BAB I
PRINSIP PENILAIAN DAN EVALUASI PENDIDIKAN

A. PENGERTIAN PENILAIAN DAN EVALUASI PENDIDIKAN

Dalam dunia pendidikan, penilaian atau asesmen (*assessment*) diartikan sebagai prosedur yang digunakan untuk mendapatkan informasi untuk mengukur taraf pengetahuan dan keterampilan subjek didik yang hasilnya akan digunakan untuk keperluan evaluasi. Informasi adalah data yang diperoleh melalui pengukuran dan nonpengukuran termasuk di dalamnya dengan melakukan observasi kelas, menggunakan tes yang standar atau tes buata guru, proyek, dan protofolio subjek belajar (*Assessment is a process of gathering information to measure a student's level of knowledge and/or skill. Assessment can incorporate many methods including classroom observation, standardized tests, teacher-developed tests, classroom projects, and portofolios of student work*).

Evaluasi merupakan suatu proses yang sistematis yang dilaksanakan untuk mengetahui tingkat keberhasilan dan efisiensi dari program yang bersangkutan. Dalam hal ini termasuk di dalamnya untuk mengetahui keberhasilan seluruh subjek belajar yang menempuh suatu program.

Untuk memperoleh informasi yang berupa data kuantitatif dilakukan melalui pengukuran. Jadi, pengukuran (*measurement*) merupakan proses untuk memperoleh deskripsi numerik atau kuantitatif tentang tingkatan karakteristik yang dimiliki seseorang dengan aturan tertentu. Untuk memperoleh data kuantitatif dalam dilakukan melalui tes dan nontes. Tes merupakan metode pengukuran yang menggunakan alat ukur berbentuk satu set pertanyaan untuk mengukur sampel tingkah laku, dan jawabannya dapat dikategorikan benar dan salah. Nontes merupakan metode pengukuran yang menggunakan alat ukur untuk mengukur sampel tingkah laku, tetapi jawabannya tidak dapat dikategorikan benar dan salah, misalnya positif dan negatif, setuju dan tidak setuju, suka dan tidak suka.

Dalam Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 tentang Standar Penilaian Pendidikan dinyatakan bahwa ulangan adalah proses yang dilakukan untuk mengukur pencapaian kompetensi peserta didik secara berkelanjutan dalam proses pembelajaran, untuk memantau kemajuan, melakukan perbaikan pembelajaran, dan menentukan keberhasilan belajar peserta didik. Ulangan harian adalah kegiatan yang dilakukan secara periodik untuk mengukur pencapaian kompetensi peserta didik setelah

menyelesaikan satu Kompetensi Dasar (KD) atau lebih. Ulangan tengah semester adalah kegiatan yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik setelah melaksanakan 8 – 9 minggu kegiatan pembelajaran. Cakupan ulangan meliputi seluruh indikator yang merepresentasikan seluruh KD pada periode tersebut.Ulangan akhir semester adalah kegiatan yang dilakukan oleh pendidik untuk mengukur pencapaian kompetensi peserta didik di akhir semester. Cakupan ulangan meliputi seluruh indikator yang merepresentasikan semua KD pada semester tersebut. Ulangan kenaikan kelas adalah kegiatan yang dilakukan oleh pendidik di akhir semester genap untuk mengukur pencapaian kompetensi peserta didik di akhir semester genap pada satuan pendidikan yang menggunakan sistem paket. Cakupan ulangan meliputi seluruh indikator yang merepresentasikan KD pada semester tersebut. Ujian sekolah/madrasah adalah kegiatan pengukuran pencapaian kompetensi peserta didik yang dilakukan oleh satuan pendidikan untuk memperoleh pengakuan atas prestasi belajar dan merupakan salah satu persyaratan kelulusan dari satuan pendidikan.

B. MAKNA EVALUASI

Suatu program, termasuk di dalamnya program pendidikan, adalah suatu kegiatan yang terencana yang lengkap dengan rincian tujuan beserta jenis-jenis kegiatannya. Oleh karena itu, apakah suatu program yang diimplementasikan benar-benar berharga, diperlukan adanya evaluasi. Evaluasi yang dimaksud adalah suatu proses yang sistematis yang dilaksanakan untuk mengetahui tingkat keberhasilan dan efisiensi dari program yang bersangkutan. Evaluasi terhadap tingkat efisiensi terutama ditujukan kepada program yang dilaksanakan berulang-ulang. Karena keberhasilan suatu program tidak dapat terlepas dari segi pelaksanaannya, maka evaluasi terhadap suatu program menyangkut berbagai hal yang terkait, baik yang menyangkut kualitas masukan, kualitas proses maupun kualitas hasil pelaksanaannya. Selain itu, evaluasi dapat dilaksanakan atas dasar sekuen implementasi program, dapat pula dilakukan terhadap komponen program. Dalam program pendidikan misalnya, evaluasi menjadi sangat kompleks karena dapat dilakukan terhadap kurikulumnya, sarana dan prasarana, tenaga yang terlibat baik edukatif maupun administratif, kelancaran pelaksanaan program, efisiensi waktu penyelenggaraan program, dan tentunya seberapa jauh efektifnya program yang telah diselenggarakan.

Evaluasi sebagai suatu bentuk penetapan sangat tergantung kepada perspektif yang digunakan. Perspektif tersebut dapat menyangkut hal-hal berikut.

1. Perspektif alat-tujuan, yang lebih menekankan kepada pengukuran, yang kadang-kadang hasilnya bias.
2. Perspektif situasional, yang menekankan kepada sosok dikaitkan dengan penghayatan semiia pihak yang terkait.
3. Perpektif kritis, yang dikembalikan kepada asumsi dasar dan digunakan dalam penyelenggaraan program.

Karena evaluasi program pendidikan merupakan suatu bentuk penetapan untuk menyatakan berharga tidaknya suatu implementasi program, maka perlu adanya kriteria yang dapat dipertanggungjawabkan. Kriteria tersebut adalah sebagai berikut.

1. Kriteria internal, yang dijabarkan dari dalam rancangan program pendidikan itu sendiri, yang dapat ditinjau dari sudut:
   a. koherensi (konsistensi), baik koherensi antara:
      1) tujuan dengan penilaian;
      2) tujuan dengan pengalaman proses pembelajaran yang diselenggarakan;
      3) pengalaman prosespembelajaran dengan penilaiannya;
      4) tujuan dengan bahan ajarnya, dll.
   b. pengetahuan penempatan *resource* yakni mencakup pemilihan staf;
   c. reaksi pemakai program (kelompok sasaran) yang dapat ditinjau dari
      1) kepuasan;
      2) pencapaian tujuan pribadi;
      3) minat;
      4) wawasan, dll.
   d. reaksi pelaksana program, dalam hal ini adalah tenaga pengajar, yang dapat ditinjau dari sudut
      1) sikapnya terhadap program;
      2) cara penerimaan terhadap program;
      3) kepuasan;
      4) minat;
      5) wawasan;
      6) kepentingan/tujuan pribadi, dll.
      7) efektivitas penggunaan dana;
      8) kemampuan generatif atau pengembangan diri dari program (*side effect*).

2. Kriteria eksternal, yang mencakup
   a. kemampuan pengarahan kebijakan, maksudnya adalah sejauh mana pelaksanaan atau implementasi program sesuai dengan garis kebijakan yang telah ditetapkan;
   b. analisis *cost-benefid* untuk membandingkan antara biaya dengan keuntungan secara keseluruhan;
   c. efek *multiplier* (melipat ganda), baik yang berupa imbasan langsung ataupun imbasan yang tidak langsung.

C. MANFAAT EVALUASI

Suatu program ditujukan agar sasaran-sasaran yang telah ditetapkan dapat tercapai sesuai harapan. Oleh karena itu, manfaat atau kegunaan evaluasi berupa pengambilan keputusan atau untuk pertanggungjawaban terhadap kegiatan yang telah dilaksanakan. Pertanggungjawaban ditujukan kepada pihak yang menghendaki atau pihak sponsor. Dalam hal program pendidikan, maka pihak yang terkait adalah pemerintah dan masyarakat, khususnya pihak orang tua subjek belajar.

Pengambilan keputusan didasarkan pada hasil penilaian/asesmen, dan dilakukan dalam upaya untuk pengendalian kegiatan jika program masih berlangsung atau untuk upaya penyempumaan untuk pelaksanaan selanjutnya. Dengan demikian, dengan adanya hasil evaluasi akan dapat diambil kebijaksanaan apakah suatu program akan dilaksanakan lagi pada periode berikutnya, ataukah perlu dirivisi terlebih dahulu, atau jika perlu digantikan dengan program lain jika dari hasil penilaian sama sekali tidak berharga. Jadi, setiap program yang diimplementasikan hendaknya dilengkapi dengan kegiatan pengukuran, dan penilaian, dan dari hasil penilaian dilakukan evaluasi agar dapat diambil suatu kebijaksanaan untuk menentukan apakah program yang dimaksud dilaksanakan ulang, direvisi atau diubah.

Dalam dunia pendidikan, program yang ada dapat berbeda-beda tingkatannya, yaitu mulai dari tingkat kementerian, wilayah, sekolah, dan kelas. Dilihat dari segi penyelenggaraannya ada yang negeri dan ada pula yang swasta. Oleh karena ltu, pihak-pihak itulah yang memerlukan hasil penilaian. Sementara dari segi subjek belajar sendiri juga ada pihak orang tua yang terlibat di belakangnya yang memerlukan hasil penulaian untuk *self-evaluation*. Dari segi pelaksanaanya, penilai suatu program pendidikan dapat dilakukan oleh perencana dan pelaksana program, dan dapat pula diserahkan kepada pihak

lain yang dianggap ahli dan tidak terlibat daiam pelaksanaan. Jika evaluasi dilakukan terhadap setiap satuan kecil dari suatu program pendidikan yang lebih besar yang masih berjalan dalam upaya untuk pengendalian pelaksanaan program , maka evaluasi dilakukan oleh pihak pelaksana program. Dalam hal ini dikenal dengan evaluasi pada skala mikro.

Untuk menilai program pembelajaran di kelas secara periodik dalam waktu yang relatif singkat, yang paling tepat adalah dilakukannya evaluasi formatif berdasarkan hasil penilaian selama berlangsungnya proses pembelajaran oleh guru yang bersangkutan. Penilaian yang dilakukan selama proses pembelajaran disebut penilaian formatif. Hasil penilaian formatif itulah yang dipakai untuk mengevaluasi program yang sedang dijalankan. Dengan demikian, selama proses pembelajaran berlangsung, guru dapat menyempurnakan program pembelajarannya agar sasaran/target pembelajaran yang dapat tercapai sesuai harapan. Dalam skala mikro/sempit, orientasi utama evaluasi program ditujukan kepada masalah metode pembelajaran. Sebaliknya, evalauasi juga dilakukan pada skala makro/luas yang dititikberatkan pada masalah efisiensi pelaksanaan, yaitu berkenaan dengan strategi dan pelaksanaan program. Oleh karena itu, evaluasi pada skala makro akan lebih baik jika dilakukan oleh pihak luar/pihak independen.

D. KONSEP DASAR EVALUASI PROSES DAN HASIL BELAJAR

Pencapaian prestasi atau hasil belajar diperoleh setelah dilaksanakannya suatu program pembelajaran. Evaluasi pencapaian hasil belajar merupakan langkah untuk mengetahui seberapa jauh target pembelajaran dapat dicapai. Kumulatif pencapaian hasil belajar dari seluruh bidang studi akan mencerminkan keberhasilan suatu program pendidikan atau suatu kurikulum sekolah. Dengan demikian, kedudukan evaluasi hasil belajar suatu bidang studi hanyalah sebagai bagian kecil dari evaluasi pencapaian program pendidikan. Oleh karena itu, setiap pengampu bidang studi setiap saat dituntut kesadarannya bahwa dirinya sebagai bagian dari sistern pendidikan. Seorang guru bukanlah pribadi yang lepas, ia tidak dapat lepas sistem pendidikan tempat ia mengabdikan dirinya. Oleh karena itu, apa yang diambil sebagai keputusan saat ia melakukan evaluasi atas program yang diselenggarakan merupakan bagian dari sistem pendidikan.

Kalau dalam evaluasi hasil belajar pada penyelenggaraan pembelajaran masa lalu hanya sekedar untuk mengetahui sejauh mana isi/materi atau bahan ajar sudah dikuasai oleh subjek belajar, maka sekarang evaluasi hasil belajar lebih diorientasikan kepada

seberapa jauh sasaran-sasaran pendidikan telah dapat dicapai dan disertai pula dengan pelacakan peran berbagai faktor penentu aktualisasi proses pembelajarn. Sebagaimana telah diketahui bersama bahwa ada banyak faktor penentu aktualisasi proses pembelajaran. Faktor-faktor tersebut antara lain berupa faktor tujuan pendidikan, tenaga pengajar (minat, sikap, pandangan, langgam mengajar, stabilitas emosional, dll.), isi/materi pelajaran atau bahan ajar, organisasi materi, metodologi atau sistem metodologi yang dikembangkan, pendekatan yang digunakan, pengelolaan kelas yang diselenggarakan, bimbingan dan penyuluhan yang diberikan, sarana dan prasarana pendidikan, sistern administrasi, serta kondisi subjek belajar maupun sistem penilaian dan sistem evaluasi yang digunakan.

Selain itu, aktualisasi proses pembelajaran tidak dapat lepas dari faktor lingkungan yang melingkupinya, baik yang bersifat fisik maupun sosial. Jika dikaitkan dengan fungsi guru, maka aktualisasi proses pembelajaran sebagai ujud nyata implementasi program, tidak dapat lepas dari kualitas program itu sendiri. Kualitas program tidak dapat lepas dari kualitas pendidik dan tenaga pendidikan lainnya yang terlibat dalam perancangannya. Dalam skala mikro, baik buruknya penyelenggaraan program pembelajaran yang dirancang oleh guru sebagai seorang pendidik sangat tergantung kepada sejauh mana ia memahami kurikulum yang sedang berlaku, karena dalam hal ini dituntut kemampuannya dalam melakukan *restatement* kurikulum. Sementara, banyak guru yang sekedar melaksanakan kurikulum dengan asal menyusun tanpa memperhatikan potensi subjek belajar dan spesifikasi lingkungan dimana satuan pendidikan berada. Dalam implementasi kurikulum operasionalsebagai tuntutan Kurikulum Tingkat satuan Pendidikan, sudah saatnya dikembangkan program pembelajaran yang menjadi karakteristik satuan pendidikan yang bersangkutan tanpa meninggalkan standar isi yang menjadi tuntutan minimal dalam implementasi kurikulum di tingkat satuan pendidikan.

Karena tujuan belajar adalah terjadinya perubahan tingkah laku ke arah yang diinginkan, maka evaluasi pencapaian hasil belajar juga merupakan usaha untuk menetapkan tercapainya perubahan perilaku, kecakapan dan status subjek belajar selama dan setelah melakukan kegiatan belajar dari peserta program yang diselenggarakan. Perubahan tingkah laku tersebut sudah dirumuskan dalam tujuan pembelajaran, yang dalam kurikulum berbasis kompetensi tujuan pembelajaran dirumuskan dalam bentuk Standar Kompetensi Lulusan (SKL), Standar Kompetensi (SK), dan Kompetensi Dasar.

Berkait dengan kurikulum berbasis kompetensi, evaluasi pencapaian hasil belajar bukan semata-mata ditujukan untuk menetapkan apa yang telah berhasil diketahui peserta

program, tetapi lebih diorientasikan kepada apa yang dapat dikerjakan dalam bentuk kinerja (*performance*), apa sikap positif yang tumbuh di dalam diri subjek belajar, juga bagaimana kemampuan peserta program mengaktualisasikan diri dalam hidup bersama sesuai dengan hakekat belajar yakni *learning to know, learning to do, learning to be dan learning to live together*. Oleh karena itu, dewasa ini di banyak negara, penilaian menjadi bergeser kearah asesmen yang autentik dalam betuk *performance assessment* sebagai suatu *alternative assessment*.

Selama proses evaluasi formatif berlangsung guru sebagai perancang dan pelaksana program dapat melakukan penyempurnaan program pembelajarannya. Dengan demikian, diharapakan akan dapat meningkatkan prestasi belajar tiap subjek didik. Hasil penilaian yang formatif sebagai dasar melakukan evaluasi formatif bukanlah hasil yang final. Oleh karena itu, nilai formatif merupakan "nilai antara". Nilai formatif seorang subjek belajar digunakan untuk memantau ada tidaknya kesulitan yang dialaminya selama terlibat dalam proses pembelajaran. Atas dasar hasil penilaian formatif itulah guru melakukan evaluasi formatif atas program yang dirancang agar ia dapat memutuskan perlu tidaknya ia memperbaiki program pembelajarannya.

Pada akhir program, guru melakukan penilaian sumatif untuk mengetahui keberhasilan atau prestasi akhir setiap subjek didik. Evaluasi sumatif merupakan keputusan untuk menentukan hasil akhir keberhasilan subjek belajar dalam mennempuh program belajar. Evaluasi akhir semester untuk memutuskan keberhasilan subjek belajar menempuh program selama satu semester. Evaluasi akhir tahun untuk memutuskan apakah subjek belajar harus gagal sehingga tinggal kelas ataukan dapat dinyatakan sukses untuk naik kelas. peserta diik mengformapeserdengan

Dari uraian di di atas tampak bahwa demikian evaluasi pencapaian belajar atau evaluasi hasil belajar meliputi evaluasi sumatif (*sumative evaluation*) yang dikenakan pada subjek belajar sebagai penempuh program pembelajaran, guru sebagai perancang dan pelaksana program (yang rancangannya diujudkan dalam bentuk kurikulum, silabus sampai dengan RPP), dalam hubungannya dengan efektivitas kegiatan belajar dalam suatu program tertentu, sedangkan evaluasi formatif (*formative evaluation*) yang dilakukan untuk mengetahui keberhasilan subjek belajar sebagai peserta program pembelajaran selama berlangsungnya proses pembelajaran.

