



BAB 1

DEFINISI PENGUKURAN, PENILAIAN, DAN EVALUASI

A. Pengukuran

Scriven (1994) menyebutkan dimana pengukuran adalah penentuan besaran kuantitas, biasanya pada skala uji yang dirujuk kriteria atau skala numerik kontinyu. Apapun yang digunakan untuk melakukan pengukuran disebut alat ukur. Dalam konteks tertentu, kita memperlakukan pengamat instrumen yang membutuhkan kalibrasi atau validasi. Pengukuran adalah komponen evaluasi standar yang umum dan terkadang besar, namun merupakan bagian yang sangat kecil dari logikanya, yaitu pembenaran untuk kesimpulan evaluatif.

Menurut Oriondo (1998) Pengukuran didefinisikan proses dimana informasi tentang atribut atau karakteristik benda ditentukan dan dibedakan.

B. Penilaian

Dalam Peraturan Menteri Pendidikan dan Kebudayaan Indonesia telah disebutkan tentang

berbagai macam definisi tentang penilaian dan seperangkat hal yang menyangkut tentang hal tersebut. Berikut ini akan dipaparkan tentang pengertian tersebut.

Penilaian Hasil Belajar oleh Pendidik adalah proses pengumpulan informasi/bukti tentang capaian pembelajaran peserta didik dalam kompetensi sikap spiritual dan sikap sosial, kompetensi pengetahuan, dan kompetensi keterampilan yang dilakukan secara terencana dan sistematis, selama dan setelah proses pembelajaran.

Pendekatan Penilaian adalah proses atau jalan yang ditempuh dalam melakukan penilaian hasil belajar peserta didik.

Bentuk Penilaian adalah cara yang dilakukan dalam menilai capaian pembelajaran peserta didik, misalnya: penilaian unjuk kerja, penilaian proyek, dan penilaian tertulis.

Instrumen Penilaian adalah alat yang digunakan untuk menilai capaian pembelajaran peserta didik, misalnya: tes dan skala sikap

Penilaian Autentik adalah bentuk penilaian yang menghendaki peserta didik menampilkan sikap, menggunakan pengetahuan dan keterampilan yang diperoleh dari pembelajaran dalam melakukan tugas pada situasi yang sesungguhnya.

Penilaian Diri adalah teknik penilaian sikap, pengetahuan, dan keterampilan yang dilakukan sendiri oleh peserta didik secara reflektif.

Penilaian Tugas adalah penilaian atas proses dan hasil pengerjaan tugas yang dilakukan secara mandiri dan/atau kelompok.

Penilaian Proyek adalah penilaian terhadap suatu tugas berupa suatu investigasi sejak dari perencanaan, pelaksanaan, pengolahan data, sampai pelaporan.

Penilaian berdasarkan Pengamatan adalah penilaian terhadap kegiatan peserta didik selama mengikuti proses pembelajaran.

Ulangan Harian adalah penilaian yang dilakukan setiap menyelesaikan satu muatan pembelajaran.

Ulangan Tengah Semester adalah penilaian yang dilakukan untuk semua muatan pembelajaran yang diselesaikan dalam paruh pertama semester.

Ulangan Akhir Semester adalah penilaian yang dilakukan untuk semua muatan pembelajaran yang diselesaikan dalam satu semester.

C. Evaluasi

Menurut pendapat William, dkk kata kunci dari istilah 'evaluasi' mengacu pada proses penentuan nilai, nilai, atau nilai sesuatu, atau produk dari proses itu. Istilah yang digunakan untuk merujuk pada proses ini atau sebagian

darinya meliputi: menilai, menganalisis, menilai, mengkritik, memeriksa, menilai, memeriksa, menilai, menilai, menilai, meninjau, mempelajari, menguji, mengukur. Proses evaluasi biasanya melibatkan beberapa identifikasi standar kelayakan, nilai, atau nilai yang relevan; beberapa penyelidikan terhadap kinerja evaluasi terhadap standar ini; dan beberapa integrasi atau sintesis hasil untuk mencapai evaluasi keseluruhan. Ini kontras dengan proses pengukuran, yang juga melibatkan perbandingan pengamatan terhadap standar, dalam hal itu (i) pengukuran secara khas tidak berkaitan dengan prestasi, hanya dengan sifat 'deskriptif murni', dan (ii) sifat-sifat tersebut bersifat unidimensional, yang menghindari kebutuhan untuk langkah integrasi. Proses integrasi terkadang menghakimi, terkadang akibat perhitungan kompleks, sangat umum hibrida keduanya. Singkatnya [evaluasi] adalah sine qua non pemikiran cerdas dan tindakan, dan khususnya praktik profesional.

Ditambah lagi dengan pendapat dari Mertenz yang mengatakan bahwa evaluasi adalah proses penyelidikan terapan untuk mengumpulkan bukti yang berpuncak pada kesimpulan tentang keadaan, nilai, manfaat, nilai, kepentingan, atau kualitas program, produk, orang, kebijakan, proposal, atau rencana. Kesimpulan yang dibuat dalam evaluasi mencakup aspek empiris (ada sesuatu yang terjadi) dan aspek normative (penilaian tentang nilai sesuatu). Ini adalah nilai murni yang membedakan evaluasi dari jenis penyelidikan lainnya, seperti penelitian sains dasar, epidemiologi klinis, jurnalisme investigatif, atau *polling public*.

Kesimpulan:

Pengukuran adalah proses meng-angka-kan suatu objek. Dalam hal ini produk dari pengukuran adalah berupa angka.

Penilaian adalah proses intepretasi atau mengartikan dari hasil pengukuran yang telah

dilakukan. Penilaian ini terdapat dua jenis yaitu penilaian yang berbentuk kualitatif yaitu yang berupa deskripsi ataupun narasi dan satu bentuk lainnya yaitu pengukuran kuantitatif yaitu yang hasilnya berupa data atau angka.

Evaluasi adalah proses menyimpulkan akhir yang produknya berupa keputusan terkait dengan proses yang telah dijalani sebelumnya, yaitu berupa pengukuran dan penilaian.

Seorang guru mengadakan posttest untuk melihat kemampuan belajar matematika peserta didik dengan cara mengujikan 10 soal tes dengan ketentuan skor 1 untuk tiap-tiap jawaban benar dan skor 0 untuk tiap jawaban salah atau tidak menjawab.

Dari test tersebut diperoleh hasil :

- a. Dinda mendapatkan skor 8
- b. Sani mendapatkan skor 6
- c. Khriel mendapatkan skor 9
- d. Andrew mendapatkan skor 7
- e. Soimah mendapatkan skor 4

Lalu langkah selanjutnya yang terkait dengan penilaian adalah ketika guru memberikan sebuah batasan minimal atau yang biasa kita sebut dengan KKM yaitu berupa nilai minimal sebesar 70. Sehingga penilaian yang dilakukan oleh guru dari hasil pengukuran tersebut, antara lain:

- a. Dinda menguasai materi dengan baik dan dapat lulus KKM
- b. Sani belum menguasai materi dengan baik dan belum lulus KKM
- c. Khiel sangat menguasai materi dengan baik dan mendapat melampaui KKM
- d. Andrew menguasai materi dengan cukup baik dan mendapat skor sesuai KKM
- e. Soimah belum mampu menguasai materi dengan baik dan belum lulus KKM

Evaluasi dari hasil pengukuran dan penilaian yang telah dilakukan dapat berupa keputusan terkait dengan kegiatan pembelajaran yang akan diterapkan pada waktu yang akan datang. Berikut ini akan diberikan contoh keputusan dari tahap evaluasi

berdasarkan hasil pengukuran dan penilaian yang telah dilakukan oleh guru. Terdapat dua anak yang belum mencapai KKM dengan skor terendah 4. Dikarenakan terdapat angka selisih yang cukup besar pada anak yang mendapatkan skor terendah dengan skor KKM maka dapat diidentifikasi bahwa ia memerlukan perhatian khusus mengenai pembelajaran matematika ini sehingga guru perlu meninjau ulang metode, media, dan sumber pembelajarannya apakah sudah mencakup semua (holistik) apa belum. Bila belum maka pembelajaran selanjutnya diperbaiki (evaluasi sumatif).



BAB 2

MACAM-MACAM PENILAIAN

Penilaian yang ada di sekolah dasar dapat dibedakan menjadi macam-macam penilaian yang ada. Berikut ini akan dipaparkan tentang macam-macam penilaian berdasarkan sumber dari buku:

A. Penilaian Formatif

Penilaian formatif merupakan bagian integral dari proses belajar mengajar. Ini tidak memberikan kontribusi pada nilai akhir yang diberikan untuk modul; melainkan berkontribusi untuk belajar melalui pemberian umpan balik. Ini harus menunjukkan apa yang baik tentang sebuah karya dan mengapa karya tersebut dianggap bagus; itu juga harus menunjukkan apa yang tidak begitu bagus dan bagaimana pekerjaan bisa diperbaiki. Umpan balik formatif yang efektif akan mempengaruhi apa yang dilakukan oleh siswa dan guru selanjutnya.

Fungsi dari penilaian formatif ini adalah untuk memperbaiki proses pembelajaran kearah yang lebih baik dan efisien atau

memperbaiki satuan atau rencana pembelajaran. Sejalan dengan yang dikatakan oleh Mark D. Shermis dan Frances J. Di vesta bahwa penilaian formatif mengacu pada penafsiran yang berguna untuk meningkatkan kegiatan pembelajaran (*assessment for learning*).

Menurut Lorin W Anderson dan David R Katwohl, asesmen formatif bertujuan untuk memonitor pembelajaran siswa dan memperbaiki pembelajarannya, demi kepentingan individual dan kolektif siswa (h.370-371). Sehingga dapat mengetahui hingga dimana penguasaan peserta didik tentang materi yang diajarkan dalam satu rencana atau satuan pelajaran.

Aspek-aspek yang dinilai pada penailain formatif adalah : hasil kemajuan belajar siswa yang meliputi : pengetahuan, keterampilan, sikap terhadap materi ajar agama yang di sajikan.

B. Penilaian Sumatif

Digunakan untuk mengukur prestasi, memberi penghargaan prestasi, memberikan data seleksi (ke tahap selanjutnya dalam pendidikan atau ke pekerjaan). Untuk semua alasan ini, validitas dan reliabilitas penilaian sumatif sangat penting. Penilaian sumatif dapat memberikan informasi yang memiliki nilai formatif/diagnostik. Penilaian sumatif ini biasanya dilakukan pada akhir tahun pelajaran atau pada akhir semester yang sedang berjalan.

Untuk mengetahui angka atau nilai murid setelah mengikuti program belajar dalam satu semester. Hal ini dikuatkan dengan pernyataan Scriven (1967: 371) bahwa fungsi utama asesmen sumatif adalah menyimpulkan pembelajaran siswa pada akhir periode pembelajaran.

Menurut Lorin W Anderson dan David R Katwoh, asesmen somatif bertujuan untuk

memberi nilai siswa yang telah mengikuti rangkaian pembelajaran (374). Jadi, asesmen sumatif dapat mengetahui taraf hasil belajar yang dicapai oleh peserta didik setelah melakukan program pembelajaran dalam satu semester, akhir tahun atau akhir program pembelajaran pada suatu unit pendidikan tertentu.

Aspek-aspek yang dinilai ialah kemajuan hasil belajar meliputi: pengetahuan, keterampilan, sikap dan penguasaan murid tentang materi pembelajaran yang diberikan.

Penilaian sumatif inilah yang begitu sering dipraktikkan dalam kegiatan pembelajaran yang ada di Indonesia. Hal ini membuat pembelajaran begitu terfokus pada penilaian akhir pembelajaran.

Berikut ini akan dipaparkan tabel tentang perbedaan penilaian formatif dengan penilaian sumatif.

Penilaian Formatif	Penilaian Sumatif
Kuis, tes, berbagai bentuk (esai dan objektif) dalam ulangan harian	Ujian penghabisan (Ujian akhir semester, ujian tengah semester)
Tes diagnostik	Ujian nasional
Laporan laboratorium	Laporan proyek dalam 1 tahun
Bersifat informal	Lebih bersifat formal
Sumber penilaian variatif	Sumber penilaian terbatas

Tabel 1. Perbedaan Penilaian Formatif dan Penilaian Sumatif

C. Penilaian Autentik

Penilaian di mana tugas dan kondisinya lebih sesuai dengan apa yang akan dialami dalam pekerjaan. Bentuk penilaian ini dirancang untuk mengembangkan keterampilan dan kompetensi siswa disamping pengembangan akademik.

D. Penilaian Diagnostik

Penilaian diagnostik ditujukan untuk pengalaman pelajar dan tingkatan prestasi mereka. Namun, penilaian ini mengarah ke belakang daripada ke depan. Ini menilai apa yang pelajar sudah tahu dan / atau sifat yang mungkin dimiliki pelajar, yang jika tidak terdiagnosis maka hal ini akan memberikan batasan kepada siswa dalam hal kegiatan pembelajaran.

E. Penilaian Dinamik

Penilaian dinamis mengukur apa yang dicapai siswa saat diberi beberapa pengajaran dalam topik atau bidang yang tidak mereka kenal. Contohnya mungkin penilaian tentang seberapa banyak orang Swedia belajar dalam blok pendek pengajaran kepada siswa yang tidak memiliki pengetahuan sebelumnya tentang bahasa tersebut. Hal ini dapat berguna untuk menilai

potensi pembelajaran spesifik tanpa adanya pencapaian sebelumnya yang relevan, atau untuk menilai potensi pembelajaran umum bagi siswa yang memiliki latar belakang yang sangat tidak beruntung. Hal ini sering digunakan saat permulaan pembelajaran inti.

F. Penilaian Sinoptik

Penilaian sinoptik mendorong siswa untuk menggabungkan elemen pembelajaran mereka dari berbagai bagian program dan untuk menunjukkan akumulasi pengetahuan dan pemahaman mereka tentang topik atau bidang subjek. Penilaian sinoptik biasanya memungkinkan siswa untuk menunjukkan kemampuan mereka untuk mengintegrasikan dan menerapkan keterampilan, pengetahuan dan pemahaman mereka dengan luas dan mendalam dalam subjek. Ini dapat membantu untuk menguji kemampuan siswa dalam menerapkan pengetahuan dan pemahaman

yang diperoleh di salah satu bagian program untuk meningkatkan pemahaman mereka di bagian lain program ini, atau di seluruh program secara keseluruhan. Penilaian sinoptik dapat menjadi bagian dari bentuk penilaian lainnya.

G. Penilaian Berdasarkan Kriteria

Setiap prestasi siswa dinilai berdasarkan kriteria tertentu. Pada prinsipnya tidak ada perhitungan yang diambil tentang bagaimana siswa lain melakukannya. Dalam praktiknya, pemikiran normatif dapat mempengaruhi penilaian apakah kriteria tertentu telah dipenuhi atau tidak. Keandalan dan validitas harus diyakinkan melalui proses seperti moderasi, uji coba, dan pengumpulan contoh.

H. Penilaian Ipsatif

Ini adalah penilaian terhadap standar siswa sebelumnya. Ini dapat mengukur seberapa

baik tugas tertentu telah dilakukan terhadap pencapaian rata-rata siswa, terhadap pekerjaan terbaik mereka, atau terhadap karya terbaru mereka. Penilaian Ipsatif cenderung berkorelasi dengan usaha, untuk mempromosikan atribusi keberhasilan berbasis usaha, dan untuk meningkatkan motivasi belajar.

I. Penilaian Unjuk Kerja

Merupakan jenis penilaian alternatif. Apakah latihan di mana seorang siswa menunjukkan keahlian dan kompetensi spesifik sehubungan dengan serangkaian standar atau kualitas yang disepakati. Merefleksikan kinerja siswa pada tugas instruksional dan bergantung pada penilaian profesional penilai dalam disain dan interpretasinya.

Kesimpulan:

Bahwa terdapat 4 (empat) macam penilaian yang secara umum diterapkan pada pendidikan di Indonesia, yakni:

1. Penilaian Diagnostik yang merupakan penilaian pada waktu awal proses pembelajaran dan atau saat peserta didik mengalami kesulitan belajar diadakan guna mengembangkan proses pembelajaran yang efektif.
2. Penilaian Formatif yang merupakan penilaian yang dilakukan selama proses pembelajaran berlangsung, hal ini dilakukan guna memperbaiki setiap kekurangan pada proses pembelajaran sebelumnya.
3. Penilaian Sumatif yang merupakan penilaian akhir dari sebuah proses pembelajaran untuk mengetahui tingkat keberhasilan proses pembelajaran tersebut.
4. Penilaian Ipsatif yang merupakan penilaian kemajuan belajar peserta didik dengan

memperhatikan hasil kinerja sebelumnya untuk mengetahui tingkat perkembangan siswa.

5. Penilaian autentik yang merupakan penilaian penting yang diterapkan pada Kurikulum 2013 di Indonesia saat ini.



BAB 3

TUJUAN, TEKNIK, DAN ALAT PENILAIAN

A. Tujuan Penilaian

Dalam Lampiran Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 104 Tahun 2014 menyebutkan empat tujuan penilaian, yaitu:

1. Mengetahui tingkat penguasaan kompetensi dalam sikap, pengetahuan, dan keterampilan yang sudah dan belum dikuasai seseorang/kelompok peserta didik untuk ditingkatkan dalam pembelajaran remedial dan program pengayaan.
2. Menetapkan ketuntasan penguasaan kompetensi belajar peserta didik dalam kurun waktu tertentu, yaitu harian, tengah semester, satu semesteran, satu tahunan, dan masa studi satuan pendidikan.
3. Menetapkan program perbaikan atau pengayaan berdasarkan tingkat penguasaan kompetensi bagi mereka yang diidentifikasi sebagai peserta didik yang lambat atau cepat dalam belajar dan pencapaian hasil belajar.

4. Memperbaiki proses pembelajaran pada pertemuan semester berikutnya.

Dalam buku *Classroom Assessment in Action* menyebutkan bahwa tujuan asesmen selama belajar adalah untuk menyediakan data yang akan berguna dalam memantau kemajuan siswa dan, dengan berbuat demikian, menginformasikan efektivitas pengajaran.

Pada tahun 2006, Western and Northern Canada Protocol for Collaboration in Education (WNCPE) mengemukakan sebuah dokumen yang merangkum tiga tujuan penilaian, yaitu:

1. *Assessment For Learning*

Istilah "penilaian untuk belajar" sering digunakan secara bergantian dengan "penilaian formatif" sebagaimana dibuktikan oleh bab Black dan William (sekitar tahun 2006) seputar praktik penilaian formatif yang berjudul "penilaian untuk belajar di kelas" dan Harlen's (2006) menyatakan penggunaan istilah. Menurut Harlen, yang paling penting

bukan jenis penilaian yang digunakan, tapi tujuan penggunaannya. Refleksi Black dan William menangkap pembelajaran yang terfokus pada penilaian formatif sambil memberikan rincian tentang proses penilaian dan sifat konstruktivisnya. Taras (2010) tidak setuju dan menjelaskan bahwa terlepas dari bagaimana pekerjaan itu digunakan, "proses penilaian formatif hanya dapat dikatakan telah terjadi ketika umpan balik telah digunakan untuk memperbaiki pekerjaan" (P.3021). Dia merasa bahwa asesmen formatif (penilaian untuk belajar) adalah proses interaksi antara apa yang dia sebut sebagai penilaian sumatif dan umpan balik namun ketika dia mengacu pada penilaian sumatif, dia merujuk secara khusus pada fakta bahwa sebuah keputusan harus dibuat tentang melukis dan Pengetahuan, sama seperti tahap evaluasi dalam siklus kita. Ini mungkin sangat membantu untuk

mengkonseptualisasikan penilaian untuk belajar sebagai mewujudkan tujuan utama penilaian (peningkatan pembelajaran), namun terutama sebagai proses penilaian umum yang diberlakukan karena didukung oleh tujuan penilaian lainnya seperti penilaian pembelajaran atau penilaian sumatif. Karena semua penilaian seharusnya untuk pembelajaran, diharapkan kebutuhan penggunaan terminologi pada akhirnya akan kuno.

2. *Assessment As Learning*

Penilaian sebagai pembelajaran bersifat formatif, namun secara khusus mengacu pada tugas belajar metakognitif. Secara keseluruhan, tujuan umum penilaian sebagai pembelajaran adalah untuk akhirnya menjembatani dan mengalihkan kepemilikan pembelajaran dari guru ke siswa dengan kata lain dari eksternal ke pemantauan internal. Contoh penilaian yang penting seperti

pembelajaran mencakup unsur-unsur apa yang oleh Boekaerts dan Corno (2005) dan Clark (2012) diskusikan sebagai pembelajaran mandiri dan apa yang Brookhart (2001) deskripsikan sebagai penilaian diri siswa. Bockaerts dan Corno melukis potret student yang mengatur diri sendiri sebagai berikut: "Semua teoretikus menganggap bahwa siswa yang mengatur pembelajaran mereka bergerak secara aktif dan konstruktif dalam proses pembangkitan makna dan mereka menyesuaikan pemikiran, perasaan, dan tindakan mereka sesuai kebutuhan. mempengaruhi motivasi belajar mereka " (hal.201).

Meskipun mereka mengeksplorasi peraturan diri siswa jauh lebih mendalam daripada yang bisa kita dapatkan di sini, deskripsi ini dengan indah menggambarkan internet asesmen sebagai pembelajaran. Dalam kata-kata Black dan William (1998), "pengguna

akhir informasi penilaian yang diyakinkan untuk memperbaiki pembelajaran adalah muridnya" (hal.142). Untuk alasan ini, penilaian sebagai pembelajaran sangat penting bagi siswa karena memengaruhi siswa, memotivasi mereka, dan mendukung transisi akhirnya ke pelajar seumur hidup.

3. *Assessment Of Learning*

Penilaian dari sebuah proses pembelajaran sering digunakan sebagai label lain untuk penilaian sumatif. Dengan asumsi bahwa semua penilaian harus bersifat formatif (yaitu, penilaian untuk pembelajaran), apakah penilaian pembelajaran masih layak dijadikan tujuan penilaian? Dalam peran tradisionalnya, penilaian pembelajaran digunakan terutama untuk pengambilan keputusan politik dan pendidikan, sehingga berkonotasi pada tingkat kontrol yang besar terhadap siswa (Taras, 2008). Oleh karena itu, ada protes keras terhadap penilaian

sumatif secara umum, khususnya karena penelitian yang benar-benar dirasakan dan aktual misusedespite yang menunjukkan bahwa penilaian sumatif berskala besar masih merupakan kebutuhan (Brookhart, 2003). Newton (2007) menunjukkan bahwa dengan cara istilah "sumatif" digunakan, ia tidak memiliki tujuan dan hanya mencirikan jenis penilaian. Kebenaran yang melekat dalam pernyataan ini terletak pada pendekatan tradisional yang bermasalah terhadap penilaian yang mengganggu siklus penilaian pada tahap evaluasi atau mencocok informasi tersebut ke lembaga organisasi eksternal tanpa umpan balik yang dikembalikan kepada siswa. Seperti Taras (2008) membedah masalah ini, dia menarik perhatian pada fakta bahwa penghakiman secara alami terjadi dan perlu, walaupun terkadang disalahgunakan. Namun jika didekati dengan benar, penilaian

pembelajaran bisa menjadi alat yang berharga untuk merangkum kumpulan rotasi siklus asesmen. Konseptualisasikannya sebagai sampel inti geologi yang mengkomunikasikan pembelajaran siswa selama periode waktu tertentu. Ini berhasil mencapai makroskop yang terjadi pada tingkat mikrokosmos di kelas setiap saat. Bila digunakan dengan tepat, umpan balik ini dapat memberi siswa pemahaman yang lebih baik, pada tingkat organisasi dapat memberi kesempatan belajar bagi guru dan administrator untuk belajar bagaimana mendukung pembelajaran siswa lebih lanjut, dan pada tingkat politik / masyarakat dapat memberikan informasi potensialnya, Bila digunakan dengan benar dapat berkontribusi pada proses pembuatan keputusan yang penting.

Pertanyaan yang kemudian muncul untuk Anda adalah apakah Anda tahu secara persis

apakah sebenarnya tujuan dari penilaian kelas. Secara rinci tujuan tersebut dapat dijabarkan sebagai berikut:

1. Dengan melakukan asesmen berbasis kelas ini pendidik dapat mengetahui seberapa jauh siswa dapat mencapai tingkat pencapaian kompetensi yang dipersyaratkan, baik selama mengikuti pembelajaran dan setelah proses pembelajaran berlangsung.
2. Saat melaksanakan asesmen ini, Anda sebagai pendidik juga akan bisa langsung memberikan umpan balik kepada peserta didik, sehingga tidak perlu lagi menunda atau menunggu ulangan semester untuk bisa mengetahui kekuatan dan kelemahannya dalam proses pencapaian kompetensi.
3. Dalam asesmen berbasis kelas ini, Anda juga secara terus menerus dapat melakukan pemantauan kemajuan belajar yang dicapai setiap peserta didik, sekaligus Anda dapat mendiagnosis kesulitan belajar yang dialami

peserta didik sehingga secara tepat dapat menentukan siswa mana yang perlu pengayaan dan siswa yang perlu pembelajaran remedial untuk mencapai kompetensi yang dipersyaratkan.

4. Hasil pemantauan kemajuan proses dan hasil pembelajaran yang dilakukan terus menerus tersebut juga akan dapat dipakai sebagai umpan balik bagi Anda untuk memperbaiki metode, pendekatan, kegiatan, dan sumber belajar yang digunakan, sesuai dengan kebutuhan materi dan juga kebutuhan siswa.
5. Hasil-hasil pemantauan tersebut, kemudian dapat Anda jadikan sebagai landasan untuk memilih alternatif jenis dan model penilaian mana yang tepat untuk digunakan pada materi tertentu dan pada mata pelajaran tertentu, yang sudah barang tentu akan berbeda. Anda sebagai pendidik yang tahu persis pertimbangan pemilihannya.

6. Hasil dari asesmen ini dapat pula memberikan informasi kepada orang tua dan komite sekolah tentang efektivitas pendidikan, tidak perlu menunggu akhir semester atau akhir tahun. Komunikasi antara pendidik, orang tua dan komite harus dijalin dan dilakukan terus menerus sesuai kebutuhan (Poerwanti, 2015)

Sudjana (2005) juga menyebutkan beberapa tujuan penilaian yang akan dipaparkan sebagai berikut:

1. Mendeskripsikan kecakapan belajar pada siswa sehingga dapat diketahui kelebihan dan kekurangan dalam berbagai bidang studi atau mata pelajaran yang ditempuhnya.
2. Mengetahui keberhasilan proses pendidikan dan pengajaran di sekolah, yakni seberapa jauh keefektifannya dalam mengubah tingkah laku para siswa ke arah tujuan pendidikan yang diharapkan.

3. Menentukan tindak lanjut hasil penilaian, yakni melakukan perbaikan dan penyempurnaan dalam hal program pendidikan dan pengajaran serta strategi pelaksanaannya.
4. Memberikan pertanggungjawaban (accountability) dari pihak sekolah kepada pihak-pihak yang berkepentingan.

Sedangkan menurut Chittenden (1994) juga mengemukakan bahwa tujuan penilaian adalah untuk "*keeping track, finding-out, summing-up, checking-up*" yang akan dipaparkan sebagai berikut:

1. *Keeping Track*

Yang merupakan untuk menelusuri dan melacak proses belajar peserta didik sesuai dengan rencana pelaksanaan pembelajaran yang telah ditetapkan. Untuk itu guru harus mengumpulkan data dan informasi dalam kurun waktu tertentu melalui berbagai jenis dan teknik penilaian untuk memperoleh

gambaran tentang pencapaian kemajuan belajar peserta didik.

2. *Checking-Up*

Yang merupakan untuk mengecek ketercapaian kemampuan peserta didik dalam proses pembelajaran dan kekurangan-kekuarangan peserta didik selama mengikuti proses pembelajaran. Dengan kata lain, guru perlu melakukan penilaian untuk mengetahui bagian mana dari materi yang sudah dikuasai peserta didik dan bagaimana dari materi yang belum dikuasai.

3. *Finding-Out*

Yang merupakan untuk mencari, menemukan dan mendeteksi kekurangan kesalahan atau kelemahan peserta didik dalam proses pembelajaran, sehingga guru dapat dengan cepat mencari alternatif solusinya.

4. *Summing-Up*

Yang merupakan untuk menyimpulkan tingkat penguasaan peserta didik terhadap

kompetensi yang telah ditetapkan. Hasil penyimpulan ini dapat digunakan guru untuk menyusun laporan kemajuan belajar ke berbagai pihak yang berkepentingan.

Metter (2003) juga mengemukakan tentang berbagai macam tujuan dari penilaian pada pembelajaran seperti berikut ini:

1. *Planning, Conducting, and Evaluating Instruction*

Salah satu tujuan penilaian yang paling mendasar dan penting adalah membantu dalam membuat keputusan mengenai instruksi. Ini dapat dibagi menjadi tiga jenis keputusan instruksional: yang terjadi sebelum instruksi, instruksi yang terjadi selama instruksi, dan instruksi yang mengikuti.

2. *Diagnosing Student Difficulties*

Penilaian diagnostik, merupakan kegiatan penting dalam membantu guru merencanakan pengajaran, jika siswa sudah

mengetahui beberapa gagasan dan konsep dasar pada satu unit pengajaran, guru dapat memutuskan untuk tidak meliputi materi tersebut atau mungkin hanya menawarkan tinjauan singkat kepada siswa, sebelum beralih konsep baru. Guru mungkin juga memperhatikan bahwa siswa tertentu mengalami jenis kesulitan yang umum, apakah masalah yang bersifat akademis, emosional, atau sosial.

3. *Placing The Student*

Guru kelas sering membuat keputusan tentang penempatan siswa. Keputusan penempatan dapat dilakukan berdasarkan kemampuan akademis dan sosial (airasian, 2000).

4. *Providing Feedback To The Student (Formative)*

Tujuan penilaian lain yang sangat penting adalah memberikan umpan balik kepada siswa mengenai kemajuan akademis mereka.

Evaluasi formatif didefinisikan sebagai pengambilan keputusan yang terjadi selama instruksi untuk menentukan penyesuaian terhadap instruksi yang harus dilakukan.

5. *Grading and Evaluating Academic (Summative)*

Ingat bahwa evaluasi sumatif berarti guru menilai dan mengevaluasi pembelajaran siswa di akhir matapelajaran, periode naik kelas, atau tahun ajaran. Penugasan kelas ini biasanya bagaimana guru berkomunikasi dengan orang lain (yaitu, orang tua, administrator, dan guru lainnya, serta siswa mereka sendiri) seberapa baik siswa telah melakukan di bawah asuhan mereka.

