

PENGEMBANGAN INSTRUMEN NONTES

Ani Rusilowati
Program Pascasarjana Universitas Negeri Semarang
rusilowati@yahoo.com

Kurikulum 2013 menuntut penilaian berbasis kompetensi, dan bergeser dari penilaian tes menjadi penilaian otentik, yang meliputi pengetahuan, sikap, dan keterampilan. Beragam teknik dapat dilakukan untuk mengumpulkan informasi tentang kemajuan belajar peserta didik, baik yang berhubungan dengan proses belajar maupun hasil belajar. Teknik pengumpulan informasi tersebut pada prinsipnya adalah cara penilaian kemajuan belajar peserta didik berdasarkan standar kompetensi (kompetensi inti pada kurikulum 2013) dan kompetensi dasar yang harus dicapai. Penilaian kompetensi dasar dilakukan berdasarkan indikator-indikator pencapaian kompetensi yang memuat satu ranah atau lebih. Berdasarkan indikator-indikator ini dapat ditentukan cara penilaian yang sesuai, apakah dengan tes tertulis, observasi, tes praktek, dan penugasan perseorangan atau kelompok. Pada penilaian kelas terdapat tujuh teknik yang dapat digunakan, yaitu penilaian tes tertulis/lisan, penilaian unjuk kerja, penilaian sikap, penilaian proyek, penilaian produk, penggunaan portofolio, dan penilaian diri. Penilaian tes tertulis/lisan menggunakan tes, sedangkan keenam teknik penilaian lain (nontes) menggunakan lembar observasi, kuesioner. Bentuk instrumen nontes dapat berupa checklist, skala, atau catatan pengamatan. Pada kesempatan ini akan dibahas tentang penilaian nontes meliputi: (1) langkah pengembangan instrumen nontes, (2) bentuk instrumen nontes, dan (3) penentuan kualitas instrumen nontes yang dikembangkan.

1. Langkah Pengembangan instrumen nontes

Seperti halnya pengembangan instrumen tes, pengembangan instrumen nontes juga memiliki langkah- langkah yang harus diikuti, yaitu: menentukan spesifikasi instrumen; menulis instrumen; menentukan skala instrumen; menentukan sistem penskoran; menelaah instrumen; merakit instrumen; melakukan ujicoba; menganalisis hasil ujicoba; memperbaiki instrumen; melaksanakan pengukuran; dan menafsirkan hasil pengukuran.

a. Spesifikasi Instrumen

Spesifikasi intrumen terdiri atas tujuan, dan kisi-kisi instrumen. Tujuan pengembangan instrumen nontes sangat tergantung pada data yang akan dihimpun. Instrumen nontes mencakup afektif dan psikomotorik. Ditinjau dari tujuannya, instrumen ranah afektif dibedakan menjadi lima, yaitu instrumen sikap, minat, konsep diri, nilai, dan moral. Ada empat hal yang perlu diperhatikan ketika menyusun spesifikasi instrumen, yaitu: tujuan pengukuran, kisi-kisi instrumen, bentuk dan format instrumen, dan panjang instrumen.

Instrumen minat bertujuan untuk memperoleh informasi tentang minat siswa terhadap mata pelajaran. Selanjutnya hasil pengukuran terhadap minat digunakan untuk meningkatkan minat siswa terhadap mata pelajaran. **Instrumen sikap** bertujuan untuk mengetahui sikap siswa terhadap suatu objek. Misalnya, sikap siswa terhadap kegiatan sekolah, guru, dll. Sikap terhadap mata pelajaran bisa positif bisa negatif. Hasil pengukuran sikap berguna untuk menentukan strategi pembelajaran yang tepat bagi siswa. **Instrumen konsep diri** bertujuan untuk mengetahui kekuatan dan kelemahan diri sendiri. Siswa melakukan evaluasi secara objektif terhadap potensi yang ada dalam dirinya. Karakteristik potensi siswa sangat penting untuk menentukan jenjang karirnya. Informasi kekuatan dan

kelemahan siswa digunakan untuk menentukan program yang sebaiknya ditempuh oleh siswa. **Instrumen nilai** bertujuan untuk mengungkap nilai dan keyakinan individu. Informasi yang diperoleh bisa positif bisa negatif. Hal-hal yang positif diperkuat, sedangkan yang negatif diperlemah dan akhirnya dihilangkan. **Instrumen moral** bertujuan untuk mengungkap moral. Informasi moral seseorang diperoleh melalui pengamatan atas perbuatan yang ditampil-kan dan laporan diri, yaitu dengan mengisi kuesioner. Informasi hasil pengamatan bersamaan dengan hasil kuesioner menjadi informasi penting tentang moral seseorang.

Setelah tujuan pengukuran afektif ditetapkan, kegiatan berikutnya adalah menyusun kisi-kisi instrumen. Kisi-kisi juga disebut *blueprint*. Kisi-kisi berupa tabel matriks yang berisi spesifikasi instrumen yang akan ditulis. Langkah pertama menentukan kisi-kisi adalah menentukan definisi konseptual yang berasal dari teori-teori yang diambil dari referensi. Selanjutnya, mengembangkan definisi operasional berdasarkan definisi konseptual. Kemudian tentukan aspek/dimensi yang mengkonstruksi instrumen yang dikembangkan. Aspek atau dimensi ini kemudian dijabarkan menjadi sejumlah indikator, yang digunakan sebagai pedoman dalam menulis instrumen. Tiap indikator dapat terdiri atas dua atau lebih butir instrumen. Salah satu contoh format kisi-kisi instrumen minat dapat dilihat pada tabel berikut.