Jika benar-benar dapat didudukkan sebagai evaluasi proses belajar, maka akan ada upaya-upaya yang melekat atau yang menyertainya yaitu upaya untuk meningkatkan

efektivitas proses belajar sehingga dapat mencapai hasil yang lebih baik. Akan sangat ideal pula jika sebelum pelaksanaan proses pembelajaran dilaksanakan evaluasi penempatan (*placement evaluation*) dengan menghimpun data melalui penilaian/asesmen penempatan (*placement assessment*) untuk menjawab pertanyaan

1) sejauh mana subjek belajar sebagai peserta program menguasai kemampuan prasyarat yang diperlukan untuk proses pembelajaran yang akan diselenggarakan;
2) seberapa jauh subjek belajar sebagai peserta program telah menguasai kompetensi yang ditargetkan, ataupun
3) seberapa jauh seorang subjek belajar sebagai peserta program berminat terhadap program pembelajaran yang akan diselenggarakan.

Dari ragam evaluasi yang ada, maka tujuan evaluasi menjadi spesifik. Berikut ini adalah tujuan masing-masing evaluasi.

1. **Evaluasi sumatif** didasarkan pada kumulatif hasil **penilaian sumatif** subjek belajar dalam menempuh program. Dalam hal ini pengertian penilaian sumatif adalah hasil final dari subjek belajar menempuh suatu program. Misalnya, nilai sumatif dalam suatu program semester diperoleh melalui ulangan akhir suatu pokok bahasan (setelah dilakukan proses remediasi bagi yang mengalami kegagalan dan dilakukan program pengayaan bagi yang sudah menguasai target pembelajaran berdasarkan penilaian formatif), ulangan tengah semester , dan ulangan akhir semester. Adapun tujuan evaluasi sumatif adalah untuk
   a. menentukan nilai akhir seluruh peserta penempuh program pembelajaran, agar dapat dinyatakan berhasil atau gagal. Bila berhasil maka akan dapat diberi sertifikat karena ia telah menguasai kecakapan ataupun keterampilan tertentu yang ditargetkan dalam program pembelajaran yang dirancang;
   b. meramalkan kecakapan subjek belajar dalam menyelesaikan program/ semester berikutnya;
   c. menetapkan efektivitas dan efisiensi penyelenggaraan suatu program pembelajaran;
   d. dalam konteks untuk seleksi, seperti seleksi masuk berarti untuk menetapkan siap yang layak lolos seleksi, bila untuk seleksi untuk menetapkan juara untuk mewakili satuan pendidikan yang bersangkutan berarti untuk menetapkan siapa yang layak lolos menjadi wakil satua pendidikan yang bersangkutan.

2. **Evaluasi formatif** didasarkan pada hasil penilaian formatif selama subjek belajar sebagai penempuh program pembelajaran mengikuti proses pembelajaran dalam kaitannya dengan penyelenggaraan program, dan tujuannya untuk:
    a. menetapkan langkah-langkah/urutan kegiatan belajar selanjutnya agar supaya lebih efektif dan efisien;
    b. pendalaman dan pernantapan penguasaan perilaku yang ditargetkan;
    c. mendiagnosis kesulitan belajar, dalam arti bahwa subjek belajar yang mendapat nilai jelek identik belum menguasai perilaku yang ditargetkan;
    d. mencari cara mengatasi kesulitan belajar jika subjek belajar dinyatakan gagal, berdasar kegiatan belajar yang telah dilakukan;
    e. umpan balik bagi guru dalam mengelola kegiatan pembelajaran sehingga mengetahui seberapa jauh tujuan yang ditetapkan sudah dapat dicapai;
    f. meramalkan seberapa jauh keberhasilan peserta program belajar dalam mengikuti penilaian sumatif;
    g. mengetahui seberapa jauh seluruh subjek belajar sebagai penempuh program pembelajaran akan berhasil dalam mengikuti proses pembelajaran sampai akhir program, berdasar kecakapan dan keterampilan yang dikuasainya sekarang, dalam konteks bahwa subjek belajar sebagai masukan;
    h. mengetahui subjek belajar yang mana yang harus dibantu melalui program remediasi agar ia dapat berhasil menempuh program yang ditempuh;
    i. mediagnosis penyebab kegagalan subjek belajar dalam dalam menguasai kemampuan yang ditargetkandari program pembelajaran yang diselenggarakan.

3. **Evaluasi penempatan** didasarkan pada hasil penilaian terhadap subjek sebelum menempun program pembelajaran, dan bertujuan untuk:
    a. mengetahui penguasaan kemampuan prasyarat yang diperlukan dalam KBM yang akan diseienggarakan;
    b. mmenjajagi penguasaan subjek belajar sebagai peserta program terhadap kemampuan yang ditargetkan;
    c. meneliti interes, langgam belajar, ataupun karakteristik personal subjek belajar sebagai peserta program pembelajaran;

d. mediagnosis kemampuan subjek belajar yang mengalami kegagalan dalam menguasai kemampuan prasyarat yang diperlukan.

E. PRINSIP PENILAIAN DAN EVALUASI DALAM PENDIDIKAN

Ada beberapa prinsip penilaian proses dan pencapaian hasil belajar yang harus dijadikan landasannya. Prinsaip-prinsip tersebut antara lain:

1. Prinsip integralitas/menyeluruh

Bahwa penilaian terhadap setiap subjek belajar harus komprehensif, mencakup seluruh aspek, baik yang menyangkut kemampuan (ability) dan personalitas (*aptitude*), atau menyangkut aspek pengetahuan, keterampilan, sikap, dan perilaku. Untuk itu diperlukan banyak teknik evaluasi yang harus diterapkan karena setiap macarn penilaian memerlukan teknik tersendiri dan setiap teknik penilaian memiliki kelemahan.

2. Prinsip kontinuitas/berkeninambungan

Pelaksanaan penilaian terhadap setiap subjek belajar harus dilakukan secara kontinyu dan periodik, dengan harapan bahwa adanya kegiatan penilaian dapat berfungsi untuk membimbing perkembangan subjek belajar.

3. Prinsip objektivitas

Penilaian terhadap setiap subjek belajar harus bebas dari unsur yang bersifat subjektif, harus dapat dimaknakan/ditafsirkan dengan jelas dan tegas. Semakin banyak data yang dijadikan dasar penilaian, maka hasil penilaian akan semakin objektif.

4. Prinsip berorientasi pada tujuan

Penilaian hasil belajar terhadap setiap subjek belajar hendaknya dimaksudkan untuk mengetahui seberapa jauh subjek belajar menguasai tujuan yang ditargetkan.

5. Prinsip terbuka

Proses dan hasil belajar terhadap setiap subjek belajar perlu diketahui oleh semua pihak. Oleh karena itu hasil penilaian harus disebarluaskan (dapat diketahui dan

diterima) oleh pihak-pihak yang terkait (siswa, orang tua, sekolah, pemerintah dan masyarakat).

6. Prinsip kebermaknaan

Hasil penilaian terhadap setiap subjek belajar harus memiliki kebermaknaan bagi orang yang menggunakan. Bagi guru, selain harus berguna untuk meningkatkan hasil belajar siswa juga untuk umpan balik dalam upaya memperbaiki proses pembelajaran. Bagi siswa juga harus berguna untuk memperbaiki diri dalam hal cara belajarnya agar pada penilaian berikutnya hasilnya akan lebih baik.

7. Prinsip kesesuaian

Penilaian terhadap setiap subjek belajar harus sesuai dengan pendekatan/ strategi/metode kegiatan pembelajaran yang diterapkan dalam rangka pelaksanaan kurikulum. Apabila dalam pelaksanaan kurikulumnya menggunakan pendekatan induktif, maka dalam penilaiannya juga harus menjadikan pendekatan induktif menjadi salah satu aspek yang dinilai. Kalau dalam pembelajarannya menerapkan metode eksperimen, maka kemampuan bereksperimen harus menjadi salah satu aspek yang dinilai.

8. Prinsip determinasi dan klarifikasi

Dalam melakukan penilaian terhadap setiap subjek belajar harus jelas apa yang akan dinilai, apakah tentang kemajuan belajamya ataukah pencapaian hasil akhirnya.

9. Prinsip mendidik

Hasil penilaian terhadap setiap subjek belajar hendaknya dapat digunakan untuk membina dan memberikan motivasi pada subjek belajar agar dapat meningkatkan hasil belajarnya. Hasil penilaian harus dinyatakan dan dapat dirasakan sebagai suatu penghargaan bagi subjek belajar yang berhasil dan merupakan peringatan bagi subjek belajar yang gagal Hasil penilaian yang dicantumkan dalam rapor hendaknya merupakan pertanggungjawaban subjek belajar yang bersangkutan kepada orang tuanya yang telah mempercayakan pendidikan anaknya kepada pihak sekolah/guru. Dengan demikian, penilaian dapat memperkuat perilaku dan sikap subjek belajar.

Prinsip penilaian menurut Grounlund dan Linn adalah

1. harus ada spesifikasi yang jelas apa yang mau dinilai: penempatan, formatif, ataukah sumatif;
2. harus komprehensif: afektif, psikomotor, dan kognitif;
3. butuh berbagai ragam teknik/metode penilaian, baik metode tes maupun nontes;
4. harus dapat memilih instrumen penilaian yang sesuai;
5. harus jelas apa maksud dan tujuan diadakan penilaian, jadi akan jelas pula apa tindak lanjutnya.

Prinsip penilaian hasil belajar peserta didik pada jenjang pendidikan dasar dan menengah menurut Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 adalah sebagai berikut.

1. sahih, berarti penilaian didasarkan pada data yang mencerminkan kemampuan yang diukur.
2. objektif, berarti penilaian didasarkan pada prosedur dan kriteria yang jelas, tidak dipengaruhi subjektivitas penilai.
3. adil, berarti penilaian tidak menguntungkan atau merugikan peserta didik karena berkebutuhan khusus serta perbedaan latar belakang agama, suku, budaya, adat istiadat, status sosial ekonomi, dan gender.
4. terpadu, berarti penilaian oleh pendidik merupakan salah satu komponen yang tak terpisahkan dari kegiatan pembelajaran.
5. terbuka, berarti prosedur penilaian, kriteria penilaian, dan dasar pengambilan keputusan dapat diketahui oleh pihak yang berkepentingan.
6. menyeluruh dan berkesinambungan, berarti penilaian oleh pendidik mencakup semua aspek kompetensi dengan menggunakan berbagai teknik penilaian yang sesuai, untuk memantau perkembangan kemampuan peserta didik.
7. sistematis, berarti penilaian dilakukan secara berencana dan bertahap dengan mengikuti langkah-langkah baku.
8. beracuan kriteria, berarti penilaian didasarkan pada ukuran pencapaian kompetensi yang ditetapkan.
9. akuntabel, berarti penilaian dapat dipertanggungjawabkan, baik dari segi teknik, prosedur, maupun hasilnya.

Prinsip evaluasi menurut Cronbach adalah sebagai berikut.

1. Kebijaksanaan pendidikan adalah kebijaksanaan pemerintah, sehingga evaluator harus membantu pemerintah.
2. Evaluasi adalah seni, evaluasi yang baik bukan mengandalkan hasil penilaian yang tunggal.
3. Evaluasi bukan keputusan yang absolut, tujuannya memilihkan altematif bagi pengambil kebijaksanaan untuk mengambil keputusan.
4. Tidak ada orang yang mampu membuat seluruh pertimbangan dalam merancang evaluasi dan menafsirkan hasil evaluasi.
5. Evaluator tidak hanya mengacu pada salah satu aliran evaluasi dalam melaksanakan evaluasi. Metode objektif kuantitatif dan humanistik kualitatif harus saling melengkapi.
6. Evaluasi sebagai suatu proses harus bersifat kontinyu dan luwes.
7. Evaluator harus mengidentifikasi setiap permasalahan yang berkait dengan evaluasi.
8. Program pendidikan bukan perlakuan tunggal, banyak faktor yang berpengaruh. Oleh karena itu, evaluator harus melihat bagian dalam proses, antar perlakuan, dan dalam populasi unfuk melihat besarnya pengaruh tindakan yang diberikan.
9. Dalam melakukan evaluasi, aspek psikomotor dan afektif tak boleh dikesampingkan
10. Evaluasi formatif dan sumatif harus menjadi satu kesatuan yang utuh, jadi harus melihat seluruh hasil dari pelaksanaan program, jangan hanya melihat hasil evaluasi sumatif.
11. Analisis keseluruhan lebih dapat dipertanggungjawabkan daripada hanya berdasar pada satu pengukuran.
12. Evaluasi sebagai suatu simpulan berdasarkan hasil penilaian harus inferensial, bukan hanya berdasar data tetapi juga berdasarkan asumsi.

Prinsip penilaian menurut BSNP (2007) untuk pendidikan dasar dan menengah mengacu kepada standar penilaian pendidikan jenjang pendidikan dasar dan menengah. Prinsip tersebut mencakup:

1. ***Sahih,*** yakni penilaian didasarkan pada data yang mencerminkan kemampuan yang diukur. Oleh karena itu, instrumen yang digunakan perlu disusun melalui prosedur sebagaimana dijelaskan dalam panduan agar memiliki bukti kesahihan dan keandalan.
2. ***Objektif***, yakni penilaian didasarkan pada prosedur dan kriteria yang jelas tanpa dipengaruhi oleh subjektivitas penilai. Oleh karena itu, dalam rangka meningkatkan objektivitas penilaian, pendidik menggunakan rubrik atau pedoman dalam memberikan skor terhadap jawaban peserta didik atas butir soal uraian dan tes praktik atau kinerja.
3. ***Adil***, yakni penilaian tidak menguntungkan atau merugikan peserta didik karena berkebutuhan khusus serta perbedaan latar belakang agama, suku, budaya, adat istiadat, status sosial ekonomi, dan gender. Faktor-faktor tersebut tidak relevan di dalam penilaian, sehingga perlu dihindari agar tidak berpengaruh terhadap hasil penilaian.
4. ***Terpadu***, yakni penilaian oleh pendidik merupakan salah satu komponen kegiatan pembelajaran. Dalam hal ini hasil penilaian benar-benar dijadikan dasar untuk memperbaiki proses pembelajaran yang diselenggarakan oleh peserta didik. Jika hasil penilaian menunjukkan banyak peserta didik yang gagal, sementara instrumen yang digunakan sudah memenuhi persyaratan secara kualitatif, berarti proses pembelajaran kurang baik. Dalam hal demikian, pendidik harus memperbaiki rencana dan/atau pelaksanaan pembelajarannya.
5. ***Terbuka,*** yakni prosedur penilaian, kriteria penilaian, dan dasar pengambilan keputusan dapat diketahui oleh pihak yang berkepentingan. Oleh karena itu, pendidik menginformasikan prosedur dan kriteria penilaian kepada peserta didik. Selain itu, pihak yang berkepentingan dapat mengakses prosedur dan kriteria penilaian serta dasar penilaian yang digunakan.
6. ***Menyeluruh dan berkesinambungan***, yakni penilaian mencakup semua aspek kompetensi dengan menggunakan berbagai teknik penilaian yang sesuai, untuk memantau perkembangan kemampuan peserta didik. Oleh karena itu, penilaian bukan semata-mata untuk menilai prestasi peserta didik melainkan harus mencakup semua aspek hasil belajar untuk tujuan pembimbingan dan pembinaan.
7. ***Sistematis***, yakni penilaian dilakukan secara berencana dan bertahap dengan mengikuti langkah-langkah baku. Oleh karena itu, penilaian dirancang dan

dilakukan dengan mengikuti prosedur dan prinsip-prinsip yang ditetapkan. Dalam penilaian kelas, misalnya, guru mata pelajaran agama menyiapkan rencana penilaian bersamaan dengan menyusun silabus dan RPP.

8. ***Beracuan kriteria***, yakni penilaian didasarkan pada ukuran pencapaian kompetensi yang ditetapkan. Oleh karena itu, instrumen penilaian disusun dengan merujuk pada kompetensi (SKL, SK, dan KD). Selain itu, pengambilan keputusan didasarkan pada kriteria pencapaian yang telah ditetapkan.
9. ***Akuntabel***, yakni penilaian dapat dipertanggungjawabkan, baik dari segi teknik, prosedur, maupun hasilnya. Oleh karena itu, penilaian dilakukan dengan mengikuti prinsip-prinsip keilmuan dalam penilaian dan keputusan yang diambil memiliki dasar yang objektif.

F. FUNGSI DAN TINDAK LANJUT EVALUASI

Evaluasi proses dan hasil belajar merupakan serangkaian kegiatan yang dilaksanakan tahap demi tahap berdasarkan keseluruhan hasil penilaian yang dilakukan. Adapun manfaat hasil evaluasi bagi subjek belajar adalah untuk bimbingan belajar, bimbingan pribadi, dan kebutuhan subjek belajar yang berkaitan dengan studinya. Jadi, meliputi aspek bimbingan dan aspek pembelajaran. Dengan demikian evaluasi proses dan hasil belajar akan berfungsi untuk memberi

1. arah dan petunjuk dalam pelaksanaan pembelajaran, baik guru maupun subjek belajar;
2. gambaran tentang diri subjek belajar mengenai perkembangan baik kemampuan maupun personalitasnya, sehingga mereka mampu mengenali diri/mawas diri serta seberapa jauh produktivitasnya, sehingga mampu menentukan langkah/keputusan lebih lanjut guna peningkatan prestasi.
3. dorongan/motivasi subjek belajar agar mampu berusaha untuk meraih prestasi yang lebih baik.
4. masukan untuk perbaikan dan pelaksanaan program guna memperbaiki proses pembelajaran yang akan diselenggarakan saat berikutnya.

