\
 &= 57
 \end{aligned}$$

Jika skor-skor dan *mean* tersebut kita tempatkan dalam sebuah garis *kontinum*, maka akan tampak sebagai berikut:

62	60	58	57	56	54	52

Terlihat bahwa *mean* terletak tepat di tengah garis kontinum dan membagi rangkaian skor tersebut ke dalam dua bagian yang persis sama. Jumlah selisih antara masing-masing skor dengan *mean* baik yang terletak di bawah maupun di atas *mean* adalah sama, yaitu $5+3+1 = 1+3+5$. Situasi semacam ini disebut berada dalam *ekuilibrium*, dengan *mean* sebagai pusat gravitas distribusi. Inilah ciri khas pertama *mean*, yaitu dalam jenis distribusi skor tertentu, *mean* terletak tepat pada *fulcrum* (bahasa Latin, berarti kaki atau landasan) dan berfungsi sebagai pusat gravitas distribusi skor. Hal ini membawa kita ke ciri khas kedua dan ketiga *mean*.

Jika data yang sama di atas kita tempatkan dalam tabel seraya kita tambahkan data deviasi masing-masing skor dari *mean* ($x = X - \text{Mean}$) dan kuadrat masing-masing deviasi, maka akan kita peroleh tampilan data seperti pada Tabel 7.10.

Tabel 7.10: Data Enam Skor Beserta Deviasi Masing-masing dari Mean dan Kuadratnya

X	x	x ²
62	+5	25
60	+3	9
58	+1	1
56	-1	1
54	-3	9
52	-5	25
$\Sigma = 342$	$\Sigma = 0$	$\Sigma = 70$

Keterangan:

x = jarak atau selisih masing-masing skor X dari *mean*.

Dari Tabel 7.10. di atas tampak dua hal. Pertama seperti terlihat pada kolom kedua, dalam jenis distribusi skor seperti di atas, jumlah jarak atau selisih masing-masing skor dari *mean* adalah = 0. Inilah pula definisi matematis dari *mean*, yaitu “*that point in a distribution about*

which the sum of the deviations is zero" (Downie & Heath, 1984; h. 38). Artinya, *mean* adalah titik dalam sebuah distribusi di mana jumlah deviasi atau selisih antara masing-masing skor di sekitarnya dengan *mean* tersebut adalah nol. Inilah ciri khas kedua *mean*.

Kedua, kuadrat masing-masing deviasi ditempatkan pada kolom ketiga. Jika dijumlahkan, maka jumlah kuadrat masing-masing deviasi ini akan merupakan "jumlah kuadrat terkecil" yang akan diperoleh jika dibandingkan dengan jumlah kuadrat yang mungkin diperoleh dari selisih skor mana pun lainnya dalam distribusi skor yang bersangkutan. Jadi, 70 merupakan *jumlah kuadrat terkecil* yang akan diperoleh dari sekalian jumlah kuadrat yang mungkin dihitung dari distribusi skor di atas. Inilah ciri khas ketiga *mean*.

Median. *Median* (dilambangkan *Md*) adalah titik dalam distribusi yang memisahkan skor-skor ke dalam dua paroh (50%) yang persis sama, masing-masing di bawah dan di atas *median*. Dengan kata lain, *median* adalah titik tengah setiap distribusi, dan sebagaimana sudah kita lihat, *median* adalah juga C_{50} alias *Persentil* ke-50, D_5 alias *Desil* ke-5, atau Q_2 alias *Kuartil* ke-2. Cara menghitung titik *median* mengikuti rumus perhitungan *persentil*. Seperti *Persentil*, *Desil*, dan *Kuartil*, *Median* sering juga dijuluki *positional measures* alias ukuran penunjuk posisi, sebab memang berfungsi menunjukkan posisi sebuah kasus atau seseorang dibandingkan kelompoknya.

Mode. *Mode* (dilambangkan *Mo*) adalah nilai yang paling sering muncul dalam sebuah distribusi. Pengertian tersebut secara khusus berlaku dalam distribusi frekuensi tunggal. Dalam distribusi frekuensi bergolong, *mode* adalah titik tengah interval yang memiliki frekuensi paling banyak. Pada contoh distribusi frekuensi yang disajikan di Tabel 7.7. *mode*-nya adalah titik tengah interval (55-59), yaitu 57.

Sedikit catatan, ketiga ukuran tendensi sentral tersebut memiliki sifat yang berlainan, khususnya dari segi sifat stabilnya. Berdasarkan

pengalaman, penggunaan ketiga ukuran tersebut akan memberikan gambaran yang mendekati akurat apabila mengikuti pedoman sebagai berikut (Downie & Heath, 1984):

1. *Mode* lebih cocok digunakan pada data nominal atau bisa juga pada setiap jenis distribusi jika terdesak oleh waktu.
2. *Median* lebih cocok digunakan pada data ordinal atau yang bertaraf lebih tinggi, khususnya jika distribusi datanya tidak normal.
3. *Mean* lebih cocok digunakan pada data interval atau rasio, khususnya jika distribusinya bersifat simetris atau normal.

Variabilitas

Variabilitas adalah gejala kecenderungan skor menyebar di sekitar, artinya di bawah dan di atas, *mean*. Pentingnya memeriksa variabilitas selain tendensi sentral skor, bisa ditunjukkan lewat contoh sebagaimana dikutip dari Downie dan Heath (1984) berikut. Misal, seorang guru memiliki data skor ulangan Matematika murid-murid kelas IV di tiga kelas paralel. Ringkasan data skor Matematika tiga kelompok murid kelas IV tersebut disajikan di Tabel 7.11.

Tabel 7.11: Rekapitulasi Skor Ulangan Matematika Tiga Kelas IV SD Paralel

Kelas	X_R	X_T	Mean
IVA	20	160	90
IVB	80	100	90
IVC	50	130	90

Keterangan:

X_R = Skor terendah.

X_T = Skor tertinggi.

Dari Tabel 7.11. tampak, kendati *mean* atau rerata skor ketiga kelas paralel tersebut sama, namun distribusi skor masing-masing

sebagai kelompok berlainan sebagaimana terlihat dari perbedaan skor terendah dan skor tertinggi di antara mereka. Hal ini menunjukkan bahwa kemampuan Matematika murid di tiga kelas paralel tersebut berlainan, kendati skor rerata mereka sama. Maka, dalam mencermati distribusi frekuensi sebuah kelompok selain memperhatikan gejala tendensi sentralnya juga penting memperhatikan gejala variabilitas atau dispersinya. Ada tiga ukuran variabilitas penting yang lazim dipakai untuk mendeskripsikan distribusi frekuensi, yaitu *range*, *standard deviation*, dan *variance*.