KISI-KISI INSTRUMEN MINAT MAHASISWA DALAM MENGIKUTI PERKULIAHAN

Aspek/Dimensi	Indikator	Nomor Butir	Jumlah
Keterlibatan dalam perkuliahan	1. Memiliki catatan	1, 2	2
	2. Berusaha memahami materi perkuliahan	3, 4, 5	3
Pemenuhan sarana/prasarana	3. Memiliki buku referensi	6, 7, 12	3
Usaha yang dilakukan	4. Kehadiran dalam perkuliahan	8, 9	2
	5. Melakukan diskusi dengan teman	10	1
	6. Kunjungan ke perpustakaan	11	1
Jumlah			12

b. Menulis Instrumen

Instrumen disusun berdasarkan kisi-kisi yang telah dibuat. Instrumen dapat berbentuk pernyataan atau pertanyaan. Kaidah yang perlu diperhatikan ketika menulis butir instrumen adalah:

- a. Hindari kalimat yang mengandung banyak interpretasi
- b. Rumusan pernyataan/pertanyaan singkat
- c. Satu pernyataan hanya mengandung satu pikiran yang lengkap
- d. Pernyataan dirumuskan dengan kalimat sederhana
- e. Hindari penggunaan kata-kata selalu, semua, tidak pernah, dan sejenisnya
- f. Hindari pernyataan tentang fakta, atau yang dapat diinterpretasikan sebagai fakta.

Hal yang perlu diingat ketika menyusun instrumen afektif adalah penentuan kalimat pernyataan. Ada dua macam pernyataan, *favorable* dan *unfavorable*. Kedua pernyataan ini berhubungan dengan penetapan skala. Skala untuk pernyataan *favorable* berlawanan

dengan *unfavorable*. Jika salah dalam menentukan skala, maka kesimpulan yang dihasilkan juga akan salah.

c. Menentukan skala instrumen

Ada beberapa skala yang biasa digunakan dalam mengukur ranah afektif, di antaranya adalah skala Likert, Thrustone, dan Beda Semantik. Langkah-langkah pengembangan skala:

- a. Menentukan objek sikap yang akan dikembangkan skalanya
- b. Menyusun kisi-kisi instrumen (skala sikap)
- c. Menulis butir pernyataan
- d. Melengkapi butir pernyataan dengan skala sikap (bisa genap, 4 atau 6, dan bisa ganjil 5 atau 7)

Contoh Skala Likert: Sikap terhadap pelajaran IPA

1) Pelajaran IPA bermanfaat	SS	S	N	TS
2) Pelajaran IPA sulit	SS	S	N	TS
3) Tidak semua orang harus belajar IPA	SS	S	N	TS
4) Pelajaran IPA harus dibuat mudah	SS	S	N	TS
5) Pelajaran IPA menyenangkan	SS	S	N	TS

Keterangan:

- SS = sangat setuju, skor 4
- S = setuju, skor 3
- N = netral/tidak berpendapat, skor 2
- TS = tidak setuju, skor 1

Contoh skala Thrustone:

No	Pernyataan	Skor						
		7	6	5	4	3	2	1
1	Saya senang belajar IPA							
2	Pelajaran IPA bermanfaat							
3	Saya berusaha hadir pada pelajaran IPA							
4	Saya berusaha memiliki buku-buku IPA							
5	Pelajaran IPA membosankan							

Contoh skala beda semantik:

Pelajaran IPA

	7	6	5	4	3	2	1	
Menyenangkan								Membosankan
Sulit								Mudah
Bermanfaat								Sia-sia
Menantang								Biasa-biasa saja
Banyak								Sedikit
Rumit								Sederhana

d. Sistem Penskoran

Sistem penskoran yang digunakan tergantung pada skala yang digunakan. Misalnya, apabila digunakan skala Thrustone, maka skor tertinggi tiap butir adalah 7 dan terendah 1. Selanjutnya dilakukan analisis untuk tingkat siswa dan tingkat kelas, yaitu dengan mencari rerata dan simpangan baku skor. Hasil analisis digunakan untuk menafsirkan ranah afektif dari setiap siswa dan kelas terhadap suatu objek. Hasil tafsiran perlu ditindak lanjuti oleh guru dengan melakukan perbaikan-perbaikan, seperti perbaikan metode pembelajaran, penggunaan alat peraga, dll.

e. **Telaah Instrumen**

Kegiatan pada telaah instrumen adalah meneliti tentang: (a) kesesuaian antara butir pertanyaan/ Pernyataan dengan indikator, (b) kekomunikatifan bahasa yang digunakan, (c) kebenaran dari tata bahasa yang digunakan, (d) ada tidaknya bias pada pertanyaan/ pernyataan, (e) kemenarikan format instrumen, (f) kecukupan butir instrumen, sehingga tidak membosankan.