Kesimpulan:

Dari berbagai pendapat yang telah dipaparkan sebelumnya maka tujuan dari penilaian terdapat berbagai macam, penilaian untuk pembelajaran bertujuan untuk mengukur

tingkat kemampuan pemahaman siswa ketika mereka sedang melakukan proses pembelajaran, selain itu juga dapat untuk menempatkan posisi 'kecerdasan' siswa yang ada di kelas tersebut. Penilaian sebagai pembelajaran adalah penilaian yang bertujuan untuk mengiring proses pembelajaran yang ada di kelas, sehingga proses pembelajaran yang ada pun juga memiliki regulasi atau aturan terkait dengan penilaian terhadap siswa-siswa yang ada di kelas. Penilaian sebagai dari proses pembelajaran adalah penilaian akhir untuk mengetahui tingkat pemahaman siswa dan memetakan serta memberikan umpan balik kepada siswa terkait dengan keputusan dan tindakan apa yang akan dilakukan selanjutnya dalam proses pembelajaran di kelas.

B. Alat Penilaian

Dalam pengertian umum, alat adalah sesuatu yang dapat digunakan untuk mempermudah

seseorang untuk melaksanakan tugas atau mencapai tujuan secara efektif dan efisien. Kata "alat" bisa disebut juga dengan istilah "instrument". Dengan demikian, alat evaluasi dikenal dengan instrument evaluasi. Macam-macam alat penilaian:

1. *Standardised ability tests* (kadang-kadang disebut tes psikometri) dirancang untuk menjadi kurikulum yang mandiri dan mengukur sifat yang lebih tahan lama dan jangka panjang daripada pencapaian pembelajaran, yaitu mengukur apa yang dapat diketahui siswa daripada pengetahuan yang diketahui.
2. *Standardised achievement tests* (kadang-kadang disebut sebagai tes pencapaian) mengukur apa yang diketahui dan dapat digunakan siswa untuk menilai, misalnya, pengetahuan dan keterampilan siswa dalam melek huruf dan berhitung, dan untuk menentukan kemajuan di bidang ini. Dengan

berjalan kaki informasi yang dikumpulkan dari berbagai praktik penilaian termasuk tes kemampuan dan pencapaian, informasi lebih lanjut dapat dikumpulkan melalui administrasi tes diagnostik untuk memberikan pandangan yang lebih rinci tentang kekuatan dan kebutuhan belajar siswa.

3. *A diagnostic test* dirancang untuk memberikan informasi spesifik tentang kekuatan dan kebutuhan siswa dalam beberapa aspek pembelajaran, misalnya, keterampilan identifikasi kata atau pemahaman tentang konsep bilangan. Selain itu, konselor bimbingan dapat menggunakan hasil tes kemampuan dan minat untuk membimbing dan mendukung pengambilan keputusan siswa mengenai opsi pilihan subjek dan opsi siklus senior dan untuk mendukung pengembangan karir mereka, termasuk perencanaan untuk pendidikan

tinggi dan / atau pendidikan lanjutan dan latihan.

Pemilihan instrumen penilaian

Sekolah hanya boleh menggunakan instrumen penilaian yang sesuai dan memiliki tujuan yang jelas dan pasti. Secara khusus, perawatan harus dilakukan sehubungan dengan kesesuaian instrumen penilaian yang digunakan dengan siswa dengan kebutuhan pendidikan khusus (SEN) atau dengan siswa yang belajar bahasa Inggris sebagai Bahasa Tambahan (EAL).

Hasil penilaian seharusnya tidak dianggap definitif. Kebutuhan siswa dan informasi kontekstual lainnya harus selalu dipertimbangkan. Perlu dicatat bahwa tidak semua instrumen yang tersedia untuk sekolah memiliki norma Irlandia, dan beberapa belum direvisi selama beberapa tahun. Perhatian harus diberikan oleh pengguna untuk meneliti secara akurat kesesuaian, reliabilitas dan validitas semua instrumen penilaian yang dipilih untuk

digunakan, dan untuk memastikan bahwa versi yang digunakan paling mutakhir dan sesuai untuk diberikan kepada kohort siswa tertentu.

Penggunaan instrumen penilaian di sekolah untuk penilaian kepribadian tidak sesuai dan oleh karena itu instrumen semacam itu tidak tercantum dalam daftar di bawah ini. Petugas uji coba disarankan untuk memeriksa secara dekat manual tes dan informasi pendukung lainnya sebelum memilih tes. Dalam memilih instrumen penilaian sekolah harus menyadari kualifikasi yang dibutuhkan untuk mengelola, mencetak, menafsirkan dan memberikan umpan balik mengenai tes tertentu. Bila perlu, penerbit tes dapat dihubungi untuk mendapatkan informasi lebih rinci mengenai hal ini. National Educational Psychological Service (NEPS) dan National Centre for Guidance in Education (NCGE) juga memberikan informasi dan saran tentang penggunaan dan kesesuaian instrumen penilaian

tertentu. Lihat daftar referensi yang berguna di bawah ini.

Keterlibatan antara sekolah dasar dan pasca sekolah dasar akan menginformasikan keputusan seputar pemilihan dan waktu penilaian bagi siswa yang baru mengikuti pendidikan pasca sekolah dasar. Sekolah harus memanfaatkan Paspur Pendidikan dengan sebaik-baiknya untuk mendukung pembagian data yang relevan, termasuk hasil penilaian, saat siswa mentransfer dari sekolah dasar ke sekolah pasca sekolah dasar.

Menurut Yessy (2015) yang dimaksud dengan alat penilaian adalah instrument, perangkat, dokumen, dan hal-hal lainnya yang dapat dipakai guru sebagai alat untuk mengumpulkan hasil informasi. Dari berbagai alat penilaian ada alat yang digunakan mengukur hasil belajar siswa secara lebih tepat yang dapat dinamakan alat pengukur. Alat pengukur ini dapat dianalogikan dengan meteran yang digunakan untuk mengukur

panjang, lebar dan tinggi. Selain itu untuk menilai pengetahuan siswa sebagai belajar adalah perangkat tes presensi belajar selain itu untuk menilai apakah keadaan belajar mahasiswa dapat juga digunakan dokumen-dokumen lain, misalnya:

1. Buku catatan pribadi mahasiswa

Buku ini berisi data lengkap data pribadi peserta didik termasuk di dalamnya data minat dan bakat mereka, sikap dan kepribadian, watak dan kebiasaanya.

2. Karangan, laporan, karya tulis

Hasil karya tulis, karangan dan laporan dari peserta didik sangat bermanfaat bagi dosen terkait penilaian verbal (kemampuan peserta didik dalam menggunakan tata bahasa dalam mengerjakan tugas), kemampuan ide-ide cemerlang, dan mengekspresikan perasaannya.

3. Hasil penyelesaian tugas

Hasil kognitif merupakan hasil yang mana dapat dinilai. Dan hasil penilaian tersebut dapat dijadikan sebagai bahan masukan untuk mengevaluasi peserta didik.

4. Hasil wawancara

Pada sesekali waktu pendidik dapat melakukan wawancara dengan peserta didik maupun dengan wali untuk mengetahui hal hal tambahan yang dibutuhkan.

Kesimpulan:

Alat penilaian adalah berbagai macam benda atau sesuatu yang dapat digunakan untuk melakukan penilaian terhadap siswa di kelas. Alat penilaian juga sering disebut dengan instrumen penilaian. Instrumen penilaian terdiri dari berbagai macam yang jenisnya juga digunakan sesuai dengan kebutuhan penilai atau guru yang ada di dalam kelas. Hal ini menjadi hal

yang sangat penting untuk guru dan peserta didik yang ada di kelas saat pembelajaran.

C. Teknik Penilaian

Teknik penilaian memungkinkan memperoleh informasi yang andal dan valid. Meskipun demikian, jika pengamatan dasar, baik informal maupun formal, tidak mendapat informasi karena kurangnya pengetahuan, prosedur, atau pemahaman, hal ini secara fungsional tidak berguna.

Dalam memperoleh data, pendidik dapat menggunakan berbagai teknik penilaian secara komplementer (saling melengkapi) sesuai dengan kompetensi yang dinilai, sebagaimana diuraikan dalam panduan penilaian masing-masing kelompok mata pelajaran. Teknik penilaian adalah metode atau cara penilaian yang dapat digunakan guru untuk mendapatkan informasi tentang proses dan produk belajar peserta didik. Penggunaan berbagai teknik dan alat tersebut harus disesuaikan dengan tujuan

penilaian, waktu yang tersedia, sifat tugas yang dilakukan peserta didik, dan jumlah materi pembelajaran yang sudah disampaikan (Rosana, 2014).

Menurut Brookhart, Susan M. and Anthony, Nitko J. (2007 : 5) dilihat dari tekniknya, asesmen proses dan hasil belajar dibedakan menjadi dua macam yaitu dengan teknik tes dan non tes namun pada umumnya pengajar lebih banyak menggunakan tes sebagai alat ukur dengan rasional bahwa tingkat obyektivitas evaluasi lebih terjamin, hal ini tidak sepenuhnya benar. Anda bisa lebih jauh mencermati pada unit-unit selanjutnya.

- a. Teknik tes adalah seperangkat tugas yang harus dikerjakan oleh orang yang dites, dan berdasarkan hasil menunaikan tugas-tugas tersebut, akan dapat ditarik kesimpulan tentang aspek tertentu pada orang tersebut. Tes sebagai alat ukur sangat banyak macamnya dan luas penggunaannya.

b. Teknik non tes dapat dilakukan dengan observasi baik secara langsung ataupun tak langsung, angket ataupun wawancara. Dapat pula dilakukan dengan Sosiometri, teknik non tes digunakan sebagai pelengkap dan digunakan sebagai pertimbangan tambahan dalam pengambilan keputusan penentuan kualitas hasil belajar, teknik ini dapat bersifat lebih menyeluruh pada semua aspek kehidupan anak. Dalam KBK teknik non tes disarankan untuk banyak digunakan.

Ada beberapa teknik dan alat penilaian yang dapat digunakan guru sebagai sarana untuk memperoleh informasi tentang keadaan belajar siswa. Penggunaan berbagai teknik dan alat itu harus disesuaikan dengan tujuan penilaian, waktu yang tersedia, sifat tugas yang dilakukan siswa dan banyaknya/jumlah materi pelajaran yang sudah disampaikan. Teknik penilaian adalah metode atau cara penilaian yang dapat digunakan guru untuk mendapatkan informasi mengenai

keadaan belajar dan prestasi peserta didik. Teknik penilaian yang memungkinkan dan dapat dengan mudah digunakan oleh guru antara lain:

1. Teknik Tes

Istilah "tes" berasal dari bahasa Perancis, yaitu "*testum*", berarti piring yang digunakan untuk memilih logam mulia dari benda-benda lain, seperti pasir, batu, tanah, dan sebagainya. Tes merupakan suatu teknik atau cara yang digunakan dalam rangka melaksanakan kegiatan pengukuran, yang di dalamnya terdapat serangkaian pertanyaan atau latihan atau alat lain yang digunakan untuk mengukur keterampilan, pengetahuan, kecerdasan, kemampuan, atau bakat yang dimiliki oleh seseorang atau kelompok. Dilihat dari bentuknya, maka penilaian jenis tes ini dapat diklasifikasikan menjadi 3 bagian, yaitu:

a. Tes Tertulis

Tes tertulis adalah tes yang soal-soalnya harus dijawab peserta didik dengan memberikan jawaban tertulis. Jenis tes tertulis secara umum dapat dikelompokkan menjadi dua yaitu:

1.1.Uraian

Bentuk uraian dapat digunakan untuk mengatur kegiatan-kegiatan belajar yang sulit diukur oleh bentuk objektif. Disebut bentuk uraian, karena menuntut peserta didik untuk menguraikan, mengorganisasikan dan menyatakan jawaban dengan kata-katanya sendiri dalam bentuk, teknik, dan gaya yang berbeda satu dengan yang lainnya. Dilihat dari luas-sempitnya materi yang ditanyakan, maka tes bentuk uraian ini dapat dibagi menjadi 2 bentuk, yaitu:

1.1.a. Uraian Bebas (*Extended Respons Items*)

Dalam bentuk ini peserta didik bebas untuk menjawab soal dengan cara dan sistematika sendiri. Peserta didik bebas mengemukakan pendapat sesuai dengan kemampuannya. Oleh karena itu, setiap peserta didik mempunyai cara dan sistematika yang berbeda-beda. Namun, guru tetap mempunyai acuan atau patokan dalam mengoreksi jawaban peserta didik nanti. Contoh:

- Bagaimana perkembangan komputer di Indonesia, jelaskan secara singkat!
- Bagaimana peranan komputer dalam pendidikan!

1.1.b. Uraian Terbatas (*Restricted Response Items*)

Dalam menjawab soal bentuk uraian ini, peserta didik harus mengemukakan hal-hal tertentu sebagai batas-batasnya. Walaupun kalimat jawaban peserta didik itu beraneka ragam, tetap harus ada

pokok-pokok penting yang terdapat dalam sistematika jawabannya sesuai dengan batas-batas yang telah ditentukan dan dikendaki dalam soalnya. Contoh:

- Jelaskan bagaimana prosedur operasional sebuah pesawat komputer!
- Sebutkan lima komponen dalam sistem komputer!

1.2.Objektif

Tes objektif sering juga disebut tes dikotomi (*dichotomously scored item*) karena jawabannya antara benar atau salah dan skornya antara 1 atau 0. Tes objektif terdiri dari beberapa bentuk, antara lain:

1.2.a. Benar-Salah (*True-False*)

Bentuk tes benar-salah (B-S) adalah pernyataan yang mengandung dua kemungkinan jawaban, yaitu benar atau salah. Salah satu fungsi bentuk soal benar-salah adalah untuk mengukur kemampuan

peserta didik dalam membedakan antara fakta dengan pendapat. Bentuk soal seperti ini lebih banyak digunakan untuk mengukur kemampuan mengidentifikasi informasi berdasarkan hubungan yang sederhana.

- Ada beberapa teknik/petunjuk praktis dalam penyusunan soal bentuk B-S, yaitu: Jumlah item yang benar dan salah hendaknya sama.
- Berilah petunjuk cara mengerjakan soal yang jelas dan memakai kalimat sederhana.
- Hendaknya jumlah item cukup banyak, sehingga dapat dipertanggungjawabkan.

1.2.b. Pilihan Ganda (*Multiple Choices*)

Soal tes bentuk pilihan ganda dapat digunakan untuk mengukur hasil belajar yang lebih kompleks dan berkenaan dengan aspek ingatan, pengertian, aplikasi, analisis, sintesis, dan evaluasi.

Pilihan jawaban (*option*) terdiri atas jawaban yang benar atau paling benar, selanjutnya disebut kunci jawaban dan kemungkinan jawaban salah yang dinamakan pengecoh (*distractor/decoy/fails*). Beberapa petunjuk praktis dalam menyusun soal bentuk pilihan ganda, yaitu:

- Harus mengacu pada kompetensi dasar dan indikator soal.
- Jangan memasukkan materi soal yang tidak relevan dengan apa yang sudah dipelajari peserta didik.
- Pernyataan dan pilihan hendaknya merupakan kesatuan kalimat yang tidak terputus.
- Harus diyakini bahwa hanya ada satu jawaban yang benar. Bila perlu beri jawaban pengecohnya.

Kebaikan soal bentuk pilihan-ganda, antara lain: (1) cara penilaian dapat

dilakukan dengan mudah, cepat, dan objektif, (2) dapat mencakup ruang lingkup bahan/materi yang luas, (3) mampu mengungkap tingkat kognitif rendah sampai tinggi, dan (4) dapat digunakan berulang kali. Sedangkan kelemahannya antara lain: (1) proses penyusunan soal benar-benar membutuhkan waktu yang lama, (2) memberi peluang siswa untuk menebak jawaban, dan (3) kurang mampu meningkatkan daya nalar siswa.

1.2.c. Menjodohkan (*Matching*)

Soal tes bentuk menjodohkan terdiri atas kumpulan soal dan kumpulan jawaban yang keduanya dikumpulkan pada dua kolom berbeda, yaitu kolom sebelah kiri menunjukkan kumpulan persoalan, dan kolom sebelah kanan menunjukkan kumpulan jawaban. Bentuk soal seperti ini sangat baik untuk mengukur kemampuan

peserta didik dalam mengidentifikasi hubungan antara dua hal.

Untuk penyusunan soal bentuk ini perlu memperhatikan teknik berikut:

- Sesuaikan dengan kompetensi dasar dan indikator.
- Kumpulan soal diletakkan di sebelah kiri, dan jawaban di sebelah kanan.
- Gunakan kalimat singkat dan terarah pada pokok persoalan.

1.2.d. Melengkapi (*Completion*)

Soal bentuk melengkapi (*completion*) dikemukakan dalam kalimat yang tidak lengkap.

Contoh: Tempat sampah daur ulang dalam komputer disebut.....

b. Tes Lisan

Tes lisan yakni tes yang pelaksanaannya dilakukan dengan mengadakan tanya jawab secara langsung antara pendidik dan peserta didik. Tes ini memiliki

kelebihan dan kelemahan. Kelebihannya antara sebagai berikut:

- Dapat menilai kemampuan dan tingkat pengetahuan yang dimiliki peserta didik, sikap, serta kepribadiannya karena dilakukan secara berhadapan langsung.
- Bagi peserta didik yang kemampuan berpikirnya relatif lambat sehingga sering mengalami kesukaran dalam memahami pernyataan soal, tes bentuk ini dapat menolong sebab peserta didik dapat menanyakan langsung kejelasan pertanyaan yang dimaksud.
- Hasil tes dapat langsung diketahui peserta didik.

Sedangkan kelemahan dari tes lisan adalah sebagai berikut.

- Subjektivitas guru sering mencemari hasil tes
- Waktu pelaksanaan yang diperlukan relatif cukup lama.

c. Tes Perbuatan

Tes perbuatan yakni tes yang penugasannya disampaikan dalam bentuk lisan atau tertulis dan pelaksanaan tugasnya dinyatakan dengan perbuatan atau unjuk kerja. Penilaian tes perbuatan dilakukan sejak peserta didik melakukan persiapan, melaksanakan tugas, sampai dengan hasil yang dicapainya. Untuk menilai tes perbuatan pada umumnya diperlukan sebuah format pengamatan, yang bentuknya dibuat sedemikian rupa agar pendidik dapat menuliskan angka-angka yang diperolehnya pada tempat yang sudah disediakan. Bentuk formatnya dapat disesuaikan menurut keperluan. Untuk tes perbuatan yang sifatnya individual, sebaiknya menggunakan format pengamatan individual. Untuk tes perbuatan yang dilaksanakan secara kelompok digunakan format tertentu yang

sudah disesuaikan untuk keperluan pengamatan kelompok.

2. Teknik Non-tes

Teknik non-tes sangat penting dalam mengevaluasi siswa pada ranah afektif dan psikomotor, berbeda dengan teknik tes yang lebih menekankan aspek kognitif. Ada beberapa macam teknik non-tes, yakni: pengamatan (*observation*), wawancara (*interview*), kuesioner/angket (*questionnaire*).

a. Observasi

Observasi adalah suatu proses pengamatan dan pencatatan secara sistematis, logis, objektif, dan rasional mengenai berbagai fenomena untuk mencapai tujuan tertentu. Dalam evaluasi pembelajaran, observasi dapat digunakan untuk menilai proses dan hasil belajar peserta didik, seperti tingkah laku peserta didik pada waktu belajar, berdiskusi,

mengerjakan tugas, dan lain-lain. Alat yang digunakan untuk melakukan observasi disebut pedoman observasi.

b. Wawancara

Wawancara merupakan salah satu bentuk alat evaluasi jenis non-tes yang dilakukan melalui percakapan dan tanya jawab, baik secara langsung maupun tidak langsung. Ada beberapa teknik atau cara yang harus diperhatikan dalam melaksanakan wawancara, yaitu:

- Pewawancara harus mempunyai *background* atau latar belakang tentang apa yang akan ditanyakan.
- Dalam mewawancarai jangan terlalu kaku, tunjukkan sikap yang bersahabat, bebas, ramah, terbuka, dan dapat menyesuaikan diri.
- Hilangkan prasangka-prasangka yang tidak baik.

- Pertanyaan hendaknya jelas, tepat, dan denan bahasa yang sederhana.
- Hindari kevakuman pembicaraan yang terlalu lama.
- Batasi waktu wawancara.

c. Angket

Angket adalah sebuah daftar pertanyaan yang harus diisi oleh orang yang akan diukur (responden). Angket adalah alat penilaian hasil belajar yang berupa daftar pertanyaan tertulis untuk menjaring informasi tentang sesuatu, misalnya tentang latar belakang keluarga siswa, kesehatan siswa, tanggapan siswa terhadap metode pembelajaran, media, dan lain- lain. Angket umumnya dipergunakan pada ranah afektif

d. Daftar Cek

Daftar cek adalah deretan pertanyaan singkat dimana responden yang dievaluasi tinggal membubukan tanda centang (√)

pada aspek yang diamati sesuai dengan hasil penilaiannya.

Ikhtisar Teknik Penilaian

1. Alternative Assessment Techniques

Pada 1980-an, kritik terhadap tes pilihan ganda dan jenis pengujian "pensil-dan-kertas" lainnya mulai meningkat secara dramatis. Pendidik mulai mencari berbagai bentuk penilaian yang akan membantu mengarahkan pengajaran pada pembelajaran siswa yang lebih bermakna. Pendidik ini ingin memasukkan ke dalam penilaian pengaturan kelas yang akan (1) memungkinkan siswa menunjukkan kemampuan untuk secara langsung menerapkan pengetahuan dan keterampilan terpelajar dan (2) mencerminkan tingkat aplikasi dunia nyata yang lebih tinggi.

Jenis utama penilaian alternatif meliputi penilaian informal, penilaian berdasarkan kinerja, dan penilaian portofolio.

a. Informal assessments

Penilaian informal bersifat spontan dan sering terjadi tanpa disadari. Jenis penilaian umum ini biasanya menghasilkan informasi lisan. Penilaian informal mencakup metode penilaian khusus pengamatan guru dan pertanyaan guru.

b. Performance-based assessments

PBA memungkinkan siswa untuk menunjukkan apa yang mereka ketahui dan benar-benar dapat mereka lakukan dalam situasi nyata; mereka melibatkan penerapan pengetahuan dan keterampilan yang benar (airasian, 2000). Karakteristik pba yang paling penting adalah observasi langsung terhadap kinerja siswa (Oosterhof, 1999).

c. Portofolio assessments

Penilaian portofolio memungkinkan guru dan siswa mengumpulkan bukti kemampuan siswa dan skills dalam jangka waktu yang lama. Portofolios berisi contoh karya siswa,

biasanya dipilih oleh siswa berdasarkan pedoman yang disiapkan oleh guru. Ini adalah pengumpulan informasi yang sangat terarah (chase, 1999). Selanjutnya, portofolios berfokus pada pencapaian siswa, dengan penekanan khusus pada pekerjaan terbaik mereka atau pertumbuhan selama periode waktu tertentu. Gagasan portofolio adalah untuk memamerkan jenis barang di mana siswa unggul, bukan untuk menunjukkan hal-hal yang tidak dapat mereka lakukan. Selain itu, fokusnya adalah pada penilaian produk, bukan proses (Oosterhof, 1999).

2. *Traditional Assessment Techniques*

Selama bertahun-tahun dalam pendidikan amerika, hanya ada satu pendekatan untuk menilai apa yang dapat dan tidak dapat dilakukan oleh siswa. Alat penilaian utama ini, yang melibatkan penggunaan uji pensil dan kertas, dikenal sebagai penilaian

tradisional. Dua kategori utama dari tes tradisional adalah tes objektif dan subyektif.

a. *Objective test items* (Tes obyektif)

Tes obyektif hanya memiliki satu jawaban yang benar, yang ditentukan sebelum waktu tes siswa berlangsung (chase, 1999). Tujuan tes obyektif dapat bervariasi dalam format mereka, namun karakteristik utama adalah bahwa mereka mengajukan pertanyaan atau masalah, beserta beberapa opsi. Tugas siswa untuk memilih jawaban yang benar dari pilihan yang diberikan. Untuk itu, tes objektif juga disebut sebagai item respons terpilih. Tes obyektif meliputi pilihan ganda, pilihan alternatif, dan tes mencocokkan. Tes pilihan ganda adalah tipe tes objektif yang paling populer digunakan (chase, 1999). Item pilihan ganda terdiri dari soal dan seperangkat jawaban yang dikenal sebagai pilihan. Salah satu pilihan menjadi jawaban yang benar, sementara yang lain disebut distraktor, atau

respons yang salah. Item pilihan alternatif serupa dalam struktur pada item pilihan ganda, namun hanya berisi dua kemungkinan tanggapan. Tipe yang paling populer adalah true-false item. Mencocokkan biasanya terdiri dari dua daftar istilah. Daftar di sebelah kiri disebut rangsangan, dan daftar di sebelah kanan dikenal sebagai tanggapannya.

b. *Subjective test items* (Uji subjektif)

Uji subyektif dinamai demikian karena, karena konstruksinya, subjektivitas masuk ke dalam proses penilaian. Item subyektif mengharuskan siswa mengembangkan respons mereka terhadap sebuah pertanyaan. Jawaban yang benar tidak diberikan; mereka harus membangunnya sendiri. Untuk alasan ini, item uji subjektif terkadang juga disebut sebagai item respons yang dibangun atau item penawaran.



BAB 4

PENILAIAN AUTENTIK

Penilaian merupakan serangkaian proses yang sistematis dan sistemik, mengumpulkan data dan atau informasi, menganalisis dan selanjutnya menarik kesimpulan tentang tingkat pencapaian hasil dan tingkat efektifitas serta efisiensi suatu.

Diberlakukannya kurikulum 2013 yang menekankan pada pembelajaran berbasis aktivitas, maka penilaiannya lebih menekankan pada penilaian proses baik pada aspek sikap, pengetahuan, dan keterampilan. Maka diperlukan suatu pergeseran dari penilaian melalui tes (mengukur semua kompetensi pengetahuan hanya berdasarkan hasil), menuju penilaian autentik (mengukur semua kompetensi sikap, keterampilan, dan pengetahuan berdasarkan proses dan hasil).

Penilaian autentik merupakan ciri khas Kurikulum 2013. Penilaian autentik ini merupakan penilaian yang dilakukan secara komprehensif untuk menilai dari masukan (*input*), proses, dan keluaran (*output*) pembelajaran. Penilaian autentik harus mencerminkan masalah dunia nyata, bukan dunia

sekolah. Menggunakan berbagai cara dan kriteria yang holistik (kompetensi utuh merefleksikan pengetahuan, ketrampilan, dan sikap).

Penilaian autentik tidak hanya mengukur apa yang diketahui oleh peserta didik, tetapi lebih menekankan mengukur apa yang dapat dilakukan oleh peserta didik. Guru dalam penilaian autentik ini harus memiliki wawasan yang luas tentang pengalaman maupun permasalahan-permasalahan kehidupan. Melalui penilaian dapat dilihat apakah tujuan pembelajaran sebagaimana yang termuat dalam kurikulum sudah tercapai atau belum. Model penilaian ini selalu berkembang dan disempurkan seiring dengan perkembangan dan perobhan kurikulum yang berlaku. Perobahan kurikulum di Indonesia sudah terjadi sebanyak 9 kali yang dimulai dari tahun 1947 yang dikenal dengan rentjana pelajaran hingga kurikulum 2013 yang dikenal dengan kurikulum berkarakter.

Penilaian otentik mengharuskan siswa untuk berdemonstrasi tingkat keterampilan atau prosedur

dalam setting sehari-hari, atau sejenisnya konteks rinci yang melibatkan simulasi tugas kehidupan nyata atau reallife masalah. Untuk menilai unit sains, Anda bisa mengevaluasi melakukan percobaan untuk melihat efek cahaya menyala pertumbuhan tanaman. Untuk mengukur ketrampilan dalam menulis jurnalistik, Anda Bisa saja siswa menulis laporan tentang kejadian terkini.

Gerakan penilaian otentik sesuai dengan filsafat John Dewey (1923) dan desanya tentang "belajar dengan melakukan. "Dia merasa bahwa siswa akan dilayani dengan lebih baik sekolah yang memberikan instruksi tentang tugas yang dipraktekkan di kemudian hari.

Perhatikan bahwa untuk banyak topik, menulis mungkin merupakan tugas asli pilihan. Seorang calon wartawan kemudian akan menghabiskan sebagian besar Kehidupan profesionalnya menyampaikan deskripsi dan gagasan melalui media penulisan. Kebanyakan profesional berkomunikasi melalui beberapa bentuk ekspresi tertulis. Ini sangat penting

keterampilan, tapi satu dari banyak siswa yang perlu dikuasai.

A. Pengertian Penilaian Autentik

Penilaian autentik adalah proses pengumpulan informasi oleh guru tentang perkembangan dan pencapaian pembelajaran yang dilakukan oleh peserta didik melalui berbagai teknik yang mampu mengungkapkan, membuktikan, atau menunjukkan secara tepat bahwa tujuan pembelajaran telah benar-benar dikuasai dan dicapai.

Penilaian autentik merupakan suatu bentuk tugas yang menghendaki pebelajar untuk menunjukkan kinerja di dunia nyata secara bermakna yang merupakan penerapan esensi pengetahuan dan keterampilan.

Menurut Stiggins, penilaian autentik merupakan penilaian kinerja (performansi) yang meminta pebelajar untuk mendemonstrasikan keterampilan dan kompetensi tertentu yang

merupakan penerapan pengetahuan yang dikuasainya.

Ada banyak buku dan artikel tentang penilaian autentik. Banyak penulis telah mengabdikan hidup mereka untuk penilaian *examining*, menawarkan kritik kuat terhadap tes standar, dan memberikan alternatif yang masuk akal untuk tes standar yang seragam. Penulis seperti Grant Wiggins (1993, 1999), Alfie Kohn (2000, 2004), Ray Horn dan Joe Kincheloe (2001), Mark Goldberg (2005), dan Susan Ohanian (1999) telah menerbitkan banyak buku dan artikel tentang topik ini. Wiggins (1998) menyarankan standar penilaian autentik berikut ini. Tugas penilaian adalah autentik bila:

1. Ini realistis. Tugas penilaian harus mengikuti dengan cermat cara-cara di mana kemampuan seseorang "diuji" dalam kata sebenarnya. Misalnya, sebagai mantan penari, kami mempraktikkan latihan menari

seperti plies, jetes, turns, dan sebagainya di kelas balet. Ini hanyalah latihan. Tugas penilaian realistis akan ditemukan dalam kinerja ballet sebenarnya. Di Ballber Nutcracker Suite, anggaplah pemeran, saya terpaksa menunjukkan apa yang bisa saya lakukan. Ini adalah tes yang realistis, ukuran penilaian yang autentik.