Range. Menurut pengertian umum, *range* adalah selisih antara skor tertinggi dan skor terendah dalam distribusi. Dinyatakan dalam rumus,

$$\text{Range} = (X_T - X_R)$$

Namun ada yang memberikan definisi yang agak berbeda, yaitu bahwa *range* adalah “*the high score minus the low score plus one of the smallest units of measurement*” (Downie & Heath, 1984; h. 45). Menurut pengertian kedua, maka,

$$\text{Range} = (X_T - X_R) + 1$$

Pengertian atau rumus mana pun yang diikuti, *Range* sebagai ukuran variabilitas atau dispersi skor bersifat sangat labil atau tidak reliabel. Perubahan signifikan skor pada salah satu atau kedua ujung distribusi akan mengubah *range* secara signifikan pula. Untuk menstabilkan atau menjadikannya lebih reliabel, sering orang menghilangkan 5% skor yang terletak di kedua ujung distribusi. Dengan hanya menggunakan 90% skor sisa dalam distribusi

frekuensi, lazimnya *range* akan menjadi lebih stabil atau reliabel (Downie & Heath, 1984).

Standard deviation. *Standard deviation* atau deviasi standar yang lazim disingkat *SD* atau dilambangkan *s* adalah akar rerata deviasi skor dari *mean* yang dikuadratkan. Deviasi skor dari *mean* perlu dikuadratkan sebab sudah kita lihat, khususnya jika distribusinya bersifat simetris atau normal, maka jumlah deviasi skor dari *mean* = 0. Dinyatakan dalam rumus, deviasi standar adalah sebagai berikut:

$$s = \sqrt{\frac{\sum x^2}{N}}$$

Varians. Jika dikuadratkan, *standard deviation* menjadi *variance* atau diindonesiakan menjadi *varians*. Dinyatakan dalam rumus, *varians* adalah:

$$s^2 = \frac{\sum x^2}{N}$$

Ketiga ukuran variabilitas atau dispersi skor dalam sebuah distribusi tersebut akan menunjukkan homogenitas versus heterogenitas kelompok. Jika *Range*, *s*, atau s^2 tinggi atau besar, berarti kelompok yang bersangkutan bersifat *heterogen*. Artinya, variasi skor dalam kelompok tersebut tinggi. Sebaliknya, jika *range*, *s*, atau s^2 rendah atau kecil, berarti kelompok yang bersangkutan *homogen*, dalam arti bahwa variasi skor dalam kelompok tersebut rendah.

Kurve Normal

Terkait berbagai gejala atau atribut fisik maupun psikologis murid seperti berat badan dan tinggi badan, skor hasil belajar pada berbagai

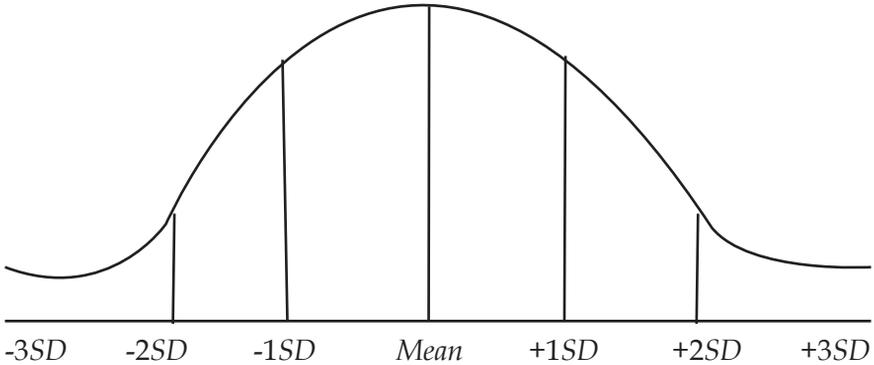
mata pelajaran, taraf inteligensi umum, taraf motivasi belajar, dan sebagainya, jika kita berhasil mengumpulkan data dari populasi murid dalam jumlah besar, katakanlah ribuan, maka distribusi frekuensi skor atau nilai untuk masing-masing atribut tersebut jika disajikan dalam sebuah grafik akan membentuk suatu *kurve normal*. Kurve semacam ini sering juga disebut *kurve berbentuk genta* (Sutrisno Hadi, 2001b) sesuai bentuknya, atau juga disebut *kurve DeMoivre* sesuai nama pakar yang pertama kali menemukannya pada 1733, atau *kurve Gaussian* sesuai nama pakar yang menemukan manfaat praktisnya untuk mengolah dan menganalisis data (Guilford, 1954).

Kurve normal memiliki sejumlah ciri, namun kita hanya akan membatasi menyinggung dua ciri yang lazim dimanfaatkan untuk menyusun norma dalam penilaian beracuan norma. Kedua ciri yang dimaksud adalah sebagai berikut (Downie & Heath, 1984; Sutrisno Hadi, 2001b; Blalock, Jr., 1979).

1. Ordinat atau garis vertikal maksimum alias ketinggian maksimum kurve normal terletak pada *mean* dan berada tepat di tengah aksis atau garis horisontal distribusi. Semua ordinat lainnya, baik di bawah (di sebelah kiri) atau di atas (di sebelah kanan) *mean*, lebih pendek dari ordinat maksimum atau *mean* dan terdistribusi secara kontinyu dan simetris dengan sejenis ekor di kiri dan kanan *mean* sehingga menyerupai bentuk genta.
2. Kurve normal bersifat asimtomatik, artinya kedua ekor kurve normal sebenarnya tidak akan pernah menyentuh aksis melainkan berlanjut *ad infinitum* alias tanpa batas ke kedua arah. Namun dalam praktek, tiga *SD* di masing-masing dari kedua sisinya, yaitu di sebelah kiri dan kanan *mean*, akan mencakup seluruh kasus atau data.

Berdasarkan kedua ciri di atas, kita bisa mengetahui bahwa sebuah kurve normal terbagi ke dalam sejumlah wilayah dengan

mean sebagai jangkar di tengah dan *SD* sebagai satuan pembagiannya, seperti kurang lebih disajikan pada Gambar 7.1.



Gambar 7.1. Daerah di Bawah Kurve Normal pada Berbagai Satuan *SD* dari *Mean*.