Dengan mengetahui hasil evaluasi yang diperoleh, maka dapat diarnbil keputusan sebagai tindak lanjut. Keputusan tersebut sangat tergantung kepada jenis evaluasi yang telah dilaksanakan, yakni:

1. Keputusan atas dasar evaluasi sumatif

Jika sistemnya menggunakan sistem kenaikan, maka kegagalan subjek belajar dalam evaluasi sumatif dapat dinyatakan tinggal kelas. Di perguruan tinggi misalnya, setelah empat semester pertama dilakukan peninjauan terhadap hasil belajamya, jika tidak mencapai jatah yang ditetapkan, maka subjek belajar disuruh mengundurkan diri.

2. Keputusan atas dasar evaluasi formatif

Karena evaluasi formatif berfungsi untuk memberikan umpan balik, baik kepada subjek belajar ataupun guru, maka kegagalan kelas berarti merupakan masukan baik bagi guru maupun subjek belajar.

a. Bagi guru, kegagalan kelas berarti masukan untuk rnerevisi program pembelajaran yang telah disusun, baik dari segi metode, media, urutan kegiatan, guru, kegiatan pembelajaran yang dilakukan subjek, belajar, serta alokasi waktunya, sehingga dalam proses pembelajaran selanjutnya akan ada perbaikan hasil. Dengan demikian, akan dapat disusun suatu program pembelajaran yang lebih efektif yang lebih sesuai dengan kondisi, minat, bakat dan kecerdasan subjek belajar sebagai peserta program.

b. Bagi subjek belajar maka kegagalan dalam evaluasi formatif hendaknya digunakan untuk mawas diri dan mengambil langkah-langkah guna mengatasi kegagalan. Bagi yang belum memuaskan prestasinya dapat mengambil langkah-langkah baru guna meningkatkan prestasi.

3. Keputusan yang dianrbil berdasarkan evaluasi diagnostik

a. Jika evaluasi diagnostik dilakukan sebelum proses pembelajaran, maka tujuannya guru dapat mengambil sikap perlu tidaknya subjek belajar diberi pelajaran ekstra agar mampu menguasai prasyarat yang belum dikuasainya. Jika tanpa penguasaan prasyarat subjek belajar akan mengalami kesulitan dalam mengikuti proses pembelajaran , maka mau tidak mau guru harus meningkatkan penguasaan subjek belajar terhadap prasyarat tersebut.

b. Jika dilakukan selama proses, maka tugas guru untuk melakukan remediasi terhadap subjek belajar yang mengalami kesulitan belajar, sehingga tidak terjadi

kegagalan pada evaluasi sumatifnya. Dalam hal ini dapat dilakukan dengan berbagai cara, seperti remediasi akademik dilakukan dengan menggunakan sistem modul ataupun tugas terstruktur kalau tidak ada waktu khusus untuk tatap muka di kelas. Kendala utama adalah kalau beban mengajar guru terlalu banyak.

4. Keputusan yang harus diambil berdasar evaluasi penempatan

a. Subjek belajar yang tidak menguasai kemampuan prasyarat harus diremediasi (sebagai penilaian diagnosis sebelum proses pembelajaran).
b. Kemampuan yang sudah dikuasai subjek belajar tanpa harus melalui proses pembelajaran jangan ditargetkan sebagai kemampuan yang harus dikuasai melalui proses pembelajaran yang akan diselenggarakan.
c. Subjek belajar yang tidak interes atau berbeda karakteristik personalnya perlu mendapat perhatian khusus.

G. KEDUDUKAN PENGUKURAN DALAM EVALUASI

Evaluasi merupakan suatu proses penetapan untuk menyatakan baik atau buruk, berhasil atau gagal, sukses atau tidaknya sesuatu. Dengan kata lain evaluasi adalah proses untuk menentukan kualitas atau mutu sesuatu. Agar dapat dilakukan proses tersebut, maka harus ada data yang dijadikan dasar untuk penetapannya. Data harus benar-benar dapat dipercaya/akurat agar ketetapan yang diambil tidak salah. Jika salah datanya salah pula hasil penilaiannya dan akibatnya salah pula keputusannya. Data yang diolah untuk penilaian dan yang selajutnya untuk dasar evaluasi guna menetapkan apakah berhasil atau gagal atau apakah baik atau buruk itu harus data kuantitatif yang diperoleh melalui proses pengukuran. Jadi pengukuran merupakan proses untuk memperoleh kuantitas tentang sesuatu. Oleh karena itu diperlukan adanya alat atau instrumen penilaian yang dipakai dalam kegiatan pengukuran atau dapat pula disebut sebagai alat ukur. Ada alat ukur yang berupa tes dan ada pula alat ukur yang tergolong nontes. Agar dapat diperoleh alat penilaian atau alat ukur yang baik perlu dikembangkan suatu prosedur yang benar, dan disebut dengan prosedur pengembangan alat penilaian yang metriputi:

1. perencanaan penilaian yang memuat maksud dan tujuan penilaian,;
2. penyusunan kisi-kisi;
3. penyusunan instrumen/alat ukur;

4. penelahan (*review*) untuk menilai kualitas alat ukur/instrumen secara kualitatif, yakni sebelum digunakan;
5. uji coba alat ukur, untuk menyelidiki kesahihan dan keandalan secara empiris;
6. pelaksanaan pengukuran;
7. penilaian yang merupakan interpretasi hasil pengukuran;
8. pemanfaatan hasil penilaian.

Dalam hal ini, jika pengukuran dilakukan dengan menggunakan instrumen yang berupa tes, maka penyusunan instrumen harus mengacu pada cara-cara penyusunan suatu tes, demikian pula apabila instrumen yang digunakan berupa instrumen nontes maka cara penyusunan instrumen harus mengacu pada cara-cara penyusunan instrument nontes.

H. MACAM SKALA PENGUKURAN

1. Skala nominal

Skala nominal yaitu ukuran yang hanya mendasarkan pada prinsip tak tumpang tindih (*mutually exclusive*) dan tuntas (*exhaustive*), serta tidak ada asumsi tentang jarak maupun urutan/jenjang antar kategori yang ada di dalamnya. Angka-angka yang ada sekedar label untuk memisahkan kategori-kategori yang ada. Misalnya skala 1: SD Negeri, skala 2: SD Swasta atau sebaliknya; skala 1: Islam, skala 2: Kristen, skala 3: Katolik, skala 4: Hindu, dan skala 5: Budha, dapat pula sebaliknya; skala 1: laki-laki dan skala 2:perempuan atau sebaliknya, dan sebagainya.

2. Skala ordianal

Skala ordianal yaitu ukuran yang sudah mengurutkan objek yang diukur dari jenjang atau peringkat terendah sampai dengan yang tertinggi pada suatu atribut tertentu, tetapi tanpa petunjuk yang jelas tentang berapa jumlah/nilai absolut yang dimiliki oleh objek pada tiap atribut, dan jarak atau interval antara satu objek dengan yang lainnya juga tidak diketahui. Misalnya, untuk variabel tingkat pendidikan digunakan skala 1: SD, 2: SLTP, 3: SLTA, 4: PT Diploma sampai S1, 5: PT S2, dan 6: PT S3, tak berarti bahwa responden yang memiliki ijazah SD sekaligus MI sama dengan memiliki ijazah SLTP. Seseorang yang memiliki ijazah SMP sekaligus MTs setara

dengan memiliki ijazah S-1. Tingkat pendidikan SLTA juga tidak sama tiga kalinya tingkat pendidikan SD.

3. Skala interval

Skala interval yaitu ukuran yang menunjukkan kedudukan tiap subjek yang diukur, tetapi titik nolnya arbriter, misalnya skor hasil ulangan/ujian dalam penguasaan hasil belajar seperti penguasaan konsep untuk aspek kognitif.

4. Skala rasio

Skala rasio yaitu ukuran yang menunjukkan kedudukan tiap subjek yang diukur dan titik nolnya absolut, misal berat benda, panjang benda, volume benda dan sebagainya.

I. KESALAHAN PENGUKURAN

Adanya ketidaktepatan data atau data yang tidak dapat dipercaya kebenarannya dapat ditimbulkan oleh beberapa faktor berikut.

1. Kesalahan pada alat ukur

Seperti halnya dalam Ilmu Sosial pada umumnya, variabel-variabel dalam bidang pendidikan banyak yang bersifat abstrak fenomenanya. Oleh karena itu terhadap variabel yang demikian, sukar dilakukan pengukuran secara langsung. Jadi, tidak seperti mengukur tinggi peserta didik, panjang jangkauannya atau berat tubuhnya, karena fenomenanya benar-benar kongkrit. Guru tidak dapat mengukur kecakapan subjek belajar dalam mata pelajaran Biologi secara langsung misalnya. Kita juga tidak dapat secara langsung mengukur intelegensi subjek belajar. Dengan demikian jika fenomenanya sendiri abstrak, maka hanya dapat dilakukan pengukuran secara tak langsung. Persoalannya adalah bagaimana kita mampu menyusun alat ukur yang sahih (*valid*) dalam arti bahwa hasil pengukuran mencerminkan apa yang diukur. Kedua, bagaimana kita dapat menghasilkan alat ukur yang dapat memberikan data yang ajeg atau konsisten (*reliable*), yakni yang menghasilkan skor yang sama jika dilakukan pengukuran berulang-ulang. Variabel yang memiliki fenomena abstrak diukur secara tidak langsung yakni dengan menetapkan indikan-indikan/indikator-indikator dari variabel yang bersangkutan. Agar mampu mencari indikan yang tepat,

maka variabel yang akan diukur harus didefinisikan dengan tegas. Misalnya, apa yang dimaksudsubjek belajar telah menguasai Perlindungan Khusus dalam bidang Perawatan Gigi? Apakah subjek belajar harus hafal semua teks tentang Perlindungan Khusus atau mampu memecahkan berbagai persoalan perlindungan khusus terhadap gigi, atau terampil menguasai kemampuan dalam melakukan perlindungan khusus terhadap gigi, atau ketiganya? Demikian pula apa indikan-indikannya siswa berminat terhadap Biologi? Adanya definisi akan menegaskan apa sebenamya indikan-indikan dari variabel tersebut. Dengan adanya indikan yang tepat kemudian disusun pertanyaan-pertanyaan yang sesuai. Persoalan yang muncul adalah apakah pertanyaan yang dibuat sudah benar-benar mengukur indikan yang dimaksud. Kadang-kadang variabel yang akan kita ukur bersifat multidooensi. Sementara, dalam melakukan pengukuran seharusnya dilakukan terhadap satu dimensi atau harus bersifat unidimensional. Selain itu, suatu variabel harus bersifat unidimensional, dalam arti jika dilakukan pengukuran maka setiap hasil pengukuran dapat ditentukan letaknya pada garis yang menggambarkan keseluruhan harga dari variabel yang bersangkutan. Dengan kata lain hasil pengukuran harus dapat diplotkan pada garis abstrak (*abstract continuum*) yang merupakan garis variabel tersebut. Namun demikian, pada kenyataannya sering suatu variabel dapat mengandung banyak dimensi, sehingga pada saat melakukan pengukuran terhadap variabel tersebut harus mencakup pengukuran terhadap seluruh dimensinya. Untuk itu perlu dicari terlebih dahulu apa saja dimensi dari variabel yang akan kita ukur. Setelah diperoleh dimensinya baru kita jabarkan ke dalam indikan-indikar/indikator-indikatornya.

2. Kesalahan dalam proses pengukuran

Kesalahan ini ditimbulkan ketidaktelitian fihak yang melakukan pengukuran, seperti kesalahan dalam mengoreksi jawaban, kesalahan dalam memberikan skor. Kesalahan tersebut dalam proses pengukuran juga dapat diakibatkan karena kesalahan dalam menerapkan alat ukur terhadap pihak yang akan diukur; maksudnya, apakah pertanyaan-pertanyaan yang ada dalam alat ukur tersebut benar-benar cocok untuk pihak yang akan diukur.

3. Kondisi pihak yang diukur

Karena yang diukur adalah manusia, maka situasi pada saat dilakukan pengukuran akan sangat menentukan skor hasil pengukuran. Hasil pengukuran daiam kondisi stress akan lain dengan dalam kondisi yang bugar. Anak yang penakut, dalam kondisi cemas, akan memberikan hasil pengukuran yang lain jika dilakukan pengukuran terhadapnya pada situasi yang berbeda.

4. Kondisi/situasi selama dilakukan pengukuran

Karena yang diukur seseorang, maka situasi yang gaduh akan menghasilkan skor yang berbeda dibandingkan dengan situasi yang tertib dan tenang pada waktu dilakukan pengukuran, khususnya untuk anak-anak yang mudah terganggu oleh situasi tersebut.

5. Pengaruh pengulangan pengukuran

Untuk keperluan tertentu maka pengukuran dengan menggunakan alat ukur yang sama dilakukan secara berulang. Jika jarak antara pengukuran pertama dengan pengukuran uiang terlalu dekat, maka pihak ada efek *testing*. Efek *testing* merupakan efek yang ditimbulkan akibat subjek yang diukur masih memiliki ingatan yang baik terhadap alat ukur tersebut, sehingga ia masih ingat apa jawaban-jawaban yang diberikan pada pengukuran pertama. Sebaliknya, jika pengulangan pengukuran terlampau lama selang waktunya, maka kemampuan seseorang akan menurun kalau ia tidak memiliki kesempatan untuk berlatih kembali, sehingga akan terjadi hasil yang berbeda antara pengukuran pertama dengan pengukuran ulangnya.

J. KESALAHAN PENILAIAN

Setelah diperoleh data, kemudian dilakukan proses pengolahan data untuk mengambil keputusan akhir dalam menilai subjek belajar. Dalam hal ini, meskipun datanya benar dapat terjadi kesalahan pengambilan keputusan akibat oleh beberapa faktor. Faktor-faktor tersebut adalah sebagai berikut.

1. Kekeliruan dalam mengolah data

Hal ini akibat ketidaktelitian dalam mengubah skor menjadi nilai atau karena kesalahan cara pengolahan datanya.

2. Pengaruh penilaian sebelumnya

Penilai dapat terpengaruh oleh keadaan-keadaan yang mendahului. Misalnya, seorang anak biasanya mendapat nilai jelek, kemudian suatu saat mendapat nilai baik. Keadaan ini dapat mempengaruhi penilai untuk menurunkan nilai anak tersebut.

3. Menilai mahal atau murah

Kecenderungan seorang penilai untuk memberi nilai lebih rendah atau lebih tinggi, sehingga keluar istilah ada guru yang murah ada guru yang mahal tanpa ada dasar yang dapat percaya.

4. Pengaruh kesan luar

Penampilan pihak yang dinilai dapat berpengaruh terhadap si penilai. Seorang subjek belajar yang "suka mendebat" oleh guru tertentu dinilai jelek, walaupun sebenamya skor yang diperoleh selalu tinggi, karena guru tersebut tidak suka didebat. Sebaliknya, anak yang berpenampilan manis, penurut, akan diberi nilai yang lebih tinggi walaupun skornya rendah.

5. Pengaruh *hallo effect*

Kesan jelek terhadap seseorang yang dinilai akan dapat mempengaruhi pihak penilai. Seorang guru yang merangkap dua mata pelajaran, dapat terpengaruh kesan pada nilai dari saiah satu mata pelajaran. Jika dalam mata pelajaran A yang satu anak tersebut baik, maka ia dapat terpengaruh sehingga akan memberikan nilai yang baik pula dalam mata pelajaran B walaupun skor yang diperoleh dalam mata pelajaran B selalu rendah.

BAB II
OBJEK PENILAIAN

Dalam penilaian pencapaian hasil belajar, kedudukan subjek belajar merupakan subjek yang dinilai. Adapun sebagai objek penilaian adalah apa-apa yang melekat pada diri subjek belajar yang dijadikan sasaran penilaian. Adapun yang melekat pada diri subjek belajar selain variabel-variabel yang berhubungan dengan fisik subjek belajar juga yang berhubungan dengan nonfisik. Dalam hal ini berupa kemampuan (*ability*) dan personalitas (*aptitude*). Menurut Gorman kemampuan seseorang meliputi kemampuan intelektual dan kemampuan psikomotor. Kemampuan intelektual mencakup kemampuan penguasaan produk dan proses. Adapun yang termasuk produk adalah fakta, konsep dan struktur ilmu pengetahuan, sedangkan yang termasuk proses adalah kreativitas, pemecahan masalah, dan komprehensif. Kemampuan psikomotor berupa keterampilan (*skill*), sementara personalitas mencakup temperamen yang berupa tingkah laku (b*ehavior*), karakter, moral dan sikap (*attitude*).

Karena fungsi penilaian hasil belajar untuk mengetahui seberapa jauh “kompetensi yang ditargetkan” telah tercapai, maka kunci utama dalam melakukan penilaian adalah ketepatan dalam merumuskan indikator pencapaian kompetensi. Indikator tersebut menjadi kesatuan dengan Kompetensi Dasar yang ditargetkan untuk dicapai. Dengan kata lain, dalam melakukan penilaian harus diawali dengan perencanaan berupa menyusun kisi-kisi penilaian.

Dalam panduan penilaian yang diterbitkan oleh BSNP tahun 2007 dinyatakan bahwa kisi-kisi penilaian merupakan bagian yang tak terpisahkan dari kegiatan perencanaan pembelajaran dalam bentuk silabus dan rencana pelaksanaan pembelajaran (RPP). Di dalam silabus, pendidik menunjukkan keterkaitan antara SK, KD, materi pokok/materi pembelajaran, alokasi waktu, sumber belajar di satu sisi, dengan indikator pencapaian KD yang bersangkutan beserta teknik penilaian dan bentuk instrumen yang digunakan. Teknik penilaian dan bentuk instrumen dapat dituliskan dalam satu kolom, dan dapat pula dituliskan pada kolom yang berbeda. Berikut ini disajikan contoh format kisi-kisi penilaian yang menyatu dengan silabus.