2. Hal ini membutuhkan penilaian dan inovasi. Disini siswa harus menggunakan pengetahuan dan keterampilan untuk memecahkan masalah.
3. Meminta siswa untuk melakukan subjek. Kembali ke penari balet, penari harus meletakkan semua langkah bersama dan berperan dalam balet yang sebenarnya.
4. Ini mereplikasi atau menstimulasi "tes" aktual di tempat kerja, kehidupan pribadi, dan kehidupan sipil. Karena setiap siswa berada pada tahap pertumbuhan dan

perkembangan yang unik pada waktu tertentu, mengapa penilaian autentik lebih masuk akal daripada tes standar yang dibuat dengan mudah dapat diapresiasi. Untuk satu hal, satu ukuran tidak sesuai untuk semua siswa. Akal sehat pasti menunjukkan hal ini.

5. Ini menilai kemampuan dan keterampilan siswa untuk secara efektif dan efisien menggunakan repertoar banyak keterampilan untuk menyelesaikan masalah atau tugas. Dalam hal mengakses lebih dari sekedar keterampilan verbal atau matematis, penilaian autentik bergantung pada semua kecerdasan yang dapat dikembangkan seseorang.
6. Ini memungkinkan banyak kesempatan untuk berlatih, berlatih, berkonsultasi, mendapatkan umpan balik, dan memperbaiki pertunjukan dan produksi aktual. Dengan demikian kita memiliki kinerja, umpan balik,

revisi kinerja, umpan balik, kinerja, dan sebagainya. Dengan kata lain, siswa harus belajar sesuatu dan menjadi lebih baik dalam melakukan tugas yang ada. Dalam banyak hal, ini adalah like para seniman yang memiliki kritik konstan untuk perbaikan. Untuk menggunakan tarian sebagai contoh sekali lagi, setelah setiap pertunjukan, direktur bagian pertunjukan biasanya membaca "catatan kritik" di akhir setiap pertunjukan. Dengan demikian konsep yang dibangun ke dalam feedback adalah asumsi bahwa siswa akan bekerja untuk meningkatkan performance test selanjutnya.

Penilaian autentik berbeda dari ukuran standar dan alternatif karena mereka berbasis kinerja dan mencakup keputusan dan perilaku nyata dalam profesi dalam sebuah disiplin. Armstrong (1994) mencirikan penilaian autentik sebagai hal yang sah karena mereka secara intelektual menantang namun responsif terhadap

siswa dan sekolah. Penilaian autentik tidak berfokus pada pengetahuan faktual sebagai tujuan akhir. Sebaliknya, ini berfokus pada kemampuan untuk menggunakan pengetahuan, keterampilan, dan proses yang relevan untuk memecahkan masalah terbuka selama tugas yang berarti. Faktor kunci lain yang membedakan penilaian autentik dari yang tradisional adalah bahwa mereka memberi kesempatan kepada siswa untuk mengintegrasikan berbagai jenis pembelajaran.

Baik formatif, sumatif, standar, autentik, alternatif, atau diciptakan oleh guru atau spesialis penilaian, tidak ada ukuran penilaian yang dapat menangkap semua jenis dan tingkat pembelajaran yang terjadi di kelas seni. Memang, jika setiap jenis ukuran yang mungkin digunakan pada setiap produk siswa, pembelajaran yang signifikan akan tetap tidak terukur karena sifat pembelajaran yang sangat kompleks yang rutin terjadi di kelas seni.

Membuat penilaian itu autentik juga melibatkan pemahaman bahwa usaha perlu difokuskan untuk membantu siswa belajar. Kita harus jelas tentang apa yang ingin kita ketahui dan dapat lakukan. Kita juga perlu memastikan bahwa tindakan yang kita gunakan dan hasil yang kita dapatkan secara akurat mencerminkan apa yang benar-benar diketahui dan dapat dilakukan oleh siswa. Dalam penilaian autentik, kita harus ingat bahwa tujuan kita adalah untuk meningkatkan kualitas pembelajaran siswa dan memanfaatkan kekuatan siswa daripada kelemahan mereka. Intinya dalam penilaian autentik benar-benar tentang membantu siswa menetapkan standar prestasi mereka sendiri yang tinggi.

Karena instrumen yang dirancang untuk proyek ini terutama dibutuhkan untuk penilaian kinerja, diputuskan bahwa pendekatan autentik akan lebih sesuai dengan tujuan penilaian proyek. Penilaian autentik memerlukan konstruksi item

penilaian alternatif (Armstrong, 1994). Penilaian alternatif dianggap salah satu alternatif untuk apa yang tradisional (tes objektif dan esai). Hal ini juga difokuskan pada kinerja siswa, yang merupakan bukti nyata dari apa yang dapat diketahui dan dapat dilakukan siswa. Penilaian autentik memerlukan kinerja yang benar, termasuk keputusan hidup nyata, seperti perilaku estetika, arsitek, sejarawan seni dan kritikus, seniman seperti seniman rakyat, orang-orang yang bekerja dalam segala bentuk yang menghadapi seni dalam kehidupan sehari-hari mereka, dan orang-orang yang avokasionalnya. kegiatan berhubungan dengan seni. Pembelajaran autentik dalam seni menyiratkan aplikasi informasi yang relevan dan bermanfaat, berlawanan dengan perolehan pengetahuan faktual untuk kepentingannya sendiri. Ini juga mengilhami perubahan dalam praktik kurikuler dalam proses penilaian.

Penilaian autentik bukan tanpa kritiknya. Sebagian besar kritik terhadap penilaian alternatif berasal dari keinginan pengembang uji untuk memiliki populasi yang stabil agar memperoleh data keras yang dapat ditangani secara statistik dan dapat dilaporkan sebagai prediktif atau sebagai skor yang diacu norma. Ini mengharuskan siswa dan materi pelajaran untuk diukur juga stabil dan mudah ditebak. Masalahnya adalah, dalam usaha menilai pertunjukan dalam kehidupan nyata, tujuan konten biasanya.

B. Prinsip-Prinsip Penilaian Autentik

Dalam Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 14 Tahun 2014 pasal 4 ayat 6 disebutkan tentang prinsip khusus mengenai penilaian autentik, antara lain:

1. Prinsip khusus untuk Penilaian Autentik meliputi: Materi penilaian dikembangkan dari kurikulum;

2. Bersifat lintas muatan atau mata pelajaran;
3. Berkaitan dengan kemampuan peserta didik;
4. Berbasis kinerja peserta didik;
5. Memotivasi belajar peserta didik;
6. Menekankan pada kegiatan dan pengalaman belajar peserta didik;
7. Memberi kebebasan peserta didik untuk mengkonstruksi responnya;
8. Menekankan keterpaduan sikap, pengetahuan, dan keterampilan;
9. Mengembangkan kemampuan berpikir divergen;
10. Menjadi bagian yang tidak terpisahkan dari pembelajaran;
11. Menghendaki balikan yang segera dan terus menerus;
12. Menekankan konteks yang mencerminkan dunia nyata;
13. Terkait dengan dunia kerja;
14. Menggunakan data yang diperoleh langsung dari dunia nyata; dan

15. Menggunakan berbagai cara dan instrumen;

C. Fitur Penilaian Autentik

Siswa harus menjadi peserta aktif dalam menilai pembelajaran mereka daripada memberi jawaban terhadap serangkaian tes.

Contoh siswa kelas tiga kami yang terlibat dalam menilai kinerja mereka saat mereka belajar dimaksudkan untuk menggambarkan praktik penilaian autentik dalam tindakan. Kami secara *havedeliberately* memilih untuk menunjukkan dengan contoh seperti penilaian autentik, karena definisi sederhana tidak menangkap dinamika interpersonal dan kekuatan penilaian yang autentik. Lebih jauh lagi, *weknow* bahwa apa yang "asli" dalam bahasa anak sekolah atau dalam satu keadaan mungkin tidak terlalu autentik di sekolah atau negara lain. Mengapa? Karena penilaian autentik ditentukan oleh kesesuaian situasional dari praktik pengajaran dan pembelajaran. Tugas semua guru adalah

untuk membuat penilaian di kelas mereka mencerminkan nilai harapan dalam kurikulum mereka dan disesuaikan dengan metode pembelajaran mereka sehingga siswa menganggap penilaian tersebut asli dan adil. Meski mengalami kesulitan mendefinisikan penilaian autentik dengan satu cara, penting untuk mengidentifikasi beberapa prinsip yang menjadi dasar pendekatan. Valencia, Hiebert, dan Afflerbach (1994) memberikan diskusi panjang dan contoh kelas tambahan tentang penilaian autentik yang menampilkan beberapa fitur berikut:

1. Penilaian autentik konsisten dengan penerapan kelas. Ini memiliki validitas instruksional dan kurikuler karena prosedur penilaian dan konten berasal dari pembelajaran sehari-hari siswa di sekolah. Dalam praktiknya, ini berarti bahwa siswa ditanyai tentang informasi yang benar dan

diminta memecahkan masalah yang relevan dengan pengalaman pendidikan mereka.

2. Penilaian autentik mengumpulkan beragam bukti tentang kemampuan belajar siswa dari berbagai aktivitas. Alih-alih mengandalkan satu tes tunggal atau sampel sisipan pengetahuan siswa yang sempit, penilaian autentik melibatkan pengumpulan waktu dari banyak kegiatan akademik yang berbeda (Calfee & Hiebert, 1990). Pengukuran kinerja ini mencakup pembacaan dan retellings lisan, multiple choice tests tentang teks, tanggapan tertulis terhadap literatur, penulisan kreatif, buku catatan, dan jurnal (Winograd, Paris, & Bridge, 1991).
3. Penilaian autentik mendorong pembelajaran dan pengajaran di antara para peserta. Penilaian bersifat fungsional, pragmatis, dan bermanfaat. Messick (1989) berpendapat bahwa validitas untuk memasukkan sebuah akun konsekuensi dari penilaian sehingga

menghasilkan efek yang diinginkan dan tidak memiliki konsekuensi yang tidak disengaja (bandingkan Linn, Baker, QDunbar, 1994). Dengan demikian, penilaian autentik berupaya mendorong pembelajaran dan motivasi para promotor secara langsung dan dinilai berdasarkan patokan tersebut.

4. Penilaian autentik mencerminkan nilai, standar, dan kontrol lokal. Hal ini tidak dipaksakan secara eksternal dengan norma dan harapan dari populasi yang tidak diketahui, juga kewenangan kontroversial yang dikeluarkan dari para peserta. Penilaian kebenaran dapat dimodifikasi oleh para guru untuk memastikan kinerja optimal dari siswa dan memberikan informasi yang berguna kepada orang tua dan administrator. Apa yang diukur dinilai di masyarakat, dan bagaimana dasarnya meyakinkan bahwa siswa memberikan indikator kinerja mereka yang andal.

D. Mengapa Harus Penilaian Autentik?

Di sekolah, masalah keadilan diterapkan hanya pada aktivitas belajar yang paling dasar, bagaimana kita menilai apa yang diketahui siswa. Dibutuhkan lebih banyak waktu, usaha, keterlibatan, pelatihan guru, dan pengembangan profesional guru untuk menerapkan teknik penilaian autentik. Sejak tahun 1980 pendidik telah berteriak untuk mengetahui lebih banyak tentang penilaian autentik di setiap negara bagian di negara ini. Mengapa? Sederhananya, karena itu masuk akal. Apakah kita ingin anak-anak kita dinilai dan dinilai dengan tes satu pukulan yang memberikan skor sederhana, atau lebih baik kita menilai siswa dinilai dan dinilai berdasarkan tugas kinerja yang tidak hanya menunjukkan apa yang mereka langgeng tapi juga apa yang dapat mereka lakukan? Alasan untuk memilih penilaian autentik adalah solid dan jelas:

1. Penilaian autentik secara keseluruhan adalah adil. Tidak ada satu kelompok ras atau etnis yang dihukum dengan skor satu tembakan;
2. Penilaian autentik menceritakan banyak hal tentang bagaimana siswa menghubungkan pengetahuan konten ke masalah tertentu di dunia siswa;
3. Penilaian autentik secara keseluruhan memberikan umpan balik mengenai kemajuan siswa;
4. penilaian autentik menunjukkan bagaimana seorang siswa mengkontruksi suatu produk atau kinerja sehingga pertumbuhan siswa dapat dilihat;
5. Penilaian autentik secara keseluruhan memberikan umpan balik yang berkesinambungan, memungkinkan siswa untuk menyesuaikan dan meningkatkan kinerja;

6. Penilaian autentik memberi siswa suatu kepentingan nyata dan partisipatif dalam proses pembelajaran;
7. Penilaian autentik menghilangkan penekanan pada menghafal fakta dan mengulangnya; dan
8. Penilaian autentik memungkinkan tugas kinerja untuk evaluasi.



BAB 5

PERENCANAAN PENILAIAN

A. Tentang Perencanaan Penilaian

Perencanaan penilaian mencakup penyusunan kisi-kisi yang memuat indikator dan strategi penilaian. Strategi penilaian meliputi pemilihan metode dan teknik penilaian, serta pemilihan bentuk instrumen penilaian.

1. Perencanaan penilaian oleh pendidik

Secara teknis kegiatan pada tahap perencanaan penilaian oleh pendidik sebagai berikut:

- a. Menjelang awal tahun pelajaran, guru mata pelajaran sejenis pada satuan pendidikan (MGMP sekolah) melakukan:
 - a.1 Pengembangan indikator pencapaian KD,
 - a.2 Penyusunan rancangan penilaian (teknik dan bentuk penilaian) yang sesuai,
 - a.3 Pembuatan rancangan program remedial dan pengayaan setiap KD,
 - a.4 Penetapan Kriteria Ketuntasan Minimal (KKM) masing-masing mata pelajaran melalui analisis indikator dengan memperhatikan

karakteristik peserta didik (kemampuan rata-rata peserta didik/intake), karakteristik setiap indikator (kesulitan/kerumitan atau kompleksitas), dan kondisi satuan pendidikan (daya dukung, misalnya kompetensi guru, fasilitas sarana dan prasarana).

- b. Pada awal semester pendidik menginformasikan KKM dan silabus mata pelajaran yang di dalamnya memuat rancangan dan kriteria penilaian kepada peserta didik.
 - c. Pendidik mengembangkan indikator penilaian, kisi-kisi, instrumen penilaian (berupa tes, pengamatan, penugasan, dan sebagainya) dan pedoman penskoran.
2. Perencanaan penilaian oleh satuan pendidikan
- Perencanaan penilaian oleh satuan pendidikan meliputi kegiatan sebagai berikut:

- a. Melalui rapat dewan pendidik, satuan pendidikan melakukan:
 - a.1 Pendataan KKM setiap mata pelajaran
 - a.2 Penentuan kriteria kenaikan kelas (bagi satuan pendidikan yang menggunakan sistem paket) atau penetapan kriteria program pembelajaran (untuk satuan pendidikan yang melaksanakan Sistem Kredit Semester)
 - a.3 Penentuan kriteria nilai akhir kelompok mata pelajaran agama dan akhlak mulia, kelompok mata pelajaran kewarganegaraan dan kepribadian, kelompok mata pelajaran estetika, dan kelompok mata pelajaran jasmani, olahraga, dan kesehatan, dengan mempertimbangkan hasil penilaian oleh pendidik
 - a.4 Penentuan kriteria kelulusan ujian sekolah
 - a.5 Koordinasi ulangan tengah semester, ulangan akhir semester, dan ulangan kenaikan kelas.

- b. Membentuk tim untuk menyusun instrumen penilaian (untuk ulangan tengah semester, ulangan akhir semester, dan ujian sekolah) yang meliputi:
 - b.1 Membentuk tim untuk menyusun instrumen penilaian (untuk ulangan tengah semester, ulangan akhir semester, dan ujian sekolah) yang meliputi:
 - b.2 Pengembangan kisi-kisi penulisan soal (di dalamnya terdapat indikator soal),
 - b.3 Penyusunan butir soal sesuai dengan indikator dan bentuk soal, serta mengikuti kaidah penulisan butir soal,
 - b.4 Penelaahan butir soal secara kualitatif, dilakukan oleh pendidik lain (bukan penyusun butir soal) pengampu mata pelajaran yang sama dengan butir soal yang ditelaahnya,
 - b.5 Perakitan butir-butir soal menjadi perangkat tes

3. Perencanaan penilaian oleh pemerintah

Perencanaan penilaian oleh pemerintah meliputi kegiatan sebagai berikut:

- a. Mengembangkan SKL untuk mata pelajaran yang diujikan dalam UN;
- b. Menyusun dan menetapkan spesifikasi tes UN berdasarkan SKL;
- c. Mengembangkan dan memvalidasi perangkat tes UN;
- d. Menentukan kriteria kelulusan UN.

B. Alasan Pentingnya Perencanaan Penilaian

Terdapat berbagai macam alasan yang mengungkapkan mengapa penting membuat perencanaan penilaian di dalam sebuah program. Berikut ini akan disampaikan beberapa alasan tentang pentingnya membuat perencanaan penilaian:

1. Memiliki rencana atau rancangan penilaian yang tertulis dengan benar akan mengurangi waktu pendidik dalam menghubungkan

berbagai macam proyek yang berbeda dengan total penilaian yang berbeda pula. Sehingga dengan perencanaan penilaian, maka seorang programmer akan lebih jelas dalam mengelompokkan kinerja dan programnya.

2. Sebuah perencanaan penilaian menyediakan sebuah bingkai untuk membimbing proyek penilaian dan memastikan setiap orang atau siswa terdapat pada halaman yang sama dan jelas mengenai tanggung jawab dan semua yang berhubungan dengan jadwal.
3. Rancangan atau perencanaan penilaian akan memudahkan guru dalam berhubungan atau menjalin komunikasi dengan pemangku kebijakan yang berhubungan dengan sekolah. Tak hanya penilaian saja, namun juga berlaku pada pembelajaran, sekolah, kurikulum, dll.

C. Tahapan Umum Perencanaan Penilaian

Perencanaan penilaian memuat secara rinci tentang bagaimana Anda bekerja selama siklus

tahapan penilaian untuk mencapai satu atau dua capaian hasil pembelajaran. Selama Anda membaca deskripsi berikut ini tentang macam-macam rencana penilaian apa yang termasuk didalamnya maka Anda mungkin akan menemukan bahwa Anda sudah lengkap dalam beberapa pekerjaan yang telah diuraikan sebelumnya. Jika Anda menemukan hal ini, maka pastikan Anda memiliki catatan tentang apa yang sudah Anda buat dan kembangkan perencanaan Anda sesuai dengan hal tersebut.

Tidak ada aturan yang pasti untuk menciptakan perencanaan penilaian pada program Anda. Bimbingan berikut ini akan mengenalkan tentang pertanyaan-pertanyaan yang dijawab oleh perencanaan tersebut untuk menilai capaian hasil pembelajaran, membantu Anda dalam mengembangkan jawaban dari pertanyaan-pertanyaan tersebut, bertukar ide untuk mendokumentasikan perencanaan Anda, dan menyediakan beberapa saran untuk

membuat perencanaan penilaian Anda berjalan secara efektif.

Berikut ini pertanyaan yang harus dijawab oleh perencanaan penilaian. Pada dasarnya perencanaan penilaian menjawab tiga pertanyaan berikut ini:

- a. Capaian hasil pembelajaran mana yang Anda fokuskan?
- b. Dasar apa yang akan Anda gunakan untuk menentukan bagaimana siswa menerima capaian hasil pembelajaran?
- c. Bagaimana Anda menggunakan informasi untuk meningkatkan program Anda?

Mengembangkan perencanaan penilaian. Berikut ini pernyataan jika-maka yang dapat digunakan untuk menuntun Anda menjawab tiga pertanyaan dari perencanaan penilaian di atas.

Jika program Anda tidak memiliki capaian hasil pembelajaran, maka perencanaan tersebut harus mengembangkan ini sebagai langkah awalnya. Jika program Anda memiliki hasil

capaian pembelajaran, maka perencanaan harus berfokus pada refleksi program Anda sekarang atau bahkan hal tersebut mengharuskan perbaikan. Jika Anda merasa puas dengan capaian hasil pembelajarannya, maka perencanaan harus memilih satu atau dua hasil capaian pembelajaran agar bisa lebih fokus lagi.

D. Kegiatan Evaluasi Hasil Belajar

Sekalipun tidak selalu sama, namun pada umumnya para pakar dalam bidang evaluasi/ penilaian pendidikan merinci kegiatan evaluasi hasil belajar ke dalam 6 (enam) langkah pokok, yakni:

1. Menyusun rencana evaluasi hasil belajar

Sebelum evaluasi hasil belajar dilaksanakan, harus disusun lebih dahulu perencanaannya secara baik dan matang. Perencanaan evaluasi hasil belajar itu umumnya oleh Sudijono (2003:59) mencakup enam jenis kegiatan, yakni: (a) Merumuskan tujuan

dilaksanakannya evaluasi. (b) menetapkan aspek-aspek yang akan dievaluasi, (c) memilih dan menentukan teknik yang akan dipergunakan di dalam pelaksanaan evaluasi, (d) Menyusun alat-alat pengukur dan penilaian hasil belajar peserta didik, (e) Menentukan tolak ukur, norma atau kriteria yang akan dijadikan pegangan atau patokan dalam memberikan interpretasi terhadap data hasil evaluasi dan (f) Menentukan frekuensi dari kegiatan evaluasi hasil belajar itu sendiri (kapan dan seberapa kali evaluasi hasil belajar itu akan dilaksanakan).

2. Menghimpun data

Dalam evaluasi hasil belajar, wujud nyata dari kegiatan menghimpun data adalah melaksanakan pengukuran, misalnya dengan menyelenggarakan tes hasil belajar (apabila evaluasi hasil belajar itu menggunakan teknik tes), atau melakukan pengamatan, wawancara, atau angket dengan

menggunakan instrumen-instrumen tertentu berupa rating scale, check list, interview guide, atau questionnaire (apabila evaluasi hasil belajar menggunakan teknis non tes).

3. Melakukan verifikasi data

Data yang telah berhasil dihimpun harus disaring lebih dahulu sebelum diolah lebih lanjut. Proses penyaringan itu dikenal dengan istilah penelitian data atau verifikasi data. Verifikasi data dimaksudkan untuk dapat memisahkan data yang "baik" (yaitu data yang dapat memperjelas gambaran yang akan diperoleh mengenai diri individu atau sekelompok individu yang sedang dievaluasi) dari data yang "kurang baik" (yaitu data yang akan menguburkan gambaran yang akan diperoleh apabila data itu ikut serta diolah).

4. Mengolah dan menganalisis data

Mengolah dan menganalisis hasil evaluasi dilakukan dengan maksud untuk memberikan makna terhadap data yang telah berhasil

dihimpun dalam kegiatan evaluasi. Untuk keperluan itu, maka data hasil evaluasi perlu disusun dan diatur sedemikian rupa sehingga "dapat berbicara". Dalam mengolah dan menganalisis data hasil evaluasi itu dapat dipergunakan teknik statistik dan atau teknik non statistik, tergantung kepada jenis data yang akan diolah atau dianalisis. Dengan analisis statistic misalnya, penyusunan atau pengaturan dan penyajian data lewat tabel-tabel, grafik, atau diagram, perhitungan-perhitungan rata-rata, standar deviasi, pengukuran korelasi, uji benda mean, atau uji benda frekuensi dan sebagainya akan dapat menghasilkan informasi-informasi yang lebih lengkap dan amat berharga.

5. Memberikan interpretasi dan menarik kesimpulan

Memberikan interpretasi terhadap data hasil evaluasi belajar pada hakikatnya adalah merupakan verbalisasi dari makna yang

terkandung dalam data yang telah mengalami pengolahan dan penganalisisan itu. Atas dasar interpretasi terhadap data hasil evaluasi itu pada akhirnya dapat dikemukakan kesimpulan-kesimpulan tertentu. Kesimpulan-kesimpulan hasil evaluasi itu sudah barang tentu harus mengacu kepada tujuan dilakukannya evaluasi itu sendiri.

6. Tindak lanjut hasil evaluasi

Bertitik tolak dari hasil evaluasi yang telah disusun, diatur, diolah, dianalisis dan disimpulkan sehingga dapat diketahui apa makna yang terkandung di dalamnya, maka pada akhirnya evaluator akan mengambil keputusan dan merumuskan kebijakan-kebijakan yang dipandang perlu sebagai tindak lanjut dari kegiatan hasil evaluasi tersebut. Harus senantiasa diingat bahwa setiap kegiatan evaluasi menuntut adanya tindak lanjut yang konkrit. Tanpa diikuti oleh

tindak lanjut yang konkrit, maka pekerjaan evaluasi itu hanya akan sampai kepada pernyataan, yang menyatakan bahwa; "saya tahu, bahwa begini dan itu begitu". Apabila hal seperti itu terjadi, maka kegiatan evaluasi itu sebenarnya tidak banyak membawa manfaat bagi evaluator.

E. Rancangan Penilaian Berbasis Kelas

Dalam membuat rancangan penilaian atau evaluasi pembelajaran, perlu diperhatikan hal-hal berikut:

1. Memberikan ukuran yang dipakai, seperti bagaimana mengukur, menilai dan mengevaluasi sebagai kata-kata kunci yang sering digunakan dalam diskusi materi evaluasi hasil belajar.
2. Mendiskusikan tentang fungsi penilaian untuk memperoleh pemahaman tentang hal-hal apa saja yang dapat dinilai melalui pelaksanaan suatu ujian, apakah sekedar

memberi nilai untuk menentukan lulus tidaknya mahasiswa atau siswa dari ujian tersebut atautkah ada tujuan-tujuan lain yang ingin dicapai melalui ujian tersebut.

3. Melaksanakan standar penilaian ujian
4. Merancang soal-soal ujian dalam struktur soal sedemikian rupaa sehingga jumlah maupun derajat kesukaran soal tetap relevan dengan pencapaian sasaran belajar (saasbel) yang telah ditetapkan dalam rancangan kegiatan belajar mengajar (RKBM).
5. Mengingat derajat kesukaran soal dapat berbeda satu dengan yang lainnya, tiap-tiap soal perlu mendapat bobot soal menurut relevansinya dengan sasaran belajar.
6. Sesudah pross membuat, menstrukturkan, dan menentukan bobot soal, soal-soal tersebut dapatlah disajikan melalui ujian.
7. Pengambilan keputusan atas hasil evaluasi ujian.

Menurut Kartika (2013), adapun ciri penilaian berbasis kelas yakni sebagai berikut:

1. Belajar tuntas, yakni peserta didik tidak diperkenankan mengerjakan pekerjaan berikutnya, sebelum mampu menyelesaikan pekerjaan dengan prosedur yang benar, dan hasil yang baik. "Jika peserta didik dikelompokkan berdasarkan tingkat kemampuannya untuk beberapa mata pelajaran dan diajarkan sesuai dengan karakteristik mereka, maka sebagian besar dari mereka akan mencapai ketuntasan".
2. Autentik, yakni memandang penilaian dan pembelajaran secara terpadu, mencerminkan masalah dunia nyata bukan dunia sekolah menggunakan berbagai cara dan kriteria holistik (kompetensi utuh merefleksikan pengetahuan, keterampilan, dan sikap)
3. Berkesinambungan, adalah memantau proses, kemajuan, dan perbaikan hasil terus menerus dalam bentuk Ulangan Harian,

Ulangan Tengah Semester, Ulangan Akhir Semester, dan Ulangan Kenaikan Kelas.

4. Berdasarkan acuan kriteria/patokan Prestasi kemampuan peserta didik tidak dibandingkan dengan peserta kelompok, tetapi dengan kemampuan yang dimiliki sebelumnya dan patokan yang ditetapkan
5. Menggunakan berbagai cara & alat penilaian. Seperti mengembangkan dan menyediakan sistem pencatatan serta penilaian yang bervariasi: Tertulis, Lisan, Produk, Portofolio, Unjuk Kerja, Proyek, Pengamatan, dan Penilaian diri.

Penilaian berbasis kelas terdiri dari beberapa macam, yaitu penilaian melalui portofolio (*portfolio*), penilaian melalui unjuk kerja, penilaian melalui penugasan (*project*), penilaian melalui hasil kerja, penilaian melalui tes tertulis (*paper and pen test*), dan penilaian afektif siswa (Muslich, 2007).

F. Langkah-Langkah Perencanaan Evaluasi Hasil Belajar

Secara umum, perencanaan evaluasi hasil belajar dapat digambarkan dalam langkah-langkah berikut:

1. Merumuskan tujuan dilaksanakannya evaluasi

Seorang guru yang akan mengevaluasi haruslah merumuskan tujuan dilaksanakannya evaluasi, misalnya untuk mengetahui penguasaan peserta didik dalam kompetensi/subkompetensi tertentu setelah mengikuti proses pembelajaran. Dapat pula evaluasi tersebut bertujuan untuk mengetahui kesulitan belajar peserta didik (diagnostic tes). Tujuan evaluasi tersebut harus jelas sehingga dapat memberikan arah dan lingkup pengembangan evaluasi selanjutnya.

2. Menetapkan aspek-aspek yang akan dievaluasi

Aspek tersebut adalah aspek kognitif, aspek afektif atau aspek psikomotorik.

3. Memilih dan menentukan teknik yang akan digunakan dalam evaluasi

Langkah selanjutnya adalah menentukan teknik yang akan digunakan seperti teknik tes atau teknik non tes.

4. Menyusun alat evaluasi yang akan digunakan dalam evaluasi

Contohnya adalah menyusun butir-butir soal esay atau *multiple choise* (teknik tes) dan daftar chek (*check list*), *rating scale*, panduan wawancara, daftar angket (teknik non tes).

5. Menentukan tolok ukur, norma, atau kriteria yang akan dijadikan patokan dalam evaluasi

Misalnya, apakah akan menggunakan PAP (Penilaian Acuan Patokan) atau PAN (Penilaian Acuan Norma).

6. Menentukan frekuensi kegiatan evaluasi

Langkah terakhir adalah menentukan berapa kali kegiatan evaluasi akan dilakukan dalam pembelajaran.

7. Penyajian tes

Setelah tes tersusun, naskah (tes) siap diberikan atau disajikan kepada peserta didik. Hal-hal yang perlu diperhatikan dalam penyajian tes ini adalah waktu penyajian, petunjuk yang jelas mengenai cara menjawab atau mengerjakan tes, ruangan dan tempat duduk peserta didik. Pada prinsipnya, hal-hal yang menyangkut segi administrasi penyajian tes harus diperhatikan sehingga evaluasi dapat terselenggara dengan benar dan baik.