Telaah dilakukan oleh pakar dalam bidang yang diukur dan akan lebih baik bila ada pakar penilaian. Telaah bisa juga dilakukan oleh teman sejawat. Panjang instrumen berhubungan dengan masalah kebosanan. Lama pengisian instrumen sebaiknya tidak lebih dari 30 menit. Pertanyaan/ pernyataan yang diajukan jangan sampai bias, yaitu mengarahkan jawaban responden pada arah tertentu, positif atau negatif. Contoh pernyataan bias: Sebagian besar responden setuju bahwa masyarakat berhak menerima layanan kesehatan Apakah Anda setuju bila semua masyarakat menerima layanan kesehatan?

Hasil telaah selanjutnya digunakan untuk memperbaiki instrumen. Perbaikan dilakukan terhadap konstruksi instrumen, yaitu kalimat yang digunakan, waktu yang diperlukan untuk mengisi instrumen, cara pengisian, dll.

f. **Merakit Instrumen**

Setelah instrumen diperbaiki, selanjutnya dirakit dengan memperhatikan format, tata letak, urutan pernyataan dan pertanyaan. Format harus menarik. Urutan pernyataan sesuai dengan aspek yang akan diukur.

g. **Ujicoba Instrumen**

Setelah dirakit, instrumen diujicobakan. Sampel ujicoba dipilih yang karakteristiknya mewakili populasi yang ingin dinilai. Ukuran sampel minimal 30 orang, bisa berasal dari satu sekolah atau lebih. Pada saat ujicoba, yang perlu dicatat adalah saran-saran dari responden atas kejelasan pedoman pengisian instrumen, kejelasan kalimat, waktu yang digunakan, dll.

h. **Analisis Hasil Ujicoba**

Analisis hasil ujicoba meliputi variasi jawaban tiap butir pertanyaan/ pernyataan. Apabila skala instrumen 1 sampai 5, maka bila jawaban bervariasi dari 1 sampai 5 berarti instrumen tersebut baik. Namun apabila jawaban semua responden sama, misalnya 3 semua, maka instrumen tergolong tidak baik. Indikator yang digunakan adalah besarnya daya beda atau korelasi antara skor butir dengan skor total. Bila daya beda butir lebih dari 0,3 maka instrumen tergolong baik.

Indikator lain yang diperhatikan adalah indeks kehandalan atau reliabilitas. Besarnya indeks reliabilitas sebaiknya minimal 0,7.

i. **Perbaikan Instrumen**

Perbaikan dilakukan terhadap butir-butir pertanyaan/pernyataan yang tidak baik. Perbaikan berdasarkan hasil ujicoba dan saran masukan dari responden.

j. Pelaksanaan Pengukuran

Pelaksanaan pengukuran sebaiknya dilakukan pada saat responden tidak lelah. Ruang untuk pelaksanaan pengukuran harus representatif, baik kondisi ruang, tempat duduk, ataupun yang lain. Diusahakan responden tidak saling bertanya ketika pengukuran dilaksanakan. Pengisian instrumen dimulai dengan penjelasan tujuan pengisian, manfaat bagi responden, dan pedoman pengisian instrumen.

k. Penafsiran Hasil Pengukuran

Hasil pengukuran berupa skor atau angka. Menafsirkan hasil pengukuran disebut dengan penilaian. Untuk menafsirkan hasil pengukuran diperlukan suatu kriteria. Kriteria yang digunakan tergantung pada skala dan jumlah butir yang digunakan. Misalnya, digunakan skala Likert dengan 5 pilihan untuk mengukur sikap siswa yaitu:

Favorable:	SS	S	N	TS	STS
	(5)	(4)	(3)	(2)	(1)
Unfavorable :	SS	S	N	TS	STS
	(1)	(2)	(3)	(4)	(5)

Hasil pengukuran sikap atau minat, dapat dikategorikan ke dalam beberapa kategori. Misaln akan dikategorikan dalam 4 kategori, yaitu sangat tinggi, tinggi, rendah, dan sangat rendah. Penentuan skor tiap kategori dapat dilihat pada tabel berikut.

Kategori Sikap atau Minat Siswa untuk 10 Butir Pertanyaan

No	Skor Siswa	Kategori Sikap atau Minat
1	$33 \leq X \leq 40$	Sangat positif/Sangat tinggi
2	$25 \leq X < 33$	Positif/Tinggi
3	$18 \leq X < 25$	Negatif/rendah
4	$10 \leq X < 18$	Sangat negatif/Sangat Rendah

Kategori tersebut juga dapat digunakan untuk menentukan sikap atau minat kelas. Skor yang digunakan adalah skor rata-rata kelas. Hasil pengukuran minat kelas untuk semua mata pelajaran dapat digunakan untuk menentukan profil minat kelas. Selanjutnya, profil ini dikaitkan dengan profil prestasi belajar.

2. Jenis instrumen nontes

Telah disampaikan pada pendahuluan, bahwa teknik penilaian nontes mengacu pada penilaian kelas yang meliputi unjuk kerja, sikap, proyek., produk, portofolio, dan penilaian diri.

a. Penilaian Unjuk Kerja

Penilaian unjuk kerja merupakan penilaian yang dilakukan dengan mengamati kegiatan peserta didik dalam melakukan sesuatu. Penilaian ini cocok digunakan untuk menilai ketercapaian kompetensi yang menuntut peserta didik melakukan tugas tertentu seperti: praktik di laboratorium, praktek 5sholat, praktek OR, presentasi, diskusi, bermain

peran, memainkan alat musik, bernyanyi, membaca puisi/deklamasi dll. Cara penilaian ini dianggap lebih otentik daripada tes tertulis karena apa yang dinilai lebih mencerminkan kemampuan peserta didik yang sebenarnya.