Wilayah yang terbentang antara (*Mean-1SD*) dan (*Mean+1SD*) mencakup 68,26% dari kurve atau sekitar 2/3 dari seluruh kasus atau data. Wilayah yang terletak antara (*Mean-1SD*) dan (*Mean-2SD*) maupun antara (*Mean+1SD*) dan (*Mean+2SD*) masing-masing mencakup 13,59% dari kurve atau sekitar 1/7 dari seluruh kasus atau data. Sedangkan wilayah yang terletak antara (*Mean-2SD*) dan (*Mean-3SD*) maupun antara (*Mean+2SD*) dan (*Mean+3SD*) masing-masing mencakup 2,15% dari kurve atau sekitar 1/46 dari seluruh kasus atau data. Sebagaimana sudah disinggung dan akan kita lihat, pembagian wilayah di bawah kurve normal dengan satuan jarak *SD* dari *Mean* seperti disajikan di atas lazim dipakai sebagai salah satu dasar menyusun norma dalam penilaian beracuan norma.

Pembahasan tentang pengolahan data hasil penilaian nontes baik secara kualitatif maupun kuantitatif kita cukupkan sampai di sini.

Bab 8

Penentuan dan Pelaporan Nilai Akhir

Sesudah memiliki sejumlah data hasil asesmen hasil belajar murid yang dikumpulkan dengan berbagai metode asesmen baik tes maupun nontes dan yang dilakukan pada berbagai kesempatan selama semester kegiatan belajar-mengajar berlangsung, pada akhir semester kegiatan belajar-mengajar guru harus menentukan nilai akhir masing-masing murid untuk setiap mata pelajaran dan melaporkannya dalam buku rapor.

Penentuan Nilai Akhir

Langkah pertama yang harus ditempuh guru tentu saja adalah menentukan *grade* atau nilai akhir masing-masing murid pada setiap mata pelajaran yang diberikan. Menurut Shermis dan Di Viesta (2011), *grade* atau nilai akhir semacam ini memiliki aneka manfaat bagi berbagai pihak, yaitu:

1. *Credit point*. Nilai akhir suatu mata pelajaran berperan sebagai sejenis sertifikasi atau keterangan untuk memberikan bukti atau kesaksian tentang dua hal, yaitu (a) bahwa murid telah menempuh mata pelajaran tersebut, dan (b) taraf penguasaan yang dicapai

oleh murid atas materi atau kemampuan yang dipelajari dalam mata pelajaran tersebut.

2. *Umpan balik bagi guru dan murid.* Nilai akhir suatu mata pelajaran bermanfaat memberikan umpan balik kepada guru tentang sejauh mana pembelajaran yang dilakukannya (pemilihan materi, alokasi waktu untuk mengolah setiap materi, strategi pembelajaran yang diterapkan, dan sebagainya) berlangsung efektif, serta memberikan umpan balik kepada murid tentang efektivitas usaha belajarnya.
3. *Umpan balik bagi orang tua murid.* Orang tua perlu mengikuti perkembangan belajar anak mereka. Berdasarkan umpan balik yang diterima mereka bisa melakukan langkah-langkah baik secara sendiri maupun bekerja sama dengan guru untuk melakukan perbaikan, seandainya hasil belajar sang anak kurang sesuai dengan harapan mereka.
4. *Umpan balik bagi para administrator pendidikan baik di dalam maupun di luar sekolah.* Umpan balik yang diterima bisa dipakai sebagai dasar untuk melakukan perbaikan kebijakan pendidikan, berupa revisi kurikulum, pelatihan guru di bidang perbaikan strategi pembelajaran, perbaikan strategi penilaian hasil belajar, dan sebagainya.
5. *Dasar pertimbangan dalam pemilihan jurusan sekolah.* Pola sebaran tertentu dalam nilai rapor seringkali efektif dipakai sebagai dasar untuk membantu orang tua memilihkan jurusan yang paling sesuai dengan kondisi murid.
6. *Dasar pengambilan keputusan dalam penerimaan murid baru.* Penerimaan murid baru di jenis dan jenjang sekolah yang lebih tinggi seringkali juga efektif didasarkan pada pola sebaran tertentu dalam nilai rapor.

7. *Dasar pertimbangan untuk memberikan hadiah dan/atau penghargaan.* Sekolah lazim memberikan hadiah atau predikat pujian bagi murid-murid yang mencapai nilai tinggi.
8. *Dasar pertimbangan dalam seleksi untuk berbagai program kegiatan.* Penentuan murid untuk dilibatkan dalam aneka program kegiatan tertentu, misal penyelenggaraan kelas akselerasi, seringkali juga efektif didasarkan pada pola sebaran tertentu dalam nilai rapor yang dicapai murid di tingkat kelas sebelumnya.
9. *Dasar untuk melakukan konseling akademik dan karir.* Nilai rapor lazim dipakai sebagai dasar memberikan konseling akademik, misal pemilihan jurusan pendidikan, atau konseling karir, misal pemilihan jenis pendidikan lanjutan.
10. *Sarana untuk memberikan motivasi belajar.* Nilai akhir mata pelajaran bisa berfungsi sebagai sumber motivasi ekstrinsik untuk memacu murid belajar lebih giat lagi jika belum baik, atau mempertahankan tingkat belajarnya jika sudah baik.
11. *Dasar pertimbangan dalam seleksi calon pegawai.* Nilai rapor seringkali juga dipakai sebagai dasar penerimaan calon pegawai.

Jenis Komponen Penilaian sebagai Dasar Penentuan Nilai Akhir

Dalam penentuan nilai akhir suatu mata pelajaran, guru pasti akan mendasarkan pada hasil penilaian pengerjaan berbagai tugas-ulangan-ujian yang dilakukan selama semester pembelajaran berlangsung. Hasil penilaian guru terhadap pengerjaan aneka tugas, ulangan, ujian oleh murid selama berlangsungnya kegiatan pembelajaran tersebut merupakan komponen penilaian yang dipandang mencerminkan perubahan perilaku murid sebagai bukti pencapaian tujuan pengajaran. Namun tingkat relevansi atau keabsahan aneka komponen penilaian tersebut sebagai indikator pencapaian tujuan pengajaran sesungguhnya tidak sama, sehingga

guru perlu selektif demi mendapatkan penilaian yang sah, tepat dan adil.