8. *Scorsing*

Scorsing atau pemeriksaan terhadap lembar jawaban dan pemberian angka merupakan langkah untuk mendapatkan informasi kuantitatif dari masing-masing peserta didik. Pada prinsipnya, scorsing ini harus

diusahakan agar dapat dilakukan secara objektif. Artinya, apabila *scorsing* dilakukan oleh dua orang atau lebih, yang sama tingkat kompetensinya, akan menghasilkan scor atau angka yang sama. Atau jika orang yang sama mengulangi proses pengscoran, akan dihasilkan scor yang sama.

9. Pengelolaan hasil tes

Setelah dilakukan *scorsing*, hasilnya perlu dipilah dengan mencari konvermasi nilai. Dalam proses konversi ini ada norma dan ada pula skala, yaitu norma relatif dan penilaian Acuan norma (PAP), dan norma mutlak dengan penilaian Acuan patokan (PAP), masing-masingnya dengan skala 5 (A, B, C, D, E), skala 9 (1-9), skala 100. Kemudian dilakukan prosedur statistic mencari ranking (*rank order mean*).

10. Pelaporan hasil tes

Setelah tes dilaksanakan dan dilakukan *scorsing*, hasil pengetesan tersebut perlu

dilaporkan. Laporan tersebut dapat diberikan kepada peserta didik yang bersangkutan, kepada orang tua peserta didik, kepada kepala sekolah, dan sebagainya. Laporan kepada masing-masing yang berkepentingan dengan hasil tes ini sangat penting karena dapat memberikan informasi yang sangat berguna dalam rangka penentuan kebijaksanaan. Selanjutnya pelaporan hasil penilaian tersebut harus diketahui oleh siswa yang melakukan penilaian, guru untuk mendapat umpan balik terhadap pembelajaran yang telah dilakukan, pihak sekolah untuk mengetahui mutu pembelajaran yang telah dilaksanakan guru-guru, dan juga orang tua sebagai stake holder dari jasa yang ditawarkan sekolah dalam menyelenggarakan pendidikan.

11. Pemanfaatan hasil tes

Hasil pengukuran yang diperoleh melalui ujian sangat berguna sesuai dengan tujuan

ujian. Informasi atau data hasil pengukuran dapat dimanfaatkan untuk perbaikan atau penyempurnaan system , proses atau kegiatan belajar mengajar, maupun sebagai data untuk mengambil keputusan atau menentukan kebijakan.

G. Perencanaan Penilaian Autentik

Berdasarkan olahan dari Permendiknas Nomor 20 tahun 2007 dan Permendikbud Nomor 66 Tahun 2013 standar penilaian autentik menurut adalah sebagai berikut:

1. Guru harus membuat rencana penilaian secara terpadu dengan mengacu kepada silabus dan rencana pembelajarannya. Perencanaan penilaian setidaknya-tidaknya meliputi komponen yang akan dinilai, teknik yang akan digunakan serta kriteria pencapaian kompetensi.

2. Guru harus mengembangkan kriteria pencapaian Kompetensi Dasar (KD) sebagai dasar hukum penilaian.
3. Guru menentukan teknik dan instrumen penilaian sesuai indikator pencapaian KD.
4. Guru harus menginformasikan seawal mungkin kepada peserta didik tentang aspek-aspek yang dinilai dan kriteria pencapaiannya.
5. Guru menuangkan seluruh komponen penilaian ke dalam kisi-kisi penilaian.
6. Guru membuat instrumen berdasarkan kisi-kisi yang telah dibuat dan dilengkapi dengan penskoran sesuai dengan teknik penilaian yang digunakan.
7. Guru menganalisis kualitas instrumen penilaian dengan mengacu pada persyaratan instrumen serta menggunakan acuan kriteria.
8. Guru menetapkan bobot untuk tiap-tiap teknik/jenis penilaian baik untuk KI 1 dan 2 dan KI 3 dan 4 dan menetapkan rumus

penentuan nilai akhir hasil belajar peserta didik.

9. Guru menetapkan acuan kriteria yang akan digunakan berupa nilai kriteria ketuntasan minimal (KKM) untuk dijadikan rujukan dalam pengambilan keputusan.

H. Perencanaan Penilaian Aspek Psikomotor

Aspek perencanaan penilaian ranah psikomotor meliputi hal-hal berikut ini:

1. Tingkatan penguasaan gerakan awal berisi kemampuan peserta didik dalam menggerakkan sebagian anggota badan.
2. Tingkatan gerakan semirutin meliputi kemampuan melakukan atau menirukan gerakan yang melibatkan seluruh anggota badan.
3. Tingkatan gerakan rutin berisi kemampuan melakukan gerakan secara menyeluruh dengan sempurna dan sampai pada tingkatan otomatis.

Alat penilaian yang di gunakan untuk mengukur domain psikomotor adalah tes penampilan atau kinerja (*performance*) yang telah di kuasi peserta didik, seperti:

1. Tes paper dan pencil. Walaupun bentuknya seperti tes tertulis, tetapi sasarannya adalah kemampuan peserta didik dalam menampilkan karya, misalnya berupa desain alat, desain grafis, dan sebagainya.
2. Tes identifikasi. Tes ini ditentukan untuk mengukur kemampuan peserta didik dalam mengidentifikasi sesuatu. Misalnya, menemukan bagian yang rusak atau yang tidak berfungsi dari suatu alat.
3. Tes simulasi. Tes ini dilakukan jika tidak ada alat yang sesungguhnya yang dipakai untuk mempergakan penampilan peserta didik. Dengan demikian, melalui simulasi peserta didik tetap dapat dinilai, apakah dia sudah menguasai keterampilan dengan bantuan

peralatan tiruan atau mempergunakan seolah-olah menggunakan alat.

4. Tes petik kerja (*work sample*). Tes ini dilakukan dengan alat yang sesungguhnya. Tujuannya adalah untuk mengetahui apakah peserta didik sudah menguasai atau terampil menggunakan alat tersebut.

Tes penampilan atau perbuatan, berupa tes identifikasi, tes simulasi maupun untuk kerja datanya dapat diperoleh dengan menggunakan daftar cek (*check list*) ataupun skala penilaian (*rating scale*). Daftar cek lebih praktis jika digunakan untuk menghadapi subjek dalam jumlah yang lebih besar, atau jika perbuatan yang dinilai memiliki resiko tinggi. Skala penilaian cocok untuk menghadapi peserta didik dengan jumlah terbatas.

I. Perencanaan Penilaian Aspek Afektif

Berkenaan dengan ranah afektif, ada dua hal yang harus dinilai:

1. Kompetensi afektif yang ingin dicapai dalam pembelajaran meliputi tingkatan pemberian respons, apresiasi, penilaian, dan internalisasi.
2. Sikap dan minat peserta didik terhadap mata pelajaran dan proses pembelajaran. Sikap peserta didik terhadap pelajaran bisa positif, bisa negatif, atau netral. Hal ini tidak dapat dikategorikan benar atau salah. Guru memiliki tugas untuk membangkitkan dan meningkatkan minat peserta didik terhadap mata pelajaran, serta mengubah sikap peserta didik, dari sikap negatif ke sikap positif.

Adapun tingkatan aspek afektif yang dinilai adalah kemampuan peserta didik dalam:

1. Memberikan respons atau reaksi terhadap nilai-nilai yang dihadapkan kepadanya.
2. Menikmati atau menerima nilai, norma serta objek yang mempunyai nilai etika dan estetika.

3. Menilai (*valuing*) ditinjau dari segi baik buruk, adil-tidak adil, indah-tidak indah terhadap objek studi.
4. Menerapkan atau mempraktikkan nilai, norma, etika, dan estetika dalam perilaku kehidupan sehari-hari.

J. Perencanaan Penilaian Aspek Sikap

Perencanaan penilaian sikap dilakukan berdasarkan KI-1 dan KI-2. Guru merencanakan dan menetapkan sikap yang akan dinilai dalam pembelajaran sesuai dengan kegiatan pembelajaran. Pada penilaian sikap di luar pembelajaran guru dapat mengamati sikap lain yang muncul secara natural.

Langkah-langkah perencanaan penilaian sikap adalah sebagai berikut:

1. Menentukan sikap yang akan dikembangkan di sekolah mengacu pada KI-1 dan KI-2.
2. Menentukan indikator sesuai dengan kompetensi sikap yang akan dikembangkan.

Sebagai contoh, sikap pada KI-1 beserta indikator-indikatornya yang dapat dikembangkan oleh sekolah sebagai berikut:

1. Ketaatan beribadah
 - a. Perilaku patuh dalam melaksanakan ajaran agama yang dianutnya,
 - b. Mau mengajak teman seagamanya untuk melakukan ibadah bersama,
 - c. Mengikuti kegiatan keagamaan yang diselenggarakan sekolah,
 - d. Melaksanakan ibadah sesuai ajaran agama, misalnya: sholat, puasa.
 - e. Merayakan hari besar agama,
 - f. Melaksanakan ibadah tepat waktu.
2. Berperilaku syukur
 - a. Perilaku menerima perbedaan karakteristik sebagai anugerah Tuhan,
 - b. Selalu menerima penugasan dengan sikap terbuka,
 - c. Bersyukur atas pemberian orang lain,

- d. Mengakui kebesaran Tuhan dalam menciptakan alam semesta,
 - e. Menjaga kelestarian alam, tidak merusak tanaman, tidak mengeluh,
 - f. Selalu merasa gembira dalam segala hal,
 - g. Tidak berkecil hati dengan keadaannya,
 - h. Suka memberi atau menolong sesama,
 - i. Selalu berterima kasih bila menerima pertolongan.
3. Berdoa sebelum dan sesudah melakukan kegiatan
- a. Perilaku yang menunjukkan selalu berdoa sebelum atau sesudah melakukan tugas atau pekerjaan,
 - b. Berdoa sebelum makan,
 - c. Berdoa ketika pelajaran selesai,
 - d. Mengajak teman berdoa saat memulai kegiatan,
 - e. Mengingatkan teman untuk selalu berdoa.
4. Toleransi dalam beribadah

- a. Tindakan yang menghargai perbedaan dalam beribadah,
- b. Menghormati teman yang berbeda agama,
- c. Berteman tanpa membedakan agama,
- d. Tidak mengganggu teman yang sedang beribadah,
- e. Menghormati hari besar keagamaan lain,
- f. Tidak menjelekkkan ajaran agama lain.

Sebagai contoh, sikap pada KI-2 beserta indikator-indikatornya yang dapat dikembangkan oleh sekolah sebagai berikut:

- 1. Jujur
 - a. Tidak mau berbohong atau tidak mencontek,
 - b. Mengerjakan sendiri tugas yang diberikan guru, tanpa menjiplak tugas orang lain,
 - c. Mengerjakan soal penilaian tanpa mencontek,
 - d. Mengatakan dengan sesungguhnya apa yang terjadi atau yang dialaminya dalam kehidupan sehari-hari,
 - e. Mau mengakui kesalahan atau kekeliruan,

- f. Mengembalikan barang yang dipinjam atau ditemukan,
- g. Mengemukakan pendapat sesuai dengan apa yang diyakininya, walaupun berbeda dengan pendapat teman,
- h. Mengemukakan ketidaknyamanan belajar yang dirasakannya di sekolah,
- i. Membuat laporan kegiatan kelas secara terbuka (transparan).

2. Peduli

- a. Ingin tahu dan ingin membantu teman yang kesulitan dalam pembelajaran, perhatian kepada orang lain,
- b. Berpartisipasi dalam kegiatan sosial di sekolah, misal: mengumpulkan sumbangan untuk membantu yang sakit atau kemalangan,
- c. Meminjamkan alat kepada teman yang tidak membawa/memiliki,
- d. Menolong teman yang mengalami kesulitan,

- e. Menjaga keasrian, keindahan, dan kebersihan lingkungan sekolah,
- f. Melerai teman yang berselisih (bertengkar),
- g. Menjenguk teman atau guru yang sakit,
- h. Menunjukkan perhatian terhadap kebersihan kelas dan lingkungan sekolah.

3. Tanggung Jawab

- a. Menyelesaikan tugas yang diberikan ,
- b. Mengakui kesalahan,
- c. Melaksanakan tugas yang menjadi kewajibannya di kelas seperti piket kebersihan,
- d. Melaksanakan peraturan sekolah dengan baik,
- e. Mengerjakan tugas/pekerjaan rumah sekolah dengan baik,
- f. Mengumpulkan tugas/pekerjaan rumah tepat waktu,mengakui kesalahan, tidak melemparkan kesalahan kepada teman,
- g. Berpartisipasi dalam kegiatan sosial di sekolah,

- h. Menunjukkan prakarsa untuk mengatasi masalah dalam kelompok di kelas/sekolah,
 - i. Membuat laporan setelah selesai melakukan kegiatan.
4. Disiplin
- a. Mengikuti peraturan yang ada di sekolah,
 - b. Tertib dalam melaks peserta didikan tugas,
 - c. Hadir di sekolah tepat waktu,
 - d. Masuk kelas tepat waktu,
 - e. Memakai pakaian seragam lengkap dan rapi,
 - f. Tertib mentaati peraturan sekolah,
 - g. Melaksanakan piket kebersihan kelas,
 - h. Mengumpulkan tugas/pekerjaan rumah tepat waktu,
 - i. Mengerjakan tugas/pekerjaan rumah dengan baik,
 - j. Membagi waktu belajar dan bermain dengan baik,
 - k. Mengambil dan mengembalikan peralatan belajar pada tempatnya,
 - l. Tidak pernah terlambat masuk kelas.

5. Percaya Diri

- a. Berani tampil di depan kelas,
- b. Berani mengemukakan pendapat,
- c. Berani mencoba hal baru,
- d. Mengemukakan pendapat terhadap suatu topik atau masalah,
- e. Mengajukan diri menjadi ketua kelas atau pengurus kelas lainnya,
- f. Mengajukan diri untuk mengerjakan tugas atau soal di papan tulis,
- g. Mencoba hal-hal baru yang bermanfaat,
- h. Mengungkapkan kritikan membangun terhadap karya orang lain,
- i. Memberikan argumen yang kuat untuk mempertahankan pendapat.

6. Santun

- a. Menghormati orang lain dan menghormati cara bicara yang tepat,
- b. Menghormati guru, pegawai sekolah, penjaga kebun, dan orang yang lebih tua,

- c. Berbicara atau bertutur kata halus tidak kasar,
- d. Berpakaian rapi dan pantas,
- e. Dapat mengendalikan emosi dalam menghadapi masalah, tidak marah-marah
- f. Mengucapkan salam ketika bertemu guru, teman, dan orang-orang di sekolah,
- g. Menunjukkan wajah ramah, bersahabat, dan tidak cemberut,
- h. Mengucapkan terima kasih apabila menerima bantuan dalam bentuk jasa atau barang dari orang lain.



BAB 6

KONSTRUKSI TES PRESTASI HASIL BELAJAR SISWA

A. Pengantar Mengonstruksi Tes Prestasi Hasil Belajar

Konsepsi Prestasi Belajar Prestasi belajar adalah hasil atau taraf kemampuan yang telah dicapai siswa setelah mengikuti proses belajar mengajar dalam waktu tertentu baik berupa perubahan tingkah laku, keterampilan dan pengetahuan dan kemudian akan diukur dan dinilai yang kemudian diwujudkan dalam angka atau pernyataan. Perwujudan bentuk hasil proses belajar tersebut dapat berupa pemecahan lisan maupun tulisan, dan keterampilan serta pemecahan masalah langsung dapat diukur atau dinilai, baik dengan tes atau pun non tes. Benyamin Bloom (Sudjana, 1989) membagi hasil belajar atau prestasi belajar menjadi tiga ranah, yakni ranah kognitif, ranah afektif, dan ranah psikomotor. Ranah kognitif tersebut terdiri dari enam (6) tipe hasil belajar, yaitu: tipe hasil belajar pengetahuan, tipe hasil belajar pemahaman, tipe hasil belajar aplikasi, tipe hasil belajar analisis,

tipe hasil belajar sintesis, dan tipe hasil belajar evaluasi. Dapat disimpulkan hasil belajar atau prestasi belajar diklasifikasikan menjadi tiga ranah atau aspek yaitu kognitif, afektif, dan psikomotor. Ranah kognitif terdiri dari pengetahuan/ingatan (C1), pemahaman (C2), aplikasi (C3), analisis (C4), sintesis (C5), dan evaluasi (C6). Dalam pembahasan ini, dibatasi pada prestasi belajar dalam ranah kognitif.

E. Gronlund, Norman dalam bukunya *Constructing Achievement Test* memberikan definisi tes prestasi adalah prosedur sistematis untuk menentukan jumlah yang telah dipelajari siswa. Meski penekanannya adalah pada pengukuran hasil belajar, seharusnya tidak disiratkan bahwa pengujian harus dilakukan hanya pada akhir instruksi. Terlalu sering, pengujian prestasi dipandang sebagai aktivitas akhir-akhir atau aktivitas akhir-kursus yang terutama digunakan untuk menugaskan nilai kelas. Meski ini adalah fungsi yang perlu dan

berguna dari pengujian, hal ini hanya satu dari banyak fungsi. Seperti pengajaran, tujuan utama pengujian adalah untuk memperbaiki pembelajaran, dan dalam konteks yang lebih luas di sana adalah sejumlah kontribusi spesifik yang dapat dibuatnya.

Dalam bidang pendidikan, yang digunakan sebagai ukuran untuk mengetahui sejauh mana anak didik telah menguasai materi pelajaran yang sudah diajarkan dan dipelajari adalah hasil belajar atau prestasi belajar (Masrun dan Martaniah, 1973). Hasil belajar ini diukur sebagian besar dengan menggunakan tes hasil belajar. Suryabrata (1987, b) menyatakan bahwa untuk mengetahui proses belajar anak didik, pendidik harus melakukan pengukuran dan evaluasi. Hal ini mutlak dilakukan karena pada saat tertentu pendidik harus membuat keputusan pendidikan. Pada waktu membuat keputusan yang bijaksana diperlukan suatu informasi yang akurat dan akurat. Tes prestasi hasil belajar

mempunyai banyak fungsi yang antara lain adalah untuk memperbesar pemahaman diri peserta, dan menyediakan umpan balik tentang efektivitas pengetahuan (Depdikbud, 1987). Selain itu menurut Anastasi (1990) tes prestasi belajar adalah tes yang mengukur pengetahuan yang dimiliki seseorang sebagai akibat adanya program pendidikan maupun program pelatihan. Melalui tes prestasi belajar dapat diperoleh informasi mengenai perbedaan kemajuan atau tambahan pengetahuan antar peserta didik. Informasi yang diperoleh melalui kegiatan tes prestasi sangat berguna untuk menentukan tahap proses belajar berikutnya, baik ditinjau dari daya serap peserta didik sehubungan dengan pokok bahasan atau bahan pelajaran yang diberikan pendidik kepada mereka.

B. Fungsi Tes Prestasi Hasil Belajar

Berbagai macam keputusan pendidikan itu menempatkan tes prestasi belajar dalam beberapa fungsi, antara lain:

1. Fungsi penempatan: penggunaan hasil tes prestasi belajar untuk klasifikasi individu kedalam bidang jurusan yang sesuai dengan kemampuan yang telah diperlihatkan pada hasil belajar yang lalu.
2. Fungsi formatif: penggunaan tes prestasi belajar untuk melihat sejauhmana kemajuan belajar yang telah dicapai oleh siswa dalam suatu program pelajaran.
3. Fungsi diagnostik: dilakukan tes prestasi apabila hasil yang tes bersangkutan digunakan untuk mengdiagnosis kesukaran dalam belajar, mendeteksi kelemahan siswa agar segera diperbaiki.
4. Fungsi sumatif: penggunaan hasil tes prestasi untuk memperoleh informasi mengenai penguasaan pelajaran yang telah

direncanakan sebelumnya dalam suatu program belajar. Ini adalah tes akhir dalam suatu program dan hasilnya dipakai untuk menentukan kelulusan siswa dalam program itu, dan apakah siswa itu dapat melanjutkan ke jenjang yang lebih tinggi.

C. Kategori Penilaian Tes Prestasi Hasil Belajar

Pengujian bisa diklasifikasikan dengan cara yang berbeda sesuai dengan berbagai kriteria. Tidak ada keseragaman dalam klasifikasi yang ditawarkan oleh penulis yang berbeda. Menurut Stoyanova, F., & C, Amuka dalam buku *The Test for Teachers* (Стоянова, Ф. Тестология за учители tahun 1996 halaman 277) tes dapat dibagi menjadi delapan kategori dasar. Berikut ini akan dipaparkan mengenai kategori dasar penilaian tes prestasi belajar siswa yang dapat diaplikasikan pada saat melakukan pembelajaran di sekolah, antara lain:

1. Menurut tujuan pengujian
 - b. Tes proficiency digunakan untuk mengukur kemampuan bahasa siswa terlepas dari pelatihan apapun;
 - c. Tes prestasi, untuk mengukur seberapa banyak bahasa yang diajarkan selama periode waktu tertentu telah dipelajari;
 - d. Tes diagnostik, untuk mengidentifikasi kekuatan dan kelemahan siswa dan merencanakan pengajaran lebih lanjut;
 - e. Tes penempatan, untuk menempatkan siswa dalam kelompok sesuai dengan kriteria mereka kemampuan bahasa;
 - f. Tes bakat, untuk memprediksi keberhasilan atau potensi masa depan seorang siswa di lingkungan bahasa.
2. Menurut waktu pengujian
 - a. Tes waktu terbatas
 - b. Tes waktu tak terbatas
3. Menurut administrasi ujian
 - a. Tes individu

- b. Tes kelompok
- 4. Menurut jenis jawaban
 - a. Tes tertulis
 - b. Tes komputer
 - c. Tes kinerja
- 5. Menurut tipe keputusan
 - a. Pendahuluan
 - b. Tes saat ini
 - c. Tes akhir
 - d. Tes diagnostik
- 6. Menurut jenis format item
 - a. Obyektif
 - b. Esai
- 7. Menurut metode evaluasi
 - a. Tes yang diacu oleh referensi
 - b. Tes yang direkomendasikan kriteria
- 8. Menurut kualitas tes
 - a. Tes standar
 - b. Tes non-standar

D. Prinsip-Prinsip Tes Prestasi Hasil Belajar

Sejauh mana tes prestasi berkontribusi pada peningkatan pembelajaran dan instruksi sebagian besar ditentukan oleh prinsip-prinsip yang mendasari mereka gunakan dan kembangkan. Tes dapat mengarahkan perhatian siswa terhadap tujuan pengajaran atau jauh dari waktu. Mereka bisa mendorong siswa untuk fokus pada aspek terbatas dari isi kursus atau mengarahkan perhatiannya pada area yang tidak penting. Mereka bisa memberi hadiah pembelajaran dangkal atau membutuhkan pemahaman mendalam. Mereka dapat memberikan informasi yang dapat diandalkan untuk keputusan instruksional, atau bisa memberikan bias dan informasi terdistorsi. Berikut prinsip-prinsip pengujian prestasi memberikan dasar yang kuat untuk membangun dan menggunakan tes di kelas sebagai sesuatu yang kekuatan positif dalam proses belajar mengajar (E. Gronlund, Norman, 1977: 7).

1. Tes prestasi memaparkan hasil belajar yang jelas selaras dengan tujuan instruksional.

Tes prestasi dapat dirancang untuk mengukur berbagai kehebatan belajar, seperti pengetahuan tentang fakta-fakta tertentu, pengetahuan tentang istilah, pengertian konsep dan prinsip, kemampuan untuk menerapkan fakta dan prinsip, dan berbagai kemampuan berpikir. Urutan pertama bisnis dalam membangun dengan membuat tes prestasi, bukan konstruksi tes item, melainkan identifikasi dan definisi hasil belajar untuk diukur. Ini harus secara logis tumbuh dari pembelajaran tujuan kursus di mana tes ini akan digunakan. Urutan langkah untuk menentukan hasil belajar adalah sebagai berikut:

- a. Identifikasi tujuan instruksional kursus.
- b. Sebutkan tujuan dalam hal hasil pembelajaran umum (misalnya, "Siswa memahami arti istilah umum").

- c. Di bawah objektif, tuliskan hasil belajar spesifik yang Anda inginkan untuk menerima sebagai bukti pencapaian tujuan itu. Seharusnya juga dinyatakan dalam hal perilaku yang dapat diamati (misalnya, "Siswa dapat (a) mendefinisikan istilah dalam kata-katanya sendiri, (b) membedakan antara istilah di dasar makna, dan (c) menggunakan istilah secara efektif dalam bahasa kalimat aslinya").

Jika urutan langkah ini dilakukan dengan benar, pembelajaran yang hasilnya spesifik akan menunjukkan dengan jelas perilaku apa yang harus diukur oleh tes untuk mencerminkan tujuan instruksi secara memadai. Kapan dinyatakan dengan jelas dalam hal perilaku, langkah-langkah ini juga menyarankan jenis tes item yang akan digunakan jika Anda mengharapkan siswa untuk dapat "mendefinisikan istilah dalam bahasa mereka menggunakan kata-kata

sendiri, "misalnya, berikan mereka istilah" dan arahkan mereka untuk menulis definisi. Tes benar-salah atau pilihan ganda jelas tidak memunculkan perilaku spesifik yang dijelaskan dalam hasil pembelajaran ini.

Singkatnya, tes prestasi mengukur perilaku spesifik siswa diharapkan untuk menunjukkan pada akhir pengalaman belajar.

Agar tes menjadi sangat berguna, hasil pembelajaran ini semestinya didefinisikan dengan jelas dan harus secara setia mencerminkan tujuan instruksional tentu saja. Prosedur ini merupakan tahapan pencapaian yang penting menguji bahwa sebagian dari bab berikutnya akan dikhususkan untuk penjelasan rinci deskripsi hal tersebut.

2. Tes penggabungan harus mengukur sampel representatif dari hasil belajar dan materi pelajaran termasuk dalam instruksi.

Pengujian selalu masalah sampling. Kami tidak pernah bisa mengajukan pertanyaan kami ingin bertanya dalam sebuah ujian. Di akhir sebuah pengalaman belajar di sana mungkin ratusan fakta dan istilah yang kita harapkan bisa diketahui oleh siswa, tapi karena waktu yang tersedia untuk pengujian, dan pertimbangan lainnya, kita hanya bisa memasukkan item yang hanya cukup untuk mengukur sebagian kecil dari fakta-fakta ini. Atau, kita bisa menyelesaikan sebuah unit instruksi yang berisi prinsip-prinsip yang dimilikinya aplikasi yang tak terhitung banyaknya untuk situasi tertentu, tapi kita hanya bisa mengukurnya sejumlah aplikasi ini. Dengan demikian, tes selalu memberikan sedikit terbatasnya sampel perilaku yang kita minati. Ini adalah tanggung jawab pengguna tes untuk menentukan cukup sampelnya dalam tes itu mencerminkan situasi alam semesta yang seharusnya diwakilinya.

Salah satu cara untuk lebih yakin bahwa tes prestasi memberikan contoh representatif dari perilaku yang diinginkan adalah menggunakan spesifikasi tabel. Ini adalah tabel dua kali di mana hasil pembelajaran terdaftar di satu sisi dan topik-topik materi di sisi lain. Mereka berpotongan sel di meja memungkinkan untuk menunjukkan proporsi tes dikhususkan untuk setiap hasil belajar dan topik subjek masing-masing.

Perhatikan tabel berikut ini untuk memahami bagaimana tes prestasi dapat menilai kemampuan peserta didik. Selain hal tersebut, guru sebagai penilai juga harus memperhatikan bagaimana penilaian yang cocok untuk peserta didik. Dengan tabel perencanaan tes prestasi yang tepat, maka pelaksanaan tes prestasi dapat dilakukan sesuai dengan rencana dan terstruktur.

<div> <div>LEARNING OUTCOMES</div> <div>SUBJECT MATTER TOPICS</div> </div>	KNOWLEDGE OF FACTS AND PRINCIPLES	UNDERSTANDING OF FACTS AND PRINCIPLES	APPLICATION OF FACTS AND PRINCIPLES	TOTAL NUMBER OF ITEMS
1. Planning the test	5*	5		10
2. Constructing objective test items	5	10	10	25
3. Constructing essay questions	5	5	5	15
4. Assembling the test	10	5	15	30
5. Using and appraising the test	5	5	10	20
<i>Total number of items</i>	30	30	40	100

* Either the number of items or the percentage of items may be recorded in each cell.

Tabel 2. Spesifikasi Unit Pada Konstruksi Tes

3. Tes prestasi harus mencakup jenis item tes yang ada paling tepat untuk mengukur hasil belajar yang diinginkan.

Hasil pembelajaran untuk kursus menentukan jenis perilaku kita yang akan diterima sebagai bukti pencapaian tujuan

instruksional. Tes prestasi hanyalah alat untuk memanggil yang telah ditentukan perilaku sehingga penilaian dapat dibuat mengenai sejauh mana pembelajaran yang telah terjadi. Kunci keefektifan pengujian, kemudian, pilih jenis item yang paling sesuai dan buatlah itu. Dengan hati-hati hal itu memunculkan respons yang diinginkan dan menghalangi yang lain tanggapan yang tidak relevan. Ini adalah situasi ideal, tentu saja, dan kita bisa hanya perkiraan di kelas pengujian.

Sejauh mana tanggapan siswa terhadap item tes dapat dikendalikan bervariasi dengan jenis barang yang digunakan. Klasifikasi utama dari item tes prestasi adalah sebagai berikut:

- a. *Supply-type* (jawaban perlengkapan siswa)
 - 1) Jawaban esai terbuka
 - 2) Jawaban esai terbatas
 - 3) Jawaban singkat (kata atau frasa)

- 4) Penyelesaian (mengisi kekosongan)
- b. *Selection-type* (siswa memilih jawaban)
 - 1) Benar-salah (atau respons alternatif)
 - 2) Menjodohkan
 - 3) Pilihan ganda
4. Tes prestasi harus dirancang agar sesuai dengan kegunaan tertentu dibuat dari hasil. Seperti yang telah kita catat sebelumnya, tes prestasi dapat digunakan untuk berbagai keperluan. Mereka dapat digunakan untuk mengukur masuk siswa (1) perilaku di awal pengajaran (tes penempatan), (2) belajar kemajuan selama pengajaran (tes formatif), (3) penyebab kesulitan belajar selama instruksi (tes diagnostik), dan (4) prestasi umum di akhir instruksi (tes sumatif).
5. Tes prestasi harus dilakukan seefisien mungkin dan kemudian harus ditafsirkan dengan hati-hati. Jika nilai yang diterima siswa pada tes prestasi sesuai dengan yang mereka terima

pada administrasi kedua dari tes yang sama, atau bentuk tes yang setara, skor akan dianggap sangat dapat diandalkan. Skor semua tes mengandung beberapa kesalahan (karena variasi dalam kondisi pengujian, tanggapan siswa, dan begitu juga), tetapi jumlah kesalahan yang proporsional dapat terjadi, dikurangi dengan konstruksi uji yang tepat

Secara umum, reliabilitasnya dari sebuah tes dapat ditingkatkan dengan memperpanjang tes dan dengan memperbaiki kualitas item uji tetangga. Tes yang lebih panjang memberi lebih banyak tcontoh perilaku yang diukur, dan uji yang benar. Item memberikan deskripsi yang lebih dapat dipuji dari setiap tes kinerja individu. Bila tujuan pengujian prestasi adalah untuk mendapatkan kerabat peringkat siswa (misalnya, untuk penilaian relatif), keandalan yang lebih besar rapat diperoleh dengan membuat tes yang

memberikan berbagai macam tes skor. Hal ini biasanya dilakukan dengan menulis barang pada tingkat 50 persen. Level kesulitan (50 persen siswa mendapatkan jawaban yang benar) dan dengan memperhatikan seberapa baik setiap item mendiskriminasi antara prestasi tinggi dan rendah.