Silabus Pembelajaran

Sekolah : .....
Mata Pelajaran : .....
Kelas/Semester : .....
Standar Kompetensi : .....

Kompetensi Dasar	Materi Pokok/ Materi Pembelajaran	Kegiatan Pembelajaran	Indikator Pencapaian	Penilaian		Alokasi Waktu	Sumber Belajar
				Teknik Penilaian	Bentuk Instrumen		

Perencanaan penilaian yang sudah dilengkapi dengan contoh instrumen disajikan secara menyatu dengan Rencana Pelaksanaan Pembelajaran (RPP). Berikut ini adalah contoh kisi-kisi penilaian yang menyatu dengan RPP.

Rencana Pelaksanaan Pembelajaran (RPP)

Sekolah : .....
Mata Pelajaran : .....
Kelas/Semester : .....
Alokasi Waktu : … jam pelajaran (… x pertemuan)
A. SK :.....
B. KD : .....
C. Materi Pembelajaran : .....
D. Model/Metode Pembelajaran : .....
E. Skenario/Langkah-langkah Kegiatan Pembelajaran
Pertemuan 1 : .....
Pertemuan 2 : .....
dst.
F. Sumber Belajar : .....

G. Penilaian

Indikator Pencapaian	Teknik Penilaian	Bentuk Instrumen	Contoh Instrumen
Berupa indikator yang ada di dalam rumusan silabus sesuai dengan KD yang bersangkutan	Dipilih sesuai dengan karakteristik indikator pencapaian, seperti tes tertulis, tes lisan, tes kinerja, dan portofolio.	Dipilih sesuai dengan teknik penilaian yang dipilih, misalnya memilih bentuk pilihan ganda untuk teknik penilaian tertulis atau memilih bentuk instrumen lembar penilaian portofolio untuk teknik penilaian portofolio.	Disusun sesuai dengan bentuk instrumen yang telah dipilih.

Untuk menilai pencapaian standar kompetensi dalam satu semester, pendidik merancang penilaian untuk semester yang bersangkutan. Kisi-kisi ulangan akhir semester memuat SK, KD, dan indikator pencapaiannya yang dapat dijadikan dasar penyusunan tes pada akhir semester. Kisi-kisi ulangan akhir semester dapat dirancang dengan memuat tes tertulis dan tes praktik yang formatnya disajikan sebagai berikut.

Kisi-Kisi Ulangan Akhir Semester

Sekolah : .....
Mata Pelajaran : .....
Kelas/Semester : .....
Alokasi waktu : .....

Standar Kompetensi	Kompetensi Dasar	Indikator Pencapaian	Teknik Penilaian	
			Tes Tertulis	Tes Praktik
Dituliskan seluruh SK dalam semester bersangkutan	Dituliskan KD yang esensial dari SK yang bersangkutan	Dituliskan indikator pencapaian yang esensial dari KD yang bersangkutan.	Dicantumkan bentuk butir tes yang dipilih, seperti benar-salah, menjodohkan, dan pilihan ganda	Dituliskan bentuk tes yang dipilih seperti tes keterampilan tertulis, tes identifikasi, tes simulasi, atau tes contoh kerja

Untuk tes tertulis, guru dapat membuat kisi-kisi tes tertulis untuk ulangan akhir semester seperti contoh berikut.

Kisi-Kisi Tes Tertulis Ulangan Akhir Semester

Sekolah : .....
Mata Pelajaran : .....
Kelas/Semester : .....
Alokasi waktu : .....

Standar Kompetensi	Kompetensi Dasar	Indikator Pencapaian	Bentuk Butir Tes			
			Pilihan Ganda	Uraian	.....	.....
Dituliskan seluruh SK dalam semester bersangkutan	Dituliskan KD yang esensial dari SK yang bersangkutan	Dituliskan indikator pencapaian yang esensial dari KD yang bersangkutan.	... butir	...butir	...butir	...butir

*) Keterangan: di bawah kolom bentuk butir tes diisi bentuk butir tes yang akan digunakan seperti pilihan ganda, uraian, dan menjodohkan

A. RANAH KOGNITIF

Menurut taksonomi Bloom, kemampuan intelektual atau kognitif meliputi jenjang sebagai berikut:

1. Ingatan (k*nowledge*):
   a. Ingatan tetang hal yang spesifik, baik ingatan tentang peristilahan (terminologi) maupun kejadian yang spesifik, misal menyebutkan bagian-bagian, menyebutkan istilah, nama, sifat, contoh, dan sebagainya; mengingat definisi, bagian-bagian, kejadian, tempat, dan sebagainya
   b. Ingatan tentang jalur-jalur dan arti dari hubungan-hubungan yang spesifik, baik ingatan tentang konvensi, kecenderungan (*trend*) dan urutan (s*equence*), klasifikasi dan kategori, kriteria serta metodologi.
   c. Ingatan tentang universalitas dan abstraksi di lapangan, misal mengingat/menyebutkan tentang prinsip-prinsip dan generalisasi-generalisasi, maupun teori-teori dan struktur-skturktur.

2. Pemahaman (comprehension) merupakan kemampuan terendah dari mengerti (*understanding*), meliputi:

   a. Translasi (penerjemahan), yakni kemampuan menterjemahkan/menjelaskan suatu maksud atau informasi, misal menyatakan kembali dengan kata-katanya sendiri tentang suatu definisi, maksud, contoh dan sebagainya.

b. Interpretasi (penafsiran), yakni kemampuan menafsirkan atau mengartikan suatu informasi, misal menjelaskan hal yang berhubungan atau yang ada relevansinya, mengurutkan ataupun menyusun kembali sesuai dengan urutannya, dan sebagainya.
c. Ekstrapolasi atau estimasi, yakni kemampuan untuk meramalkan kemungkinan-kemungkinan dari suatuinformasi,misalmendugaakibatlefek yang mungkin terjadi, memperkirakan faktor-faktor yang berpengaruh, menarik kesimpulan, dan sebagainya.
d. Jastifikasi, yakni kemampuan membenarkan, misal membenarkan suatu prosedur atau metode. Semuanya tanpa dihubungkan dengan penerapannya ataupun dihubungkan dengan hal-hal atau informasi yang lain.

3. Penerapan (*aplication*) meliputi kemampuan:
a. Menerapkan prinsip pada situasi yang baru.
b. Menerapkan teori ke dalam praktek.
c. Menerapkan rumus untuk pemecahan soal.
d. Menyrsun skema atau diagram dari datalinformasi yang tersedia.
e. Mendemonstrasikan suatu prosedur dengan benar.

4. Analisis (*analysis*) meliputi :
a. Analisis unsur-unsur, rnisal menemukan asumsi yang belum ada/belum dinyatakan dalam suatu informasi, membedakan kesimpulan yang berdasarkan fakta dan yang bukan, membedakan antara fakta dan pendapat.
b. Analisis hubungan-hubungan, misal dapat menemukan hubungan sebab-akibat, dapat membedakan antara alasan yang relevan dan yang tak relevan.
c. Analisis prinsip-prinsip yang terorganisasi, misal menemukan bentuk-bentuk, formula, pola atau struktur dalam suatu hal.

5. Sintesis (*synthesis*) meliputi:
a. Produksi/hasil suatu komunikasi yang unik/khas, misal membuat ringkasan, menyusun suatu alat dan sebagainya.
b. Produksi/hasil suatu rencana atau seperangkat usulan kegiatan, misal menyusun suatu rencan a kegiatan/rencana percobaan
c. Menurunkan/mencari derivat seperangkat hubungan abstrak, misal merumuskan hipotesis berdasar kajian pustaka yang ada.

6. Evaluasi (e*valuation*) meliputi:
a. Evaluasi berupa pertimbangan internal dari suatu kejadian, missal evaluasi dari segi ketepatan, kecermatan, konsistensi ataupun urutan logis.
b. Evaluasi berupa pertimbangan eksternal dari kejadian yang ada, missal evaluasi dari segi efisiensi, efektifitas, nilai ekonomis, atau dari segi makna.

B. RANAH AFEKTIF

Ranah afektif menurut Krathwohl (1964) meliputi jenjang sebagai berikut.

1. Kemampuan menerima (*receiving*)
   a. Kesadaran (*awareness*), misal membedakan suara, memilah kejadian, memilih rencana, menunjukkan kesadaran tentang pentingnya belajar, menunjukkan sensivitas terhadap problem-problem sosial.
   b. Kemauan untuk menerima (*willingness to receive*), misal memilih contoh, mengkombinasi bentuk, mengumpulkan model, mendengarkan dengan perhatian penuh, menerima perbedaan suku serta budaya, melibatkan diri secara penuh terhadap aktivitas kelas.
   c. Perhatian yang terkontrol atau terseleksi (*controlled or selected attention*), misal memilih alternatif, mengontrol jawaban.

2. Kemampuan merespon (*responding*)
   a. Persetujuan pada dirinya untuk sepenuhnya merespon (*acquiescence in responding*) dan kemauan untuk merespon (*willingness to respond*), misalnya mengikuti/menyetujui aturan yang berlaku, menghargai pendapat atau kebijaksanaan, menyelesaikan tugas rumah ataupun tugas laboratorium, ikut dalam suatu kegiatan secara sukarela (sebagai volunteer), aktif dalam diskusi kelas.
   b. Kepuasan dalam respon (*satisfaction in response*), misai menyambut dengan gembira keputusan yang diambil bersama, dengan tulus memuji karya/penampilan orang lain

3. Kemampuan menilai atau memaknakan (*valuing*)
   a. Kernarnpuan menerima secara baik suatu nilai (*acceptance of a value*), misalnya meningkatkan kecakapannya dalam hubungan personal ataupun dalam klubnya, mengkhususkan diri menetapkan pilihan pada suatu hal, menghargai peran serta ilmu pengetahuan dalam kehidupan sehari-hari, menghargai literature yang baik.
   b. Menentukan pilihan terhadap suatu nilai (*preference for a value*), missalnya memberikan bantuan terhadap suatu proyek/tencana, medukung argumen orang lain.
   c. Tanggung jawab, misal mendebat hal-hal yang tak relevan, mengajukan argumentasi atas jawaban yang diberikan, memprotes hal-hal yang tidak benar, menunjukkan tanggung jawabnya terhadap pembaharuan sosial/masyarakal, menunjukkan sikap mau memecahkan masalah, menunjukkan perhatian yang besar terhadap kesejahteraan sesamanya,

4. Kemampuan mengorganisasi (*organizing*)
   a. Konsepsualisasi nilai (*conceptualization of a value*), misal membandingkan dengan suatu standar, mendiskusikan parameter-parameter, menghargai kebutuhan yang seimbang antara kebebasan dan tanggung jawab, menghargai

peran perencanaan yang sistematik dalam pemecahan masalah, mengakui kelebihan dan kelemahan diri.

b. Organisasi sistem nilai (*organization of a value system*), misalnya menyusun kriteria, mengorganisasi sistem, menyusun rencana sesuai dengan minat, tanggung jawab, serta keyakinannya.

5. Kemampuan yang dikarakterisasi oleh suatu nilai atau gabungan nilai (*value complex*) yang akan terbentuk suatu *life stile*.
   a. Generalized set, misalnya menyusun rencana, mengubah perilaku, melengkapi cara, memilih prosedur.
   b. Karakterisasi (*characterizing*), misalnya dinilai baik oleh teman-teman, oleh guru ataupun oleh anggota kelompoknya, menghindari konflik, menentang tindakan yang boros, mengatasi akibat yang tak dikehendaki, menunjukkan kepercayaan diri dalam kerja individual, menggunakan pendekatan objektif dalam memecahkan masalah, menunjukkan disiplin dan produktivitas yang tinggi, mampu bekerjasama dalam kerja kelompok, memelihara cara hidup yang sehat, menunjukkan kesadaran yang tinggi.

Hubungan antara ranah kognitif dan afektif secara rasional dapat direlasikan sebagai berikut.

1. Kontinum kognitif dimulai dengan kemampuan mengingat dan mengenal kembali pengetahuan yang dimiliki	1. Kontinum afektif dimulai dengan penerirnaan belaka stimuli dan secara pasif mengikuti sesuatu hal dan dilanjutkan dengan lebih aktifnya mengikuti hal tersebut
2. Dilanjutkan dengan pemahaman terhadap pengetahuan yang dimiliki	2. Dilanjutkan dengan respon terhadap suatu stimulus atas dasar permintaan, kemudian munculnya kemauan untuk merespons dan timbulnya kepuasan dalam merespon
3. Keterampilan dalam menerapkan pengetahuan dimiliki	3. Penilaian/penghargaanterhadap fenomena-fenomena atau aktivitas agar secara sukarela merespons dan mencari jalan/cara untuk merespon.
4. Keterampilan dalam menganalisis situasi dengan menggunakan pengetahuan yang dimiliki dan menyintesis pengetahuan ke dalam organisasi nengetahuan yang baru	4. Konsepsualisasi tiap penilaian atau penghargaan terhadap sesuatu yang ditampilkan
5. Kemampuan menilai/mengevaluasi dalam ruang lingkup pengetahuannya untuk mengritik/mengadili/menetapkan nilai/bobot dari sesuatu hal dan metode/cara untuk menyampaikan keinginan	5. Kemampuan mengorganisasi atau menyusun nilai yang kompleks ke dalam suatu nilai yang utuh sebagai kar akter (*life style*) seseorang

C. RANAH PSIKOMOTOR

Ranah psikomotor menurut Harrow mencakup:

1. Gerak refleks (*reflex movements*): merupakan gerak yang otomatis, yang tidak dapat dilatih, terdiri atas
   a. Refleks segmental (*segmental reflexes*) yang melibatkan segmen spinalis:
      1) Refleks fleksi, berupa refleks yang melibatkan aiat gerak (kaki dan tangan).
      2) Refleks miotatik, berupa refleks yang menyebabkan peregangan otot extensor (otot antigravitasi) pada saat tubuh menjaga keseimbangan badan.
      3) Refleks ekstensor, berupa reaksi anggota badan saat tubuh melawan gaya berat ketika berjalan/berlari.
      4) Reaksi ekstensi silang, berupa gerakan tangan yang berlawanan dengan gerakan kaki pada saat berjalan.
   b. ReIIeks intersegmental: yaitu gerak refleks yang melibatkan lebih dari sebuah segmen spinalis:
      1) Refleks kooperatif, jika dua atau lebih gerak refleks saling membantu satu sama lain dalam pola lebih memperlancar.

2) Refleks kompetitif, jika satu gerak refleks menghambat gerak refleks yang lain atau dua gerak refleks yang terjadi berlawanan.
3) Refleks induksi suksesif, jika dua gerak refleks yang berlawanan terjadi saling menyusul/berebut, misal refleks fleksi berganti dengan refleks ekstensi pada saat orang berlari.
4) Figur refleks (*reflex figure*) yaitu interaksi dari banyak refleks, misalnya kombinasi refleks keempat angota badan pada saat orang berlari.

c. Refleks suprasegmental (*suprasegmental reflexes*) yaitu reflex yang memerlukan peran serta pusat-pusat di otak, sepanjang medula spinalis, dan otot-otot anggota gerak maupun tubuh yang mendukung suatu gerakan:
1) Refleks ekstensor yang mengeras/tegar, yang ditunjukkan oleh seluruh otot ekstensor/antigravitasi pada lengan dan kaki untuk menjaga tubuh tetap tegak.
2) Reaksi-reaksi melentur, yaitu refleks yang mengerut atau menjulur sebagai kebalikan dari otot ekstensor.
3) Refleks postur, yaitu refleks yang mendukung kedudukan tubuh saat tubuh/anggota badan bergerak.

2. Gerak dasar pokok (*basic-foundamental movements*):
a. Gerak-gerak lokomotor (*locomotor movements*) yaitu gerakan yang mengakibatkan tubuh berpindah tempat seperti berjalan, berlari, merangkak dan sebagainya,
b. Gerak-gerak nonlokomotor (*nonlocomotor movements*) yaitu gerak-gerak yang terjadi pada sebagian tubuh/anggota badan, seperti membungkuk, menengadah, merentang tangan dan sebagainya.
c. Gerak-gerak manipulatif (*manipulative movements*) yaitu gerak-gerak kombinasi dari bagian tubuh/anggota badan, seperti gerak orang mengetik, bermain biola dan sebagainya.

3. Kemampuan perseptual (*perceptual abilities*) :
a. Pembedaan kinestetik (*kinesthetic discrimination*):
1) Kesadaran posisi tubuh (*body awareness*):

a) secara bilateral, yakni gerak yang menampilkan kedua sisi tubuh, misalnya gerak menangkap bola berukuran besar.
b) secara lateral, yakni gerak yang menonjolkan satu sisi tubuh, misal mengoper bola dari tangan kiri ke tangan kanan atau sebaliknya.
c) keunggulan (*dominance*), yakni gerak yang lebih dominan antar anggota gerak, misal antara tangan kanan dan tangan kiri, antara tangan dan kaki.
d) Keseimbangan (*balance*), yakni kemampuan menjaga posisi tubuh agar tetap seimbang pada saat melakukan berbagai gerakan bagian tubuh/anggota badan.