Secara garis besar aneka komponen penilaian yang sering digunakan guru sebagai dasar penentuan nilai akhir mata pelajaran dapat digolongkan menjadi tiga kategori berdasarkan relevansi atau keabsahannya (Haladyna, 1999, dalam Shermis & Di Viesta, 2011), yaitu: (a) bisa dipertanggungjawabkan, artinya jenis-jenis komponen penilaian dalam kategori ini sah digunakan sebagai dasar penentuan nilai akhir mata pelajaran, (b) patut dipertanyakan, artinya jenis-jenis komponen penilaian dalam kategori ini bisa menimbulkan masalah seperti ketidaktepatan atau ketidakadilan jika digunakan sebagai dasar penentuan nilai akhir mata pelajaran; sebagai contoh, keaktifan di kelas yang ditandai dengan mengajukan banyak pertanyaan, bisa jadi tidak mencerminkan penguasaan murid atas materi pelajaran melainkan lebih mencerminkan sifat kepribadiannya yang terbuka, dan sejenisnya; jika hal ini dipakai sebagai kriteria, maka murid lain yang sekalipun pandai namun memiliki sifat pendiam jelas akan dirugikan, dan (c) tidak bisa dipertanggungjawabkan, artinya jenis-jenis komponen penilaian dalam kategori ini tidak sah dan akan menimbulkan ketidakadilan atau masalah lain jika dipakai sebagai dasar penentuan nilai akhir mata pelajaran. Ringkasan ketiga kategori komponen penilaian yang lazim dipakai sebagai dasar penentuan nilai akhir mata pelajaran sebagaimana dimaksudkan oleh Haladyna disajikan dalam Tabel 8.1.

Tabel 8.1: Aneka Komponen Penilaian dalam Penentuan Nilai Akhir Mata Pelajaran

Bisa dipertanggung-jawabkan	Patut dipertanyakan	Tidak bisa dipertanggung-jawabkan
Pekerjaan rumah (PR).	Pelanggaran <i>deadline</i> atau batas waktu penyerahan tugas.	Kepatuhan pada perintah guru.
Aktivitas atau latihan di kelas.	Partisipasi atau keaktifan di kelas.	Usaha, entusiasme, sikap.
Tugas mandiri atau kelompok.	Tugas tambahan (<i>extra credit</i>).	Kerapihan.
Kuis atau tes harian.	Peningkatan hasil belajar selama periode penilaian berlangsung.	Kemampuan intelektual.
Tes tengah semester dan tes akhir semester.	Kehadiran atau presensi.	Kemampuan berbahasa.
Makalah dan laporan.	Penilaian subjektif guru.	Skor dalam tes baku.
Kinerja (<i>performance</i>).		Kreativitas.
Tugas melakukan demonstrasi atau presentasi di kelas.		Cara berpakaian.
Tugas melakukan ekshibisi.		Kebersihan.
Tugas melakukan eksperimen.		Kepribadian.
		Ketrampilan interpersonal.
		Kebutuhan emosional.
		Reputasi.
		Gender, etnisitas, ras, kelainan (fisik, psikis), agama.

Sumber: Haladyna, 1999, dalam Shermis & Di Viesta, 2011, h. 354.

Pembobotan Komponen Penilaian dalam Penentuan Nilai Akhir Mata Pelajaran

Pada awal pelaksanaan pembelajaran, lazimnya guru menentukan komponen penilaian apa saja yang akan dia gunakan dalam penentuan nilai akhir kelak untuk mata pelajaran yang diampunya. Sebagaimana sudah disinggung, dalam memilih komponen-komponen penilaian guru harus cermat mempertimbangkan relevansi atau keabsahannya. Selanjutnya, pelaksanaan penilaian terhadap masing-masing murid pada berbagai komponen penilaian tersebut akan dilakukan oleh guru selama berlangsungnya kegiatan pembelajaran mata pelajaran yang diampunya itu, meliputi misalnya pekerjaan rumah, latihan atau ulangan harian, tugas membuat laporan, tes tengah semester, dan sebagainya. Hasil penilaian guru pada masing-masing komponen penilaian tersebut akan menjadi unsur untuk menentukan nilai akhir yang akan dilaporkan sebagai nilai rapor untuk mata pelajaran yang bersangkutan bagi masing-masing murid. Tentu, bobot masing-masing komponen penilaian dalam penentuan nilai akhir berlainan, dan guru perlu menentukannya. Penentuan bobot masing-masing komponen ini bisa dilakukan sejak awal sebelum kegiatan pembelajaran dimulai atau cukup pada saat akan menentukan atau menghitung nilai akhir. Dasar atau pedoman yang bisa dipakai oleh guru dalam penentuan bobot masing-masing komponen penilaian yang sudah dipilihnya dalam menentukan nilai akhir suatu mata pelajaran bisa bermacam-macam, beberapa di antaranya yang lazim digunakan atau diterapkan oleh guru adalah sebagai berikut:

1. Nilai penting atau kurang pentingnya materi atau kompetensi yang dijadikan sasaran penilaian. Makin penting materi atau kompetensi tersebut dalam keseluruhan silabus, maka perlu diberi bobot yang makin tinggi.

2. Banyaknya materi atau kompetensi yang dijadikan sasaran penilaian. Makin banyak materi atau kompetensi yang dijadikan sasaran penilaian, maka perlu diberi bobot yang makin tinggi.
3. Sifat pengerjaan tugas dalam rangka penilaian, dikerjakan secara mandiri atau kelompok. Hasil pengerjaan tugas mandiri lazim perlu diberi bobot yang lebih tinggi dibandingkan hasil pengerjaan tugas secara kelompok.
4. Tempat pengerjaan tugas dalam rangka penilaian, dikerjakan di kelas atau di luar kelas khususnya di rumah. Pengerjaan tugas di kelas diandaikan lebih menuntut kemandirian, sedangkan pengerjaan tugas di luar kelas, lebih-lebih di rumah, diandaikan lebih terbuka bagi pemberian bantuan oleh orang lain. Maka, hasil pengerjaan tugas di kelas (mandiri) perlu diberi bobot yang lebih tinggi dibandingkan hasil pengerjaan tugas di rumah.

Sebagai contoh, untuk mata pelajaran Bahasa Indonesia di kelas V SD dalam satu semester guru memberikan tugas-tugas dalam rangka penilaian hasil belajar sebagai berikut: (1) tugas kelompok di kelas membuat sinopsis sebuah cerita pendek, (2) PR membuat sebuah puisi bertema lingkungan hidup, (3) tes formatif tengah semester, (4) tugas mandiri di kelas membuat resensi sebuah cerita pendek, dan (5) tes sumatif akhir semester. Hasil yang dicapai oleh murid dalam masing-masing tugas tersebut akan dijadikan komponen untuk menentukan nilai akhir mata pelajaran Bahasa Indonesia masing-masing murid. Memperhatikan nilai penting, sifat serta tempat pengerjaan masing-masing tugas, guru tersebut kiranya akan membuat pembobotan seperti disajikan dalam Tabel 8.2.