Bila tujuan pengujian adalah untuk mendeskripsikan tugas pembelajaran mana setiap siswa dapat dan tidak dapat melakukan, bukan untuk memberi peringkat pada siswa sesuai urutan pencapaian, kisaran nilai tes tidak relevan. Dalam situasi ini, kesulitan item ditentukan oleh sulitnya tugas belajar mengajar diukur, dan tidak ada upaya yang dilakukan untuk memanipulasi kesulitan item untuk mendapatkan penyebaran nilai. Perkiraan reliabilitas tes tradisional didasarkan pada variabilitas skor, masalah khusus ditemui saat pembuatan tes

yang andal tidak memerlukan variabilitas skor. Disini kita harus lebih bergantung pada pencocokan item uji sedekat mungkin untuk tugas belajar yang spesifik, dengan menggunakan jumlah item yang cukup untuk setiap tugas belajar, dan menulis item yang jelas sesuai keinginan tingkah laku.

Meskipun perawatan dilakukan selama konstruksi uji, hasil tesnya akan mengandung sejumlah kesalahan pengukuran. Instruktur harus mempertimbangkan kesalahan ini selama interpretasi tes.

6. Tes prestasi harus digunakan untuk meningkatkan pembelajaran siswa.

Ini, tentu saja, telah menjadi tema dari keseluruhan bab ini. Sebelumnya, kita menunjukkan bahwa tes prestasi dapat membantu dalam keputusan instruksional, dapat secara langsung mempengaruhi pembelajaran siswa, dan bisa memberikan informasi tentang efektivitas pengajaran.

Pada bagian ini, kami telah menekankan prinsip-prinsip pengujian prestasi yang seharusnya meningkatkan nilainya sebagai alat belajar. Singkatnya, tes prestasi akan memiliki yang terbaik pengaruh positif pada pembelajaran saat mereka dengan setia merefleksikan instruksionalnya tujuan, ketika mereka mengukur sampel yang memadai dari yang dimaksudkan hasil belajar, ketika mereka memasukkan jenis barang yang paling sesuai untuknya hasil belajar, kapan mereka disesuaikan dengan kegunaan tertentu dibuat dari hasil, dan kapan mereka dikelompokkan untuk menghasilkan reliabel hasil.

Selain faktor di atas, penting untuk berkomunikasi kepada para siswa, dengan kata-kata dan perbuatan, yaitu tujuan utama pengujian adalah untuk meningkatkan pembelajaran mereka. Menunjukkan hubungan antara tujuan instruksional dan

jenis tes yang digunakan, mengklarifikasi sifat dan ruang lingkup semua tes sebelum diberikan, memberikan alas kaki pada kinerja tes siswa sesegera mungkin setelah pengujian, dan membuat saran spesifik mengenai perbaikan yang dilakukan baru-baru ini adalah langkah positif yang bisa ditempuh. Prosedur ini dan yang serupa harus dilakukan menyebabkan siswa melihat tes sebagai alat bantu dan harus meningkatkan jangkauannya pengujian mana yang memperkuat pengalaman lain dalam pembelajaran proses.

E. Langkah-Langkah Mengonstruksi Tes Prestasi Hasil Belajar

Dalam pengembangan tes prestasi belajar ada delapan langkah yang harus ditempuh, yaitu:

1. Menyusun spesifikasi tes
2. Menulis soal tes
3. Menentukan soal tes

4. Melakukan uji coba tes
5. Menganalisis butir soal
6. Merakit tes
7. Administrasi tes
8. Interpretasi hasil tes

Ada beberapa hal yang harus diperhatikan dalam merencanakan suatu tes, yaitu seperti berikut ini (Mehren & Lehman, 1984):

1. Tujuan tes
2. Pengetahuan, keterampilan, sikap, atau lainnya yang ingin diukur
3. Tabel spesifikasi
4. Kesesuaian butir tes dengan tujuan
5. Format butir tes
6. Lama waktu untuk tes
7. Tingkat kesukaran tes
8. Tingkat pembedaan tes
9. Susunan format tes (bila lebih dari satu)
10. Susunan butir tes untuk tiap format
11. Persiapan mahasiswa
12. Tempat menulis jawaban tes

13. Cara penskoran
14. Penskoran tes esei dan pilihan ganda
15. Tabulasi skor tes
16. Laporan hasil tes

Saat membangun sebuah tes, perlu dilakukan beberapa tahapan untuk memastikan kualitasnya yang baik:

1. Keseluruhan rencana.

Ini adalah tahap awal yang sangat penting, perlu mempertimbangkan secara rinci apa sebenarnya yang ingin diukur, manifestasinya, dan faktor mana yang tidak dapat mempengaruhi hasil pengukuran. Sangat penting untuk mendefinisikan dengan jelas tujuan pengujian karena ini meningkatkan kemungkinan untuk mencapai validitas tinggi. Perlu juga membuat keputusan tentang format tes yang mana yang paling tepat untuk tujuan yang akan dicapai.

2. Definisi isi.

Pada tahap ini harus menentukan konten apa yang akan diuji.

3. Spesifikasi uji.

Spesifikasi uji merupakan rencana pengujian. Ini adalah dokumen praktis terperinci yang menunjukkan apa yang akan berisi tes, dan dimaksudkan untuk membantu konstruksi uji. Menurut Alderson, J. C., C. Clapham, D. Wall, spesifikasi uji meliputi informasi berikut:

- a. Tujuan pengujian - apakah itu penempatan, prestasi, kemahiran, atau tes diagnostik.
- b. Tipe pelajar yang akan mengikuti ujian, termasuk usia mereka, jenis kelamin, tingkat kemahiran, bahasa pertama, negara asal, tingkat pendidikan, alasan untuk mengikuti ujian, dll;
- c. Jumlah bagian tes akan memiliki, berapa lama mereka akan kerjakan dan dengan cara apa mereka akan dibedakan;

- d. Kemampuan apa yang harus diuji
- e. Jenis tugas apa yang dibutuhkan - dapat dinilai secara objektif, integratif, disimulasikan "asli", dll;
- f. Berapa banyak barang yang dibutuhkan untuk setiap bagian, dan berapa berat relatifnya - bobot yang sama atau bobot ekstra untuk item yang lebih sulit;
- g. Metode uji apa yang akan digunakan - pilihan ganda, menjodohkan, transformasi, deskripsi gambar, penulisan esai, dll;
- h. Rubrik apa yang harus digunakan sebagai petunjuk bagi siswa - akan disertakan contoh untuk membantu siswa mengetahui apa yang diharapkan, dan jika kriteria penilaian ditambahkan ke rubrik;
- i. Kriteria penilaian apa yang akan digunakan - seberapa penting akurasi, ejaan, panjang teks tertulis, dll.



BAB 7

ANALISIS BUTIR SOAL

A. Pengertian Analisis Butir Soal

Menurut Suharsimi (dalam Srikandi, 2002) menyatakan bahwa analisis soal merupakan prosedur sistematis yang akan memberikan informasi-informasi yang sangat khusus terhadap butir tes yang akan disusun.

Menurut Silverius (dalam Srikandi, 2002) menyatakan bahwa baikburuknya (mutu) butir soal ditetapkan dengan melihat taraf kesukarannya, fungsi stem (pokok soal), fungsi *distractor* (pengecoh), serta penyebaran jawaban pada pengecoh dalam total kelompok.

Kegiatan menganalisis butir soal merupakan suatu kegiatan yang harus dilakukan guru untuk meningkatkan mutu soal yang telah ditulis. Kegiatan ini merupakan proses pengumpulan, peringkasan, dan penggunaan informasi dari jawaban siswa untuk membuat keputusan tentang setiap penilaian (Nitko, 1996: 308).

Penilaian terhadap butir soal pada dasarnya merupakan analisis butir soal, dan selama ini

pada umumnya para ahli pengukuran mengatakan bahwa analisis butir soal maksudnya adalah penilaian terhadap soal. Telah diketahui bersama bahwa penyusunan tes sangat mempengaruhi kualitas butir soal.

Pendekatan untuk menganalisis butir soal yang berkembang saat ini terdiri dari dua pendekatan yaitu pendekatan klasik dan pendekatan modern. Kedua pendekatan ini masing-masing memiliki kelebihan dan kekurangan. Namun keduanya masih sering digunakan dalam analisis butir soal. Analisis butir soal dengan pendekatan klasik diantaranya dapat dilakukan menggunakan Program Iteman.

Dengan melihat beberapa persyaratan yang harus dipenuhi, penyusunan tes dituntut untuk mengikuti pedoman penyusunan tes dan melakukan ujicoba. Kemudian berdasarkan hasil ujicoba, respon peserta dianalisis menggunakan Program iteman untuk mendapatkan karakteristik butir soal.

Data hasil analisis dengan Program Iteman dianalisis kembali menggunakan instrumen penilaian butir soal yang memenuhi syarat sebagai alat ukur yang baik. Suryabrata (1999) menyatakan bahwa analisis butir soal mencakup telaah soal atau analisis kualitatif dan analisis terhadap data empirik hasil ujicoba atau analisis kuantitatif.

Analisis butir soal secara kualitatif menekankan penilaian dari ketiga segi yaitu materi, konstruksi, dan bahasa. Namun demikian dalam pembahasan ini dikhususkan untuk menjelaskan analisis butir soal secara kuantitatif.

Analisis ini dilakukan berdasarkan data yang diperoleh secara empiris melalui ujicoba dari suatu perangkat tes. Analisis kuantitatif sering disebut dengan analisis item yang menghasilkan karakteristik atau parameter butir dan tes, yaitu: tingkat kesukaran, daya beda dan distribusi jawaban dan kunci setiap butir, serta reliabilitas dan kesalahan pengukuran (SEM) dalam tes.

Telah disinggung di depan bahwa analisis soal antara lain bertujuan untuk mengadakan identifikasi soal-soal yang baik, kurang lebih atau sedang dan soal yang tidak baik. Dengan analisis soal dapat diperoleh informasi tentang kekurangan sebuah soal tes dan "petunjuk" untuk mengadakan perbaikan.

Dalam tes dan pengukuran, dikenal beberapa karakteristik butir soal. Untuk tes hasil belajar pada umumnya dipertimbangkan tiga karakteristik butir soal, yaitu: tingkat kesukaran, daya beda dan distribusi jawaban atau berfungsi tidaknya pilihan jawaban (distraktor). Ketiga karakteristik butir soal ini secara bersama-sama akan menentukan mutu butir soal. Bila salah satu dari ketiga karakteristik ini tidak memenuhi persyaratan maka mutu butir soal akan turun.

Menurut Nana Sudjana (2006: 135), "Analisis butir soal atau analisis item adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas

yang memadai". Menurut Arikunto (2009:205), "Analisis soal adalah suatu prosedur yang sistematis, yang akan memberikan informasi informasi yang sangat khusus terhadap butir tes yang kita susun". Adapun fungsi mengadakan analisi butir soal, Arikunto (2006:205) mengatakan : Membantu kita dalam mengidentifikasi butir-butir soal yang jelek, memperoleh informasi yang akan dapat digunakan untuk menyempurnakan soal-soal untuk kepentingan lanjut, memperoleh gambaran secara selintas tentang keadaan yang kita susun.

Suryabrata (1999) menyatakan bahwa analisis butir soal mencakup telaah soal atau analisis kualitatif dan analisis terhadap data empirik hasil ujicoba atau analisis kuantitatif. Dalam tes dan pengukuran, dikenal beberapa karakteristik butir soal. Untuk tes hasil belajar pada umumnya dipertimbangkan tiga karakteristik butir soal, yaitu : tingkat kesukaran, daya beda dan distribusi jawaban atau berfungsi

tidaknya pilihan jawaban (distraktor). Ketiga karakteristik butir soal ini secara bersama-sama akan menentukan mutu butir soal. Bila salah satu dari ketiga karakteristik ini tidak memenuhi persyaratan maka mutu butir soal akan turun.

B. Alasan dan Tujuan Analisis Butir Soal

Tujuan penelaahan atau analisis butir adalah untuk mengkaji dan menelaah setiap butir soal agar diperoleh soal yang bermutu sebelum soal digunakan. Di samping itu, tujuan analisis butir soal juga untuk membantu meningkatkan tes melalui revisi atau membuang soal yang tidak efektif, serta untuk mengetahui informasi diagnostik pada siswa apakah mereka sudah/belum memahami materi yang telah diajarkan (Aiken, 1994: 63).

Menurut Thordike dan Hagen (dalam Srikandi, 2002) menyatakan bahwa analisis terhadap soal-soal tes yang telah dijawab peserta didik mempunyai dua tujuan:

1. Jawaban soal-soal itu merupakan informasi diagnostik untuk meneliti pelajaran dari kelas itu dan kegagalan-kegagalan belajarnya serta selanjutnya untuk membimbing ke arah cara belajar yang lebih baik.
2. Jawaban terhadap soal-soal yang terpisah dan perbaikan (*review*) soal-soal yang didasarkan atas jawaban-jawaban itu merupakan langkah awal bagi penyiapan tes-tes yang lebih baik untuk tahun berikutnya.

Ada beberapa alasan mengapa diperlukan analisis butir soal. Menurut (Asmawi Zainul, dkk: 1997) alasan tersebut antara lain:

1. Untuk mengetahui kekuatan dan kelemahan butir tes, sehingga dapat dilakukan seleksi dan revisi butir soal.
2. Untuk menyediakan informasi tentang spesifikasi butir soal secara lengkap, sehingga akan lebih memudahkan bagi pembuat soal dalam menyusun perangkat

soal yang akan memenuhi kebutuhan ujian dalam bidang dan tingkat tertentu.

3. Untuk segera dapat mengetahui masalah yang terkandung dalam butir soal, seperti: kemenduaan butir soal, kesalahan meletakkan kunci jawaban, soal yang terlalu sukar dan terlalu mudah, atau soal yang mempunyai daya beda rendah. Masalah ini bila diketahui dengan segera akan memungkinkan bagi pembuat soal untuk mengambil keputusan apakah butir soal yang bermasalah itu akan digugurkan atau direvisi guna menentukan nilai peserta didik.
4. Untuk dijadikan alat guna menilai butir soal yang akan disimpan dalam kumpulan soal.
5. Untuk memperoleh informasi tentang butir soal sehingga memungkinkan untuk menyusun beberapa perangkat soal yang paralel. Penyusunan perangkat seperti ini sangat bermanfaat bila akan melakukan ujian ulang atau mengukur kemampuan beberapa

kelompok peserta tes dalam waktu yang berbeda.

C. Manfaat Analisis Butir Soal

Kegiatan analisis butir soal memiliki banyak manfaat, di antaranya adalah:

1. Dapat membantu para pengguna tes dalam evaluasi atas tes yang digunakan,
2. Sangat relevan bagi penyusunan tes informal dan lokal seperti tes yang disiapkan guru untuk siswa di kelas,
3. Mendukung penulisan butir soal yang efektif,
4. Secara materi dapat memperbaiki tes di kelas,
5. Meningkatkan validitas soal dan reliabilitas (Anastasi and Urbina, 1997:172).

Disamping itu manfaat lainnya menurut Nitko, 1996 adalah:

1. Menentukan apakah suatu fungsi butir soal sesuai dengan yang diharapkan,

2. Memberi masukan kepada siswa tentang kemampuan dan sebagai dasar untuk bahan diskusi di kelas,
3. Memberi masukan kepada guru tentang kesulitan siswa,
4. Memberi masukan pada aspek tertentu untuk pengembangan kurikulum,
5. Merevisi materi yang dinilai atau diukur,
6. Meningkatkan keterampilan penulisan soal.

D. Karakteristik Butir Soal

Butir-butir soal yang akan diberikan kepada peserta didik dalam rangka melakukan tes memiliki beberapa karakteristik berikut ini:

1. Tingkat Kesulitan (*Difficulty Level*)

Menurut Asmawi Zainul, dkk (1997) tingkat kesukaran butir soal adalah proporsi peserta tes menjawab benar terhadap butir soal tersebut. Tingkat kesukaran butir soal biasanya dilambangkan dengan p . Makin besar nilai p yang berarti makin besar

proporsi yang menjawab benar terhadap butir soal tersebut, makin rendah tingkat kesukaran butir soal itu. Hal ini mengandung arti bahwa soal itu makin mudah, demikian pula sebaliknya.

Soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak merangsang mahasiswa untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan mahasiswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya (Suharsimi Arikunto : 2001).

Besarnya tingkat kesukaran berkisar antara 0,00 sampai 1,00. Untuk sederhananya, tingkat kesukaran butir dan perangkat soal dapat dibagi menjadi tiga kelompok, yaitu mudah, sedang dan sukar. Sebagai patokan menurut (Asmawi Zainul, dkk : 1997) dapat digunakan tabel sebagai berikut:

Tingkat Kesukaran	Nilai p
Sukar	0,00 – 0,25
Sedang	0,26 – 0,75
Mudah	0,76 – 1,00

Tabel 3. Tingkat Kesukaran Butir Soal

Untuk menyusun suatu naskah ujian sebaiknya digunakan butir soal yang mempunyai tingkat kesukaran berimbang, yaitu : soal berkategori sukar sebanyak 25%, kategori sedang 50% dan kategori mudah 25%.

Dalam penggunaan butir soal dengan komposisi seperti di atas, maka dapat diterapkan penilaian berdasar acuan norma atau acuan patokan. Bila komposisi butir soal dalam suatu naskah ujian tidak berimbang, maka penggunaan penilaian acuan norma tidaklah tepat, karena informasi kemampuan

yang dihasilkan tidaklah akan berdistribusi normal.

Walaupun demikian ada yang berpendapat bahwa soal-soal yang dianggap baik adalah soal-soal yang sedang, yaitu soal-soal yang mempunyai indeks kesukaran berkisar antara 0,26 – 0,75. Berbagai kriteria tersebut mempunyai kecenderungan bahwa butir soal yang memiliki indeks kesukaran kurang dari 0,25 dan lebih dari 0,75 sebaiknya dihindari atau tidak digunakan, karena butir soal yang demikian terlalu sukar atau terlalu mudah, sehingga kurang mencerminkan alat ukur yang baik.

2. Daya Beda

Daya beda butir soal ialah indeks yang menunjukkan tingkat kemampuan butir soal membedakan kelompok yang berprestasi tinggi (kelompok atas) dari kelompok yang berprestasi rendah (kelompok bawah) diantara para peserta tes (Asmawi Zainul, dkk : 1997).

Suryabrata (1999) menyatakan tujuan pokok mencari daya beda adalah untuk menentukan apakah butir soal tersebut memiliki kemampuan membedakan kelompok dalam aspek yang diukur, sesuai dengan perbedaan yang ada pada kelompok itu.

Daya beda butir soal yang sering digunakan dalam tes hasil belajar adalah dengan menggunakan indeks korelasi antara skor butir dengan skor totalnya. Daya beda dengan cara ini sering disebut validitas internal, karena nilai korelasi diperoleh dari dalam tes itu sendiri. Daya beda dapat dilihat dari besarnya koefisien korelasi biserial maupun koefisien korelasi point biserial.

Dalam analisis ini digunakan nilai koefisien korelasi biserial untuk menentukan daya beda butir soal. Koefisien korelasi biserial menunjukkan hubungan antara dua skor, yaitu skor butir soal dan skor keseluruhan dari peserta tes yang sama.

Koefisien daya beda berkisar antara $-1,00$ sampai dengan $+1,00$. Daya beda $+1,00$ berarti bahwa semua anggota kelompok atas menjawab benar terhadap butir soal itu, sedangkan kelompok bawah seluruhnya menjawab salah terhadap butir soal itu. Sebaliknya daya beda $-1,00$ berarti bahwa semua anggota kelompok atas menjawab salah butir soal itu, sedangkan kelompok bawah seluruhnya menjawab benar terhadap soal itu.

Daya beda yang dianggap masih memadai untuk sebutir soal ialah apabila sama atau lebih besar dari $+0,30$. Bila lebih kecil dari itu, maka butir soal tersebut dianggap kurang mampu membedakan peserta tes yang mempersiapkan diri dalam menghadapi tes dari peserta yang tidak mempersiapkan diri. Bahkan bila daya beda itu menjadi negatif, maka butir soal itu sama sekali tidak dapat dipakai sebagai alat ukur prestasi belajar.

Oleh karena itu butir soal tersebut harus dikeluarkan dari perangkat soal. Makin tinggi daya beda suatu butir soal, maka makin baik butir soal tersebut, dan sebaliknya makin rendah daya bedanya, maka butir soal itu dianggap tidak baik (Asmawi Zainul, dkk : 1997).

Menurut Dali S Naga (1992) kriteria besarnya koefisien daya beda diklasifikasikan menjadi empat kategori. Secara lebih rinci dijelaskan dalam tabel di bawah ini.

Kategori Daya Beda	Koefisien Korelasi
Baik	0,40 – 1,00
Sedang (tidak perlu revisi)	0,30 – 0,39
Perlu direvisi	0,20 – 0,29
Tidak baik	-1,00 – 0,19

Tabel 4. Tingkat Daya Beda Butir Soal

3. Distribusi Jawaban

Dilihat dari konstruksi butir soal terdiri dari dua bagian, yaitu pokok soal dan alternatif jawaban. Alternatif jawaban jawaban juga terdiri dari dua bagian, yaitu kunci jawaban dan pengecoh. Pengecoh dikatakan berfungsi apabila semakin rendah tingkat kemampuan peserta tes semakin banyak memilih pengecoh, atau makin tinggi tingkat kemampuan peserta tes akan semakin sedikit memilih pengecoh.

Hal demikian dapat ditunjukkan dengan adanya korelasi yang tinggi, rendah atau negatif pada hasil analisis. Apabila proporsi peserta tes yang menjawab dengan salah atau memilih pengecoh kurang dari 0,025 maka pengecoh tersebut harus direvisi. Dan untuk pengecoh yang ditolak apabila tidak ada yang memilih atau proporsinya 0,00 (Depdikbud : 1997).

Proporsi alternatif jawaban masing-masing butir soal dapat dilihat pada kolom *proportion endorsing* pada hasil analisis iteman.

Selain memperhatikan fungsi daya tarik untuk dipilih oleh peserta tes, pengecoh soal juga perlu memperhatikan daya beda (koefisien korelasi) yang ditunjukkan oleh masing-masing alternatif jawaban. Setiap pengecoh diharapkan memiliki daya beda negatif, artinya suatu pengecoh diharapkan lebih sedikit dipilih oleh kelompok tinggi dibandingkan dengan kelompok bawah. Atau daya beda pengecoh tidak lebih besar dari daya beda kunci jawaban setiap butir soal.

Menurut Depdikbud (1997) untuk menilai pengecoh (distraktor) dari masing-masing butir soal dapat dikategorikan sebagai berikut:

Kategori Distraktor	Nilai Proportion Endorsing
Baik	$\geq 0,025$
Revisi	$< 0,025$
Tidak Baik / Tolak	0,000

Tabel 5. Tingkat Distraktor Butir Soal

E. Validitas

Validitas merupakan suatu ukuran yang menunjukkan kevalidan atau kesahihan suatu instrument. Jadi pengujian validitas itu mengacu pada sejauh mana suatu instrument dalam menjalankan fungsi. Instrument dikatakan valid jika instrument tersebut dapat digunakan untuk mengukur apa yang hendak diukur (Sugiyono, 2008: 363).

Macam-macam validitas:

1. Validitas Isi

Validitas isi merupakan syarat kualitas tes yang sangat penting terutama bagi para guru dalam pengajaran di kelasnya. Mengingat

pentingnya validitas isi suatu tes, maka bagaimana upaya yang seharusnya dilakukan guru agar tes yang dikonstruksinyamemiliki validitas isi. Suatu tes dikatakan memiliki validitas isi apabila tujuan-tujuanpembelajaran suatu mata pelajaran terwakili dalam tes. Hal senada dikemukakan Wiersma (1990:183), bahwa validitas isi merupakan perhatian utama suatu tes yang digunakan dalam pengajaran di kelas, dan untuk mengetahuinya melalui analisis logis. Ahli lainnya, Best mengemukakan bahwa suatu tes memiliki validitas logis apabila tesbenar-benar mengukur atau berhubungan secara spesifik dengan fenomena yang diujikan.

Validitas isi berkenaan dengan kesanggupan instrumen mengukur isi yang harus diukur. Artinya, alat ukur tersebut mampu mengungkap isi suatu konsep atau variabel yang hendak diukur. Misalnya tes hasil belajar bidang studi IPS, harus bisa mengungkap isi

bidang studi tersebut. Hal ini bisa dilakukan dengan cara menyusun tes yang bersumber dari kurikulum bidang studi yang hendak diukur. Di samping kurikulum dapat juga diperkaya dengan melihat/mengkaji buku sumber. Sehingga tes hasil belajar tidak mungkin dapat mengungkap semua materi yang ada dalam bidang studi tertentu sekalipun hanya untuk satu semester. Oleh sebab itu harus diambil sebagian dari materi dalam bentuk sampel tes. Sebagai sampel maka harus dapat mencerminkan materi yang terkandung dari seluruh materi bidang studi. Cara yang ditempuh dalam menetapkan sampel tes adalah memilih konsep-konsep yang esensial dari materi yang di dalamnya. Misalnya menetapkan sejumlah konsep dari setiap pokok bahasan yang ada. Dari setiap konsep dikembangkan beberapa pertanyaan tes (lihat bagan). Di sinilah pentingnya peranan

kisi-kisi sebagai alat untuk memenuhi validitas isi. Perhatikan contoh berikut!

TES HASIL BELAJAR

Bidang Studi :

Semester :

Kelas :

Pokok Bahasan	Konsep Materi	Jumlah Pertanyaan	Jenis Tes	Abilitas yang Diukur
Pokok Bahasan 1	1.1 ...	3 soal	Pilihan ganda, aplikasi, dan seterusnya	Aplikasi dan seterusnya
	2.1 ...	2 soal		
Pokok Bahasan 2	2.1 ...	2 soal		
	2.2 ...	3 soal		
Pokok Bahasan 3	3.1	3 soal		
	3.2	2 soal		

	Dan seterus nya			

Tabel 6. Contoh Kisi-Kisi Butir Soal

Dalam hal tertentu tes yang telah disusun sesuai dengan kurikulum (materi dan tujuannya) agar memenuhi validitas isi, peneliti atau pemakai tes dapat meminta bantuan ahli bidang studi untuk menelaah apakah konsep materi yang diajukan telah memadai atau tidak, sebagai sampel tes. Dengan demikian validitas isi tidak memerlukan uji coba dan analisis statistik atau dinyatakan dalam bentuk angka-angka. Tes yang tidak memiliki validitas isi dapat berakibat:

- a. Siswa tidak dapat menunjukkan kemampuan yang sesungguhnya,

- b. Item yang disajikan tidak relevan memungkinkan siswa salah menjawab hanya karena isinya belum diajarkan, sehingga skor tes tidak mengukur performa siswa secara akurat terhadap materi yang diukur (Wiersma & Jurs, 1990: 185).

Berdasarkan uraian tersebut, dapat ditarik pengertian bahwa yang dimaksud validitas isi dalam konteks penelitian ini adalah analisis logis atau rasional, dalam pengertian proses mempertimbangan, menilai dan mengambil keputusan tentang representasi item-item tes baik dilakukan sendirian atau melalui judgement para pakar dalam bidangnya. Setiap item tes dinilai mengenai relevansinya dengan sifat/karakteristik yang diujikan, yang dalam hal ini relevansinya dengan tujuan.

2. Validitas Bangun Pengertian

Validitas bangun pengertian (*construct validity*) berkenaan dengan kesanggupan alat ukur mengukur pengertian-pengertian yang

terkandung dalam materi yang diukurnya. Pengertian-pengertian yang terkandung dalam konsep kemampuan, minat, sebagai variabel penelitian dalam berbagai bidang kajian harus jelas apa yang hendak diukurnya. Konsep-konsep tersebut masih abstrak, memerlukan penjabaran yang lebih spesifik, sehingga mudah diukur. Ini berarti setiap konsep harus dikembangkan indikator-indikatornya. Dengan adanya indikator dari setiap konsep maka bangun pengertian akan nampak dan memudahkan dalam menetapkan cara pengukuran. Untuk variabel tertentu, dimungkinkan penggunaan alat ukur yang beraneka ragam dengan cara mengukurnya yang berlainan. Menetapkan indikator suatu konsep dapat dilakukan dalam dua cara, yakni (a) menggunakan pemahaman atau logika berpikir atas dasar teori pengetahuan ilmiah dan (b)

menggunakan pengalaman empiris, yakni apa yang terjadi dalam kehidupan nyata.

Contoh: Konsep mengenai "Hubungan Sosial", dilihat dari pengalaman, indikatornya empiris adalah keterkaitan dari:

- a. Bisa bergaul dengan orang lain
- b. Disenangi atau banyak teman-temannya
- c. Menerima pendapat orang lain
- d. Tidak memaksakan pendapatnya
- e. Bisa bekerja sama dengan siapa pun, dan lain-lain.

Mengukur indikator-indikator tersebut, berarti mengukur bangun pengertian yang terdapat dalam konsep hubungan sosial.

Contoh lain: Konsep sikap dapat dilihat dari indikatornya secara teoretik (deduksi teori) antara lain keterkaitan dari:

- a. Kesiediaan menerima stimulus objek sikap
- b. Kemauan mereaksi stimulus objek sikap
- c. Menilai stimulus objek sikap
- d. Menyusun/mengorganisasi objek sikap

- e. Internalisasi nilai yang ada dalam objek sikap.