Pengamatan unjuk kerja perlu dilakukan dalam berbagai konteks untuk menetapkan tingkat pencapaian kemampuan tertentu. Untuk menilai kemampuan melakukan kinerja ilmiah peserta didik, dilakukan pengamatan atau observasi yang beragam, seperti: mempersiapkan alat, merangkai percobaan, dan mengamati hasil percobaan. Untuk mengamati unjuk kerja peserta didik dapat menggunakan alat atau instrumen berupa daftar cek (check-list), skala bertingkat (*rating scale*), catatan pengamatan.

1) Daftar Cek (Check-list)

Penilaian unjuk kerja dapat dilakukan dengan menggunakan daftar cek (baik-tidak baik). Dengan menggunakan daftar cek, peserta didik mendapat nilai bila kriteria penguasaan kompetensi tertentu dapat diamati oleh penilai. Jika tidak dapat diamati, peserta didik tidak memperoleh nilai. Kelemahan cara ini adalah penilai hanya mempunyai dua pilihan mutlak, misalnya benar-salah, dapat diamati-tidak dapat diamati, baik-tidak baik. Dengan demikian tidak terdapat nilai tengah, namun daftar cek lebih praktis digunakan mengamati subjek dalam jumlah besar.

Contoh *check list* untuk menilai kegiatan praktikum Fisika:

Nama peserta didik :
Kelas :

No	Aspek yang dinilai	Baik	Tidak
1	Pengetahuan tentang prosedur kerja		
2	Ketepatan memilih alat dan bahan		
3	Ketepatan cara mengoperasikan alat		
4	Hasil pengamatan		
5	Ketepatan menyusun laporan		
	Skor yang dicapai		
	Skor maksimum		

Keterangan

Baik mendapat skor 1
Tidak baik mendapat skor 0

Kriteria penilaian dapat dilakukan sebagai berikut
Skor maksimum : 5

Skor yang dicapai: jumlah skor setiap aspek
Nilai = (skor yang dicapai/skor maksimum) x 10

Penentuan Kriteria:

Skor 4-5, dapat ditetapkan kompeten
Skor 3-4, dapat ditetapkan cukup kompeten
Skor 1-2, dapat ditetapkan tidak kompeten

2) Skala Penilaian (*Rating Scale*)

Penilaian unjuk kerja yang menggunakan skala penilaian memungkinkan penilai memberi nilai tengah terhadap penguasaan kompetensi tertentu, karena pemberian nilai secara kontinum di mana pilihan kategori nilai lebih dari dua. Skala penilaian terentang dari tidak sempurna sampai sangat sempurna. Untuk memperkecil faktor subjektivitas, perlu dilakukan penilaian oleh lebih dari satu orang, agar hasil penilaian lebih akurat. Agar para penilai memiliki frame yang sama maka perlu dibuat rubrik penilaiannya. Rubrik penilaian dapat berupa generik dan spesifik. Contoh *rating scale* dan rubriknya dapat dicermati pada uraian berikut.

Contoh *Rating Scale*

Format Penilaian Praktikum Fisika

Nama Siswa : _____
 Kelas : _____

No	Aspek yang dinilai	4	3	2	1
1	Pengetahuan tentang prosedur kerja				
2	Ketepatan memilih alat dan bahan				
3	Ketepatan cara mengoperasikan alat				
4	Hasil pengamatan				
5	Ketepatan menyusun laporan				
	Skor yang dicapai				
	Skor maksimum				

Rubrik penilaian:

- a. Pengetahuan tentang prosedur kerja
 - Skor 4 = memahami prosedur kerja, dan menerapkannya secara benar dan tepat
 - Skor 3 = dapat menerapkan prosedur kerja secara tepat dengan membaca petunjuk
 - Skor 2 = dapat menerapkan prosedur kerja dengan bimbingan orang lain
 - Skor 1 = tidak mampu menerapkan prosedur kerja
- b. Ketepatan memilih alat dan bahan
 - Skor 4 = memilih alat dan bahan, dan dilakukan dengan benar dan tepat
 - Skor 3 = memilih alat dan bahan secara tepat dengan bertanya kepada teman sekelompok
 - Skor 2 = memilih alat dan bahan sendiri, tetapi tidak seluruhnya benar
 - Skor 1 = memilih alat dan bahan dengan bimbingan orang lain (guru/laboran)

Dst.

Kriteria penilaian dapat dilakukan sebagai berikut

Skor maksimum : $5 \times 4 = 20$

Skor yang dicapai: jumlah skor setiap aspek

$$\text{Nilai} = (\text{skor yang dicapai} / \text{skor maksimum}) \times 10$$

Penentuan Kriteria:

Skor 16-20, dapat ditetapkan sangat kompeten

Skor 12-15, dapat ditetapkan kompeten

Skor 8-11, dapat ditetapkan cukup kompeten

Skor 4-7, dapat ditetapkan tidak kompeten

Contoh Penilaian Unjuk Kerja:

Format Penilaian Kinerja Ilmiah

No	Nama siswa	Aspek yang dinilai												Jml skor
		Mempersiapkan alat dan bahan				Merangkai percobaan				Mengamati hasil percobaan				
		1	2	3	4	1	2	3	4	1	2	3	4	
1														
2														
3														
4														
5														
6														

Rubrik Penilaian:

Mempersiapkan alat

Skor 4 : Jika alat dan bahan disiapkan dengan sempurna, sesuai dengan langkah-langkah kegiatan

Skor 3 : Jika alat dan bahan disiapkan, tetapi kurang lengkap

Skor 2 : Jika alat dan bahan disiapkan oleh laboran, tetapi ikut memandu menentukan alat dan bahan

Skor 1 : Jika tidak ikut menyiapkan alat dan bahan, hanya pasif menunggu orang lain yang menyiapkan

Merangkai percobaan

Skor 4 : Jika alat dirangkai dengan benar

Skor 3 : Jika alat dirangkai, tetapi kurang benar

Skor 2 : Jika alat dirangkai oleh orang lain, tetapi ikut aktif memperhatikan

Skor 1 : Jika tidak ikut merangkai alat, dan tidak memperhatikan

Mengamati hasil percobaan

Skor 4 : Jika melakukan percobaan, mengamati semua variabel dan mencatat hasilnya dengan cermat

Skor 3 : Jika melakukan percobaan, mengamati percobaan, tetapi ada sebagian variabel yang tidak diamati dan dicatat hasil dengan cermat

Skor 2 : Jika hanya ikut mengamati dan mencatat hasil dengan cermat

Skor 1 : Jika hanya memindahkan catatan hasil pengamatan orang lain.

Skor maksimal = 3 x 4 = 12

$$\text{Nilai} = \frac{\text{Jumlah skor yang dicapai}}{\text{Skor maksimal}} \times 10$$

b. Penilaian Sikap

Penilaian sikap merupakan bagian dari pengukuran psikologi. Karena menyangkut sikap manusia, maka hasil pengukuran tidak pernah mencapai hasil yang sempurna. Pengukuran sikap sangat sukar bahkan mungkin tidak pernah dapat dilakukan dengan validitas, reliabilitas, dan objektifitas yang tinggi. Hal ini, antara lain dikarenakan: (1) atribut psikologi bersifat tidak tampak, (2) indikator-indikator perilaku jumlahnya terbatas, (3) respons dipengaruhi oleh variabel-variabel tidak relevan seperti: suasana hati, kondisi dan situasi sekitar, dan (4) banyak sumber kesalahan, baik dari penilai, yang dinilai, alat yang digunakan, cara analisis.

Sebagai contoh akan menilai skala sikap ilmiah. Lembar observasi yang dapat digunakan untuk menilai sikap ilmiah siswa secara individu dapat dilihat pada tabel berikut.

No	Nama siswa	Indikator Sikap							Jml skor
		Keterbukaan	objektif	Teliti	Kedisiplinan	Kerjasama	Kejujuran	Tanggung jawab	
1									
2									
3									

Skor untuk masing-masing sikap di atas di rata-rata dan dikonversikan ke dalam bentuk kualitatif. Skala penilaian dibuat dengan rentangan dari 1 sampai dengan 5. Penafsiran angka-angka tersebut adalah sebagai berikut: 1 = sangat kurang, 2 = kurang, 3 = cukup, 4 = baik, dan 5 = amat baik.

Skor maksimum = skor maks setiap indikator X jumlah indikator
 = 5 x 7 = 35

Nilai sikap ilmiah dapat diberikan dalam bentuk huruf, oleh karena itu total skor yang telah diperoleh harus dikonversi.

$$\text{Konversi Nilai} = \frac{\text{Skor total jawaban siswa}}{\text{Skor maksimum}} \times 100$$

Jadi siswa yang memperoleh skor 28 setelah dikonversi nilainya menjadi:

$$\frac{28}{35} \times 100 = 80$$

Banyak cara untuk mengkonversi skor menjadi nilai, salah satunya yang sederhana yaitu menggunakan kriteria sbb.

Skor Total	NILAI KONVERSI		Kategori
	Angka	Huruf	
29 – 35	81 - 100	A	Amat Baik
21 - 28	61 - 80	B	Baik
14 - 20	41 - 60	C	Cukup
7 - 13	20 - 40	D	Kurang

c. Penilaian Proyek

Penilaian proyek merupakan kegiatan penilaian terhadap suatu tugas yang harus diselesaikan dalam periode/waktu tertentu. Tugas tersebut berupa suatu investigasi sejak dari perencanaan, pengumpulan data, pengorganisasian, pengolahan dan penyajian data. Penilaian proyek dapat digunakan untuk mengetahui pemahaman, kemampuan mengaplikasikan, kemampuan penyelidikan dan kemampuan menginformasikan peserta didik pada mata pelajaran tertentu secara jelas. Dalam penilaian proyek setidaknya ada 3 (tiga) hal yang perlu dipertimbangkan yaitu:

1) Kemampuan pengelolaan

Kemampuan peserta didik dalam memilih topik, mencari informasi dan mengelola waktu pengumpulan data serta penulisan laporan.

2) Relevansi

Kesesuaian dengan mata pelajaran, dengan mempertimbangkan tahap pengetahuan, pemahaman dan keterampilan dalam pembelajaran.

3) Keaslian

Proyek yang dilakukan peserta didik harus merupakan hasil karyanya, dengan mempertimbangkan kontribusi guru berupa petunjuk dan dukungan terhadap proyek peserta didik.