Tabel 8.2: Pembobotan Aneka Komponen Penilaian dalam Rangka Penentuan Nilai Akhir

No.	Kriteria	Bobot	Skor	Skor terbobot
1.	Tugas kelompok membuat sinopsis cerita pendek di kelas.	2		
2.	PR membuat puisi bertema lingkungan hidup.	1		
3.	Tes formatif tengah semester.	3		
4.	Tugas mandiri di kelas membuat resensi sebuah cerita pendek.	4		
5.	Tes sumatif akhir semester.	5		
	Jumlah	15		

Untuk masing-masing tugas sebagai komponen penilaian, setiap murid akan mendapat skor dari guru. Selanjutnya skor setiap tugas tersebut dikalikan bobot masing-masing, sehingga diperoleh skor terbobot untuk masing-masing tugas sebagai komponen penilaian. Jika skor terbobot masing-masing komponen penilaian tersebut dijumlahkan maka diperoleh skor terbobot total masing-masing murid untuk mata pelajaran yang bersangkutan, dalam contoh ini Bahasa Indonesia. Jika prosedur yang sama diterapkan oleh guru dalam penilaian hasil belajar untuk semua mata pelajaran, maka akan diperoleh skor terbobot total masing-masing murid untuk semua mata pelajaran.

Skor terbobot total ini masih merupakan *raw score* alias skor mentah. Tugas guru selanjutnya adalah melakukan penilaian dalam arti sempit, yaitu menentukan nilai akhir untuk setiap mata pelajaran dengan cara mengubah skor terbobot total tersebut dengan membandingkannya dengan sebuah patokan tertentu. Nilai akhir inilah kelak yang akan dilaporkan sebagai nilai rapor. Lantas, bagaimana cara menentukan nilai akhir yang dimaksud?

Dua Pendekatan dalam Penentuan Nilai Akhir

Ada dua macam pendekatan yang lazim dipakai dalam menentukan nilai akhir setiap mata pelajaran, yaitu (1) acuan norma (*norm referenced*), dan (2) acuan patokan (*criterion referenced*).

Pendekatan acuan norma. Prinsip pendekatan ini adalah bahwa nilai akhir seorang murid dalam suatu mata pelajaran ditentukan berdasarkan posisinya dalam distribusi nilai bersama murid-murid lain di kelasnya. Penggunaan strategi ini dilandasi oleh asumsi bahwa fungsi utama pemberian nilai akhir adalah memilah murid berdasarkan tingkat pencapaian hasil belajar mereka, sehingga hanya murid-murid terbaik akan diberi nilai-nilai tertinggi (Shermis & Di Viesta, 2011). Tentu strategi tersebut didasari oleh kenyataan bahwa di masa lalu kesempatan untuk melanjutkan pendidikan ke jenjang yang lebih tinggi masih terbatas, sehingga hanya mereka yang dipandang memiliki kemampuan melebihi teman-teman sebayanya diberi kesempatan untuk belajar lanjut. Tentu saja, kelemahan pendekatan ini ialah bahwa nilai akhir yang diperoleh setiap murid tidak mencerminkan apa yang benar-benar mampu dilakukannya, melainkan sekadar menunjukkan seberapa baik atau buruk kemampuannya dibandingkan teman-teman sekelas atau sebayanya.

Secara garis besar langkah-langkah penentuan nilai akhir beracuan norma adalah sebagai berikut. Pertama, guru perlu menentukan kelompok yang kinerja atau hasil belajarnya akan dipakai sebagai acuan pembandingan. Dalam bahasa Inggris, kelompok ini disebut *reference group* atau *norm group*. Dalam penilaian hasil belajar, kelompok acuan atau kelompok norma ini lazimnya adalah kelompok kelas penempuh mata pelajaran yang diampu oleh guru yang bersangkutan. Kedua, guru perlu memilih jenis nilai beracuan norma yang akan dilaporkan sebagai nilai akhir atau nilai rapor.

Di jenjang pendidikan sekolah menengah ke bawah lazim dipakai *nilai angka*, sedangkan di jenjang pendidikan tinggi lazim dipakai *nilai huruf*. Untuk itu guru perlu memilih dan mengembangkan *norma penilaian* yang sesuai. Langkah ketiga dan terakhir, sesudah memiliki norma penilaian yang sesuai guru mentransformasikan atau mengubah skor kasar masing-masing murid menjadi nilai akhir, entah berupa nilai angka atau nilai huruf, dengan menggunakan norma penilaian sebagai acuan.

Norma penilaian dalam penilaian beracuan norma disusun berdasarkan distribusi skor kelompok acuan atau kelompok norma yang dipakai. Salah satu jenis norma penilaian yang lazim dipakai untuk menentukan nilai akhir berupa nilai angka adalah *C scale* atau skala C yang memiliki 11 ruas skor atau nilai (*11-point scaled score*) sehingga seringkali juga disebut *Standard Eleven* atau *Stanel* (Guilford & Fruchter, 1978). Sebelas ruas nilai yang dimaksud bergerak mulai dari nilai terendah 0 sampai nilai tertinggi 10, dengan nilai *mean* atau rerata = 5. Salah satu cara untuk menentukan batas masing-masing nilai adalah dengan menggunakan nilai-nilai persentil tertentu (*centile-rank limits*), seperti disajikan pada Tabel 8.3. (Guilford & Fruchter, 1978).

**Tabel 8.3: Norma Penilaian Skala-C
Berdasarkan Nilai Persentil**

Nilai Skala-C	Persentil Batas	Nilai Persentil*
10	98,8	
9	96	
8	89,4	
7	77,3	
6	59,9	
5	40,1	
4	22,7	
3	10,6	
2	4	
1	1,2	
0	0,3	

***Keterangan:**

1. Nilai masing-masing persentil yang menjadi batas bawah masing-masing nilai skala-C yang akan dipakai sebagai nilai akhir murid-murid untuk suatu mata pelajaran perlu dihitung dulu berdasarkan distribusi skor murid-murid dalam mata pelajaran yang bersangkutan.
2. Sumber: Guilford & Fruchter (1978).

Dalam Tabel 8.3., nilai skala-C adalah rentang nilai yang akan menjadi nilai akhir dan yang akan dilaporkan sebagai nilai rapor untuk masing-masing mata pelajaran. Persentil batas adalah titik persentil yang menjadi batas bawah untuk mengubah skor mentah murid menjadi nilai akhir. Untuk menentukan nilai akhir murid dalam setiap mata pelajaran, masing-masing nilai persentil tersebut perlu dihitung terlebih dulu berdasarkan distribusi skor murid-murid di suatu kelas dalam masing-masing mata pelajaran yang bersangkutan. Cara menghitung nilai persentil sudah dijelaskan di Bab 7. Sesudah semua nilai persentil yang diperlukan ditemukan, maka diperolehlah norma penilaian untuk dipakai sebagai acuan dalam mengubah skor masing-masing murid dalam suatu mata pelajaran menjadi skor akhir berupa nilai angka yang berkisar 0-10.