2) Kesan posisi tubuh (body image), yakni kesadaran tentang tubuhnya saat seseorang akan memulai gerakan, sehingga mampu mengira-ngira seberapa ia harus melakukan gerakan, misal saat orang harus melangkahi parit.
3) Kesadaran posisi tubuh terhadap objek sekitar (*body relationship to surrounding objects in space*).

b. Pembedaan menurut penglihatan/secara visual (*visual discrimination*) :
1) Kemampuan membedakan berdasarkan ketajaman penglihatan (*visual acuity*).
2) Kemampuan mengikuti arah gerak berdasar penglihatannya (*visual tracking*), misal kemampuan mengikuti gerak pesawat di udara.
3) Kemampuan merekam apa yang dilihat sehingga mampu menirukan gerakan yang baru saja dilihat (*visual memory*).
4) Kemampuan menyeleksi gambaran yang dominan dari sekitamya, misalnya mampu menebak kearah mana bola akan jatuh, ia bergerak sehingga ia mampu menangkapnya (*figure-ground dffirentiation*).
5) Kemampuan membedakan objek yang berbeda dan kemampuan mengelompokkan objek yang sama (keajegan/*consistency*).

c. Pembedaan menurut pendengaran (*auditory discrimination*) :
1) Kemampuan membedakan berbagai suara berdasar nada maupun intensitasnya (*auditory acuity*).
2) Kemampuan mengikuti/menentukan sumber suara (*auditory tacking*).
3) Kemampuan merekam suara yang baru didengar, sehingga mampu menirukannya kembali, menirukan nyanyian burung (*auditory memory*).

d. Pembedaan berdasar rabaan (*tactile discrimination*)

Kemampuan-kemampuan terkoordinasi (*coordinated abilities*), yakni kemampuan/ aktivitas yang melibatkan dua atau lebih kemampuan persepsual:

1) Aktivitas yang memerlukan kemampuan tangan dan penglihatan, misalnya kemampuan memukul bola, kemampuan mengukur menggunakan alat ukur (*eye-hand coordination*).
2) Aktivitas yang memerlukan kemampuan kaki dan pengiihatan, misalnya kemampuan menggirng bola dalam bola sepak (*eye-foot coordination*).

4. Kemampuan fisik (*physical abilities*):

a. Daya tahan (*endurance*), yakni kemampuan tubuh memasok oksigen pada diri seseorang untuk dapat melakukan gerakan yang kontinyu.

1) Daya tahan otot (*muscular endurance*), yakni daya tahan untuk tetap bekerja dalam waktu yang lama.
2) Daya tahan kardiovaskuler (*cardio-vascular endurance*), yakni daya tahan bekerja berat secara terus menerus untuk periode beberapa lama yang membutuhkan kerja interaksi jantung-pembuluh darah paru-paru, misalnya berlari sambil mengumpulkan kartu sebanyak-banyaknya.

a. Kekuatan (*strength*), yakni kemampuan mengerahkan kekuatan, misal latihan angkat beban, *push-up* dan sebagainya.
b. Fleksibilitas (f*lexibility*), yakni kemampuan bergerak dengan tenaga yang efisien, misal gerak gimnastik.
c. Ketangkasan (*agility*) yaitu kemampuan untuk bergerak cepat, baik dalam menanggapi rangsang, memulai/mengakhiri sesuatu, kemampuan mengubah arah gerakan ataupun kemampuan mengganti gerakan.

5. Gerak terlatih (*skilled movements*):

Merupakan gerakan hasil suatu efisiensi pada saat seseorang disuruh melakukan tugas, termasuk di dalamnya adalah semua bentuk adaptasi pola gerak terpadu dari gerak-gerak dasar pokok (*basic-foundamental movements*) :

a. Keterampilan adaptif sederhana (*simple adaptive skill*) merupakan berbagai penyesuaian gerak-gerak dasar pokok yang diubah atau disesuaikan dengan situasi baru, misal gerakan menggergaji merupakan penyesuaian/perpaduan gerak menarik dan mendorong; berdansa adaptasi gerak bejalan, keterampilan melakukan titrasi

merupakan perpaduan menggoyang labu Erlenmeyer sambil mengawasi perubahan warna untuk segera menghentikannya, dan sebagainya. Keterampilan ini dapat dibedakan mulai dari tingkat pemula (*beginner*), menengah (*intermidiate*), lanjut (*advanted*) dan tingkat tinggi/benar-benar terampil (*highly skilled*).

b. Keterampilan adaptif gabungan (*compound adaptive skill*) yang merupakan perpaduan dua atau beberapa keterampilan adaptif sederhana, misalya keterampilan bermain tenis merupakan perpaduan gerakan memukul dan kemampuan menggunakan alat berupa raket.

c. Keterampilan adaptif kompleks (*complex adaptive skill*) merupakan perpaduan banyak keterampilan adaptif sederhana yang memerlukan penguasaan lebih besar, misal main selancar sambil bersalto, melakukan roll pada senam lantai sambil menangkap bola.

6. Komunikasi berkesinambungan (*non-discursive communication*) yaitu gerak yang dilakukan untuk komunikasi baik dalam bentuk ekspresi wajah ataupun gerak isyarat lainnya :

   a. Gerak ekspresif (*expressive movement*) yaitu gerak-gerak untuk menunjukkan suatu ekspresi seperti dalam kehidupan sehari-hari, misal berekspresi sebagai orang yang sedang marah, sedang bergembira dan sebagainya.

   b. Gerak interpretatif (*interpretative movement*): gerak dengan maksud tertentu.

   c. Gerak aestetik (*aesthetic movement*), yaitu gerakan-gerakan ditampilkan untuk menciptakan gerak y ang indah/cantik.

   d. Gerak kreatif (*creative movement*) yaitu gerakan-gerakan untuk mengkomuni-kasikan suatu pesan atau sesuatu yang lebih baru yang didukung oleh kemampuan fisik serta kemampuan persepsual.-

Pembagian ranah psikomotor menurut Simpson mencakup jenjang sebagai berikut.

1. **Persepsi**, yakni kemampuan menangkap stimulus, menyeleksi isyarat, dan kemampuan mentranslasinya ke dalam aksi yang ditampiikan, misal dapat menunjukkan adanya gangguan mesin berdasarkan suara yang didengarnya, dapat menghubungkan irama musik dengan langkah-langkah gerakan saat menari.
2. Kesiapan untuk berperan aktif (*set*) dalam suatu bagian dan kegiatan, baik secara menial, fisik maupun emosional, misal mengetahui urut-urutan langkah suatu kegiatan,

menunjukkan langkah yang efisien untuk melaksanakan suatu kegiatan, mendemonstrasikan cara berposisi yartg benar pada saat akan memulai suatu kegiatan.

3. Respons terpandu (*Guided response*), merupakan kemampuan awal dalam belajar suatu keterampilan yang bersifat kompleks, termasuk kemampuan menirukan ataupun kemampuan-mencoba berdasarkan kriteria atau instruksi, misalnya mendemonstrasikan cara memukul bola, mendemonstrasikan cara menggosok gigi geraham, mendeterminasi langkah-langkah pokok dalam rnelakukan peiawatan untuk mebuang karang gigi.
4. Mekanisme, yaitu -rnutnpilkan suatu kegiatan yang sifatnya habitual sehingga menghasilkan suatu keterampilan (skilt), misal merangkai alat laboratorium, menlgunakan mikroskop sehingga sampai dapat menemukan bayangan benda yung ingitt dilihatnya, menggunakan *slide projector.*
5. Respons yang benar-benar kompleks (*complex overt response*),yaitu menunjukkan keterampilan secara utuh, misalnya memperagakan cara menggergaji, memperagakan cara berenang menggunakan suatu gaya atau berganti gaya, memperagakan cara mengemudikan kendaraan, memperagakan cara membersihkan karang gigi,atau mendemontrasikan cara menambal gigi.
6. Adaptasi, yakni kemampuan mengubah-ubah pola gerakan karena adanya masalah yang dihadapi, misal membelokkan mobil saat menghindari kubangan, mengubah gerakan tangan saat berenang dalam menghadapi arus yang berputar.
7. Originasi, yaitu berkreasi menilorkan suatu gerakan baru yang benar-benar orisinal, misalnya menciptakan tari-tarian atau menciptakan mode baru dalam disain pakaian.

D. TAKSONOMI BLOOM TERBARU

Kemampuan pada diri manusia dalam taksonomi Bloom dipisahkan menjadi tiga domain, yakni domain (a) kognitif, (b) afektif, dan (c). psikomotor. Sekarang kemampuan manusia dalam taksonomi Bloom yang telah diperbaharui dipisahkan menjadi empat domain (Dettmer, 2006: 71-73). Empat domain tersebut adalah domain (a) kognitif, (b) afektif, (c) sensorimotor (sebagai pengganti psikomotor), dan (d) sosial. Keempat domain tersebut sebagai aktualisasi dalam pembelajaran membentuk satu kesatuan yang disebut dengan *unity*. Secara keseluruhan domain dan jenjang dari masing-masing domain dapat dilihat dalam Tabel 1. Keempat domain pada Tabel 1 memiliki jalinan satu sama lain dalam kaitannya dengan aktivitas pembelajaran dan melakukan sesuatu hal (*learning and*

doing). Kemampuan berkreasi merupakan puncak dari domain kognitif yang dapat ditumbuhkembangkan agar dimiliki seseorang. Konsep Bloom yang baru memaparkan bahwa pembelajaran dapat dibedakan menjadi pembelajaran dasar (*basic learning*), pembelajaran terapan (*applied learning*), dan pembelajaran ideasional (*ideational learning*). Ketiga bentuk pembelajaran tersebut tidak dapat terlepas dari target yang ingin dicapai (Dettmer, 2006: 70-78)

Ciri pembelajaran dasar adalah realisme (apa yang akan peserta didik ketahui). Isi/konten bersifat esensial. Perolehan aspek kognitif berupa proses mengetahui dan memahami. Pembelajarannya bersifat rudimenter. Konsep yang dipelajari sangat diperlukan dan harus dikuasai oleh peserta didik. Pendidik mengajarkan apa yang harus dipelajari peserta didik. Oleh karena itu, konsep diajarkan dalam bentuk proses yang terstruktur dan dengan domain isi yang standar. Peserta didik yang belum menguasai harus diberi waktu tambahan.

Tabel 1
Domain yang Dikembangkan dalam Pembelajaran

No	**Domain kognitif**	**Domain afektif**	**Domain sensorimotor**	**Domain sosial**	**Kesatuan (*Unity*)**
1	Mengetahui (*know*)	Menerima (*receive*)	Mengamati (*observe*)	Menghubung-kan (*relate*)	Merasa (*perceive*)
2	Memahami (*comprehend*)	Menanggapi (*respond*)	Bereaksi (*react*)	Berkomunikasi (*communicate*)	Mengerti (*understand*)
3	Menerapkan (*apply*)	Menilai (*value*)	Beraktifitas (*act*)	Berpartisipasi (*participate*)	Menangani/berbuat/ menggunakan untuk mencapai sesuatu (*use*)
4	Menganalisis (*analysis*)	Mengorganisasi (*organize*)	Beradaptasi (*adapt*)	Bernegosiasi (*negotiate*)	Memilah/memukenali penyebab perbedaan (*differentiate*)
5	Mengevaluasi (*evaluate*)	Menginternali-sasi (*internalize*)	Melakukan aktivitas yang sesungguhya (*authenticate*)	Memutuskan berdasarkan pertimbangan (*adjudicate*)	Memvalidasi/ membenarkan/ atau menunjukkan yang sebenarnya (*validate*)
6	Menyintesis (*synthesize*)	Mengkarakteri-sasi (*characterize*)	Mengharmo-nisasikan beberapa hal (*harmonize*)	Berkolaborasi (*collaborate*)	Berintegrasi (*integrate*)
7	Berimajinasi (*imagine*)	Mengagumi (*wonder*)	Berimprovi-sasi (*improvise*)	Berinisiatif (*initiate*)	Berspekulasi/berani menempuh risiko (*venture*)
8	Berkreasi (*create*)	Beraspirasi/ menggagas (*aspire*)	Berinovasi (*innovate*)	Mengkonversi ke hal baru (*convert*)	Melakukan sesuatu yang orisinal (*originate*)

(Sumber: Dettmer, 2006: 73).

Ciri pembelajaran terapan adalah pragmatisme (apa yang dapat peserta didik perbuat). Pembelajaran ini bersifat pengembangan sehingga penekanan pada penerapan, analisis dan evaluasi. Oleh karena itu, pembelaaran ini sudah bersifat kompleks dan bersifat individual bagi setiap peserta didik. Pendidik hanya membimbing (tidak mengajarkan) agar kemampuan aplikasi perserta didik dapat tumbuh. Konten/isi sangat penting, proses berlangsung luwes, dan domain isi menyesuaikan. Capaian hasil yang diharapkan dapat bervariasi dan kesempatan pembelajaran disediakan sebagai tantangan bagi setiap peserta didik.

Pembelajaran ideasional dikarakterisasi oleh idealisme atau aspirasi peserta didik. Pembelajaran bertumpu pada apa yang menjadi aspirasi peserta didik untuk mencapai perolehan sampai pada tataran inovasi atau hal-hal baru. Perolehan dari aspek kognitif mencakup: (a) proses menyintesis dari berbagai komponen untuk menghasilkan satu gabungan yang punya arti, (b) berimajinasi dalam arti menciptakan dan menjelajah gambaran mental dari situasi yang tidak tersajikan secara fisik, dan (c) berkreasi dalam arti menciptakan hal-hal yang baru yang berbeda dengan yang sudah ada. Pembelajaran ideasional menjadi bersifat personal bagi setiap peserta didik. Pendidik hanya sebagai fasilitator yang bertugas melayani peserta didik sehingga peserta didik ”terbangkitkan” untuk menemukan hal baru. Konten/isi pembelajaran berupa hal-hal yang baru. Proses pembelajaran bersifat *open-endend* dan untuk mengembangkan domain yang mendukung keunikan. Hasil belajar yang berbeda justru diharapkan, dan dorongan diberikan kepada setiap peserta didik untuk dapat memenuhinya.

BAB III
VALIDITAS DAN RELIABILITAS

Suatu alat ukur dinyatakan sahih (valid), jika alat ukur tersebut benar-benar mampu memberikan informasi empirik sesuai dengan apa yang diukur. suatu alat penimbang dinyatakan sahih jika benda yang beratnya 1 kg akan terukur seberat 1 kg pula. Alat timbangan tersebut dinyatakan benar-benar memiliki akurasi/keakuratan/ketelitian yang tinggi.

Hakekat kesahihan dalam ilmu sosial selain terkait dengan kelayakan suatu alat ukur, juga berkait dengan masalah derajat. Jadi bukan masalah sahih dan tidak sahih. Kesahihan selalu bersifat spesifik kaitannya dengan tujuan dan interpretasi. Kesahihan juga mencerminkan ketunggalan (unidimensionalitas) konsep yang akan diukur.

Selain itu alat ukur juga harus memiliki sifat andal (*reliable*). Artinya jika dipakai untuk mengukur secara berulang-ulang selalu tetap/konsisten/stabil hasilnya. Jadi, jika penimbangan pertama 50 kg, penimbangan kedua juga menunjukkan berat 50 kg, demikian pula pada ulangan yang selanjutnya. Alat ukur yang andal memiliki presisi/ketepatan/kecermatan yang tinggi.

Suatu alat ukur dapat saja memiliki keandalan yang tinggi namun tidak sahih. Misalnya, suatu benda yang beratnya 50 kg setelah ditimbang dengan alat penimbang berulang-ulang hasilnya selalu menunjukkan 49 kg.

Kecuali itu, hakekat keandalan suatu alat ukur adalah mengacu kepada hasil yang diperoleh bukan alat ukurnya itu sendiri. Keandalan alat ukur juga tidak berlaku secara umum, ganti waktu atau ganti sampel mungkin sudah dapat berubah keandalannya. Suatu alat ukur yang terandalkan tidak otomatis sahih, dan keandalan suatu alat ukur juga sangat statisTPK.

Menurut Embretson & Gorin (2001) validitas konstrak adalah sentral untuk menetapkan mutu tes. Secara tradisional, spesifikasi item sering hanya samar-samar. Pengembang tes membuat spesifikasi item yang sering kali hanya berisi pertimbangan-pertimbangan dari segi isi secara umum (seperti penetapan ruang lingkup/topik area dan ringkasan dari materi/isi) atau samar-samar menggambarkan pengolahan derajat pemrosesan berpikir (seperti abstrak melawan konkrit). Begitu keseluruhan spesifikasi item dihasilkan, dilakukan review item oleh reviewer untuk meyakinkan justifiabilas dari

kunci jawaban dan menguji konten/isi item untuk berbagai isu penyetaraan dan mutu tes. Metode-metode psikometrik diterapkan setelah item dikembangkan. Statistika item dari uji coba empiris menjadi sangat esensial untuk menentukan kualitas item, terutama untuk mengevaluasi bahwa proses-proses "konstrak-relevansi" telah terukur. Secara umum item-item yang tidak saling berkorelasi dengan item-item yang lain tidak dipilih. Prosedur tradisional yang standar tersebut sangat cocok dengan paradigma validitas konstrak menurut teori klasik.

Berdasarkan prinsip psikologi kognitif, hal pertama yang terpenting dalam pemenuhan validitas konstrak suatu tes kemampuan/abilitas adalah adanya dukungan satu set prinsip teoritis dari item-item tes yang akan disusun. Kedua, suatu set standar yang baru yang disusun harus didasarkan pada prinsip-prinsip psikologi kognitif. Mengutip pendapat Messick (1995) bahwa item-item yang disusun tidak sekedar cocok/sesuai dengan kriteria/ukuran tradisional, seperti tingkat kesulitan yang sesuai dan daya pembeda yang tinggi, tetapi item-item tersebut juga dibenarkan sebagai bagian juga relevan dengan konstrak dari proses-proses kognitif. Ketiga, bahwa tes yang disusun adalah untuk mengindikasikan kecakapan atau untuk mendiagnosis ketrampilan yang ada sehingga memerlukan bermacam informasi baru tentang item-item yang disusun.