Penilaian proyek dilakukan mulai dari perencanaan, proses pengerjaan, sampai hasil akhir proyek. Untuk itu, guru perlu menetapkan hal-hal atau tahapan yang perlu dinilai, seperti penyusunan disain, pengumpulan data, analisis data, dan menyiapkan laporan tertulis. Laporan tugas atau hasil penelitian juga dapat disajikan dalam bentuk poster. Pelaksanaan penilaian dapat menggunakan alat/instrumen penilaian berupa daftar cek ataupun skala penilaian. Beberapa contoh kegiatan peserta didik dalam penilaian proyek: (1) Penelitian sederhana tentang penggunaan listrik di rumah; (2) penelitian sederhana tentang perkembangan harga sembako, dll.

d. Penilaian Produk

Penilaian produk adalah penilaian terhadap proses pembuatan dan kualitas suatu produk. Penilaian produk meliputi penilaian kemampuan peserta didik membuat produk-produk teknologi dan seni, seperti: makanan, pakaian, hasil karya seni (patung, lukisan, gambar), barang-barang terbuat dari kayu, keramik, plastik, dan logam. Pengembangan produk meliputi 3 (tiga) tahap dan setiap tahap perlu diadakan penilaian yaitu:

- 1) Tahap persiapan, meliputi: penilaian kemampuan peserta didik dan merencanakan, menggali, dan mengembangkan gagasan, dan mendesain produk.
- 2) Tahap pembuatan produk (proses), meliputi: penilaian kemampuan peserta didik dalam menyeleksi dan menggunakan bahan, alat, dan teknik.
- 3) Tahap penilaian produk (appraisal), meliputi: penilaian produk yang dihasilkan peserta didik sesuai kriteria yang ditetapkan.

Penilaian produk biasanya menggunakan cara holistik atau analitik.

- a. Cara analitik, yaitu berdasarkan aspek-aspek produk, biasanya dilakukan terhadap semua kriteria yang terdapat pada semua tahap proses pengembangan (tahap: persiapan, pembuatan produk, penilaian produk).
- b. Cara holistik, yaitu berdasarkan kesan keseluruhan dari produk, biasanya dilakukan hanya pada tahap penilaian produk (appraisal).

e. Penilaian Portofolio

Penilaian portofolio merupakan penilaian berkelanjutan yang didasarkan pada kumpulan informasi yang menunjukkan perkembangan kemampuan peserta didik dalam satu periode tertentu. Informasi tersebut dapat berupa karya peserta didik dari proses pembelajaran yang dianggap terbaik oleh peserta didik, lembar jawaban tes yang menunjukkan soal yang mampu dan tidak mampu dijawab (bukan nilai) atau bentuk informasi lain yang terkait dengan kompetensi tertentu dalam satu mata pelajaran.

Penilaian portofolio pada dasarnya menilai karya-karya siswa secara individu pada satu periode untuk suatu mata pelajaran. Akhir suatu periode hasil karya tersebut dikumpulkan dan dinilai oleh guru dan peserta didik sendiri. Berdasarkan informasi perkembangan tersebut, guru dan peserta didik sendiri dapat menilai perkembangan kemampuan peserta didik dan terus melakukan perbaikan. Dengan demikian, portofolio dapat memperlihatkan perkembangan kemajuan belajar peserta didik melalui karyanya, antara lain: karangan, puisi, surat, komposisi musik, gambar, foto, lukisan, resensi buku/literatur, laporan penelitian, sinopsis, dsb.

Hal-hal yang perlu diperhatikan dan dijadikan pedoman dalam penggunaan penilaian portofolio di sekolah, antara lain:

- 1) Karya siswa adalah benar-benar karya peserta didik itu sendiri.
Guru melakukan penelitian atas hasil karya peserta didik yang dijadikan bahan penilaian portofolio agar karya tersebut merupakan hasil karya yang dibuat oleh peserta didik itu sendiri.
- 2) Saling percaya antara guru dan peserta didik.
Dalam proses penilaian guru dan peserta didik harus memiliki rasa saling percaya, saling memerlukan dan saling membantu sehingga terjadi proses pendidikan berlangsung dengan baik.
- 3) Kerahasiaan bersama antara guru dan peserta didik.
Kerahasiaan hasil pengumpulan informasi perkembangan peserta didik perlu dijaga dengan baik dan tidak disampaikan kepada pihak-pihak yang tidak berkepentingan sehingga memberi dampak negatif proses pendidikan
- 4) Milik bersama (joint ownership) antara peserta didik dan guru.
Guru dan peserta didik perlu mempunyai rasa memiliki berkas portofolio sehingga peserta didik akan merasa memiliki karya yang dikumpulkan dan akhirnya akan berupaya terus meningkatkannya.
- 5) Kepuasan
Hasil kerja portofolio sebaiknya berisi keterangan dan atau bukti yang memberikan dorongan peserta didik untuk lebih meningkatkan diri.
- 6) Kesesuaian
Hasil kerja yang dikumpulkan adalah hasil kerja yang sesuai dengan kompetensi yang tercantum dalam kurikulum.
- 7) Penilaian proses dan hasil
Penilaian portofolio menerapkan prinsip proses dan hasil. Proses belajar yang dinilai misalnya diperoleh dari catatan guru tentang kinerja dan karya peserta didik.
- 8) Penilaian dan pembelajaran
Penilaian portofolio merupakan hal yang tak terpisahkan dari proses pembelajaran. Manfaat utama penilaian ini sebagai diagnostik yang sangat berarti bagi guru untuk melihat kelebihan dan kekurangan peserta didik.