Apabila hasil tes menunjukkan indikator-indikator tes yang tidak berhubungan secara positif satu sama lain, berarti ukuran tersebut tidak memiliki validitas bangun pengertian. Atas dasar itu indikatornya perlu ditinjau atau diperbaiki kembali. Cara lain untuk menetapkan validitas bangun pengertian suatu alat ukur adalah menghubungkan (korelasi) antara alat ukur yang dibuat dengan alat ukur yang sudah baku/standardized, seandainya telah ada yang baku. Bila menunjukkan koefisien korelasi yang tinggi maka alat ukur tersebut memenuhi validitasnya.

3. Validitas Ramalan

Validitas ramalan artinya dikaitkan dengan kriteria tertentu. Dalam validitas ini yang diutamakan bukan isi tes tapi kriterianya, apakah alat ukur tersebut dapat digunakan

untuk meramalkan suatu ciri atau perilaku tertentu atau kriteria tertentu yang diinginkan. Misalnya alat ukur motivasi belajar, apakah dapat digunakan untuk meramal prestasi belajar yang dicapai. Artinya terdapat hubungan yang positif antara motivasi dengan prestasi. Dengan kata lain dalam validitas ini mengandung ciri adanya relevansi dan keajegan atau ketetapan (*reliability*). Motivasi dapat digunakan meramal prestasi bila skor-skor yang diperoleh dari ukuran motivasi berkorelasi positif dengan skor prestasi. Validitas ramalan ini mengandung dua makna. Pertama validitas jangka pendek dan kedua jangka panjang. Validitas jangka pendek, artinya daya ramal alat ukur tersebut hanya untuk masa yang tidak lama. Artinya, skor tersebut berkorelasi pada waktu yang sama. Misalnya, ketetapan (*reliability*) terjadi pada semester dua artinya

daya ramal berlaku pada semester dua, dan belum tentu terjadi pada semester berikutnya. Sedangkan validitas jangka panjang mengandung makna skor tersebut akan berkorelasi juga dikemudian hari. Mengingat validitas ini lebih menekankan pada adanya korelasi, maka faktor yang berkenaan dengan persyaratan terjadinya korelasi harus dipenuhi. Faktor tersebut antara lain hubungan dari konsep dan variabel dapat dijelaskan berdasarkan pengetahuan ilmiah, minimal masuk akal sehat dan tidak mengada-ada. Faktor lain adalah skor yang dikorelasikan memenuhi linieritas. Ketiga validitas yang dijelaskan di atas idealnya dapat digunakan dalam menyusun instrumen penelitian, minimal dua validitas, yakni validitas isi dan validitas bangun pengertian. Validitas isi dan bangun pengertian mutlak diperlukan dan bisa

diupayakan tanpa melakukan pengujian secara statistika.

F. Reliabilitas

Variabel penting kedua sebagai indikator kualitas tes hasil belajar adalah reliabilitas. Istilah reliabilitas bersinonim dengan keterandalan, keterpercayaan, keakuratan, ketelitian, keajegan, dan keteramalan (Kerlinger, 1995: 709-710). Dalam aplikasinya untuk pengujian dan asesmen, reliabilitas secara umum diklarifikasi sebagai berikut:

1. Merujuk kepada hasil yang diperoleh instrumen asesmen dan bukan instrumen itu sendiri,
2. Estimasi reliabilitas selalu merujuk kepada konsistensi tertentu,
3. Reliabilitas diperlukan sebagai syarat validitas tes, dan
4. Reliabilitas terutamabersifat statistik (Linn, 1995: 82).

Secara teknis terdapat beberapa teknik atau acuan untuk mengestimasi keterandalan. Menurut Kerlinger (1995), terdapat tiga acuan, yaitu:

1. Stabilitas, dependabilitas, prediktabilitas, yaitu pengukuran obyek yang sama berulang kali dengan instrumen yang sama mendapatkan hasil yang sama;
2. Akurasi, yaitu pengukuran dari suatu instrumen mengukur 'sebenarnya' sifat yang hendak diukur, dan
3. Berkaitan dengan kekeliruan pengukuran instrumen pengukur.

Sax (1978: 257) mengemukakan prosedur untuk estimasi reliabilitas, yaitu:

1. Stabilitas, yaitu korelasi antara skor pada tes yang sama dua kali;
2. Ekuivalen, yaitu korelasi antara skor pada dua atau lebih bentuk tes yang sama secara konkuren;

3. Stabilitas dan ekuivalen, yaitu korelasi antara skor pada dua atau lebih tes yang sama yang dilakukan pada waktu berbeda, dan
4. Konsistensi internal, yaitu korelasi antara item di dalam tes.

Reliabilitas adalah suatu hal yang sangat penting pada alat pengukuran standar. Reliabilitas dihubungkan dengan pengertian adanya ketepatan tes dalam pengukurannya. Reliabilitas adalah kestabilan skor yang diperoleh peserta tes yang sama ketika diuji ulang dengan tes yang sama pada situasi yang berbeda atau dari suatu pengukuran ke pengukuran lainnya. Dengan kata lain reliabilitas merupakan tingkat konsistensi atau kemantapan hasil terhadap hasil dua pengukuran hal yang sama. Dapat juga diartikan sebagai tingkat kepercayaan dari suatu alat ukur (Depdikbud: 1997).

Hasil pengukuran diharapkan akan sama apabila pengukuran itu diulangi. Dengan perangkat tes yang reliabel, apabila tes itu

diberikan dua kali pada peserta yang sama tetapi dalam selang waktu yang berbeda sepanjang tidak ada perubahan dalam kemampuan maka skor yang diperoleh akan konstan. Kriteria untuk menentukan tinggi rendahnya reliabilitas sebuah perangkat tes, menurut (Suharsimi Arikunto: 2001) dilihat pada rentangan koefesien korelasi sebagai berikut.

Kategori Reliabilitas Tes	Nilai Koefesien Korelasi
Sangat Tinggi	0,800 – 1,000
Tinggi	0,600 – 0,799
Cukup	0,400 – 0,599
Rendah	0,200 – 0,399
Sangat Rendah	0,000 – 0,199

Tabel 7. Tingkat Reliabilitas Butir Soal

Tes yang memiliki konsistensi reliabilitas tinggi adalah akurat, reproducibel, dan generalized terhadap kesempatan testing dan

instrumen tes lainnya. Secara rinci faktor yang mempengaruhi reliabilitas skor tes di antaranya:

1. Semakin banyak jumlah butir soal, semakin ajek suatu tes.
2. Semakin lama waktu tes, semakin ajek.
3. Semakin sempit range kesukaran butir soal, semakin besar keajegan.
4. Soal-soal yang saling berhubungan akan mengurangi keajegan.
5. Semakin objektif pemberian skor, semakin besar keajegan.
6. Ketidaktepatan pemberian skor.
7. Menjawab besar soal dengan cara menebak.
8. Semakin homogen materi semakin besar keajegan.
9. Pengalaman peserta ujian.
10. Salah penafsiran terhadap butir soal.
11. Menjawab soal dengan buru-buru/cepat.
12. Kesiapan mental peserta ujian.
13. Adanya gangguan dalam pelaksanaan tes.
14. Jarak antara tes pertama dengan tes kedua.

15. Mencontek dalam mengerjakan tes.
16. Posisi individu dalam belajar.
17. Kondisi fisik peserta ujian.

Untuk mengetahui koefisien reliabilitas tes soal bentuk pilihan ganda digunakan rumus Kuder Richadson 20 (KR-20) seperti berikut ini.

$$KR-20 = \frac{k}{k-1} \left[1 - \frac{\sum p(1-p)}{(SD)^2} \right]$$

Keterangan:

k : Jumlah butir soal
 $(SD)^2$: Varian



BAB 8

PENGEMBANGAN INSTRUMEN PENILAIAN RANAH AFEKTIF

A. Pengertian Penilaian Ranah Afektif

Pendidikan merupakan suatu hal yang sangat penting bagi manusia. Hal ini membuat pendidikan menjadi salah satu langkah dalam rangka mengembangkan manusia menjadi lebih baik lagi dari struktu kualitas. Kualitas yang harus diperbaiki pada setiap prosesnya tidak hanya aspek kognitifnya saja melainkan juga harus memperhatikan ranah afektif dan psikomotorik.

Dengan demikian dari Harvey dan Smith (1991) berpendapat bahwa sikap adalah kesiapan merespons secara konsisten dalam bentuk positif atau negative terhadap objek atau situasi. Dari pengertian tersebut dapat dimengerti bahwa sikap merupakan kesiapan respon, maka respon inilah yang akan menjadi bahan untuk dijadikan sebagai nilai pada siswa yang akan diamati oleh pendidik. Sikap dan afektif memiliki keterkaitan didalamnya. Seperti halnya yang telah disampaikan oleh Sudaryono (2012) sikap

merupakan variable tersembunyi yang tidak dapat diamati secara langsung, tetapi dapat disimpulkan melalui tingkah laku. Dengan demikian untuk menilai sebuah ranah afektif maka pengamat atau dalam hal ini adalah pendidik harus bisa mengamati tingkah laku (sikap) dari peserta didik itu sendiri.

Rusgiyanto (2005) berpendapat bahwa Komponen afektif adalah perasaan yang dimiliki oleh seseorang atau penilaian terhadap suatu objek, Komponen kognitif adalah kepercayaan atau keyakinan seseorang mengenai objek. Adapun komponen konatif kecendrungan untuk berperilaku atau berbuat dengan cara-cara tertentu berkenaan dengan kehadiran objek sikap.

Dalam tulisan Departemen Pendidikan Nasional tahun 2008 juga dijelaskan bahwa ranah afektif merupakan salah satu taksonomi tujuan instruksional yang berkaitan dengan kondisi psikologis atau perasaan seseorang. Ada lima

karakteristik afektif yang penting, yaitu "sikap, minat, konsep diri, nilai dan moral."

Menurut Nurul Zuriah (2008: 19) mendefinisikan afektif sebagai bentuk emosi atau perasaan yang dihubungkan dengan nilai-nilai hidup, sikap, keyakinan terkait dengan moral dan watak seseorang. Pendapat tersebut diperkuat oleh pendapat Zubaedi (2011: 26) yang mendefinisikan ranah afektif atau affective domain sebagai kawasan yang berkaitan dengan aspek-aspek emosional seperti perasaan, minat, sikap, dan kepatuhan terhadap moral.

Dari pemaparan tersebut, dapat ditarik kesimpulan bahwa ranah afektif merupakan bentuk emosional seperti perasaan, minat, sikap, dan kepatuhan terhadap moral yang dipengaruhi oleh kesadaran siswa terhadap nilai yang diyakini dan kemudian diinternalisasikan ke dalam dirinya yang akan mempengaruhi perilaku. Oleh karena ranah afektif 24 menyangkut minat, motivasi dan sikap siswa dalam pembelajaran

maka akan berdampak pada pencapaian terhadap aspek kognitif dan psikomotorik. Oleh karena itu ranah afektif memegang peranan pokok dalam menentukan pencapaian hasil belajar siswa.

B. Tujuan Penilaian Ranah Afektif

Menurut Arikunto (Amir, 2016) Tujuan penilaian afektif adalah :

1. Untuk mendapatkan umpan balik (*feed back*) baik bagi guru maupun siswa sebagai dasar untuk memperbaiki proses belajar-mengajar dan mengadakan program perbaikan (*remedial program*) bagi siswanya.
2. Untuk mengetahui tingkat perubahan tingkah laku siswa yang dicapai antara lain diperlukan sebagai bahan bagi: perbaikan tingkah laku siswa, pemberian laporan kepada orang tua, dan penentuan lulus tidaknya siswa.

3. Untuk menempatkan siswa dalam situasi belajar-mengajar yang tepat, sesuai dengan tingkat pencapaian dan kemampuan serta karakteristik siswa.
4. Untuk mengenal latar belakang kegiatan belajar-mengajar dan kelainan tingkah laku siswa.

Sebenarnya guru dalam melaksanakan proses penilaian tidak hanya mencakup penilaiain kognitif saja, namun idealnya guru juga dapat melakukan penilaian pada aspek afektif (sikap). Hasil belajar menurut Bloom (1976) mencakup prestasi belajar, kecepatan belajar, dan hasil afektif. Aderson (1981) berpendapat bahwa karakteristk manusia meliputi cara yang tipikal dari berpikir, berbuat, dan perasaan. Tipikal perasaan berkaitan dengan ranah afektif. Penilaian afektif dilakukan oleh pendidik melalui pengamatan terhadap perkembangan afeksi peserta didik. Komponen penilaian afektif

seperti yang tercantum dalam Standar Kompetensi Lulusan (Permendiknas Nomor 23 Tahun 2006 tentang Standar Kompetensi Lulusan); meliputi:

1. Memiliki keimanan dan ketaqwaan terhadap Tuhan Yang Maha Esa sesuai ajaran agama masing-masing yang tercermin dalam perilaku sehari-hari,
2. Menunjukkan sikap percaya diri dan bertanggung jawab atas perilaku, perbuatan, dan pekerjaannya,
3. Menunjukkan sikap kompetitif dan sportif untuk mendapatkan hasil yang terbaik dalam bidang pendidikan jasmani, olah raga, dan kesehatan,
4. Menganalisis sikap positif terhadap penegakan hukum, peradilan nasional, dan tindakan anti korupsi,
5. Mengevaluasi sikap berpolitik dan bermasyarakat madani sesuai dengan Pancasila dan Undang-Undang Dasar 1945,

sikap cermat dan menghargai hak atas kekayaan intelektual,

6. Menunjukkan sikap toleran dan empati terhadap keberagaman budaya yang ada di masyarakat setempat dalam kaitannya dengan budaya nasional,
7. Menunjukkan sikap peduli terhadap bahasa dan dialek, dan
8. Menunjukkan sikap kompetitif, sportif, dan etos kerja untuk mendapatkan hasil yang terbaik dalam bidang iptek.

Pengukuran ranah afektif dilakukan melalui metode observasi dan metode laporan diri. Penggunaan metode observasi berdasarkan pada asumsi bahwa karakteristik afektif dapat dilihat dari perilaku atau perbuatan yang ditampilkan dan atau reaksi psikologi. Mungkin pada KTSP, penilaian afektif belum terlalu diperhatikan, namun seiring dengan dikembangkannya pendidikan karakter bangsa, penilaian afektif menjadi

lebih penting dan harus dilakukan guru agar dapat diketahui keberhasilan pembelajaran yang dapat diwujudkan melalui internalisasi sikap yang ditunjukkan oleh peserta didik setelah mengikuti proses pembelajaran.

C. Macam-Macam Cara Penilaian Ranah Afektif

Menurut Andersen (1981), ada dua metode yang dapat digunakan untuk mengukur ranah afektif, yaitu metode observasi dan metode laporan diri. Penggunaan metode observasi berdasarkan pada asumsi bahwa karakteristik afektif dapat dilihat dari perilaku atau perbuatan yang ditampilkan dan/atau reaksi psikologi. Metode laporan diri berasumsi bahwa yang mengetahui keadaan afektif seseorang adalah dirinya sendiri. Namun hal ini menuntut kejujuran dalam mengungkap karakteristik afektif diri sendiri. Menurut Lewin (dalam Andersen, 1981), perilaku seseorang merupakan

fungsi dari watak (kognitif, afektif, dan psikomotor) dan karakteristik lingkungan saat perilaku atau perbuatan ditampilkan. Jadi tindakan atau perbuatan seseorang ditentukan oleh watak dirinya dan kondisi lingkungan.

Pengukuran ranah afektif juga dapat dilakukan dengan menggunakan skala pengukuran yang disebut skala sikap. Hasilnya berupa kategori sikap, yakni mendukung (positif), menolak (negatif), dan netral. Sikap pada hakikatnya adalah kecenderungan berperilaku pada seseorang. Ada tiga komponen sikap, yakni kognisi, afeksi, dan konasi. Kognisi berkenaan dengan pengetahuan seseorang tentang objek yang dihadapinya. Afeksi berkenaan dengan perasaan dalam menanggapi objek tersebut, sedangkan konasi berkenaan dengan kecenderungan berbuat terhadap objek tersebut. Oleh sebab itu, sikap selalu bermakna bila dihadapkan kepada objek tertentu. Skala sikap dinyatakan dalam bentuk pernyataan untuk

dinilai oleh responden, apakah pernyataan itu didukung atau ditolaknya, melalui rentangan nilai tertentu. Oleh sebab itu, pernyataan yang diajukan dibagi ke dalam dua kategori, yakni pernyataan positif dan pernyataan negative. Menurut Widoyoko (2009), ada beberapa bentuk skala sikap antara lain: Skala Likert, skala Trustone, skala Guttman, dan Semantic Differensial.

Berikut ini adalah berbagai teknik dan instrumen yang dapat digunakan untuk mengukur ranah afektif, yaitu:

1. Skala Sikap Likert

Sikap pada hakekatnya adalah kecenderungan berperilaku pada seseorang. Skala sikap Likert merupakan teknik pengukuran yang sederhana dan paling sering dijumpai dalam pengukuran ranah afektif, khususnya untuk sikap. Skala Likert menyajikan pernyataan yang harus

ditanggapi dengan memilih satu di antara beberapa alternatif.

2. Observasi

Pengamatan atau observasi adalah suatu teknik yang dilakukan dengan cara mengadakan pengamatan secara teliti serta pencatatan secara sistematis (Sulistiyorini, 2009: 85). Dengan kata lain observasi dapat mengukur atau menilai hasil dan proses belajar misalnya tingkah laku siswa pada waktu belajar, kegiatan diskusi siswa, partisipasi. Melalui pengamatan dapat diketahui bagaimana sikap dan perilaku siswa, kegiatan yang dilakukannya, partisipasi dalam kegiatan, proses kegiatan yang dilakukannya, kemampuan bahkan hasil yang diperoleh dari kegiatannya. Observasi harus dilakukan pada saat proses kegiatan itu berlangsung.

Pengamat terlebih dahulu harus menetapkan aspek-aspek tingkah laku apa yang hendak

diobservasinya. Lalu dibuat pedoman agar memudahkan dalam pengisian observasi. Pengisian hasil observasi dalam pedoman yang dibuat sebenarnya bisa diisi secara bebas (Arifin, 2009: 153).

3. Anecdotal Record

Anecdotal Record adalah catatan seketika yang berisi peristiwa atau kenyataan yang spesifik dan menarik mengenai sesuatu yang diamati atau terlihat secara kebetulan. Catatan tersebut bisa terjadi saat di luar kelas ataupun di dalam kelas. Tujuan pemberian catatan tersebut adalah untuk pembinaan peserta didik lebih lanjut (Sukardi, 2008: 176).

4. Kuesioner

Kuesioner adalah alat pengumpulan data secara tertulis yang berisi daftar pertanyaan (questions) atau pernyataan (statement) yang disusun secara khusus dan digunakan untuk menggali dan menghimpun keterangan

dan/atau informasi sebagaimana dibutuhkan dan cocok untuk dianalisis (Sudjana, 2008: 177).

Pemikiran atau perilaku harus memiliki dua kriteria untuk diklasifikasikan sebagai ranah afektif (Andersen, 1981: 4). Pertama, perilaku melibatkan perasaan dan emosi seseorang. Kedua, perilaku harus tipikal perilaku seseorang. Kriteria lain yang termasuk ranah afektif adalah intensitas, arah, dan target. Intensitas menyatakan derajat atau kekuatan dari perasaan. Beberapa perasaan lebih kuat dari yang lain, misalnya cinta lebih kuat dari senang atau suka. Sebagian orang kemungkinan memiliki perasaan yang lebih kuat dibanding yang lain. Arah perasaan berkaitan dengan orientasi positif atau negatif dari perasaan yang menunjukkan apakah perasaan itu baik atau buruk. Misalnya senang pada pelajaran dimaknai positif, sedang kecemasan dimaknai negatif. Bila intensitas dan arah perasaan ditinjau bersama-sama, maka

karakteristik afektif berada dalam suatu skala yang kontinum. Target mengacu pada objek, aktivitas, atau ide sebagai arah dari perasaan. Bila kecemasan merupakan karakteristik afektif yang ditinjau, ada beberapa kemungkinan target. Peserta didik mungkin bereaksi terhadap sekolah, matematika, situasi sosial, atau pembelajaran. Tiap unsur ini bisa merupakan target dari kecemasan. Kadang-kadang target ini diketahui oleh seseorang namun kadang-kadang tidak diketahui. Seringkali peserta didik merasa cemas bila menghadapi tes di kelas. Peserta didik tersebut cenderung sadar bahwa target kecemasannya adalah tes.

Objek ranah afektif menurut Krathwohl (1973: 24) unsur-unsurnya terdiri dari minat (*interest*), sikap (*attitude*), nilai (*value*), apresiasi (*apresiasi*), dan penyesuaian (*adjustment*). Fishbein dan Ajzen (1975) membagi dalam kepercayaan (*belief*), sikap (*attitude*), keinginan/maksud (*intention*), dan perilaku

(*behaviour*). Berbeda dengan Fishbein dan Ajzen, Hammond (Worthen dan Sanders, 1973) menyatakan bahwa objek afektif meliputi unsur perhatian, minat (*interest*), sikap (*attitude*), perasaan (*feeling*), dan emosi (*emotion*). Menurut Hopkins dan Antes (1990), unsur-unsur ranah afektif meliputi emotion, interest, attitude, value, character development, dan motivation. Mardapi (2011: 183) menambahkan bahwa karakter juga merupakan bagian dari ranah afektif. Berdasarkan uraian di atas, dapat diidentifikasi bahwa unsur-unsur ranah afektif paling tidak meliputi: perhatian/minat, sikap, nilai, apresiasi, karakter, kepercayaan, perasaan, emosi perilaku, keinginan, dan penyesuaian.

D. Instrumen Penilaian Ranah Afektif

Krathwohl (Zubaedi, 2011: 26-28) membagi level ranah afektif ke dalam lima taksonomi Krathwohl yaitu receiving, responding, valuing,

organization, dan characterization. Kelima taksonomi tersebut akan dijelaskan sebagai berikut:

1. Penerimaan (*Receiving*)

Kawasan penerimaan dijabarkan ke dalam tiga tahap yang meliputi: 1) Kesiapan untuk menerima (*awareness*), yaitu munculnya kesiapan untuk berinteraksi dengan stimulus (fenomena atau objek yang dipelajari). Ditandai dengan kehadiran dan usaha untuk memberi perhatian pada stimulus. 2) kemauan untuk menerima (*willingness to receive*), yaitu usaha untuk memusatkan perhatian pada stimulus. 3) mengkhususkan perhatian (*controlled or selected attention*), mengkhususkan perhatian pada fokus tertentu misalnya warna, suara atau kata-kata saja.

Pada tingkat *receiving*, peserta didik memiliki persepsi terhadap suatu fenomena khusus atau stimulus, yang menarik perhatiannya.

Tugas pendidik menjaga perhatian peserta didik pada fenomena yang menjadi objek pembelajaran afektif. Misalnya pendidik mengarahkan peserta didik agar senang membaca buku, senang bekerjasama, dan sebagainya. Kata kerja operasional yang dapat digunakan pada perumusan tujuan adalah menghadiri, melihat, memperhatikan.

2. Sambutan (*Responding*)

Memberikan respon terhadap stimulus yang meliputi proses sebagai berikut: 1) kesiapan menanggapi (*acquiescence of responding*). Contohnya: mengajukan pertanyaan, menaati peraturan lalu lintas dll,. 2) kemauan menanggapi (*willingness to respond*), yaitu usaha untuk melihat hal-hal khusus dalam bagian yang diperhatikan. 3) kepuasan menanggapi (*satisfaction in response*), yaitu adanya kegiatan yang berhubungan dengan usaha untuk memuaskan keinginan untuk

mengetahui. Contoh: bertanya, membuat gambar, memotret dll.

Pada tingkat ini peserta didik tidak saja memperhatikan fenomena khusus tetapi ia juga bereaksi. Hasil pembelajaran pada ranah ini menekankan pada pemerolehan respons, berkeinginan memberi respons, atau kepuasan dalam memberi respons. Pada tingkat ini siswa tidak hanya memberi respon tetapi ia sungguh-sungguh berpartisipasi aktif. Misalnya senang membaca buku, senang bertanya, senang membantu teman, senang dengan kebersihan dan kerapian, dan sebagainya. Kata kerja operasional yang dapat digunakan adalah mengikuti, mendiskusikan, berlatih, berpartisipasi, menjawab pertanyaan.

3. Penilaian (*Valuing*)

Pada tahap ini, mulai timbul proses internalisasi untuk memiliki dan menghayati nilai dari stimulus yang dihadapi. Penilaian

ini dibagi menjadi empat tahap yang meliputi:

1) menerima nilai (acceptance of value), merupakan kelanjutan dari kepuasan menanggapi yang lebih intensif. 2) menyeleksi nilai yang lebih disenangi (preference for a value) yang ditunjukkan dengan usaha untuk mencari sesuatu yang dapat memuaskan. 3) komitmen yaitu keyakinan terhadap suatu nilai dengan alasan tertentu yang muncul setelah melalui pengalaman-pengalaman.

Komitmen ditunjukkan dengan rasa senang, kagum, terpesona. Misalkan kagum atas keberanian seseorang, menunjukkan komitmen terhadap nilai keberanian yang dihargainya.

Valuing berhubungan dengan pengungkapan perasaan, keyakinan, atau anggapan bahwa suatu gagasan, benda, atau cara berpikir tertentu mempunyai nilai. Unsur yang penting pada jenjang ini adalah seseorang telah termotivasi bukan karena keinginan

atau kepatuhan tetapi lebih disebabkan karena keterkaitannya dengan nilai-nilai tertentu. Dalam tujuan pembelajaran kata kerja operasional yang dapat digunakan adalah meyakinkan, bertindak, mengemukakan argumentasi.

4. Pengorganisasian (*Organization*)

Pada tahap ini siswa tidak hanya menginternalisasikan suatu nilai tertentu namun mulai melihat beberapa nilai yang relevan untuk disusun menjadi satu sistem nilai. Proses ini terjadi dalam dua tahapan yaitu: 1) konseptualisasi nilai, yaitu keinginan untuk menilai hasil karya orang lain. 2) pengorganisasian sistem nilai, yaitu menyusun nilai dalam suatu sistem nilai dengan urutan tingkatan sesuai dengan urutan kepentingan atau kesenangan pada masing-masing siswa.

Pada tingkat *organization*, nilai satu dengan nilai lain dikaitkan, konflik antar nilai

diselesaikan, dan mulai membangun sistem nilai internal yang konsisten. Hasil pembelajaran pada tingkat ini berupa konseptualisasi nilai atau organisasi sistem nilai. Kata kerja operasional yang dapat digunakan adalah memodifikasi, membandingkan, memutuskan.

5. Karakterisasi (*Characterization*)

Karakterisasi yaitu kemampuan untuk menghayati atau menginternalisasikan sistem nilai secara konsisten. Tahap ini terdiri dari dua tahap yaitu: 1) generalisasi, yaitu kemampuan untuk melihat suatu masalah dari sudut pandang tertentu. 2) karakterisasi, yaitu mengembangkan pandangan hidup tertentu yang membentuk kepribadian. Dari pemaparan di atas, dapat disimpulkan bahwa ranah afektif memiliki lima jenjang hasil belajar. Lima jenjang tersebut dimulai dari tingkat paling sederhana hingga yang paling kompleks.

Kelima jenjang tersebut antara lain receiving atau penerimaan, responding atau sambutan, valuing atau penilaian, organization atau pengorganisasian, dan *characterization* atau karakteristik nilai.

Pada tingkat ini peserta didik memiliki system nilai yang mengendalikan perilaku sampai pada waktu tertentu hingga terbentuk gaya hidup. Hasil pembelajaran pada tingkat ini berkaitan dengan pribadi, emosi, dan sosial.

E. Langkah-Langkah Menyusun Instrumen Ranah Afektif

Ada sebelas langkah yang harus diikuti dalam mengembangkan instrumen penilaian afektif, 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.

1. Menentukan spesifikasi instrumen

Spesifikasi instrumen terdiri atas tujuan dan kisi-kisi instrumen. Dalam bidang pendidikan pada dasarnya pengukuran afektif ditinjau dari tujuannya, contohnya instrumen sikap.

2. Menulis instrumen

Pertanyaan tentang sikap meminta responden menunjukkan perasaan yang positif atau negatif terhadap suatu objek, atau suatu kebijakan. Kata-kata yang sering digunakan pada pertanyaan sikap menyatakan arah perasaan seseorang; menerima-menolak, menyenangkan-tidak menyenangkan, baik-buruk, diinginkan-tidak diinginkan.

Contoh indikator sikap terhadap mata pelajaran matematika misalnya:

- a. Membaca buku matematika
- b. Mempelajari matematika

- c. Melakukan interaksi dengan guru matematika
- d. Mengerjakan tugas matematika
- e. Melakukan diskusi tentang matematika
- f. Memiliki buku matematika

Contoh pernyataan untuk kuesioner:

- a. Saya senang membaca buku matematika
 - b. Tidak semua orang harus belajar matematika
 - c. Saya jarang bertanya pada guru tentang pelajaran matematika
 - d. Saya tidak senang pada tugas pelajaran matematika
 - e. Saya berusaha mengerjakan soal-soal matematika sebaik-baiknya
 - f. Memiliki buku matematika penting untuk semua peserta didik
 - g. Menentukan skala instrumen
3. Menentukan skala instrumen

Secara garis besar skala instrumen yang sering digunakan dalam penelitian, yaitu

Skala Thurstone, Skala Likert, dan Skala Beda semantik.

Contoh Skala Likert: Sikap terhadap pelajaran Matematika

Pelajaran matematika bermanfaat	SS	S	TS	STS
Pelajaran matematika sulit	SS	S	TS	STS
Tidak semua harus belajar matematika	SS	S	TS	STS
Pelajaran matematika harus dibuat mudah	SS	S	TS	STS
Sekolah saya menyenangkan	SS	S	TS	STS

Keterangan:

SS : Sangat Setuju

S : Setuju

TS : Tidak Setuju

STS : Sangat Tidak Setuju

4. Menentukan sistem penskoran

Sistem penskoran yang digunakan tergantung pada skala pengukuran. Apabila digunakan skala Thurstone, maka skor tertinggi untuk

tiap butir 7 dan skor terendah 1. Demikian pula untuk instrumen dengan skala beda semantik, tertinggi 7 terendah 1. Untuk skala Likert, pada awalnya skor tertinggi tiap butir 5 dan terendah 1. Dalam pengukuran sering terjadi kecenderungan responden memilih jawaban pada kategori tiga 3 (tiga) untuk skala Likert. Untuk menghindari hal tersebut skala Likert dimodifikasi dengan hanya menggunakan 4 (empat) pilihan, agar jelas sikap atau minat responden.