Dengan cara ini memang tidak akan diperoleh nilai akhir berupa pecahan. Muncul pertanyaan, apa yang harus dilakukan jika nilai akhir harus berupa bilangan puluhan yang berkisar 0-100? Salah satu solusi sederhana kiranya adalah dengan sekadar mengalikan masing-masing nilai akhir dengan bilangan konstan 10.

Salah satu jenis norma penilaian yang lazim dipakai untuk menentukan nilai akhir berupa nilai huruf adalah norma yang disusun mengikuti pembagian daerah dalam kurve normal dengan menggunakan satuan *Mean* dan *SD*. Sebagaimana kita ketahui, dalam distribusi skor berbentuk kurve normal *Mean* terletak tepat di tengah daerah sebaran skor tersebut. Selain itu, keseluruhan daerah kurve normal tersebut bisa dibagi ke dalam enam bagian dengan satuan *SD*, yaitu tiga bagian terletak di atas *Mean* dan tiga bagian lainnya terletak di bawah *Mean*. Jika distribusi skor total terbobot murid-murid di suatu kelas untuk mata pelajaran tertentu kita ketahui sehingga bisa kita hitung *Mean* dan *SD*-nya, maka norma penilaian untuk mengubah skor total terbobot menjadi nilai huruf mulai dari A (tertinggi) sampai dengan E (terendah) dapat kita susun dengan menggunakan rumus seperti disajikan dalam Tabel 8.4.

Tabel 8.4: Norma Penilaian Berdasarkan Pembagian Daerah dalam Kurve Normal

Skor Batas	Nilai
$(Mean + 1,5SD) < X$	A
$(Mean + 0,5SD) < X \leq (Mean + 1,5SD)$	B
$(Mean - 0,5SD) < X \leq (Mean + 0,5SD)$	C
$(Mean - 1,5SD) < X \leq (Mean - 0,5SD)$	D
$X \leq (Mean - 1,5SD)$	E

X adalah skor total terbobot. Seandainya *Mean* dan *SD* sudah

diketahui, maka: (a) murid yang skor total terbobotnya lebih besar dibandingkan nilai ($Mean + 1,5SD$) akan mendapatkan nilai akhir A, (b) murid yang skor total terbobotnya lebih besar dari nilai ($Mean + 0,5SD$) namun sama atau kurang dari nilai ($Mean + 1,5SD$) akan mendapatkan nilai B, (c) murid yang skor total terbobotnya lebih besar dari nilai ($Mean - 0,5SD$) namun sama atau kurang dari ($Mean + 0,5SD$) akan mendapat nilai C, (d) murid yang skor total terbobotnya lebih besar dari ($Mean - 1,5SD$) namun sama atau kurang dari ($Mean - 0,5SD$) akan mendapat nilai D, dan (e) murid yang skor total terbobotnya sama atau kurang dari ($Mean - 1,5SD$) akan mendapat nilai E.

Pendekatan acuan patokan. Prinsip pendekatan ini adalah bahwa melalui nilai akhir guru bertujuan memberikan sertifikasi atau bukti bahwa seorang murid mencapai tingkat penguasaan tertentu terhadap materi atau kemampuan yang diajarkan dalam suatu mata pelajaran. Asumsinya, aneka ketrampilan atau pengetahuan yang menjadi materi setiap mata pelajaran dapat diuraikan ke dalam unsur-unsur atau komponen-komponen sehingga pada dasarnya setiap murid akan mampu mempelajari atau menguasainya. Maka kunci dalam penilaian adalah menyusun metode penilaian yang mampu menentukan tingkat penguasaan setiap murid terhadap materi atau kemampuan tertentu, serta menentukan batas-batas penguasaan yang dimaksud. Dengan kata lain, dalam penilaian beracuan patokan, guru tidak membandingkan kinerja seorang murid dengan kinerja murid-murid lain, melainkan dengan suatu standar kinerja yang disebut *kriterion* atau yang diindonesiakan menjadi *patokan*.

Ketika akan menetapkan *patokan* tentu terlebih dulu guru harus menetapkan *frame of reference* alias kerangka acuan, yaitu objek atau materi yang dipatok, barulah menentukan *patokan*-nya. Berdasarkan kedua hal tersebut, penilaian beracuan patokan lazim dibedakan

ke dalam tiga jenis (Friedenberg, 1995), yaitu (1) *content-referenced scoring* alias penilaian beracuan patokan materi, (2) *objective-referenced scoring* alias penilaian beracuan patokan tujuan, dan (3) *pass/fail* atau *mastery scoring* alias penilaian beracuan lulus/gagal atau beracuan penguasaan.

1. *Penilaian beracuan patokan materi.* Dalam pendekatan yang dikemukakan oleh Brown, (1983, dalam Friedenberg, 1995) ini, objek yang dijadikan kerangka acuan atau yang hendak dipatok adalah *isi* atau *materi* tes. Sudah barang tentu, isi atau materi tes itu sendiri merupakan turunan langsung dari isi atau materi bidang studi yang sedang dinilai. Patokannya adalah penguasaan atas isi atau materi tes tersebut, yang tercermin dari *jumlah jawaban benar*. Salah satu jenis nilai beracuan patokan materi yang paling sederhana adalah *percent correct* alias persentase jawaban benar, yang bisa dihitung dengan rumus sebagai berikut (Friedenberg, 1995):

$$\text{Persentase jawaban benar} = \frac{\text{Jumlah jawaban benar}}{\text{Jumlah skor yang mungkin dicapai}} \times 100$$

Friedenberg (1995) memberikan dua catatan kritis. Pertama, persentase jawaban benar sekadar menunjukkan kinerja testi menghadapi item-item tes. Jika item-item tes tersebut tidak merepresentasikan ranah materi bidang studi yang dinilai secara baik, maka persentase jawaban benar tidak mencerminkan penguasaan murid atas materi bidang studi yang diujikan. Kedua, makna persentase jawaban benar juga mudah dikacaukan oleh perbedaan taraf kesukaran item-item tes. Dua murid yang sama-sama mencapai nilai 90 tidak bisa disebut memiliki prestasi yang setara jika yang satu memperolehnya lewat tes dengan item-item yang relatif mudah

sedangkan yang lain memperolehnya lewat tes dengan item-item yang relatif sukar.