Teknik penilaian portofolio di dalam kelas memerlukan langkah-langkah sebagai berikut:

- 1) Jelaskan kepada peserta didik bahwa penggunaan portofolio, tidak hanya merupakan kumpulan hasil kerja peserta didik yang digunakan oleh guru untuk penilaian, tetapi digunakan juga oleh peserta didik sendiri. Dengan melihat portofolionya peserta didik dapat mengetahui kemampuan, keterampilan, dan minatnya. Proses ini tidak akan terjadi secara spontan, tetapi membutuhkan waktu bagi peserta didik untuk belajar meyakini hasil penilaian mereka sendiri.
- 2) Tentukan bersama peserta didik sampel-sampel portofolio apa saja yang akan dibuat. Portofolio antara peserta didik yang satu dan yang lain bisa sama bisa berbeda. Misalnya, untuk kemampuan menulis peserta didik mengumpulkan karangan-karangannya. Sedangkan untuk kemampuan menggambar, peserta didik mengumpulkan gambar-gambar buaatannya.
- 3) Kumpulkan dan simpanlah karya-karya tiap peserta didik dalam satu map atau folder di rumah masing-masing atau loker masing-masing di sekolah.
- 4) Berilah tanggal pembuatan pada setiap bahan informasi perkembangan peserta didik sehingga dapat terlihat perbedaan kualitas dari waktu ke waktu.
- 5) Sebaiknya tentukan kriteria penilaian sampel portofolio dan bobotnya dengan para peserta didik sebelum mereka membuat karyanya . Diskusikan cara penilaian kualitas karya para peserta didik. Contoh, kriteria penilaian kemampuan menulis karangan yaitu: penggunaan tata bahasa, pemilihan kosa-kata, kelengkapan gagasan, dan sistematika penulisan. Dengan demikian, peserta didik mengetahui harapan (standar) guru dan berusaha mencapai standar tersebut.
- 6) Minta peserta didik menilai karyanya secara berkesinambungan. Guru dapat membimbing peserta didik, bagaimana cara menilai dengan memberi keterangan tentang kelebihan dan kekurangan karya tersebut, serta bagaimana cara memperbaikinya. Hal ini dapat dilakukan pada saat membahas portofolio.
- 7) Setelah suatu karya dinilai dan nilainya belum memuaskan, maka peserta didik diberi kesempatan untuk memperbaiki. Namun, antara peserta didik dan guru perlu dibuat “kontrak” atau perjanjian mengenai jangka waktu perbaikan, misalnya 2 minggu karya yang telah diperbaiki harus diserahkan kepada guru.
- 8) Bila perlu, jadwalkan pertemuan untuk membahas portofolio. Jika perlu, undang orang tua peserta didik dan diberi penjelasan tentang maksud serta tujuan portofolio, sehingga orangtua dapat membantu dan memotivasi anaknya.

Contoh Penilaian Portofolio

Mata Pelajaran : Fisika
 Alokasi Waktu : 1 Semester
 Sampel yang dikumpulkan : Laporan kerja proyek
 Nama Siswa : _____
 Kelas : XII/1

No	Standar Kompetensi i/ Kompetensi i Dasar	Kriteria					Keterangan
		Periode	Sistematika laporan	Kelengkapan data	Ketajaman pembahasan	Kesimpulan	

--	--	--	--	--	--	--	--

Catatan:

Setiap karya siswa sesuai Standar Kompetensi/Kompetensi Dasar yang masuk dalam daftar portofolio dikumpulkan dalam satu file (tempat) untuk setiap peserta didik sebagai bukti pekerjaannya. Skor untuk setiap kriteria menggunakan skala penilaian 0 - 10 atau 0 - 100. Semakin baik hasil yang terlihat dari tulisan peserta didik, semakin tinggi skor yang diberikan.

Kolom keterangan diisi dengan catatan guru tentang kelemahan dan kekuatan tulisan yang dinilai.

f. Penilaian Diri (*self assessment*)

Penilaian diri adalah suatu teknik penilaian di mana peserta didik diminta untuk menilai dirinya sendiri berkaitan dengan status, proses dan tingkat pencapaian kompetensi yang dipelajarinya dalam mata pelajaran tertentu didasarkan atas kriteria atau acuan yang telah disiapkan. Tujuan utama dari penilaian diri adalah untuk mendukung atau memperbaiki proses dan hasil belajar. Meskipun demikian, hasil penilaian diri dapat digunakan guru sebagai bahan pertimbangan untuk memberikan nilai. Peran penilaian diri menjadi penting bersamaan dengan bergesernya pusat pembelajaran dari guru ke siswa yang didasarkan pada konsep belajar mandiri (*autonomous learning*). Ada beberapa jenis penilaian diri, diantaranya:

- 1) Penilaian Langsung dan Spesifik, yaitu penilaian secara langsung, pada saat atau setelah selesai melakukan tugas, untuk menilai aspek-aspek kompetensi tertentu dari suatu mata pelajaran.
- 2) Penilaian Tidak Langsung dan Holistik, yaitu penilaian yang dilakukan dalam kurun waktu yang panjang, untuk memberikan penilaian secara keseluruhan.
- 3) Penilaian Sosio-Afektif, yaitu penilaian terhadap unsur-unsur afektif atau emosional. Misalnya, peserta didik dapat diminta untuk membuat tulisan yang memuat curahan perasaannya terhadap suatu objek tertentu.