A. MACAM-MACAM VALIDITAS

Kesahihan dapat ditinjau dari beberapa aspek, yakni :

1. Kesahihan isi (*Content validity*)

Yang menjadi pertanyaan pokok yang berkait dengan kesahihan isi adalah: "apakah sampel gugus alat ukur yang digunakan sudah represntatif?" Pertanyaan tersebut muncul karena tujuan pengukuran adalah untuk mengetahui kemampuan atau ranah tertentu pada diri subjek. Jika alat ukur tersebut berupa alat tes, maka butir-butir pertanyaan yang ada diharapkan mampu mencerminkan kemampuan subjek setelah ia dikenai tes.

Karena dalam proses pembelajaran tujuan yang ditargetkan sudah dirumuskan dalam bentuk kompetensi (SK dan KD), maka persoalannya bergeser menjadi: apakah item instrumen yang digunakan sudah mencerminkan seluruh KD yang ditargetkan. Jika tidak seluruh KD dapat diukur, maka pertanyaannya adalah: "apakah item instrumen yang digunakan sudah mewakili seluruh KD yang akan diukur. Jadi dalam hal ini

berbicara tentang representasi butir soal. Karena tinjauannya dikembalikan kepada KD yang ditargetkan, maka kesahihan isi juga disebut Kesahihan Kurikuler (*Curricular validity*).

Karena tujuan pengukuran juga untuk mengetahui sejauh mana kemampuan seseorang ditinjau dari segi ranahnya, maka tinjauan representasi dapat diartikan dari sisi kelengkapan ranah (*domain*) yang akan diukur. Misalnya, jika yang akan diukur kemampuan kognitif seseorang, dan berdasarkan teori-teori yang ada orang yang bersangkutan harus sudah memiliki kemampuan kognitif mulai dari ingatan (*knowledge*) sampai dengan evaluasi (*evaluation*), maka pokok-butir soal yang digunakan dinyatakan terpenuhi kesahihan isinya alat uji yang digunakan meliputi butir soal jenjang ingatan, pemahaman, aplikasi, analisis, sintesis dan evaluasi.

2. Kesahihan konstruk (*Construct validity*)

Persoalan pokok yang dikaji dalam rangka pemenuhan validitas konstruk adalah pertanyaan mengenai: apakah hasil pengukuran benar-benar makna atau benar-benar dapat diinterpretasikan, atau dengan kata lain apakah alat ukur yang digunakan benar-benar mampu mengukur. Jadi, adalah apakah setiap butir soal sudah mencerminkan indikator variabel yang hendak diukur. Jika setiap butir soal sudah mencerminkan indikator dari kemampuan yang akan diukur maka otomatis hasil pengukuran yang diperoleh benar-benar mengukur kemampuan sebagai variabel yang diukur. Jadi kalau ingin mengukur kemampuan kognitif anak SMA dalam mata pelajaran Biologi, siswa yang benar-benar memiliki kemampuan tersebut akan dapat mengerjakan seluruh butir soal yang diujikan.

Dalam bidang psikologi, termasuk di dalamnya psikologi pendidikan, pemenuhan pemenuhan kesahihan konstruk dikaitkan dengan sifat variabel yang akan diukur. Dalam hal ini, aspek psikologi memiliki variabel-variabel yang bersifat abstrak. Dengan demikian fenomenanya tidak dapat ditangkap secara langsung melalui pancaindera. Oleh karena itu perlu dicari indikator / indikan yang tepat agar variabel tersebut benar-benar dapat diukur. Hal ini akan mudah ditangkap jika dibandingkan dengan Ilmu Kimia. Fenomena reaksi kimia sulit ditangkap dengan pancaindera. Oleh karena itu cara yang dapat ditempuh adalah membubuhkan zat yang berfungsi sebagai indikator. Zat indiator tersebut tidak akan ikut berekasi, namun akan berubah warnanya

jika reaksi kimia sudah berlangsung. Dengan demikian kita akan mengatakan bahwa zat A sudah bereaksi dengan zat B karena warna zat indikator sudah berubah.

Setelah variabel yang akan diukur dijabarkan ke dalam indikator-indikatornya barulah disusun pertanyaan-pertanyaan yang mncerminkan masing-masing indikator tersebut. Dengan demikian, alat uji akan memiliki kesahihan konstruk jika pokok-butir soalnya mencerminkan indikator-indikator dari variabel yang akan diukur, juga ketepatan memilih indikator-indikator dari variabel yang diukur.

3. **Kesahihan dihubungkan dengan kriteria** (***Criterion-related validity***)

Dalam hal ini untuk melihat kesahehan suatu alat ukur justru dibandingkan dengan alat ukur yang lain yang standar atau yang baku sebagai patokan (kriteria). Alat ukur yang kita buat dinyatakan memiliki kesahihan yang tinggi jika hasil pengukurannya menunjukkan korelasi yang sangat positif dengan hasil pengukuran yang menggunakan alat ukur baku. Kesahihan yang demikian disebut kesahihan *concurrent*.

4. Kesahihan muka (*face validity*)

Kesahihan ini dibedakan menjadi dua. Pertama adalah kesahihan muka yang hanya dilihat berdasar penampakan luarnya. Misalnya, suatu butir soal mata pelajaran Biologi tentang klasifikasi mikroorganisme, maka butir soal tersebut dinyatakan memenuhi kesahihan muka jika yang ditanyakan adalah hal-hal yang menyangkut klasifikasi organisme. Kedua, kesahihan ditinjau berdasarkan pandangan orang yang lebih ahli, dapat berarti ahli dalam ilmunya ataupun ahli dari segi evaluasi atau dari segi bahasa. Oleh karena itu, minta pendapat orang yang lebih ahli akan meningkatkan kesahihan butir soal yang kita susun.

5. Kesahihan Antar Budaya (*Cultural Validity*)

Kesahihan alat ukur juga tercermin dari keseragaman daya tangkap subjek terhadap alat ukur yang bersangkutan. Suatu butir tes model kasus harus menyajikan kasus yang dipahami secara merata oleh para subjek yang diukur. Seorang yang memahami Biologi akan lebih terpandu mengerjakan soal bahasa Inggris dengan kasus Biologi daripada dengan kasus Sastra misalnya. Jenis-jenis tanaman tumpangsari sudah tidak perlu dihafal oleh anak-anak yang hidup di pedesaan, namun untuk anak kota menjadi pekerjaan rumah tersendiri.

Alat-alat ukur yang diadopsi dari luar negeri harus disimak benar apakah pertanyaan-pertanyaan yang ada di dalamnnya bebas dari faktor budaya. Tidak otomatis bahwa alat ukur yang sudah teruji kesahihannya di luar negeri dipertanyakan kembali jika digunakan di dalam negeri.

Kesahihan alat ukur juga dapat dipenuhi dengan membuat alat ukur sebaik mungkin antara lain dengan cara :

1. Petunjuk mengerjakan harus jelas. Jika alat ukurnya soal, maka perintahnya harus jelas agar subjek uji mengerti apa yang harus dikerjakan. Demikian pula jika berupa angket usahakan agar responden mengerti cara pengisiannya.
2. Struktur kalimat harus benar (lugas, komunikatif) dan kata-kata yang digunakan sesuai dengan kemampuan subjek yang menjadi sasaran.
3. Konstruksi alat ukur harus memenuhi persyaratan sesuai dengan jenis alat ukurnya. Misalnya alat ukur berupa tes obyektif/pilihan ganda harus memenuhi kaidah pembuatan butir soal jenis tersebut. Maksudnya, persyaratan stem, kunci dan pengecoh harus dipenuhi.
4. Pernyataan-pernyataannya jangan mendua arti, Karena orang yang lebih cerdas cenderung tidak akan menjawabnya.
5. Cukup waktu untuk mengerjakan.
6. Alat ukur tidak terlalu panjang atau terlalu pendek.
7. Khusus alat ukur berupa alat tes juga memperhatikan hal berikut :
   a. Tingkat kesukaran soal hendaknya disesuaikan dengan kemampuan subjek uji. Soal yang baik adalah tidak terlalu mudah dan tidak terlalu sukar sehingga mampu mebedakan mana yang pandai dan yang tidak.
   b. Setiap butir soal benar-benar untuk mengukur keberha- silan belajar, jika tesnya untuk mengukur keberhasilan belajar.
   c. Butir-butir soal diurutkan dari yang termudah ke yang paling sukar.
   d. Jawaban jangan samapai ditemukenali oleh subjek uji, misal karena urutan jawabannya terpola (tidak acak).

B. RELIABILITAS

Pengertian reliabilitas berkait dengan konsistensi. Suatu alat ukur dinyatakan reliable/andal bila memberikan hasil yang sama pada berkali-kali pengulangan

pengukuran. Reliabilitas berlaku pada tingkat suatu perangkat tes. Jadi tidak berlaku untuk masing-masing item tes penyusun suatu perangkat alat ukur.

Berkait dengan reliabilitas tes, Frisbie (2005) menyatakan bahwa reliabilitas tes asil belajar berbeda dengan reliabilitas tes untuk seleksi karena tes hasil belajar memiliki varians yang rendah manakala anak berhasil semua dalam belajarnya. Oleh karena itu secara praktis Friesbie membuat tabel pembandingan ideal antara interpretasi dalam situasi *Norm Reference* (NR) dan *Criterion Reference.*

Tabel 2. *Comparative ideals for NR and CR interpretation situations*

	Norm-Reference	Criterion-Reference
Item difficulty	Moderate	Easy-to-hard
Item discrimination	High positive	Nonnegative
Score variability	Maximize	Non-issue
Error estimate	High reliability coefficient	High decision consistency index

Berdasarkan informasi pada Tabel 2 maka tes untuk mengukur keberhasilan belajar merupakan tes yang skornya diinterpretasikan dalam situasi *criterion-reference*, sehingga item-itemnya memiliki tingkat kesulitan item bervariasi dari mudah sampai sukar (sebagai cerminan tingkat keberhasilan belajar) dan tidak boleh memiliki indeks daya beda yang negatif (sebagai cerminan bahwa tidak ada testi yang cerdas menjawab salah). Oleh karena itu estimasi error didasarkan pada tingginya indeks konsistensi (indeks yang tinggi menunjukkan semua testi pasti benar bila sudah belajar, semua testi salah bila belum belajar). Sebaliknya, tes untuk tujuan seleksi adalah tes yang dapat memisahkan kelompok yang lolos seleksi dan yang tidak lolos seleksi. Oleh karena itu, iterpretasinya dalam situasi *norm-reference*, sehingga item-itemnya memiliki indeks kesulitan yang harus moderate (sebagai cerminan bahwa kelompok ataslah yang pasti dapat mengerjakan) dan indeks daya beda harus tinggi (sebagai cerminan yakin dapat membedakan kelompok atas dan bawah). Oleh karena itu, estimasi *error* didasarkan pada tingkat tingginya reliabilitas tes (indeks yang tes tinggi mencerminkan bahwa semakin cerdas testi di dalam kelompoknya semakin tinggi pula skor yang diperolehnya).

Menurut Stark *et. al.*(2001), pemilihan item tes dalam prosedur pengembangan tes menggunakan CTT umumnya didasarkan pada: (a) nilai kesukaran item, dan (b) korelasi skor item dan skor total atau disingkat korelasi item-total. Item yang memiliki korelasi

item-total paling tinggi dipakai sebagai elemen suatu tes untuk membentuk suatu skala dengan konsistensi internal tinggi guna memperkecil sumbangan *error* acak skor-skor tes. Distribusi skor-skor tes total yang diperoleh dari lapangan dibandingkan dengan distribusi yang diinginkan oleh pengembang tes. Sejumlah item mungkin perlu diganti untuk memperoleh sedekat/semirip mungkin antara distribusi skor total yang diinginkan dan distribusi skor total yang diperoleh dari lapangan. Format-format paralel pada umumnya diciptakan untuk memperoleh distribusi-distribusi skor tes yang identik. Kesamaan dari nilai rata-rata, varians, dan *error* skor ditafsirkan sebagai bukti bahwa format tes-tes bersifat paralel.

Menurut Stark *et. al.*(2001), seharusnya langkah pertama sebelum penulisan item mulai, pengembang tes harus mempunyai suatu pemahaman yang baik tentang konstrak variabel (kemampuan) yang akan diukur. Mengacu pendapat Nunnally *et. al.*, berdasarkan ”*rule of thumb*” ia menyatakan bahwa lazimnya disepakati bahwa banyaknya item tes yang harus dibuat sedikitnya dua kali dari banyaknya item tes final yang diperlukan. Sejumlah besar item pilihan ganda diperlukan, jika format-format ganda harus dikembangkan. Item-item tersebut harus diteskan terlebih dahulu menggunakan suatu sampel yang serupa dengan populasi pelamar. Sampel ini, yang diacu selanjutnya sebagai suatu sampel yang dijadikan pedoman saat kalibrasi, harus besar, agar cukup untuk menyediakan statistika item CTT yang stabil. Item-item dengan korelasi item-total tinggi harus tercakup di dalam tes karena item-item tersebut meningkatkan konsistensi skala internal (reliabilitas), dan hal seperti itu akan mereduksi *standard error* pengukuran. Kesulitan item (nilai p) juga harus dipertimbangkan untuk membuat suatu tes dengan distribusi skor total yang diinginkan.

Untuk memperoleh distribusi skor skala yang diinginkan dilakukan penggantian item. Agar skalanya meningkat maka maka item dengan nilai p yang rendah harus digantikan dengan nilai p yang tinggi. Untuk memperkecil dampak penggantian item terhadap reliabilitas skala, yakni dengan mencoba menggantikan item-item yang memiliki korelasi item-total yang rendah sebelum menghapus item-item yang memiliki daya pembeda yang lebih tinggi. Dapat pula dalam praktik, beberapa penyeimbangan konten/isi juga diperlukan. Setelah dilakukan penggantian kemudian dianalisis lagi.

Ada keterbatasan penggunaan pendekatan CTT (Stark *et. al.*, 2001). Pertama, statistika CTT bergantung kepada subpopulasi penempuh tes. Berbeda grup penempuh tes berbeda pula nilai rata-rata skor dari atribut variabel yang diukur. Dengan demikian, para

pengembang tes harus hati-hati ketika memilih sampel untuk kalibrasi item. Jika sampel-sampel kalibrasi berbeda karakteristik/sifat dengan sampel operasional (sampel populasi yang sesungguhnya sebagai target), properti-properti psikometri hasil pengukuran akan berubah secara dramatis. Kedua, di dalam CTT, ketepatan pengukuran suatu tes (galat baku atau *standard error* pengukuran) secara implisit dirata-ratakan ke semua level kemampuan yang diukur. Dengan demikian, ketepatan pengukuran pada level-level skor yang tertentu tidak dikenal/tidak diketahui. Oleh karena itu, dikembangkan analisis item menggunakan teori respons item atau *item response theory* (IRT).

Kegiatan mengkonstruksi tes menggunakan pendekatan IRT, seperti halnya pada penggunaan pendekatan CTT, penulis harus membuat dua sampai tiga kali banyaknya item seperti yang diinginkan di dalam format final. Dalam IRT diperlukan sampel kalibrasi heterogen yang besar. Model IRT yang lebih kompleks, seperti model IRT untuk skala politomus, memerlukan sampel lebih besar untuk mengestimasi parameter. Sebelum mengestimasi parameter item, perlu untuk melakukan suatu analisis item menurut teori tes klasik untuk menghapuskan item-item yang mempunyai skor mendekati nihil (tidak atau sedikit sekali yang dapat mengerjakan), tentu saja item yang demikian akan memiliki korelasi-korelasi item-total negatif. Item ini akan menyebabkan permasalahan konvergensi/pemusatan. Demikian pula item yang mempunyai skor prefect, dimana untuk tes pilihan ganda skor prefect adalah 1 untuk setiap testi atau *person/case*.

BAB IV
TEKNIK PENILAIAN

Teknik penilaian pendidikan ada bermacam-macam. Ada yang tergolong tes bila menyangkut benar salah dan nontes bila tidak menyangkut benar salah. Grounlund (1998) mengklasifikasikan teknik penilaian tes menjadi beberapa kategori, yakni tes bentuk pilihan, tes bentuk mengkonstruksi jawaban, dan penilaian yang diperluas. Tes bentuk pilihan dapat berupa pilihan ganda, salah-benar, menjodohkan/memasangkan, tes bentuk mengkonstruksi jawaban dapat berupa tes isian, uraian terstruktur, dan uraian terbuka, asesmen yang diperluas dapat berupa proyek atau portofolio.

Dalam Buku panduan penilaian yang diterbitkan BSNP tahun 2008, teknik penilaian untuk kelompok mata pelajaran teknologi adalah sebagai berikut.

1. **Tes tertulis:** suatu teknik penilaian yang menuntut jawaban secara tertulis, baik berupa pilihan atau isian. Tes yang jawabannya berupa pilihan meliputi pilihan ganda, benar-salah dan menjodohkan, sedangkan tes yang jawabannya berupa isian berbentuk isian singkat atau uraian
2. **Observasi:** atau pengamatan adalah teknik penilaian yang dilakukan dengan menggunakan indera secara langsung. Observasi dilakukan dengan menggunakan pedoman observasi yang berisi sejumlah indikator perilaku yang akan diamati.
3. **Tes Praktik:** atau tes kinerja, adalah teknik penilaian yang menuntut peserta didik mendemonstrasikan kemahirannya. Tes praktik dapat berupa tes tulis keterampilan, tes identifikasi, tes simulasi dan tes petik kerja. Tes tulis keterampilan digunakan untuk mengukur keterampilan peserta didik yang diekspresikan dalam kertas, misalnya peserta didik diminta untuk membuat desain atau sketsa gambar. Dalam IPA, kemampuan merancang eksperimen termasuk bagaimana merancang rangkaian peralatan yang digunakan termasuk contoh tes tulis keterampilan. Tes identifikasi dilakukan untuk mengukur kemahiran mengidentifikasi sesuatu hal berdasarkan fenomena yang ditangkap melalui alat indera, misalnya mengetahui kerusakan mesin berdasar suaranya, mengetahui nama preparat berdasar bayangan benda yang dilihat di bawah mikroskop. Tes simulasi digunakan untuk mengukur kemahiran bersimulasi memperagakan suatu tindakan tanpa menggunakan peralatan/benda yang

sesungguhnya. Tes petik kerja dipakai untuk mengukur kemahiran mendemonstrasikan pekerjaan yang sesungguhnya seperti mendemosntrasikan cara memasak, cara menghidupkan mesin, atau cara menggunakan mikroskop.