5. Menelaah instrumen

Kegiatan pada telaah instrumen adalah menelaah apakah: a) butir pertanyaan/pernyataan sesuai dengan indikator, b) bahasa yang digunakan komunikatif dan menggunakan tata bahasa yang benar, c) butir pertanyaan/pernyataan tidak bias, d) format instrumen menarik untuk dibaca, e) pedoman menjawab atau mengisi instrumen jelas, dan f) jumlah butir

dan/atau panjang kalimat pertanyaan/ Pernyataan sudah tepat sehingga tidak menjemukan untuk dibaca/dijawab.

Telaah dilakukan oleh pakar dalam bidang yang diukur dan akan lebih baik bila ada pakar penilaian. Telaah bisa juga dilakukan oleh teman sejawat bila yang diinginkan adalah masukan tentang bahasa dan format instrumen. Bahasa yang digunakan adalah yang sesuai dengan tingkat pendidikan responden. Hasil telaah selanjutnya digunakan untuk memperbaiki instrumen.

6. Merakit instrumen

Setelah instrumen diperbaiki selanjutnya instrumen dirakit, yaitu menentukan format tata letak instrumen dan urutan pertanyaan/ pernyataan. Format instrumen harus dibuat menarik dan tidak terlalu panjang, sehingga responden tertarik untuk membaca dan mengisinya. Setiap sepuluh pertanyaan sebaiknya dipisahkan dengan cara memberi

spasi yang lebih, atau diberi batasan garis empat persegi panjang. Urutkan pertanyaan/ Pernyataan sesuai dengan tingkat kemudahan dalam menjawab atau mengisinya.

7. Melakukan uji coba

Setelah dirakit instrumen diujicobakan kepada responden, sesuai dengan tujuan penilaian apakah kepada peserta didik, kepada guru atau orang tua peserta didik. Untuk itu dipilih sampel yang karakteristiknya mewakili populasi yang ingin dinilai. Bila yang ingin dinilai adalah peserta didik SMA, maka sampelnya juga peserta didik SMA. Sampel yang diperlukan minimal 30 peserta didik, bisa berasal dari satu sekolah atau lebih.

8. Menganalisis instrumen

Analisis hasil ujicoba meliputi variasi jawaban tiap butir pertanyaan/ Pernyataan. Jika menggunakan skala instrumen 1 sampai

7, dan jawaban responden bervariasi dari 1 sampai 7, maka butir pertanyaan/ Pernyataan pada instrumen ini dapat dikatakan baik. Namun apabila jawabannya hanya pada satu pilihan jawaban saja, misalnya pada pilihan nomor 3, maka butir instrumen ini tergolong tidak baik. Indikator yang digunakan adalah besarnya daya beda dan indeks keandalan yang dikenal dengan indeks reliabilitas.

9. Memperbaiki instrumen

Perbaikan dilakukan terhadap butir-butir pertanyaan/ Pernyataan yang tidak baik, berdasarkan analisis hasil ujicoba. Bisa saja hasil telaah instrumen baik, namun hasil ujicoba empirik tidak baik. Untuk itu butir pertanyaan/ Pernyataan instrumen harus diperbaiki. Perbaikan termasuk mengakomodasi saran-saran dari responden ujicoba. Instrumen sebaiknya dilengkapi dengan pertanyaan terbuka.

10. Melaksanakan pengukuran

Pelaksanaan pengukuran perlu memperhatikan waktu dan ruangan yang digunakan. Waktu pelaksanaan bukan pada waktu responden sudah lelah. Ruang untuk mengisi instrumen harus memiliki cahaya (penerangan) yang cukup dan sirkulasi udara yang baik. Tempat duduk juga diatur agar responden tidak terganggu satu sama lain. Diusahakan agar responden tidak saling bertanya pada responden yang lain agar jawaban kuesioner tidak sama atau homogen. Pengisian instrumen dimulai dengan penjelasan tentang tujuan pengisian, manfaat bagi responden, dan pedoman pengisian instrumen.

11. Menafsirkan hasil pengukuran

Hasil pengukuran berupa skor atau angka. Untuk menafsirkan hasil pengukuran diperlukan suatu kriteria. Kriteria yang

digunakan tergantung pada skala dan jumlah butir pertanyaan/pernyataan yang digunakan.

F. Karakteristik Ranah Afektif

Djemari Mardapi (2011: 194) membedakan ranah afektif menjadi empat tipe karakteristik yang meliputi:

1. Sikap

Merupakan respon yang diberikan siswa terhadap suatu objek, situasi, konsep atau orang yang dipelajari baik berupa respon positif ataupun negatif.

2. Minat

Merupakan kemauan seseorang untuk memperoleh objek khusus, aktivitas, pemahaman, dan keterampilan untuk tujuan tertentu.

3. Nilai

Merupakan keyakinan yang dalam tentang perbuatan, tindakan, atau perilaku yang dianggap baik dan yang dianggap jelek.

4. Konsep diri

Merupakan evaluasi yang dilakukan siswa terhadap kemampuan dan kelemahan yang dialaminya. Berdasarkan uraian di atas, dapat disimpulkan bahwa karakteristik afektif mencakup empat karakteristik yaitu sikap, minat, nilai dan konsep diri.

Ada 5 (lima) tipe karakteristik afektif yang penting, yaitu sikap, minat, konsep diri, nilai dan moral (Depdiknas, 2008:4).

1. Sikap

Sikap yaitu suatu kecenderungan untuk bertindak secara suka atau tidak suka terhadap suatu obyek. Sikap dapat dibentuk melalui cara mengamati dan menirukan sesuatu yang positif, kemudian melalui penguatan menerima informasi verbal. Perubahan sikap dapat diamati dalam proses pembelajaran. Penilaian sikap adalah penilaian yang dilakukan untuk mengetahui sikap peserta didik terhadap mata pelajaran,

kondisi pembelajaran, pendidik dan sebagainya. Menurut Fishbein dan Ajzen (dalam Depdiknas, 2008:4) sikap adalah suatu predisposisi yang dipelajari untuk merespon secara positif dan negatif terhadap suatu obyek, situasi, konsep, atau orang. Sikap peserta didik terhadap obyek misalnya sikap terhadap sekolah atau terhadap mata pelajaran. Sikap peserta didik ini penting untuk ditingkatkan.

2. Minat

Minat adalah kecenderungan yang tetap untuk memperhatikan dan mengenang beberapa kegiatan. Kegiatan yang diminati seseorang, diperhatikan terus menerus yang disertai dengan rasa senang. Minat sangat besar pengaruhnya terhadap belajar, karena pelajaran yang dipelajari tidak sesuai dengan minat siswa, siswa tidak akan belajar dengan sebaik-baiknya, karena tidak ada daya tarik baginya (Slameto, 2010:57). Hal penting pada

minat adalah intensitasnya. Secara umum minat termasuk karakteristik afektif yang memiliki intensitas tinggi (Depdiknas, 2008:4) siswa, siswa tidak akan belajar dengan sebaik-baiknya, karena tidak ada daya tarik baginya (Slameto, 2010:57). Hal penting pada minat adalah intensitasnya. Secara umum minat termasuk karakteristik afektif yang memiliki intensitas tinggi (Depdiknas, 2008:4).

3. Konsep Diri

Menurut Smith, konsep diri adalah evaluasi yang dilakukan individu terhadap kemampuan dan kelemahan yang dimiliki. Target, arah, dan intensitas konsep diri pada dasarnya seperti ranah afektif yang lain. Target konsep diri biasanya orang tapi bisa juga institusi seperti sekolah. Arah konsep diri bisa positif atau negatif, dan intensitasnya bisa dinyatakan dalam suatu daerah kontinum, yaitu mulai dari rendah

sampai tinggi. Konsep diri ini penting untuk menentukan jenjang karir peserta didik, yaitu dengan mengetahui kekuatan dan kelemahan diri sendiri, dapat dipilih alternatif karir yang tepat bagi peserta didik. Selain itu informasi konsep diri penting bagi sekolah untuk memberikan motivasi belajar peserta didik dengan tepat (Mardapi, 2004:104)

4. Nilai

Nilai menurut Rokeach (dalam Depdiknas, 2008:5) merupakan suatu keyakinan tentang perbuatan, tindakan, atau perilaku yang dianggap baik dan yang dianggap buruk. Selanjutnya dijelaskan bahwa sikap mengacu pada suatu organisasi sejumlah keyakinan sekitar obyek spesifik atau situasi, sedangkan nilai mengacu pada keyakinan. Nilai cenderung menjadi ide, target nilai dapat pula sesuatu seperti sikap dan perilaku.

5. Moral

Moral berkenaan dengan perasaan salah atau benar terhadap kebahagiaan orang lain atau perasaan terhadap tindakan yang dilakukan diri sendiri. Moral juga sering dikaitkan dengan keyakinan agama seseorang, yaitu keyakinan akan perbuatan yang berdosa dan berpahala. Jadi moral berkaitan dengan prinsip, nilai dan keyakinan seseorang (Depdiknas, 2008:6).

Thomas Lickona (2012: 74 - 76) menyebutkan nilai pokok yang harus diajarkan di sekolah meliputi sikap hormat dan tanggung jawab. Selain itu, bentuk nilai lain yang sebaiknya diajarkan di sekolah adalah kejujuran, keadilan, toleransi, kebijaksanaan, disiplin diri, tolong menolong, peduli sesama, kerja sama, keberanian, dan sikap demokratis. Selanjutnya Paul Suparno, dkk., (2002: 94) menyebutkan penanaman nilai pada jenjang Sekolah Dasar yang meliputi:

1. Religiositas

Contoh sikap : Mampu berterima kasih dan bersyukur, menghormati dan mencintai Tuhan yang diwujudkan dalam doa.

2. Hidup bersama orang lain

Contoh sikap : Mampu bertoleransi dalam setiap kegiatan di masyarakat, menghindari tindakan mau menang sendiri, memperbaiki diri lewat sarankritik dari orang lain.

3. Gender Penghargaan terhadap perempuan

Contoh sikap : Bertindak dan bersikap positif terhadap perempuan. Selalu menghindari sikap yang meremehkan perempuan. Menunjukkan apresiasi terhadap tamu perempuan, guru, atau teman.

4. Keadilan Menghindarkan diri dari sikap memihak

Contoh sikap : Mempunyai penghargaan kepada hak-hak orang lain dan mengedepankan kewajiban diri. Tidak ingin menang sendiri.

5. Demokrasi Menghargai usaha dan pendapat orang lain

Contoh sikap : Tidak menganggap diri yang paling benar dalam setiap perbincangan. Memandang positif sikap orang lain dan menghindarkan berburuk sangka. Bisa menerima perbedaan pendapat.

6. Kejujuran

Contoh sikap : Menghindari sikap bohong, mengakui kelebihan orang lain. Mengakui kekurangan, kesalahan, atau keterbatasan diri sendiri. memilih cara-cara terpuji dalam menempuh ujian, tugas, atau kegiatan.

7. Kemandirian

Contoh sikap : Mampu berinisiatif, bertanggung jawab pada diri sendiri secara konsekuen. Tidak tergantung pada orang lain. Terbebaskan dari pengaruh ucapan atau perbuatan orang lain.

8. Daya juang Gigih dan percaya diri dalam mengerjakan setiap hal

Contoh sikap : Menghindari tindakan sia-sia baik dalam belajar maupun kegiatan. Optimal mewujudkan keinginannya dan tidak mudah putus asa. Tidak menampakan sikap malas.

9. Tanggung jawab Mengerjakan tugas-tugas dengan semestinya

Contoh sikap : Menghindarkan diri dari sikap menyalahkan orang lain atau pihak lain. Tidak melemparkan persoalan kepada orang lain. Emmahami dan menerima resiko atau akibat suatu tindakan baik terhadap diri sendiri ataupun orang lain.

10. Penghargaan terhadap alam Menjaga kebersihan kelas dan lingkungan sekolah

Contoh sikap : Menghindarkan diri dari tindakan corat-coret meja atau dinding kelas. memperhatikan sampah-sampah dan tanamantanaman di sekitarnya.

Dari pemaparan tersebut dapat disimpulkan bahwa nilai-nilai yang termasuk ke dalam ranah afektif pada penelitian ini meliputi sikap hormat,

tanggung jawab, kejujuran, keadilan, toleransi, kebijaksanaan, disiplin diri, tolong menolong, peduli sesama, kerja sama, keberanian, dan sikap demokratis. Kemudian dirumuskan pula sepuluh nilai pokok yang meliputi religiositas, hidup bersama orang lain, gender, keadilan, demokrasi, kejujuran, kemandirian, daya juang, tanggung jawab, dan 31 penghargaan terhadap alam.

**CONTOH LEMBAR OBSERVASI SIKAP
RELIGIUS TERHADAP SISWA SEKOLAH
DASAR**

Nama Siswa :

Kelas :

Tanggal Pengamatan :

Petunjuk :

Berilah tanda cek (✓) pada kolom skor sesuai sikap spiritual yang ditampilkan oleh peserta didik, dengan kriteria sebagai berikut :

4= selalu, apabila siswa selalu melakukan sesuai pernyataan

3= sering, apabila siswa sering melakukan sesuai pernyataan

2= kadang-kadang, apabila siswa kadang-kadang melakukan

1= tidak pernah, apabila siswa tidak pernah melakukan

No.	Indikator Pengamatan	Skor			
		1	2	3	4
1.	Melaksanakan rutinitas kegiatan ibadah sesuai agama yang dianutnya.				
2.	Mengikuti kegiatan keagamaan yang diselenggarakan sekolah sesuai ajaran agama yang dianutnya.				
3.	Menjalankan ibadah sesuai waktunya/tepat waktu				

Tabel 8. Contoh Angket Observasi Sikap Religius

**CONTOH LEMBAR OBSERVASI SIKAP
TANGGUNG JAWAB TERHADAP SISWA
SEKOLAH DASAR**

Nama Siswa :

Kelas :

Tanggal Pengamatan :

Petunjuk :

Berilah tanda cek (✓) pada kolom skor sesuai sikap spiritual yang ditampilkan oleh peserta didik, dengan kriteria sebagai berikut :

4 = selalu, apabila siswa selalu melakukan sesuai pernyataan

3 = sering, apabila siswa sering melakukan sesuai pernyataan

2 = kadang-kadang, apabila siswa kadang-kadang melakukan

1 = tidak pernah, apabila siswa tidak pernah melakukan

No.	Indikator Pengamatan	Skor			
		1	2	3	4
1.	Membuang sampah pada tempatnya.				
2.	Gemar melakukan kegiatan kemanusiaan.				
3.	Menggunakan dan memelihara fasilitas di sekolah baik sarana maupun prasarana.				

Tabel 9. Contoh Angket Obesrvasi Sikap Tanggungjawab



BAB 9

PENGEMBANGAN INSTRUMEN PENILAIAN UNJUK KERJA

Pada tes bentuk perbuatan (unjuk kerja), umumnya dilakukan dengan cara menyuruh peserta tes untuk melakukan sesuatu pekerjaan yang bersifat fisik (praktik). Tes bentuk perbuatan ini sangat cocok untuk melakukan penilaian dalam pelajaran praktik/keterampilan atau praktikum di laboratorium. Alat yang digunakan untuk melakukan penilaian pada umumnya berupa lembar pengamatan (lembar observasi). Tes bentuk perbuatan ini pada umumnya dapat digunakan untuk menilai proses maupun hasil (produk) dari suatu kegiatan praktik.

Mengukur dimaksudkan memberi bentuk kuantitatif dari suatu kegiatan atau kemampuan yang dimiliki, yaitu dalam bentuk angka. Pada pengukuran unjuk kerja yang digunakan adalah lembar pengamatan. Pengukuran unjuk kerja dipergunakan untuk mencocokkan kesesuaian antara pengetahuan mengenai teori dan keterampilan di dalam praktek sehingga hasil evaluasinya menjadi lebih jelas. Penilaian penguasaan kompetensi aspek keterampilan atau psikomotor yang dimiliki oleh

seseorang atau peserta didik, hanya ada satu bentuk tes yang tepat yaitu tes perbuatan (*performance assessment*), artinya orang yang akan dinilai kemampuan skillnya harus menampilkan atau melakukan skill yang dimilikinya di bawah persyaratan-persyaratan kerja yang berlaku.

A. Pengertian Penilaian Unjuk Kerja

Asesmen kinerja adalah suatu prosedur yang menggunakan berbagai bentuk tugas-tugas untuk memperoleh informasi tentang apa dan sejauhmana yang telah dipelajari siswa. Asesmen kinerja mensyaratkan siswa dalam menyelesaikan tugas-tugas kinerjanya menggunakan pengetahuan dan ketrampilannya yang diwujudkan dalam bentuk perbuatan, tindakan atau unjuk kerja. Tes Unjuk Kerja meminta siswa mewujudkan tugas sebenarnya yang mewakili keseluruhan kinerja yang akan dinilai, seperti mempersiapkan alat, menggunakan alat/merangkai alat, menuliskan data, menganalisis data, menyimpulkan,

menyusun laporan dan sebagainya. Secara khusus penilaian kinerja menjelaskan kemampuan-kemampuan siswa, pemahaman konseptual, kemampuan untuk menerapkan pengetahuan dan ketrampilan, kemampuan melaksanakan kinerja dan kemampuan melakukan suatu proses. Penilaian unjuk kerja 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 dilaboratorium, praktik sholat, presentasi, diskusi, bermain peran, memainkan alat musik, bernyanyi, membaca puisi/deklamasi, mengcangkok, berpidato, dan lain-lain.

Menurut Trespecies (Depdiknas 2003), *Performance Assessment* adalah berbagai macam tugas dan situasi dimana peserta tes diminta untuk mendemonstrasikan pemahaman dan mengaplikasikan pengetahuan yang mendalam,

serta keterampilan di dalam berbagai macam konteks sesuai dengan kriteria yang diinginkan. Berdasar pendapat Mardjuki (1988), orang yang dinilai kemampuan skillnya harus menampilkan atau melakukan skill yang dimiliki dibawah persyaratan-persyaratan kerja yang berlaku. Menurut pendapat Zainal (1990) tes unjuk kerja adalah bentuk tes yang menuntut jawaban peserta didik dalam bentuk perilaku, tindakan atau perbuatan. Peserta didik bertindak sesuai dengan apa yang diperintahkan atau ditanyakan. Jadi *Performance Assessment* adalah suatu penilaian yang meminta peserta tes untuk mendemonstrasikan dan mengaplikasikan pengetahuan unjuk kerja ke dalam berbagai macam konteks sesuai dengan yang diinginkan. Berk (1986) menyatakan bahwa asesmen unjuk kerja adalah proses mengumpulkan data dengan cara pengamatan yang sistematis untuk membuat keputusan tentang individu.

Dalam konteks pembelajaran, penilaian unjuk kerja (*performance assessment*) menurut Berk (1986: 9) merupakan salah satu teknik penilaian yang dalam proses pengumpulan data untuk membuat keputusan tentang individu dilakukan dengan cara observasi sistematis. Melengkapi definisi tersebut, ada lima definisi operasional *performance assessment*, seperti:

1. *Performance assessment* adalah proses, bukan tes atau perangkat pengukuran tunggal.
2. Fokus dari proses ini adalah pengumpulan data, menggunakan berbagai instrumen dan strategi.
3. Data dikumpulkan dengan cara observasi sistematis. Penekanannya adalah pada teknik observasi langsung bukan pada tes kertas-dan-pensil (*paper-and-pencil*) terutama bukan pilihan ganda meskipun tes tersebut juga dapat digunakan dalam penilaian.

4. Data yang terintegrasi digunakan untuk tujuan membuat keputusan tertentu yang akan memandu bentuk dan substansi penilaian.
5. Subjek dari pengambilan keputusan adalah individu, bukan program atau produk yang mencerminkan suatu kegiatan kelompok.

Melengkapi pendapat tersebut Danielson (dalam Puji Iryanti, 2004: 6) mendefinisikan bahwa, penilaian unjuk kerja (*performance assessment*) sebagai salah satu teknik penilaian meliputi semua penilaian dalam bentuk tulisan, produk, atau perilaku kecuali tes bentuk pilihan ganda, menjodohkan, benar salah, dan jawaban singkat.

Penilaian unjuk kerja (*performance assessment*) sebagai salah satu teknik penilaian jika dibandingkan dengan teknik penilaian *paper and pencil* memiliki banyak keunggulan. Teknik ini menurut Reynolds (2010: 248-249) sangat tepat

dan telah banyak diaplikasikan dalam berbagai konteks, seperti:

1. *Laboratory classes;*
2. *Mathematics classess;*
3. *English, foreign-language, debate classes;*
4. *Social studies classes;*
5. *Art classes;*
6. *Physical education classes; dan*
7. *Music classes.*

Selanjutnya ia mengungkapkan pula perbedaan antara penilaian unjuk kerja (*performance assessment*) dengan teknik penilaian tradisional yang berbasis *paper-and-pencil* seperti: (1) Penilaian unjuk kerja lebih mendekatkan pada refleksi kehidupan yang nyata dan lebih dekat kepada aturan dan penerapan pada dunia nyata daripada penilaian tradisional yang masih berbasis penilaian pensil dan tes; (2) Penilaian unjuk kerja melibatkan berbagai bentuk penilaian, dan (3) Penilaian unjuk kerja

melibatkan evaluasi yang subjektif dari unjuk kerja yang ditampilkan oleh peserta didik.

Menurut Kusnandar (2013) mengatakan bahwa "*performance assessment*" adalah berbagai macam tugas dan situasi dimana peserta tes diminta untuk mendemonstrasikan pemahaman dan pengaplikasian pengetahuan yang mendalam, serta keterampilan di dalam berbagai macam konteks

Jadi, dari beberapa pengertian diatas dapat disimpulkan bahwa *performance assessment* adalah suatu penilaian yang meminta peserta tes untuk mendemonstrasikan dan mengaplikasikan pengetahuan ke dalam berbagai macam konteks sesuai dengan kriteria yang diinginkan.

B. Karakteristik Penilaian Unjuk Kerja

Tes unjuk kerja dapat dilakukan secara kelompok dan juga dapat dilakukan secara individual. Dilakukan secara kelompok berarti guru menghadapi sekelompok *testee*, sedangkan

secara individual berarti seorang guru seorang *testee*. Tes unjuk kerja dapat digunakan untuk mengevaluasi mutu suatu pekerjaan yang telah selesai dikerjakan, keterampilan, kemampuan merencanakan sesuatu pekerjaan dan mengidentifikasi bagian-bagian sesuatu piranti mesin misalnya. Hal yang penting dalam penilaian unjuk kerja adalah cara mengamati dan menskor kemampuan kinerja peserta didik. Guna meminimumkan faktor subyektifitas keadilan dalam menilai kemampuan kinerja peserta didik, biasanya *rater* atau penilai jumlahnya lebih dari satu orang sehingga diharapkan hasil penilaian mereka menjadi lebih valid dan reliabel. Di samping itu, dalam pelaksanaan penilaian diperlukan suatu pedoman penilaian yang bertujuan untuk memudahkan penilai dalam menilai, sehingga tingkat subyektifitas bisa ditekan.

Penilaian unjuk kerja merupakan penilaian yang dilakukan dengan mengamati kegiatan

peserta didik dalam melakukan sesuatu. Penilaian unjuk kerja cocok digunakan untuk menilai ketercapaian kompetensi yang menuntut peserta didik melakukan tugas tertentu, seperti: praktek di laboratorium, praktek olah raga, presentasi, diskusi, bermain peran, memainkan alat musik, bernyanyi, membaca puisi/deklamasi, termasuk juga membuat busana. Cara penilaian ini dianggap lebih otentik daripada tes tertulis karena apa yang dinilai lebih mencerminkan kemampuan peserta didik yang sebenarnya. Tingkat penguasaan terhadap bagian-bagian yang sulit dari suatu pekerjaan. Unsur-unsur yang menjadi karakteristik inti dari suatu pekerjaan akan menjadi bagian dari suatu tes unjuk kerja.

Hal-hal yang perlu diperhatikan dalam pelaksanaan tes unjuk kerja adalah ketersediaan peralatan dan bahan-bahan lainnya yang diperlukan untuk tugas-tugas spesifik, kejelasan, dan kelengkapan instruksi. Secara garis besar

penilaian pembelajaran keterampilan pada dasarnya dapat dilakukan terhadap dua hal, yaitu : (1) proses pelaksanaan pekerjaan, yang mencakup : langkah kerja dan aspek personal; dan (2) produk atau hasil pekerjaan.

Penilaian terhadap aspek proses umumnya lebih sulit dibanding penilaian terhadap produk atau hasil kerja. Penilaian proses hanya dapat dilakukan dengan cara pengamatan (observasi), dan dilakukan seorang demi seorang. Penilaian proses pada umumnya cenderung lebih subyektif dibanding penilaian produk, karena tidak ada standar yang baku. Namun demikian, penilai dapat lebih meningkatkan obyektivitas penilaiannya dengan cara analisis tugas (*analisis skill*).

Sementara itu, penilaian produk pada umumnya lebih mudah dilakukan daripada penilaian proses, karena dapat dilakukan dengan menggunakan instrumen yang lebih valid dan reliabel, seperti alat-alat ukur mikrometer,

meteran dan sebagainya. Dalam penilaian produk, karakteristik yang digunakan sebagai standar biasanya adalah berhubungan dengan kemanfaatan, kesesuaian dengan tujuan, dimensi, nampak luar, tingkat penyimpangan, kekuatan dan sebagainya (Ahmad Jaedun, 2010).

Menurut Van Blerkom (2009: 148), penilaian unjuk kerja (*performance assessment*) terdapat tiga tipikal karakteristik yang dapat dikelompokkan berdasarkan dimensi, meliputi:

1. Menilai proses atau produk.
2. Menggunakan simulasi atau kejadian nyata (*real settings*).
3. Menggunakan peristiwa alami atau peristiwa dan situasi yang terstruktur (*structured settings*).

Aspek - aspek yang dapat dinilai atau diukur dalam penilaian unjuk kerja adalah:

1. Kualitas penyelesaian pekerjaan, yakni bagaimana kualitas dari pekerjaan dari peserta didik ketika mengerjakan tugas

tertentu, seperti harus sesuai dengan kaidah – kaidah kerja yang telah ditentukan.

2. Keterampilan menggunakan alat – alat, yakni bagaimana peserta didik mampu menggunakan alat-alat yang digunakan dalam unjuk kerja untuk menyelesaikan tugas tertentu secara baik dan sesuai dengan Prosedur Operasional Standar (POS).
3. Kemampuan menganalisis dan merencanakan prosedur kerja sampai selesai, yakni bagaimana peserta didik mampu melakukan analisis dan merencanakan prosedur kerja dari awal sampai selesai secara baik.
4. Kemampuan mengambil keputusan berdasarkan aplikasi informasi yang diberikan.
5. Kemampuan membaca, menggunakan diagram, gambar-gambar, dan simbol-simbol.

C. Kelebihan dan Kelemahan Penilaian Unjuk Kerja

Menurut H. Djaali dan Mulyono, Pudji (2008) beberapa kelebihan dari penilaian unjuk kerja adalah:

1. Dapat menilai kompetensi yang berupa keterampilan (skill)
2. Dapat digunakan untuk mencocokkan kesesuaian antara pengetahuan mengenai teori dan keterampilan didalam praktik, sehingga informasi penilaian menjadi lengkap.
3. Dalam pelaksanaan tidak ada peluang peserta didik untuk menyontek.
4. Guru dapat mengenal lebih dalam lagi masing-masing karakteristik peserta didik.
5. Memotivasi peserta didik untuk aktif.
6. Mempermudah peserta didik untuk memahami sebuah konsep yang abstrak ke konkret.

7. Kemampuan peserta didik dapat dioptimalkan.
8. Melatih keberanian peserta didik dalam mempermudah penggalan ide-ide.
9. Mampu menilai kemampuan dan keterampilan kinerja siswa dalam menggunakan alat dan sebagainya.
10. Hasil penilaian langsung dapat diketahui oleh peserta didik.

Sedangkan kelemahan dari penilaian unjuk kerja adalah:

1. Tidak semua materi pelajaran dapat dilakukan materi penilaian ini.
2. Nilai bergantung dengan hasil kerja.
3. Jika jumlah peserta didiknya banyak guru kesulitan untuk melakukan penilaian ini.
4. Waktu terbatas untuk melakukan penilaian seluruh peserta didik.
5. Peserta didik yang kurang mampu akan merasa minder.

6. Karena peserta didik terlalu banyak sehingga sulit untuk melakukan pengawasan.
7. Memerlukan sarana dan prasarana yang lengkap.
8. Memakan waktu yang lama, biaya yang besar, membosankan.
9. Harus dilakukan secara penuh dan lengkap.
10. Keterampilan yang dinilai melalui tes perbuatan mungkin sekali belum sebanding mutunya dengan keterampilan yang dituntut oleh dunia kerja, karena kemajuan ilmu pengetahuan dan teknologi selalu lebih cepat dari pada apa yang didapatkan di sekolah.

D. Kriteria Penilaian Unjuk Kerja

Penilaian kinerja digunakan untuk menilai kemampuan siswa melalui penugasan (*task*). Dalam menilai kinerja siswa tersebut, perlu disusun kriteria. Kriteria yang menyeluruh disebut *rubric*. Dengan demikian wujud asesmen kinerja yang utama adalah task (tugas) dan

rubrics (kriteria penilaian). Tugas-tugas kinerja digunakan untuk memperlihatkan kemampuan siswa dalam melakukan suatu keterampilan tentang sesuatu dalam bentuk nyata. Selanjutnya rubrik digunakan untuk memberikan keterangan tentang hasil yang diperoleh siswa (Zainul, 2001:9-11).

Ada beberapa kriteria yang dapat digunakan sebagai bahan pertimbangan dalam penilaian kinerja antara lain: *generalizability* atau keumuman, *authenticity* atau keaslian/nyata, *multiplefocus* (lebih dari satu fokus), *fairness* (keadilan), *teachability* (bisa tidaknya diajarkan), *feasibility* (kepraktisan), *scorability* atau bisa tidaknya tugas tersebut diberi skor (Popham, 1995:147).

Pengembangan penilaian unjuk kerja dilakukan melalui kegiatan analisis jabatan, penentuan skala rating numerik, pembuatan tes unjuk kerja, analisis manfaat, dan generalisasi validitas. Guna mengevaluasi apakah penilaian

unjuk kerja sudah dapat dianggap berkualitas, maka berdasar pendapat Popham (Sriyono, 2004:5) maka perlu diperhatikan tujuh kriteria, yaitu:

1. *Generalizability*, apakah kinerja peserta tes (*student performance*) dalam melakukan tugas yang diberikan tersebut sudah memadai untuk digeneralisasikan pada tugas- tugas lain. Apabila tugas-tugas yang diberikan dalam rangka penilaian keterampilan atau penilaian unjuk kerja sudah dapat digeneralisasikan, maka semakin baik tugas yang diberikan.
2. *Authenticity*, apakah tugas yang diberikan tersebut sudah serupa dengan apa yang sering dihadapinya dalam praktek kehidupan sehari-hari.
3. *Multiple focus*, apakah tugas yang diberikan kepada peserta tes sudah mengukur lebih dari satu kemampuan yang diinginkan.