Penggunaan teknik ini dapat memberi dampak positif terhadap perkembangan kepribadian seseorang. Keuntungan penggunaan penilaian diri di kelas antara lain:

- 1) Dapat menumbuhkan rasa percaya diri peserta didik, karena mereka diberi kepercayaan untuk menilai dirinya sendiri;
- 2) Peserta didik menyadari kekuatan dan kelemahan dirinya, karena ketika mereka melakukan penilaian, harus melakukan introspeksi terhadap kekuatan dan kelemahan yang dimilikinya;
- 3) Dapat mendorong, membiasakan, dan melatih peserta didik untuk berbuat jujur, karena mereka dituntut untuk jujur dan objektif dalam melakukan penilaian.

3. Penentuan kualitas instrumen nontes yang dikembangkan

Instrumen nontes yang dikembangkan hendaknya memenuhi kriteria kualitas instrumen seperti reliabilitas, validitas, daya pembeda. Karena instrumen nontes tidak menilai benar tidaknya jawaban responden, maka tidak perlu mengukur tingkat kesukaran soal seperti pada instrumen tes.

a. Reliabilitas

Teknik penentuan reliabilitas dapat menggunakan test-retest, tes paralel, tes belah dua. Rumus yang dapat digunakan antara lain: korelasi product moment, Spermans Brown,

Alpha, dll. Reliabilitas instrumen ditentukan dari besarnya koefisien korelasi. Koefisien reliabilitas mencerminkan hubungan skor skala yang diperoleh (X) dengan skor sesungguhnya (skor murni). Koefisien reliabilitas sebesar 0,9 memiliki arti perbedaan yang tampak pada skala mampu mencerminkan 90% dari variansi skor murni

b. Validitas

Kevalidan instrumen nontes yang dikembangkan dapat dilihat dari validitas isi dan konstruk teori yang mendasarinya. Ketepatan dalam menentukan definisi konseptual, definisi operasional, dan penetapan aspek serta indikator yang direncanakan. Faktor yang melemahkan validitas:

- 1) Identifikasi kawasan ukur tidak cukup jelas
- 2) Operasionalisasi konsep (perumusan indikator) tidak tepat
- 3) Penulisan butir tidak mengikuti kaidah
- 4) Administrasi skala yang tidak hati-hati (kondisi subjek, kondisi testing)
- 5) Pemberian skor tidak cermat
- 6) Interpretasi yang keliru

c. Daya Beda

Pertanyaan/ Pernyataan pada instrumen nontes harus dapat membedakan sikap positif dan negatif. Indeks daya beda juga merupakan indikator keselarasan atau konsistensi antara fungsi butir dengan fungsi skala secara keseluruhan. Rumus yang dapat digunakan antara lain skor interval, korelasi product moment Pearson.

Daftar Pustaka

- Andersen, Lorin W. (1981). *Assessing Affective Characteristic in The School*. California: Brocks Cole Publishing Company.
- Asmawi Zainul. (2001). *Alternative Assessment*. Jakarta: Proyek Universitas Terbuka
- Crocker & Algina, *Introduction to Classical & Modern Test Theory*
- Depdiknas. (2005). *Pedoman Khusus: Pengembangan Instrumen dan Penilaian Ranah Afektif*. Jakarta: Dikmenum Depdiknas
- Depdiknas. (2006). *Model Penilaian Kelas: Sekolah Menengah Pertama dan Madrasah Tsanawiyah*. Jakarta: Pusat Kurikulum, Balitbang, Diknas
- Popham, W. J. (1995). *Classroom Assessment: What Teachers Need to Know*. Needham Heights, Mass: Allyn & Bacon
- Ryan, D. C. (1980). *Characteristics of Teacher A research study: Their description, comparison, and appraisal*. Wasington, DC: American Council of Education.
- Saifuddin Azwar. 2003. *Penyusunan Skala Psikologi*. Yogyakarta: Pustaka Pelajar
- Subali B. dan Paidi (2002). *Penilaian Pencapaian Hasil Belajar*. Yogyakarta. Universitas Negeri Yogyakarta.

5x

LEMBAR TANYA JAWAB
SEMINAR NASIONAL EVALUASI PENDIDIKAN (SNEP) I
PPs UNNES, 13 JULI 2013

Ruang : G11 / Aula PPs Unnes
Moderator : Dr. Masrukhin, M.Si

Nama Penyaji : Prof. Dr. Ani Kusrawati, M.Pd
Instansi : Unnes
Judul : Pengembangan Instrumen Non tes

Nama Peserta : Leus, Anggun
Instansi : Unnes

Pertanyaan
1. Untuk mengukur aspek sosial dalam alat non tes harus dibuat se kreatifnya, agar alat yang tidak sungguh-sungguh bisa terlihat dan ada manfaat.
2. Pada mata kuliah disebutkan rentang skor 33 - 40 (+). skor didapatkan dari apa? Dasarnya apa?

Jawab
1. sebaiknya, kalau mau dibuat kreatif, jangan dicampurkan antara (+) dengan (-). kalau hasilnya kurang maksimal, bisa dicocokkan dengan wawancara, dll.
2. yang menentukan rentang skor itu si peneliti sendiri.

Pemakalah