2. *Penilaian beracuan patokan tujuan*. Untuk mengatasi problem di atas, Thorndike et al. (1991, dalam Friedenberg, 1995) mengganti objek yang dijadikan kerangka acuan, bukan isi atau materi tes melainkan *tujuan* tes atau lebih tepat tujuan pengajaran (*instructional objectives*). Dalam penilaian, tujuan kita adalah menentukan tingkat penguasaan murid atas ranah pengetahuan-ketrampilan tertentu yang bersifat laten atau tidak langsung teramati, dan bukan sekadar mengukur kemampuan murid menjawab seperangkat pertanyaan tes. Item-item tes disusun untuk mengungkap pencapaian seperangkat tujuan pengajaran tertentu. Masing-masing tujuan pengajaran dituangkan ke dalam seperangkat item. Seorang testi atau murid disebut berhasil mencapai atau menguasai tujuan pengajaran tertentu jika dia berhasil menjawab dengan benar sejumlah item yang dimaksudkan untuk mengungkap hal itu. Dalam pendekatan ini kuncinya terletak pada merumuskan sistem penskoran yang mampu menunjukkan persentase tujuan pengajaran yang berhasil dikuasai murid. Toh menurut Friedenberg (1995), pendekatan ini belum terbebas dari jerat persoalan taraf kesukaran item dan penafsiran skor seperti pada pendekatan pertama.

3. *Pass/fail* atau *mastery scoring* alias penilaian beracuan lulus/gagal atau beracuan penguasaan. Pendekatan ketiga ini bukan pendekatan baru, melainkan sejenis aplikasi dari dua pendekatan yang pertama. Maksudnya, nilai yang diperoleh baik melalui penilaian beracuan patokan materi maupun beracuan patokan tujuan sama-sama bisa ditransformasikan ke dalam jenis lain, khususnya untuk memilah testi atau murid menjadi dua kategori: (a) mereka

yang kinerjanya dalam tes memenuhi patokan yang ditentukan (*lulus*), dan (b) mereka yang kinerjanya dalam tes tidak memenuhi patokan yang ditentukan (*gagal*). Lazimnya, guru atau pimpinan sekolah memilih salah satu skor tes yang dipandang mencerminkan taraf kinerja (minimal) yang diharapkan. Skor itulah yang dijadikan patokan, semua siswa yang mencapai skor sama atau lebih besar dari patokan dinyatakan *lulus*, dan semua siswa yang mencapai skor kurang dari patokan dinyatakan *gagal*. Perbedaan skor dalam masing-masing kategori *lulus* atau *gagal* tidak lagi memiliki makna.

Kendati tetap memiliki sejumlah kekurangan, pendekatan penilaian beracuan patokan ini secara umum dipandang lebih mencerminkan apa yang benar-benar bisa dilakukan oleh murid terkait pengetahuan, ketrampilan, atau kemampuan yang diajarkan dalam setiap mata pelajaran, dibandingkan dengan pendekatan beracuan norma (Shermis & Di Viesta, 2011).

Melaporkan Nilai Akhir

Tujuan melakukan penilaian hasil belajar murid pada setiap mata pelajaran secara akurat serta menyimpan rekaman penilaian hasil belajar tersebut secara cermat mencakup setidaknya tiga hal (Shermis & Di Viesta, 2011), yaitu (1) mencermati kemajuan belajar murid, (2) membuat rekaman kemajuan belajar murid dari tingkat kelas yang rendah sampai ke tingkat kelas yang tinggi, dan (3) memperoleh dasar untuk membuat laporan tentang kemajuan belajar murid ke pihak-pihak terkait.

Nilai akhir mata pelajaran lazim dilaporkan dalam rapor sebagai *nilai rapor*. Di masa lalu hasil belajar dalam setiap mata pelajaran memang hanya dilaporkan dalam sebuah nilai sumatif tunggal lazimnya berupa nilai angka yang dikenal sebagai nilai rapor tersebut. Dalam beberapa tahun terakhir, nilai rapor berupa nilai angka

tersebut lazim dilengkapi narasi kualitatif sebagai penilaian untuk aspek-aspek lain perkembangan murid yang tidak terkait langsung dengan mata pelajaran tertentu.

Kini, khususnya di negara-negara maju dan terutama demi keperluan memberikan informasi yang lebih rinci dan akurat kepada pihak-pihak terkait tentang kemajuan dan hasil belajar murid, laporan dalam rapor lazimnya lebih detil dan berupa narasi ringkasan deskripsi kemajuan belajar murid dalam berbagai bidang kurikular atau rumpun kompetensi, bukan lagi per mata pelajaran. Tujuan utamanya adalah memberikan umpan balik dengan cara memberikan informasi tentang aspek-aspek yang masih perlu diberi remediasi serta aspek-aspek yang dipandang sudah baik sehingga perlu dipertahankan terkait penguasaan kemampuan yang diajarkan dalam berbagai bidang kurikular. Contoh pelaporan hasil belajar berupa narasi dalam berbagai bidang kurikular sebagai rapor di kelas 4 Sekolah Dasar disajikan dalam Tabel 8.5.

Tabel 8.5: Contoh Rapor Model Narasi

KARTU RAPOR SD KUNCUP MEKAR			
Kelas 4			
Murid: Guru: Kepala Sekolah:			
Kunci penilaian: B = Baik sekali P = Memuaskan M = Ada kemajuan R = Perlu perbaikan			
Budi pekerti <i>Menunjukkan tanggung jawab dalam belajar.</i>		Membaca	Matematika <i>Ketrampilan dasar</i>
	Menunjukkan kontrol diri.	Mampu membaca dan memahami materi bacaan secara mandiri.	Melakukan penambahan. Melakukan pengurangan.
	Menunjukkan perhatian pada tugas yang sedang dikerjakan.	Mampu memilih materi bacaan dan membaca sendiri dalam membaca bebas.	Melakukan perkalian. Melakukan pembagian. Melakukan estimasi.
	Mengerjakan tugas kelas secara tuntas.	Memahami makna isi bacaan.	Melakukan perhitungan. Membuat grafik.
	Mengerjakan pekerjaan rumah.	Mampu merumuskan pertanyaan tentang isi bacaan dengan bimbingan guru.	Membuat pola. Geometri. Kesadaran ruang.
	Mengerjakan tugas secara mandiri.	Mampu membuat ringkasan tentang bab atau cerita pendek.	Pengukuran. Waktu. Uang.
	<i>Menanggapi arahan guru.</i>	Membaca aneka jenis buku yang ditugaskan oleh guru.	Pecahan. Pemecahan masalah.
	Mematuhi perintah guru.	Mampu merumuskan jawaban yang baik berdasarkan isi bacaan.	Probabilitas.
	Mendengarkan secara aktif.	Mampu memimpin atau berpartisipasi dalam diskusi tentang isi bacaan.	Membuktikan konsep-konsep matematis secara tertulis.
	Berpartisipasi aktif dalam kegiatan dan diskusi.	Menulis <i>Tata bahasa</i>	Ilmu Pengetahuan Alam
	Mengerjakan tugas secara bersungguh-sungguh.	Menulis kalimat lengkap dengan struktur kalimat yang baik.	Proses penelitian ilmiah. Ilmu bumi dan geologi. Ruang angkasa.
	Menunjukkan hormat dalam berbicara dan bertindak.	Menerapkan aneka konsep tata bahasa dalam menulis.	Planet. Atmosfer dan cuaca.