4. *Teachability*, tugas yang diberikan merupakan tugas yang hasilnya semakin baik karena adanya usaha pembelajaran. Tugas yang diberikan dalam penilaian keterampilan atau penilaian kinerja adalah tugas-tugas yang relevan dengan yang dapat diajarkan guru.
5. *Fairness*, apakah tugas yang diberikan sudah adil (*fair*) untuk semua peserta tes.
6. *Feasibility*, apakah tugas yang diberikan dalam penilaian keterampilan atau penilaian kinerja memang relevan untuk dapat dilaksanakan, mengingat faktor-faktor biaya, tempat, waktu atau peralatan.
7. *Scorability*, apakah tugas yang diberikan dapat diskor dengan akurat dan reliabel?

E. Validitas Penilaian Unjuk Kerja

Validitas suatu alat ukur atau tes atau instrumen dapat diketahui atau dapat dicapai dari hasil teoritik atau pemikiran, dan dari hasil

empirik atau pengalaman. Allen dan Yen (1979) mengemukakan bahwa suatu tes dikatakan valid jika tes tersebut mengukur apa yang ingin diukur. Untuk mengetahui apakah tes yang digunakan benar-benar mengukur apa yang dimaksudkan untuk diukur, maka dilakukan validasi terhadap tes tersebut. Fernandez 1984) mengemukakan bahwa validitas tes dikategorikan menjadi tiga, yaitu validitas isi, validitas kriteria, dan validitas konstruk.

Menurut Djemari (1996), validitas tes unjuk kerja adalah penentuan evaluatif secara keseluruhan tentang derajat bukti empiris dan rasional teori mendukung ketepatan dan kesesuaian penafsiran dan tindakan berdasarkan sekor tes atau bentuk pengukuran yang lain. Validitas isi mengacu pada sejauhmana butir-butir soal tes mencakup keseluruhan isi yang hendak diukur. Hal ini berarti isi tes tersebut harus tetap relevan dan tidak menyimpang dari tujuan pengukuran. Pengkajian validitas isi

khusus pada tes unjuk kerja tidak dilakukan melalui analisis statistik, tetapi dengan menggunakan analisis rasional. Yang dianalisis secara rasional adalah validitas isi dan validitas konstruk.

Sebuah tes dikatakan mempunyai validitas validitas isi yang tinggi apabila tes tersebut berisi materi-materi yang ada pada GBPP, tolok ukur yang kedua adalah tujuan instruksional. Jadi tes prestasi belajar dapat dinyatakan sebagai tes yang mempunyai validitas isi yang tinggi apabila butir-butir soalnya selaras dengan tujuan yang diturunkan menjadi butir soal. Dengan kata lain bahwa suatu tes dikatakan valid apabila materi tes tersebut betul-betul merupakan bahan-bahan yang representatif terhadap bahan pelajaran yang diberikan.

F. Reliabilitas Penilaian Unjuk Kerja

Pengertian reliabilitas tes adalah berhubungan dengan konsistensi, kestabilan atau

ketetapan. Reliabilitas adalah derajat keajegan yang menunjukkan hasil yang sama dalam waktu yang berlainan atau orang yang berbeda dalam waktu yang sama. Tes demikian dapat dipercaya atau dapat diandalkan (Sumadi, 1992). Berdasar pendapat Djemari Mardapi (1996) pada reliabilitas suatu alat ukur, bukti yang perlu ditunjukkan adalah besarnya konsistensi antar penilai (*inter-rater*). Misalnya suatu tugas yang dikerjakan seseorang diamati atau dinilai oleh tiga orang, hasil tiga perangkat skor tersebut dikorelasikan, bila harganya tinggi berarti penilai tersebut bisa dipercaya dalam arti berhak melakukan penilaian. Bila koefisiennya rendah, maka hasil pengukuran mengandung kesalahan yang besar

G. Pengembangan Penilaian Unjuk Kerja

Langkah-langkah yang perlu diperhatikan dalam menyusun penilaian keterampilan atau penilaian kinerja, yaitu:

1. Mengidentifikasi semua langkah-langkah penting yang diperlukan atau yang akan mempengaruhi hasil akhir (*output*) yang terbaik.
2. Menuliskan perilaku kemampuan-kemampuan spesifik yang penting dan diperlukan untuk menyelesaikan tugas dan menghasilkan hasil akhir (*output*) yang terbaik.
3. Membuat kriteria kemampuan yang akan diukur tidak terlalu banyak sehingga semua kriteria tersebut dapat diobservasi selama siswa melaksanakan tugas.
4. Mendefinisikan dengan jelas kriteria kemampuan-kemampuan yang akan diukur berdasarkan kemampuan siswa yang bisa diamati (*observable*) atau karakteristik produk yang dihasilkan.
5. Mengurutkan kriteria-kriteria kemampuan yang akan diukur berdasarkan urutan yang dapat diamati.

Menurut Djemari Mardapi (2008), ada delapan langkah yang perlu ditempuh dalam mengembangkan tes hasil atau prestasi, yaitu:

1. Menyusun spesifikasi tes
2. Menulis soal tes
3. Menelaah soal tes
4. Melakukan uji coba tes
5. Menganalisis butir soal
6. Memperbaiki tes
7. Merakit tes
8. Melaksanakan tes
9. Menafsirkan hasil tes

H. Kriteria Penilaian Unjuk Kerja Yang Baik

Instrumen unjuk kerja yang baik memuat hal-hal berikut:

1. Autentik dan menarik

Hal yang penting bagi suatu instrumen unjuk kerja adalah menarik dan melibatkan siswa dalam situasi yang akrab dengan mereka sehingga siswa berusaha untuk

menyelesaikan tugas itu dengan sebaik-baiknya. Siswa cenderung lebih tertarik terhadap situasi tugas yang menyerupai kehidupan sehari-hari. Tugas ini akan membuat siswa menggunakan pengetahuan dan ketrampilan yang dikuasainya untuk menyelesaikan tugas tersebut.. Berdasarkan pengalaman dan pemahaman tentang karakteristik siswa, seorang guru dapat memperkirakan apakah aktivitas dalam tugas unjuk kerja yang dibuat akan berhasil atau tidak. (Kusnandar, 2013).

2. Memungkinkan penilaian individual

Banyak instrumen unjuk kerja yang dimaksudkan untuk dikerjakan siswa secara berkelompok. Namun perlu diingat bahwa penilaian ini sebenarnya lebih dititik beratkan untuk penilaian individu. Karena itu disain penilaian unjuk kerja sebaiknya bisa ditujukan untuk kelompok dan individu. Sebagai contoh sekelompok siswa diberi data

dan diminta untuk menganalisisnya. Untuk penilaian individunya masing-masing siswa diminta untuk memberi rangkuman dan penafsiran apa yang ditunjukkan oleh data tersebut.

3. Memuat petunjuk yang jelas

Instrumen unjuk kerja yang baik harus memuat petunjuk yang jelas, lengkap, tidak ambigu dan tidak membingungkan. Petunjuk juga harus memuat apa yang dikerjakan siswa yang nanti akan dinilai. Sebagai contoh, jika salah satu kriteria penilaian meliputi organisasi informasi, maka siswa harus diminta untuk menampilkan informasi yang diperoleh dalam bentuk yang teratur. Setelah instrumen penilaian unjuk kerja jadi, teman-teman sejawat bisa diminta pendapatnya untuk menganalisa. Mereka mungkin dapat melihat kekurangan penjelasan petunjuk siswa atau keambiguan kata-kata yang digunakan. Di samping itu mereka juga dapat melihat bila

ada informasi yang disajikan tidak lengkap sehingga ada kemungkinan siswa tidak dapat menyelesaikan tugas. Setelah mempertimbangkan masukan dan saran-saran yang diberikan, instrumen dapat diperbaiki dan kemudian bisa diujikan kepada siswa.

I. Langkah-Langkah dan Teknik Penilaian Unjuk Kerja

Langkah-langkah kinerja yang diharapkan dilakukan dilakukan peserta didik untuk menunjukkan kinerja dari satu kompetensi:

1. Kelengkapan dan ketepatan aspek yang akan dinilai dalam kinerja tersebut.
2. Kemampuan-kemampuan khusus yang diperlukan untuk menyelesaikan tugas.
3. Upayakan kemampuan yang akan dinilai tidak terlalu banyak, sehingga semua dapat diamati.

4. Kemampuan yang akan dinilai diurutkan berdasarkan urutan yang akan diamati.
5. Peserta didik telah memperoleh semua bahan, alat, instrumen, gambar-gambar atau semua peralatan penyelesaian tes.
6. Peserta didik telah mengetahui apa yang harus dikerjakan dan berapa lama waktunya serta aspek-aspek apa saja yang akan dinilai.
7. Guru sebaiknya jangan memberi bantuan kepada peserta didik, kecuali menjelaskan petunjuk-petunjuk yang telah diberikan kepadanya. (Sudjana, Nana, 1994)

Hal yang penting dalam pembelajaran keterampilan adalah diperolehnya penguasaan keterampilan praktis, serta pengetahuan dan perilaku yang berhubungan langsung dengan keterampilan tersebut. Sehubungan dengan itu, maka para ahli telah mengembangkan berbagai metode pembelajaran keterampilan yang berbeda-beda, tergantung pada sasaran atau

maksud yang hendak dicapai di dalam pembelajaran tersebut.

Model yang sederhana untuk pembelajaran keterampilan kerja adalah metode empat tahap TWI (*Training Within Industry*). Tahap-tahap tersebut meliputi :

1. Persiapan

Dalam hal ini, pendidik atau instruktur mengutarakan sasaran-sasaran latihan kerja, menjelaskan arti pentingnya latihan, membangkitkan minat para peserta pelatihan (peserta didik) untuk menerapkan pengetahuan yang sudah dimilikinya dalam situasi yang riil.

2. Peragaan

Pada tahap ini, instruktur memperagakan keterampilan yang dipelajari oleh peserta didik, menjelaskan cara kerja dan proses kerja yang benar. Dalam hal ini, instruktur harus mengambil posisi sedemikian rupa sehingga para peserta pelatihan akan dapat

mengikuti demonstrasi mengenai proses kerja dengan baik.

3. Peniruan

Pada tahap ini, peserta pelatihan menirukan aktivitas kerja yang telah diperagakan oleh instruktur. Dalam hal ini, instruktur mengamati peniruan yang dilakukan oleh peserta pelatihan, menyuruh melakukannya secara berulang-ulang dan membantu serta mendorong para peserta pelatihan agar dapat melakukan pekerjaannya dengan benar.

4. Praktik

Setelah instruktur yakin bahwa peserta pelatihan telah dapat melakukan tugas pekerjaan dengan cara kerja yang benar, maka selanjutnya instruktur memberikan tugas kepada peserta pelatihan untuk melakukan tugas pekerjaannya. Dalam hal ini, peserta pelatihan mengulangi aktivitas kerja yang baru saja dipelajarinya sampai keterampilan tersebut dapat dikuasai

sepenuhnya. Instruktur melakukan pengamatan untuk melakukan penilaian baik terhadap aktivitas atau cara kerja peserta pelatihan maupun hasil-hasil pekerjaan atau produk yang dihasilkannya.

Metode empat tahap ini mempunyai keterbatasan, karena hanya cocok untuk pembelajaran keterampilan yang bertujuan membuat barang (fabrikasi), sedangkan pembelajaran keterampilan yang memiliki karakteristik yang berbeda (seperti: *trouble shooting*, layanan/jasa) tidak tepat menggunakan langkah-langkah pembelajaran tersebut (Jaedun, 2010)

Permasalahan yang sering dihadapi dalam penilaian unjuk kerja. Menurut Popham (1995), terdapat tiga sumber kesalahan (*sources of error*) dalam *performance assessment*, yaitu:

1. *Scoring instrument flaws*, instrumen pedoman pensekoran tidak jelas sehingga sukar untuk digunakan oleh penilai,

umumnya karena komponen-komponennya sukar untuk diamati (*unobservable*).

2. *Procedural flaws*, prosedur yang digunakan dalam *performance assessment* tidak baik sehingga juga mempengaruhi hasil pensekoran.
3. *Teachers personal-bias error*, penskor (*rater*) cenderung sukar menghilangkan masalah personal bias, yakni ada kemungkinan penskor mempunyai masalah *generosity error*, artinya *rater* cenderung memberi nilai yang tinggi-tinggi, walaupun kenyataan yang sebenarnya hasil pekerjaan peserta tes tidak baik atau sebaliknya. Masalah lain adalah adanya kemungkinan terjadinya subyektifitas penskor sehingga sukar baginya untuk memberi nilai yang obyektif.

Dengan menerapkan pedoman penilaian, merupakan salah satu cara yang baik dalam memberikan penilaian pada pekerjaan siswa secara obyektif. Seorang guru tidak menggunakan

format penilaian, maka penilaiannya akan mengada-ngada, menerka-nerka, sehingga dia tidak bisa memberikan penilaian yang objektif kepada pekerjaan siswa.

Berkenaan dengan penilaian keterampilan atau penilaian unjuk kerja (*performance assessment*) untuk bidang busana pada peserta didik di sekolah menggunakan skala rentang. Penerapan skala rentang diharapkan memperoleh ketepatan proses menilai untuk memperkecil kesalahan penilai atau *rater*. Disamping itu juga dalam membuat rubrik perlu tergambar jelas, pelatihan perlu ditingkatkan untuk *rater*, dan pemantauan berkesinambungan dalam proses menilai.

Penilaian kinerja atau unjuk kerja adalah teknik pengumpulan data dengan cara pengamatan perilaku siswa secara sistematis tentang proses atau produk berdasarkan kriteria yang jelas, yang berfungsi sebagai dasar penilaian. Pengamatan unjuk kerja perlu

dilakukan dalam berbagai konteks untuk menetapkan tingkat pencapaian kemampuan tertentu.

Untuk menilai kemampuan melakukan komunikasi di tempat kerja misalnya, perlu dilakukan pengamatan atau observasi komunikasi yang beragam, seperti: (1) komunikasi dengan pelanggan eksternal dilaksanakan secara terbuka, ramah, sopan dan simpatik; (2) bahasa digunakan dengan intonasi yang cocok; (3) bahasa tubuh digunakan secara alami/natural tidak dibuat-buat ; (4) kepekaan terhadap perbedaan budaya dan sosial diperlihatkan; (5) komunikasi dua arah yang efektif digunakan secara aktif (Andono, dkk. 2003:42). Dengan cara demikian, gambaran kemampuan peserta didik akan lebih utuh.

J. Instrumen Penilaian Unjuk Kerja

Untuk mengamati penilaian unjuk kerja peserta didik dapat menggunakan alat atau

instrumen lembar pengamatan atau observasi dengan daftar cek (*check list*) dan skala penilaian (*rating scale*). Berikut penjelasan kedua alat sebagai penilaian unjuk kerja tersebut.

1. Daftar Cek (*check list*)

Menurut H. Djaali dan Mulyono, Pudji (2008) penilaian unjuk kerja dapat dilakukan dengan menggunakan daftar cek (baik atau tidak baik, bisa atau tidak bisa). Dengan menggunakan daftar cek, peserta didik mendapat nilai baik atau mampu apabila yang ditampilkan sesuai dengan kriteria yang telah ditetapkan oleh guru. Sedangkan apabila peserta didik tidak mampu menampilkan sesuatu sesuai dengan kriteria yang telah ditetapkan, maka peserta didik dinyatakan belum mampu untuk kriteria tersebut. Kelemahan cara ini adalah penilaian hanya mempunyai dua cara mutlak, misalnya benar-salah, mampu-tidak mampu, terampil-tidak terampil dan kategori sejenisnya.

Dengan demikian, skor yang diperoleh peserta didik bersifat rigid atau kaku dan tidak terdapat nilai tengah. Namun daftar cek lebih praktis digunakan mengamati subjek dalam jumlah besar dan hasilnya kontras.

2. Skala Penilaian (*rating scale*)

Penilaian unjuk kerja 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. Skala tersebut, misalnya, tidak *kompeten* – *agak kompeten* – *kompeten* – *sangat kompeten*. Untuk memperkecil faktor subjektivitas, perlu dilakukan penilaian oleh lebih dari satu orang, agar hasil penilaian lebih akurat. Terdapat tiga jenis *rating scale*, yaitu: (1) *numerical rating scale*; (2) *graphic*

rating scale; (3) *descriptive graphic rating scale* (Gronlund, 1985).

Rosnita (2007) mengemukakan bahwa penilaian unjuk kerja yang menggunakan skala penilaian memungkinkan penilai memberi nilai tengah terhadap penguasaan kompetensi tertentu, karena pemberian nilai secara kontinum dimana pilihan kategori nilai lebih dari dua. Skala penilaian terentang dari tidak sempurna sampai sangat sempurna. Misalnya: 1 = kurang kompeten, 2 = cukup kompeten, 3 = kompeten dan 4 = sangat kompeten. Untuk memperkecil faktor subjektivitas, perlu dilakukan penilaian oleh lebih dari satu orang, agar hasil lebih akurat.

Berikut ini akan dijelaskan tentang contoh instrumen atau alat penilaian unjuk kerja yang dapat digunakan guru dalam menilai aspek unjuk kerja peserta didik.

Contoh Lembar Penilaian Praktek Desain Busana

No	Aspek yang dinilai	Penilaian				Bobot	Jumlah
		4	3	2	1		
A	PERSIAPAN						
1	Kelengkapan alat						
2	Kelengkapan bahan						
Jumlah		10%					
B	PROSES						
1	Pemakaian alat dan bahan						
2	Kecepatan kerja						
3	Kebersihan tempat kerja						
Jumlah		30%					
C	HASIL						
1	Proporsi						
2	Kesatuan						
3	Komposisi						
4	Variasi						
5	Warna						
6	Teknik penyajian gambar						
7	Teknik penyelesaian gambar						
8	Kesesuaian sumber ide						
9	Kesesuaian kesempatan						
Jumlah		60%					
Jumlah		100%					

Tabel 10. Contoh Tabel Penilaian Unjuk Kerja

Contoh Lembar Observasi Untuk Mengamati Kemampuan Siswa Dalam Membuat Pola

No	Aspek yang dinilai	Hasil Observasi	
		Benar	Salah
	Menggunakan pita ukuran Menggoreskan garis Ketepatan bentuk Ketepatan ukuran Kesesuaian desain dll		

Tabel 11. Contoh Lembar Penilaian Unjuk Kerja

K. Hambatan Penilaian Unjuk Kerja

Terdapat beberapa sumber kendala yang dihadapi oleh para guru dalam menilai unjuk kerja siswa yaitu: pertama, pedoman penyekoran dalam instrument tidak jelas sehingga sukar digunakan, komponen-komponen yang dinilai sulit untuk diamati, sehingga cenderung diabaikan; kedua, penilai (*rater*) umumnya hanya satu orang yaitu guru bidang studi, sedangkan komponen-komponen yang dinilai dan jumlah siswa yang dinilai cukup banyak, sehingga sulit untuk mendapat pembanding untuk dijadikan

bahan pertimbangan mengambil keputusan; ketiga, kemungkinan ada kecenderungan untuk memberi nilai tinggi atau sebaliknya, hal ini diakibatkan oleh instrumen yang digunakan belum memenuhi persyaratan validitas, reliabilitas dan kepraktisannya. Berdasarkan uraian di atas dapat disimpulkan bahwa penerapan instrument penilaian unjuk kerja yang valid reliabel, praktis dan dapat digunakan secara berulang pada tugas kinerja yang berbeda, dapat membantu guru untuk melakukan penilaian unjuk kerja siswa pada saat melakukan praktikum di laboratorium.

Kesukaran lainnya yang paling utama ditemukan dalam penilaian keterampilan atau penilaian kinerja (*performance assessment*) adalah pensekorannya. Banyak faktor yang mempengaruhi hasil pensekoran penilaian keterampilan atau penilaian kinerja. Masalah pensekoran pada penilaian keterampilan atau

penilaian kinerja lebih kompleks dari pada pensekoran pada bentuk soal uraian.

DAFTAR PUSTAKA

- Aiken, Lewis R. (1994). *Psychological Testing and Assessment, (Eight Edition)*, Boston: Allyn and Bacon.
- Airasian, Peter.W. 1991. *Classroom Assessment*. USA: McGraw-Hill.
- Aksara Sumadi Suryabrata . 1987. *Pengembangan tes hasil belajar*. Jakarta: Rajawali.
- Alderson, J. C., C. Clapham, D. Wall. 1995 *Language Test Construction and Evaluation*, Cambridge University Press.
- Anastasi. Anne and Urbina, Susana. 1997. *Psicoholological Testing. (Seventh Edition)*. New Jersey: Prentice-Hall, Inc.
- Arends, Richard I. 2008. *Learning to Teach*. Yogyakarta : Pustaka Pelajar

- Asmawi Zainul dan Noehi Nasoetion. 1997. Penilaian Hasil Belajar. Pusat Antar Universitas, Direktorat Jenderal Pendidikan Tinggi: Departemen Pendidikan Dan kebudayaan.
- Berk, R.A. 1986. *Performance Assessment: Methods & Applications*. London: The Johns Hopkins Press Ltd.
- Brandt, Joyce A. R.N., M.A. 1987. "Constructing Achievement Tests". *Journal of Health Occupations Education*: Vol. 2: No. 1, Article 10. Available at: <http://stars.library.ucf.edu/jhoe/vol2/iss1/10>
- Brown, A. L , and Page. A. (1970). *Elements of Functional Analysis*. Butler & Tanner Ltd. Frome and London.
- D. Shermis, Mark and J. Di Vesta, Francis. 2011. *Classroom Assessment In Action*. USA: Rowman&Littlefield Publishers, Inc.
- Dali, S Naga. 1992. Pengantar Teori Sekor Pada Pengukuran Pendidikan. Gunadarma: Jakarta.

- Departemen Pendidikan Dan Kebudayaan. 1997. Manual Item And Test Analysis (Iteman). Badan Penelitian dan Pengembangan Pendidikan dan Kebudayaan: Pusat Penelitian dan Pengembangan Sistem Pengujian.
- Djemari Mardapi. 2008. *Teknik penyusunan instrument tes dan non tes*. Yogyakarta: Mitra Cendekia.
- Farlex. The Free Dictionary. Diakses pada tanggal 24 Oktober 2017 pukul 14.45 WIB dari <http://www.thefreedictionary.com/test>
- G. Paris, Scott and R. Ayres, Linda. Xxxx. *Becoming Reflective Students and Teachers With Portfolios and Authentic Assessment*. Washington: American Psychological Association I.
- Gronlund, Norman E. 1973. *Preparing Criterion-Referenced Test for Classroom Instruction*. New York: The Macmillan Publishing Company.

- Gronlund, Norman E. 1977. *Constructing Achievement Tests – Second Ed.* USA : Prentice-Hall, Inc.
- Gronlund, N.E. 1985. *Measurement and evaluation in testing (5th Ed.)* New York: Macmillan Publising Co, Inc.
- H. Djaali dan Mulyono, Pudji .2008. Pengukuran Dalam Bidang Pendidikan. Jakarta: PT Grasindo. 2008.
- Hari Setiadi. 2006. *Penilaian Kinerja*. Jakarta: Puspendik Balitbang Depdiknas
- Hutabarat, O. R. 2004. Model-model Penilaian Berbasis Kompetensi PAK. Bandung: Bina Media Informasi.
- Iryanti, Puji. 2004. Penilaian Unjuk Kerja. Yogyakarta: Depdiknas.
- J. Bagnato, Stephen. 2007. *Authentic Assessment For Early Childhood Intervention*. New York : The Guilford Press.
- Janesick, Valerie J.2006. *Authentic Assessment Primer*. Peter Lang: New York.

- 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 McKay Co., Inc.
- Kusnandar .2013. *Penilaian Autentik*, jakarta: PT Raja Grafindo Persada.
- L. Howell, Scott and Hricko, Mary. 2006. *Online Assessment and Measurement: Case Studies From Higher Education K-12 and Corporate*. USA: Idea Group, Inc.
- Lampiran Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 104 Tahun 2014.
- Lehman, H. 1990. *The Systems Approach to Education. Spesial Presentation Conveyed in The International Seminar on Educational Innovation and Technology*. Manila : Innotech Publications Vol 20 No 05

- Linda Darling-Hammond, Jacqueline Ancess, and Beverly Falk. 1951. *Authentic Assessment in Action: Studies of Schools and Students at Work*. Teacher college press, Columbia University : New York.
- Linn, L. R., and Groundlund, N. E., 2000, *Measurement and Assessment in Teaching*, Merril, USA.
- M. Dorn, Charless. 2004. *Assessing Expressive Learning_A Practical Guide for Teacher-Directed_Authentic Assessment in K-12 Visual Arts Education*. London:Lawrence Erlbaum Associates.
- Mahmudah, S.2000. Penerapan Penilaian Kinerja Siswa (*Performance Assessment*) pada Pembelajaran Sub Konsep Jaringan Hewan. Bandung:UPI.
- Majid, A. 2006. Perencanaan Pembelajaran Mengembangkan Standar Kompetensi Guru. Bandung: Remaja Rosdakarya.

- Majid, Abdul dan S. Firdaus, Aep. (2014). Penilaian autentik Proses dan Hasil Belajar. Bandung: Interes, Cet.I.
- Mehren, W.A. & Lehmann, I.J. (1984). *Measurement and Evaluation in Education and Psychology*. New York: Holt, Rinehart.
- Nitko, Anthony J. (1996). *Educational Assessment of Students, Second Edition*. Ohio: Merrill an imprint of Prentice Hall Englewood Cliffs.
- Oriondo, L. L. & Antonio, E. M.D. 1998. Evaluating Educational Outcomes (Test, Measurement and Evaluation). L orentino St: Rex Printing Company, Inc.
- Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 104 Tahun 2014
- Popham, W. James. 1996. *Classroom assessment*. Boston: Allyn & Bacon.
- Puji Iryanti. 2004. *Penilaian Unjuk Kerja*. Depdiknas, Dirjendikdasmen, Pusat Pengembangan Penataran Guru Matematika Yogyakarta.

- Remmers, HH and Gage NL. 1955. *Educational Measurement and Evaluation*. New York : Harper
- Reynolds, C.R. add.all. 2010. *Measurement and Assessment in Education*. USA: Pearson Education LTD.
- Rosnita. 2007. *Evaluasi Pendidikan*. Bandung: Cita Pustaka Media.
- Sax, Gilbert. *Principles of Educational and Psychological Measurement and Evaluation*. Second Edition, Belmont, California: Wadsworth, Inc., 1980.
- Semiawan Stamboel, Conny. *Prinsip dan Teknik Pengukuran dan Penilaian di Dalam Dunia Pendidikan*, Cetakan II, Jakarta: Mutiara Sumber Widya, 1986.
- Setyono, Budi.2005. *Penilaian Otentik dalam Kurikulum Berbasis Kompetensi (dalam jurnal pengembangan pendidikan)*. Lembaga Pembinaan dan Pengembangan Pendidikan (LP3) Universitas Jember.

- Shermis, Mark D. dan Francis J. Di Vesta. 2011. *Classroom Assessment in Action*. Maryland : Rowman & Littlefield Publishers
- Srikandi. 2002. Analisis Butir Soal Pilihan Ganda. Vol 8.
http://eprints.uny.ac.id/1699/1/ANALISIS_BU_TIR_SOAL_PILIHAN_GANDA.pdf diakses pada tanggal 6 November 2017 pukul 16.38
- Stiggin, R.J.1994. Student-Centered Classroom Assessment. New York: Mac Millan College Publishing Company.
- Stoyanova, F., & C, Amuka. 1996. *The Test for Teachers* (Стоянова, Ф. Тестология за учители.
- Sudjana, Nana .1991. Penilaian Hasil Proses Belajar Mengajar, cetakan ketiga. Bandung: PT Remaja Rosda Karya.
- Sugihartono dkk. 2013. *Psikologi Pendidikan*. Yogyakarta : UNY Press
- Sugiyono, 2006, Statistika untuk Penelitian, Bandung: Alfabeta

- Suharsimi Arikunto. 2001. *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara: Jakarta.
- Suryabrata, S. 1999. *Pengembangan Alat Ukur Psikologis*. Direktorat jenderal Pendidikan Tinggi: Departemen Pendidikan dan Kebudayaan.
- Sutrisno, Hadi. 1997. *Methodologi Research II*. Yogyakarta : Yasbit Psikologi UGM
- Van Blerkom, M.L. 2009. *Measurement and Statistics for Teachers*. New York: Routledge.
- Vanya Ivanova. CONSTRUCTION AND EVALUATION OF ACHIEVEMENT TESTS IN ENGLISH. *Национална конференция „Образованието в информационното общество*.
- Wiersma, William & Stephen G. Jurs. *Educational Measurement and Testing. Second Edition, Needham Heights, Massachusetts*. Allyn and Bacon, 1990.

- Wiggins, G. P., 1993. *Assessing Student Performance: Exploring the Limits and Purpose of Testing*. San Francisco: Jossey-Bass Publishers.
- Williams, David D. and Mary Hricko. 2006. *Online Assessment, Measurement, and Evaluation: Emerging Practices*. United States of America: Idea Group Inc (IGI)
- Zainal A. (1990) *Evaluasi instruksional*. Bandung: PT. Remaja Rosdakarya.
- Zainul, Asmawi. 2001. *Alternative Assessment*. Jakarta: Universitas Terbuka.

TENTANG PENULIS



Bernama lengkap Andrean Sandi Wicaksono, ia lahir di Kulon Progo tepatnya pada 27 Oktober 1997. Dengan nama panggilan Andre sejak kecil, ia menempuh pendidikan sekolah dasar di SD Negeri Pepen dari tahun 2002 hingga 2009. Lalu melanjutkan ke jenjang SMP di SMP Negeri 4 Wates dan lulus tepat waktu pada tahun 2012. Setelah lulus dari SMP, ia memutuskan untuk melanjutkan studinya ke SMK Negeri 2 Pengasih dengan fokus program studi kejuruan yaitu Teknik Komputer dan Jaringan. Setelah lulus di tahun 2015 perjalanan studinya tidak langsung berjalan, namun ia masuk ke perguruan tinggi pada tahun 2016. Dengan program studi PGSD ia berniat untuk menjadi agen perubahan untuk ikut berpartisipasi mengubah negara Indonesia tercinta ini menjadi lebih hebat lagi.

Majulah pendidikan Indonesia, jayalah bangsaku!