4. **Penugasan:** suatu teknik penilaian yang menuntut peserta didik melakukan kegiatan tertentu di luar kegiatan pembelajaran di kelas. Penugasan dapat diberikan dalam bentuk individual atau kelompok. Penugasan ada yang berupa pekerjaan rumah atau berupa proyek. Pekerjaan rumah adalah tugas yang harus diselesaikan peserta didik di luar kegiatan kelas, misalnya menyelesaikan soal-soal dan melakukan latihan. Proyek adalah suatu tugas yang melibatkan kegiatan perancangan, pelaksanaan, dan pelaporan secara tertulis maupun lisan dalam waktu tertentu dan umumnya menggunakan data lapangan.
5. **Tes Lisan:** dilaksanakan melalui komunikasi langsung tatap muka antara peserta didik dengan seorang atau beberapa penguji. Pertanyaan dan jawaban diberikan secara lisan dan spontan. Tes jenis ini memerlukan daftar pertanyaan dan pedoman pensekoran.
6. **Penilaian Portofolio :** merupakan penilaian yang dilakukan dengan cara menilai portofolio peserta didik. Portofolio adalah kumpulan karya-karya peserta didik dalam bidang tertentu yang diorganisasikan untuk mengetahui minat, perkembangan, prestasi, dan/atau kreativitas peserta didik dalam kurun waktu tertentu.
7. **Jurnal :** merupakan catatan pendidik selama proses pembelajaran yang berisi informasi kekuatan dan kelemahan peserta didik yang berkait dengan kinerja ataupun sikap peserta didik yang dipaparkan secara deskriptif.
8. **Penilaian Diri :** merupakan teknik penilaian dengan cara meminta peserta didik untuk mengemukakan kelebihan dan kekurangan dirinya berkaitan dengan kompetensi yang menjadi tujuan pembelajaran
9. **Penilaian Antarteman :** merupakan teknik penilaian dengan cara meminta peserta didik untuk mengemukakan kelebihan dan kekurangan temannya dalam berbagai hal. Untuk itu perlu ada pedomanan penilaian antarteman yang memuat indikator prilaku yang dinilai.

Rangkuman bentuk penilaian beserta bentuk instrumennya disajikan dalam tabel berikut.

Tabel 3. Klasifikasi teknik penilaian serta bentuk instrumen

Teknik Penilaian	Bentuk Instrumen
• Tes tertulis	• Tes pilihan: pilihan ganda, benar-salah, menjodohkan dll. • Tes isian: isian singkat dan uraian
• Observasi (pengamatan)	• Lembar observasi (lembar pengamatan)
• Tes praktik (tes kinerja)	• Tes tulis keterampilan • Tes identifikasi • Tes simulasi • Tes uji petik kerja
• Penugasan individual atau kelompok	• Pekerjaan rumah • Proyek
• Tes lisan	• Daftar pertanyaan
• Penilaian portofolio	• Lembar penilaian portofolio
• Jurnal	• Buku cacatan jurnal
• Penilaian diri	• Kuesioner/lembar penilaian diri
• Penilaian antarteman	• Lembar penilaian antarteman

Dalam memilih teknik penilaian pendidik mempertimbangkan (1) karakteristik kelompok mata pelajaran, (2) rumusan kompetensi mata pelajaran yang dikembangkan dalam silabus, dan (3) rumusan indikator pencapaian setiap KD.

BAB V
PENGEMBANGAN INSTRUMEN PENILAIAN

A. PENGEMBANGAN TES TERTULIS UNTUK MENGUKUR KEMAMPUAN KOGNITIF

Tes tertulis untuk mengukur kemampuan kognitif dapat dibedakan untuk tujuan mengukur kemampuan kognitif tingkat rendah (kemampuan mengetahui, memahami, dan menerapkan), dan kemampuan kognitif tingkat tinggi (menganalisis, mengevaluasi, menyintesis, berimajinasi, dan mengkreasi). Bentuk tes terulis dapat berupa tes pilihan (*selected response test*), baik berupa pilihan ganda, benar-salah, ataupun menjodohkan, serta bentuk mengisikan jawaban. (*supply response test*), baik berupa isian singkat, uraian terstruktur/uraian objektif, dan uraian terbuka (*open ended supply response test*).

Item tes pilihan terdiri atas soal dan kunci jawaban, sedangkan item bentuk mengisikan jawaban terdiri atas soal beserta rubrik dan atau pedoman penskoran. Item tes bentuk uraian terbuka terdiri atas soal, rubric, dan pedoman penskoran.

Dilihat dari cara berpikir dalam menyelesaikan/mengerjakan soal, ada yang bersifat konvergen, yakni mengarah ke suatu jawaban benar, yakni berupa item bentuk pilihan, isian singkat, dan uraian objektif. Sementara ada yang bersifat divergen, yakni mengarah ke lebih dari satu jawaban yang sama-sama benar. Dalam hal ini adalah item bentul uraian terbuka.

Sebagai contoh, suatu item meminta testi untuk mengemukakan tiga penyebab gagal panen.

Pertanyaan: Kemukakan tiga penyebab terjadinya gagal panen padi pada suatu tempat!

Jawaban tiga jawaban merupakan variasi dari semua jawaban benar yang mungkin yaitu:

1. Kemarau panjang sehingga air tidak tersedia
2. Terjadi kerusakan irigasi berupa bendungan jebol sehingga air tidak tersedia
3. Terjadi serangan hama yang tidak dapat diatasi
4. Terjadi banjir yang merusak seluruh tanaman padi
5. Terjadi kelangkaan pupuk
6. Terjadi hujan abu dari letusan gunung bila daerahnya dekat dengan gunung
7. Jawaban lain yang benar

Dalam menyusun soal suatu item, harus memperhatikan aspek substansi, konstruksi, dan bahasa. Persyaratan tersebut akan dapat diketahui manakala diadakan analisis secara kualitatif. Berikut ini disajikan persyaratan masing-masing aspek tersebut dalam bentuk soal pilihan ganda dan uraian. Untuk bentuk lainnya mahasiswa diharap dapat menyusunnya sendiri.

Tabel 4. Aspek yang Harus Dipenuhi dalam Menyusun Soal Pilihan Ganda

ASPEK YANG DITELAAH	NOMOR ITEM				
	1	2	3	….	N
a. Aspek materi:					
1) Butir soal sesuai indikator					
2) Hanya ada satu kunci atau jawaban yang benar					
3) Isi materi sesuai dengan tujuan pengukuran					
4) Isi materi sesuai dengan jenjang dan jenis pendidikan					
5) Pilihan benar-benar berfungsi, jika pilihan merupakan hasil perhitungan, maka pengecoh berupa pilihan yang salah rumus/salah hitung					
b. Aspek konstruksi:					
1) Pokok soal (stem) dirumuskan dengan jelas					
2) Rumusan soal dan pilihan dirumuskan dengan tegas					
3) Pokok soal tidak memberi petunjuk/mengarah kepada pilihan jawaban yang benar					
4) Pokok soal tidak mengandung pernyataan negatif genda					
5) Bila terpaksa menggunakan kata negatif, maka harus digarisbawahi atau dicetak lain					
6) Pilihan jawaban homogeny					
7) Hindari adanya alternatif jawaban : "seluruh jawaban di atas benar" atau "tak satu jawaban di atas yang benar" dan yang sejenisnya					
8) Panjang alternatif /pilihan jawaban relatif sama, jangan ada yang sangat panjang dan ada yang sangat pendek					
9) Pilihan jawaban dalam bentuk angka/waktu diurutkan.					
10) Wacana, gambar, atau grafik benar-benar berfungsi					
11) Antar butir tidak bergantung satu sama lain					
c. Aspek bahasa:					
1) Rumusan kalimat komunikatif					
2) Kalimat menggunakan bahasa yang baik dan benar, sesuai dengan jenis bahasanya					
3) Rumusan kalimat tidak menimbulkan penafsiran ganda atau salah pengertian.					
4) Menggunakan bahasa/kata yang umum (bukan bahasa lokal)					
5) Rumusan soal tidak mengandung kata-kata yang dapat menyinggung perasaan siswa.					

Tabel 5. Aspek yang Harus Dipenuh dalam Menyusun Soal Uraian

ASPEK YANG DITELAAH	NOMOR ITEM				
	1	2	3	….	N
a. Aspek materi:					
1) Butir soal sesuai indikator					
2) Batasan pertanyaan dan jawaban yang diharapkan jelas					
3) Isi materi sesuai dengan tujuan pengukuran					
4) Isi materi yang ditanyakan sesuai dengan jenjang dan jenis pendidikan					
b. Aspek konstruksi:					
1) Rumusan kalimat dalam bentuk kalimat tanya atau perintah yang menuntut jawaban terurai.					
2) Ada petunjuk yang jelas cara mengerjakan/ menyelesaikan soal					
3) Ada pedoman penskorannya					
4) Tabel, grafik, diagram, kasus, atau yang sejenisnya bermakna (jelas keterangannya atau ada hubungannya dengan masalah yang ditanyakan.					
5) Butir soal tidak bergantung pada butir soal sebelumnya					
c. Aspek bahasa:					
1) Rumusan kalimat komunikatif					
2) Kalimat menggunakan bahasa yang baik dan benar, sesuai dengan jenis bahasanya					
3) Rumusan kalimat tidak menimbulkan penafsiran ganda atau salah pengertian.					
4) Menggunakan bahasa/kata yang umum (bukan bahasa lokal)					
5) Rumusan soal tidak mengandung kata-kata yang dapat menyinggung perasaan siswa.					
12) Antar butir tidak bergantung satu sama lain					
c. Aspek bahasa:					
6) Rumusan kalimat komunikatif					
7) Kalimat menggunakan bahasa yang baik dan benar, sesuai dengan jenis bahasanya					
8) Rumusan kalimat tidak menimbulkan penafsiran ganda atau salah pengertian.					
9) Menggunakan bahasa/kata yang umum (bukan bahasa lokal)					
10) Rumusan soal tidak mengandung kata-kata yang dapat menyinggung perasaan siswa.					

B. PENYUSUNAN INSTRUMEN PENILAIAN AFEKTIF

Dalam menyusun angket harus memperhatikan skala sikap yang digunakan. Pengukuran sikap yang dapat digunakan misalnya sebagai berikut.

1. Skala Likert

Skala Likert merupakan suatu skala penilaian untuk mengukur sikap dengan skala ordinal. Rentangan yang dipilih dari yang sangat positif sampai sangat negatif, misal dengan alternatif pilihan mulai dari sangat setuju (SS), setuju (S), ragu-ragu (R), tidak setuju (T), dan sangat tidak setuju (ST), dapat pula dari yang tidak pernah sampai yang selalu siswa lakukan sehingga rentangannya mulai dengan tidak pernah (TP), jarang (J), sering (S), hampir selalu (HS), dan selalu (S).

Dalam menyusun skala Likert sangat tergantung kemampuan penilai/penyusun angket dalam merumuskan difinisi variabel yang diukur, merumuskan dimensinya (apakah bersifat multidimensi) dan merumuskan indikatornya. Jika indikatornya sudah dirumuskan kemudian disusun penyataannya. Misalnya, ingin diukur bagaimana persepsi peserta didik terhadap kemandirian dalam belajar, untuk itu harus dicari indikator-indikator yang relevan. Misal diperoleh indikator yang di antaranya tentang:

a. Keterlibatan orang lain dalam penetapan waktu belajar (2 item).
b. Keterlibatan orang lain dalam mencari sumber acuan (1 item).
c. dst.

Kemudian dibuat daftar pertanyaan sebagai berikut :

Nama Siswa: ..... Nomor presensi: .....

Setujukah Anda terhadap pernyataan di bawah ini?

Bila sangat setuju beri tanda V pada kolom pilihan SS, bila hanya setuju beri tanda V pada kolom pilihan S, bila netral atau ragu beri tanda V pada kolom pilihan R, bila tidak setuju beri tanda V pada kolom pilihan T, dan bila sangat tidak setuju beri tanda V pada kolom pilihan TS!

No.	Pernyataan	Pilihan				
		ST	T	R	S	SS
1	Saya belajar bila ada teman mengajak saya untuk belajar					
2	Saya belajar dengan sungguh-sungguh bila akan ada ulangan					
3	Saya mencari bahan di perpustakaan bila ada tugas dari guru					
4.	dst.					

Catatan:

Bila pernyataan bersifat positif maka pilihan SS diberi skor 5 dan pilihan STS diberi skor 1. Bila pernyataan bersifat negative sebaliknya.

2. Skala Perbedaan Semantik/Skala Berdeferensiasi Sematik

Skala perbedaan semantik/skala berdeferensiasi semantik merupakan suatu model skala dengan meletakkan suatu rentangan di antara dua kata atau ide yang berlawanan, sehingga berupa skala perbedaan sematik. Model skala yang bipolar ini sangat baik untuk mengungkap unsur evaluasi (baik/buruk, bersih/kotor, jujur/tidak jujur, menguntungkan/merugikan dan sejenisnya), atau untuk mengungkap unsur potensi (besar/kecil, kuat/lemah, berat/ringan, dan sejenisnya), ataupun unsur aktivitas (aktif/pasif, cepat/lambat, loyal/tak loyal, penuh perhatian/tak acuh).

Pasangan ajektif tersebut harus dicari yang sesuai dengan konsep atau obyek yang akan diukur. Pasangan ajektif tersebut perlu diuji secara empiris, yaitu dengan terlebih dahulu mencari dua kelompok yang benar-benar pro dan yang benar-benar anti terhadap hal tersebut. Pasangan ajektif yang benar-benar dapat membedakan antara kedua kelompok tersebut dapat dipakai, dan yang tak dapat membedakan yakni baik kelompok pro dan anti sama-sama memilih titik tengah (ragu-ragu) dibuang. Dalam penskorannya, semakin ke arah yang positif semakin besar, dan skor total merupakan penjumlahan skor setiap pasangan ajektif.

Misalnya, ingin mengetahui bagaimana persepsi mahasiswa terhadap mata kuliah yang diampunya.

Contoh:
Nama Siswa: ..... No presensi: ....

Menurut Anda bagaimanakah pembelajaran Matematika yang saya selenggarakan selama tengah semester ini?

Perhatikan contoh berikut ini.

Menarik .×.____.____.____. ____. Membosankan

Karena Anda memberi tanda silang pada posisi di atas berarti menurut Anda pelajaran Matematika yang telah diselenggarakan selama setengah semester menarik

Sebaliknya kalau Anda menyilang sebagai berikut.

Menarik .____.____.____.×. ____. Membosankan

Berarti menurut Anda pelajaran Matematika yang telah diselenggarakan selama setengah semester sangat membosankan.

Pembelajaran Biologi yang telah berjalan dalam setengah semester I

Menarik	.____.____.____.____.	Membosankan
Mudah	.____.____.____.____.	Sukar
Ringan	.____.____.____.____.	Berat
Menguntungkan	.____.____.____.____.	Merugikan
Bermanfaat	.____.____.____.____.	Merugikan
Menantang	.____.____.____.____.	Tidak menantang
Mengasyikkan	.____.____.____.____.	Menjemukan

3. Skala Thurstone

Tahapan dalam pengembangan instrumen skala sikap memakai skala Thurstone adalah sebagai berikut.

a. Pengembangan daftar pernyataan yang ditawarkan pada panelis yakni dengan menyusun minimal 50 pernyataan dari yang sangat positif sampai sangat negatif yang berkait dengan mata pelajaran Biologi.
b. Menyiapkan anggota panelis, misalnya dengan memilih sekurang-kurangnya 80 guru biologi atau mahasiswa yang menempuh program Pendidikan Biologi.
c. Meminta panelis untuk memberik-an skor terhadap setiap pernyataan yang ditawarkan. Meminta panelis untuk memberikan skor sangat rendah terhadap pernyataan yang bersifat negatif dan skor yang sangat tinggi untuk pernyataan yang sangat positif. Kisaran skor 1 sampai 11.

Contoh:

Berikut ini pernyataan-pernyataan yang berkait dengan bidang Biologi. Bila suatu pernyataan dinilai positif di beri skor besar, sedangkan bila negatif diberi skor kecil. Skor untuk pernyataan yang paling positif 11, sedangkan skor untuk penyataan yang paling negatif diberi skor 1.

No	Pernyataan	Pilihan skor										
		1	2	3	4	5	6	7	8	9	10	11

1	Menguasai bidang biologi dengan baik sangat mendukung wirausaha di masa depan											
2.	Berusaha di bidang yang ada hubungannya dengan biologi merupakan usaha yang sia-sia											
3.	Keahlian dalam bidang biologi memiliki prospek yang baik sepanjang yang bersangkutan menekuninya dengan sungguh-sungguh											
4.	Biologi tidak bedanya dengan mata pelajaran yang lain di sekolah											
Dst												

d. Langkah selanjutnya menghitung nilai median untuk tiap pernyataan berdasarkan skor yang diberikan panelis

e. Menghitung nilai kuartil (Q_3 dan Q_1) dan deviasi kuartil (Q_3-Q_1) untuk tiap pernyataan berdasarkan skor yang diberikan panelis (dengan menggunakan program excel).

f. Memilih pernyataan yang memenuhi syarat.
Menyisakan 15 sampai 20 pernyataan yang bersifat positif, netral, sampai yang negatif dengan cara membuang pernyataan yang memiliki nilai deviasi kuartil yang besar.

g. Menentukan besarnya skor untuk setiap pernyataan
Skor tiap pernyataan merupakan besarnya median yang diberikan oleh panelis

Contoh:

No.	Pernyataan	Skor
1	Menguasai bidang biologi dengan baik sangat mendukung wirausaha di masa depan	10,5
2	Berusaha di bidang yang ada hubungannya dengan biologi merupakan usaha yang sia-sia	1,2
3	Keahlian dalam bidang biologi memiliki prospek yang baik sepanjang yang bersangkutan menekuni dengan dengan sungguh-sungguh	5,75
4	Biologi tidak bedanya dengan mata pelajaran yang lain di sekolah	6,5
dst.		

h. Menyiapkan angket siap dipakai

Contoh sajian

Pilihlah dengan cara membubuhkan tanda X pada kolom TS jika Anda tidak setuju dan pada kolom S jika setuju terhadap pernyataan-pernyataan di bawah ini!