Penentuan dan Pelaporan Nilai Akhir

Ilmu Pengetahuan Sosial		Menggunakan tanda baca secara tepat.	Fauna.
	Berpartisipasi dalam diskusi-diskusi tentang masalah-masalah sosial.	<i>Proses dan gaya menulis</i>	Tubuh manusia.
	Memahami konsep-konsep atau mengajukan pertanyaan-pertanyaan.	Membuat kerangka sebelum menulis.	Makanan.
	Menyelesaikan aktivitas-aktivitas pengkajian masalah-masalah sosial.	Mampu mengorganisasikan gagasan melalui pembagian paragraf.	Teknologi.
Kesenian <i>Seni rupa</i>		Mampu mengoreksi hasil tulisan sendiri.	Mampu melukiskan atau menjelaskan konsep-konsep ilmiah.
	Pengembangan ketrampilan.	Mampu memperbaiki tulisan sendiri berdasarkan masukan tertulis maupun lisan.	Pendidikan jasmani dan kesehatan
	Sikap dan tingkah laku.	<i>Ejaan</i>	Ketrampilan.
	Usaha	Menggunakan ejaan yang benar dalam membaca dan menulis.	Sikap dan tingkah laku.
<i>Seni musik</i>		<i>Tulisan tangan</i>	Pengetahuan tentang konsep-konsep kesehatan.
	Pengembangan ketrampilan	Tulisan terbaca, rapi, dan tegak bersambung.	Mampu bekerja sama dalam tim.
	Sikap dan tingkah laku.		Usaha.
	Usaha.		

Catatan guru:

Tanda tangan guru:

Tanggal:

Sumber: Shermis & Di Viesta, 2011, h. 363.

Untuk konteks persekolahan di Tanah Air, penggunaan cara pelaporan hasil belajar naratif berbasis rumpun kompetensi sebagai satu-satunya format rapor kiranya belum bisa diterima sepenuhnya. Laporan naratif berbasis rumpun kompetensi semacam itu tetap perlu disertai laporan nilai akhir masing-masing mata pelajaran. Bahkan narasinya pun perlu diberikan untuk masing-masing mata pelajaran.

Daftar Pustaka

- Azwar, S. (1997). *Tes prestasi. Fungsi dan pengembangan pengukuran prestasi belajar* (ed. II). Yogyakarta: Pustaka Pelajar.
- Blalock, Jr., H.M. (1979). *Social statistics* (rev. 2nd ed.). New York: McGraw-Hill.
- Bloom, B.S. (1956, Ed.). *Taxonomy of educational objectives. The classification of educational goals. Handbook I: Cognitive domain*. New York: David MacKay.
- Broadfoot, P. (2009). Signs of change: Assessment past, present and future. Dalam C. Wyatt-Smith & J.J. Cumming (Eds.), *Educational assessment in the 21st century* (h. v-xi). Dordrecht: Springer.
- Chatterji, M. (2003). *Designing and using tools for educational assessment*. Boston: Allyn and Bacon.
- De Landsheere, V. (1990). Taxonomies of educational objectives. Dalam H.J. Walberg & G.D. Haertel (Eds.), *The international encyclopedia of educational evaluation* (h. 179-189). Oxford: Pergamon.
- Downie, N.M., & Heath, R.W. (1984). *Basic statistical methods* (5th ed.). New York: Harper & Row.
- Friedenberg, L. (1995). *Psychological testing. Design, analysis, and use*. Boston: Allyn and Bacon.

- Guilford, J.P. (1954). *Psychometric methods* (2nd ed.). New York: McGraw-Hill.
- Guilford, J.P., & Fruchter, B. (1978). *Fundamental statistics in psychology and education* (6th ed.). Tokyo: McGraw-Hill Kogakusha.
- Krathwohl, D.R., Bloom, B.S., & Masia, B.B. (1973). *Taxonomy of educational objectives. The classification of educational goals. Handbook II: Affective domain*. New York: David MacKay.
- Kristi Poerwandari. (2005). *Pendekatan kualitatif untuk penelitian perilaku manusia*. Jakarta: LPSP3, Fakultas Psikologi, Universitas Indonesia.
- Layton, C.A., & Lock, R.H. (2008). *Assessing students with special needs to produce quality outcomes*. Upper Saddle River, NJ: Pearson.
- McCormick, E.J., & Ilgen, D. (1980). *Industrial psychology* (7th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Peraturan Menteri Pendidikan Nasional Republik Indonesia Nomor 16 Tahun 2007 tentang Standar Kualifikasi Akademik dan Kompetensi Guru.
- Rotter, J.B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 80(1 Whole No. 609).
- Sudjana, Nana (2010). *Penilaian hasil proses belajar mengajar*. Bandung: Remaja Rosdakarya.
- Sutrisno Hadi (2001a). *Statistik* (Jilid 1). Yogyakarta: Andi.
- Sutrisno Hadi (2001b). *Statistik* (Jilid 2). Yogyakarta: Andi.
- Suwandi, Sarwiji (2011). *Model-model asesmen dalam pembelajaran*. Surakarta: Yuma Pustaka.
- Tarlinton, D. (2003). *Bloom's revised taxonomy*. Diunduh dari <http://coe.sdsu.edu/eet/articles/bloomrev/index.htm>
- Wiggins, G. (1998). *Educative assessment. Designing assessments to inform and improve student performance*. San Francisco: Josey-Bass.

Tentang Penulis

A. Supratiknya, profesor pada Fakultas Psikologi Universitas Sanata Dharma, Yogyakarta. Tamat dari Fakultas Psikologi Universitas Gadjah Mada (B.A., 1976; Drs., 1980) dan dari *Department of Psychology, College of Social Sciences and Philosophy, University of the Philippines, Diliman (Ph.D., 1992)*. Pernah mengikuti *Fulbright Visiting Scholar Program* di *Center for Cross-Cultural Research, Department of Psychology, Western Washington University, Bellingham*, dan *School of Psychology, Florida Institute of Technology, Melbourne, Amerika Serikat (2003-2004)*. Menjadi anggota Himpunan Psikologi Indonesia (HIMPSI), *American Psychological Association*, dan *International Association for Cross-Cultural Psychology*. Menulis dan menerjemahkan sejumlah buku, menulis artikel, dan melakukan penelitian tentang psikologi, dengan perhatian khusus pada psikologi budaya dan pendidikan.