Contoh:

No.	Pernyataan	Pilihan	
		TS	S
1	Bidang biologi menjamin masa depan		X

Artinya: Anda *setuju* bidang biologi *menjamin* masa depan wirausaha di bidang tersebut.

Berikut ini daftar pernyataan yang harus Anda pilih!

No.	Pernyataan	Pilihan	
		TS	S
1	Menguasai bidang biologi dengan baik sangat mendukung wirausaha di masa depan		
2	Berusaha di bidang yang ada hubungannya dengan biologi merupakan usaha yang sia-sia		
3	Keahlian dalam bidang biologi memiliki prospek yang baik sepanjang yang bersangkutan menekuni dengan sungguh-sungguh		
4	Biologi tidak bedanya dengan mata pelajaran yang lain di sekolah		
Dst.			

4. Lembar Observasi/Lembar Penilaian Antarteman (*Peer Assessment*)

Lembar observasi dapat digunakan untuk melihat sikap peserta didik selama kegiatan berlangsung. Dalam hal ini perlu dikembangkan indikator kegiatan yang bersangkutan. Lembar ini juga dapat dikembangkan sebagai instrument penilaian antarteman. Misalnya kerja kelompok mengerjakan kegiatan lab digunakan lembar observasi sebagai berikut.

a. Contoh Lembar penilaian antar teman dalam kerja kelompok

Nilailah setiap anggota dalam kelompok Anda! Berilah nilai 10 bila sangat baik, sebaliknya berilah nilai 0 bila sangat jelek! Selanjutnya jumlahkan hasil penilaian Anda untuk memperoleh nilai masing-masing anggota dalam kelompok Anda!

No.	Nama	Nomor Presensi	Hal yang dinilai					
			1	2	3	4	5	Jumlah
1								
2								
3								
Dst								

Keterangan:

No.	Hal yang dinilai
1	Mendengarkan dengan perhatian penuh saat temannya berbicara,
2	mengajukan usul, atau memberikan arahan tentang tugas yang harus diselesaikan
3	Menyambut baik terhadap tugas yang diberikan kepadanya
4	Menyelesaikan tugas dengan baik dan sesuai waktu yang ditetapkan
5	Menegur dengan sopan bila ada temannya yang tidak serius dalam berkerja

C. PENGEMBANGAN INSTRUMEN PENILAIAN KINERJA

Penilaian kinerja adalah penilaian yang memfokuskan aspek keterampilan yang berkait dengan ranah psikomotor yang dapat didemonstrasikan oleh peserta didik. Dilihat dari kinerja atau kemampuan yang didemonstrasikan, kinerja dapat digradasi dari kinerja yang paling rendah sampai yang paling tinggi. Kinerja yang paling rendah misalnya kemampuan siswa mampu menjawab saat ditanya besarnya uang kembalian jika membayar dengan mata uang yang nilainya lebih besar dari harga barang. Kinerja agak tinggi misalnya siswa diminta mendemonstrasikan besarnya uang yang harus ia kembalikan menggunakan mata uang yang sesungguhnya. Kinerja yang lebih tinggi lagi misalnya siswa diminta bersimulasi dengan pasangannya mendemonstrasikan besarnya uang kembalian menggunakan mata uang yang sesungguhnya dengan nilai yang berbeda-beda, dan kinerja yang sangat tinggi jika siswa mampu berbelanja di toko dengan membawa sejumlah uang dan memperoleh uang kembalian/sisa uang sebesar nilai uang yang dibawa dikurangi harga barang yang dibelinya.

Dalam konteks di atas, maka kinerja mencakup ranah kognitif dan sekaligus mencerminkan ranah psikomotor. Ranah kognitif adalah tingkat kebenaran dari aspek berpikir yang mendasari tindakannya, dan keterampilan psikomotor yang didemonstrasikan berupa kemampuan membedakan mata uang sesuai dengan nilainya dan memilih mata uang yang sesuai/lebih besar dari harga barang saat ia sebagai pembeli, dan memilih mata uang yang nilainya sama dengan nilai pengembalian saat ia berperan sebagai penjual. Dengan kata lain aspek psikomotor menyangkut penguasan prosedur. Akan tetapi, dapat pula aspek psikomotor dapat dinilai dari produk yang dihasilkan oleh suatu tindakan tertentu yang dilakukan peserta didik. Penilaian terhadap prosedur berarti lebih mengarah kepada aktualisasi aspek psikomotor yang ditampilkan dalam suatu kinerja (*performance*).

Dari taksonomi ranah psikomotor dapat diidentifikasikan bahwa ada aspek dari ranah psikomotor yang murni sebagai gerak bagian tubuh dan kemampuan fisik tubuh dan ada pula gerak dari bagian tubuh yang berkait dengan pemakaian alat. Gerak yang dilakukan juga ada gerak yang tidak dipelajari yaitu gerak refleks, dan gerak yang dipelajari. Gerak yang dipelajari ada yang berupa gerak yang sederhana dan ada pula gerak yang kompleks.

Penguasaan teoretik tentang suatu prosedur pun oleh Simson dimasukkan sebagai aspek dari ranah psikomotor yakni termasuk dalam kesiapan untuk berperan aktif dalam

melakukan aktivitas motorik. Oleh karena itu, dalam pembelajaran motorik, tahapan yang ditempuh adalah penguasaan teori tentang teknik/prosedur yang berupa tahapan-tahapan dalam melakukan aktivitas, dilanjutkan dengan artikulasi berupa latihan untuk menguasai suatu teknik/prosedur yang dipelajari, dan otomatisasi untuk menguasai teknik secara terlatih dan spontan.

Sebagai contoh agar seseorang mampu bermain tennis maka ia harus menguasai berbagai teori tentang teknik baik yang berkait dengan teknik memukul dan mengembalikan bola serta peraturan bermain tennis. Kemudian melalui tahapan artikulasi ia berlatih berlatih cara memukul, mengembalikan bola, dan menerapkan aturan bermain. Selanjutnya melalui tahapan otomatisasi ia harus berlatih berulang-ulang (*drilling*), termasuk uji coba, dan bertanding, sehingga ia dapat melakukan gerak-gerak otomatis dalam memukul dan mengembalikan bola, serta mampu bermain secara sportif.

Penilaian terhadap prosedur dilakukan dengan pertimbangan: (1) tidak ada produk yang bisa dinilai, (2) prosedurnya memiliki langkah-langkah yang urut dan dapat diamati, (3) langkah-langkah yang benar dari suatu prosedur menunjukkan suatu keberhasilan, dan/atau (4) analisis terhadap prossedur dapat meningkatkan mutu produk. Penilaian terhadap produk dilakukan dengan pertimbangan: (1) berbeda prosedur berbeda produk, (2) prosedur sudah dikuasai, (3) prosedurnya tidak dapat dinilai, (4) prosedur tidak perlu dinilai (misalnya pekerjaan rumah), dan/atau (5) produk memiliki kejelasan aspek yang dinilai

PENGEMBANGAN ITEM TES KINERJA

1. Pengembangan Item Tes Kinerja untuk Penguasaan Prosedur

Untuk mengembangkan item tes kinerja dalam bentuk prosedur harus memperhatikan hal-hal berikut.

a. Menyesuaikan dengan jenis kinerja/*performance* yang harus ditampilkan, apakah merupakan gerak dasar yang fundamental, kemampuan perceptual, kemampuan fisik, gerak terlatih ataukah gerak yang mengekspresikan komunikasi.
b. Menyesuaikan dengan tehnik penilaian yang dipilih, yaitu:
    1) tes identifikasi untuk mengukur kinerja seseorang atas dasar tanda-tanda atau sinyal yang diberikan saat diberikan tes.

2) tes simulasi untuk mengukur kinerja dalam situasi yang mirip dengan situasi yang sebenarnya.
3) uji petik kerja (*work sampel test*) untuk mengukur kinerja dalam situasi yang sebenarnya.

c. Menyusun rubrik/pedoman penskoran

Di dalam penyusunan rubrik/pedoman penskoran ada beberapa hal yang perlu diperhatikan tergantung pada bentuk instrumen.

1) Tes identifikasi: (a) menentukan jenis kemampuan kinerja yang akan diidentifikasi, (b) menentukan banyaknya hal/aspek yang akan diidentifikasi, dan (c) membuat rubrik untuk penskoran yang dilengkap dengan kategorisasi keberhasilan identifikasi.
2) Uji petik kerja/simulasi: (a) mengidentifikasi aspek kinerja yang diskor, (b) menentukan model skala yang dipakai untuk menyekor, yakni skala penilaian (*rating scale*) atau daftar cek (*check list*), dan (c) membuat rubrik penskoran yang dilengkapi dengan kategorisasi keberhasilan kinerja.

2. Pengembangan Item Tes Kinerja untuk Penguasaan Produk

Untuk mengembangkan item tes kinerja dalam bentuk penguasan produk maka harus memperhatikan hal-hal berikut.

a. Menyesuaikan dengan jenis produk yang harus dihasilkan, apakah produk dua dimensi ataukah produk tiga dimensi.
b. Memperhatikan tehnik penilaian yang dipakai, yaitu
   1) Tes tertulis (*paper and penci test*) untuk menilai produk dua dimensi yang diujudkan dalam bentuk sketsa, tulisan, gambar, lukisan, atau bentuk dua dimensi lainnya.
   2) Penugasan produk tiga dimensi untuk menilai produk tiga dimensi yang diujudkan dalam bentuk rangkaian, pahatan, dan produk tiga dimensi lainnya.
c. Menyusun rubrik/pedoman penskoran

   Di dalam penyusunan rubrik/pedoman penskoran ada beberapa hal yang perlu diperhatikan tergantung pada bentuk instrumen.

   1) Tes *paper and pencil*: (a) menentukan cara penskoran secara holistik atau analitik, (b) menentukan aspek-aspek yang dinilai atau kata kunci, (c) menentukan bobot skor, dan (d) menentukan klasifikasi peringkat penilaian.

2) Penugasan produk tiga dimensi: (a) menentukan aspek produk yang akan dinilai, (b) menentukan bobot skor, dan (c) menentukan klasifikasi peringkat penilaian.

a. Contoh Instrumen Pengukuran Kinerja untuk ranah Kognitif

1) Pengukuran Aspek komunikasi

(a) Lembar Observasi untuk Kinerja Umum (Bentuk *Check List*)

Lembar observasi untuk kinerja umum dari aspek kognitif dalam bentuk *check list* berarti performan yang benar harus muncul dalam hal yang diamati.

No.	Nama Siswa	Nomor presensi	Aspek yang Diukur					
			1	2	3	4	5	Jmlh
1								
2								
3								
dst.								

Aspek yang dinilai:
1. Pertanyaan mengungkap kemampuan berpikir
2. Penjelasan lengkap dan jelas
3. Argumen logis dan kuat
4. Saran jelas dan logis
5. Bahasa baik/benar

Catatan: Beri skor 1 untuk setiap aspek jika sesuai

(b) Lembar Observasi untuk Kinerja Diskusi Kelompok Bentuk *Rating Scale*

Nama:: ……..... Nomor Presensi : ….....……..

No.	Aspek yang Diukur	Skala		
		1	2	3
1	Pola berpikir saat menyampaikan informasi/pendapat			
2	Pola berpikir saat memberikan argumentasi			
3	Pola berpikir saat memberikan kritikan			
4	Kejelasan fokus dan arah pertanyaan			
5	Kemampuan dalam berbicara (dalam memberikan informasi, berpendapat, berargumentasi)			
6	Penguasan bahasa (saat menyampaikan informasi, kritikan, ataupun argumentasi)			
	Jumlah skor			

Rubrik:
Aspek 1: pola berpikir saat menyampaikan informasi
1 = sulit dimengerti dan dipaparkan dengan tidak runtut/teratur
2 = dapat dimengerti tetapi tidak dipaparkan secara runtut/teratur
3 = dipaparkan secara runtut/teratur dan mudah dimengerti
Aspek 2: pola berpikir saat menyampaikan argumentasi
1 = sulit dimengerti dan dipaparkan dengan tidak runtut/teratur
2 = dapat dimengerti tetapi tidak dipaparkan secara runtut/teratur
3 = dipaparkan secara runtut/teratur dan mudah dimengerti
Aspek 3: pola berpikir saat menyampaikan kritik
1 = sulit dimengerti dan dipaparkan dengan tidak runtut/teratur
2 = dapat dimengerti tetapi tidak dipaparkan secara runtut/teratur
3 = dipaparkan secara runtut/teratur dan mudah dimengerti
Aspek 5: kemampuan berbicara
1 = tergagap-gagap, sulit berbicara
2 = kalimat diganti/diulang
3 = lancer
Aspek 6: Penguasaan bahasa
1 = tidak menggunakan EYD
2 = menggunakan EYD tetapi tidak komunikatif (tidak jelas subjek predikatnya)
3 = menggunakan EYD dan komunikatif (jelas subjek predikatnya)

2. Contoh Instrumen untuk Pengukuran Kinerja Psikomotor

Misalnya akan mengukur kinerja keterampilan dengan memerintahkan peserta didik menggunakan neraca, maka langkah pertama invetarisasilah langkah-langkah yanag harus dikerjakan jika seseorang menggunakan neraca untuk menimbang suatu benda. Setelah diinventarisasi langkah/tahapannya misalnya diperoleh hasil sebagai betrikut.

a. Mengecek posisi kedua meja neraca
b. Mengatur kembali posisi kedua meja neraca jika tidak dalam posisi seimbang
c. Memilih anak neraca sesuai dengan hasil pengukuran yang diinginkan
d. Meletakkan anak neraca pada meja neraca pada tempatnya
e. Memasukkan bbarang yang ditimbang sedikit demi sedikit ke dalam wadah sampai ujung meja benda dan ujung meja anak neraca sama tingginya.
f. Menambah atau mengurangi sehingga ujung meja neraca yang berhadap-hadapan sama tingginya.
g. Mengecek kembali dengan ditunggu bahwa kedua ujung meja neraca yang berhadap-hadapan benar-benar sama tingginya
h. Menurunkan dengan hati-hati wadah neraca yang berisi benda yang ditimbaqng.
i. Menurunkan anak timbangan dari mja neraca.

a. Lembar observasi bentuk *check list*

Bila keterampilan menimbang menggunakan neraca akan dibuat dalam bentuk *check list* maka dapat dibuat sajian sebagai berikut.

Nama:: ……..... Nomor Presensi : …....……..

..... a. Mengecek posisi kedua meja neraca dalam posisi seimbang
..... b. Mengatur kembali posisi kedua meja neraca jika tidak dalam posisi seimbang dengan menambah beban di salah satu meja sampai posisi seimbang
..... c. Memilih anak neraca seberat 1 kg
..... d. Meletakkan anak neraca pada meja neraca pada tempatnya
..... e. Memasukkan beras sedikit demi sedikit ke dalam wadah sampai ujung meja benda dan ujung meja anak neraca sama tingginya.
..... f. Menambah atau mengurangi dengan hati-hati beras pada wadah sampai ujung meja neraca yang berhadap-hadapan sama tingginya.
..... g. Mengecek kembali dengan ditunggu bahwa kedua ujung meja neraca yang berhadap-hadapan benar-benar sama tingginya
..... h. Menurunkan dengan hati-hati wadah neraca yang berisi beras, dan menuangkan beras kedalam kantung plastik yang telah disediakan.
..... i. Menurunkan anak timbangan dari meja neraca.

Catatan: Beri tanda V untuk setiap kinerja berikut ini! yang dinyatakan benar dari setiap tindakan yang dilakukan siswa untuk melakukan penimbangan beras seberat 1 kg!

b. Lembar Observasi Bentuk *Rating Scale*

Nama:: ……..... Nomor Presensi : …....……..

1	2	3	4	a.	Cara mengontrol bahwa posisi kedua meja neraca sudah seimbang
1	2	3	4	b.	Cara memilih anak timbangan sesuai dengan berat benda yang diinginkan.
1	2	3	4	c.	Cara meletakkan anak timbangan pada meja neraca.
1	2	3	4	d.	Cara memasukkan benda yang akan ditimbang ke dalam wadah
1	2	3	4	e.	Cara menambah atau mengurangi benda yang ditimbang agar kedua meja neraca pada posisi setimbang.
1	2	3	4	f.	Cara mengontrol posisi kedua meja/bagian neraca setelah diberi anak timbangan dan di sisi lain diberi benda yang ditimbang benar-benar sudah setimbang.
1	2	3	4	g.	Cara mengambil wadah untuk menuangkan/mengambil benda yang ditimbang
1	2	3	4	h.	Cara menuang/mewadahi benda yang telah selesai ditimbang dengan wadah yang telah disediakan.
1	2	3	4	i.	Cara mengambil anak neraca dari meja neraca.

Catatan: 1. Lingkari angka 5 jika sangat tepat, angka 4 jika tepat, angka 3 jika agak tepat, angka 2 jika tidak tepat dan angka 1 jika sangat tidak tepat.
2. *Agar lebih objektif dibuat kriteria untuk tiap butir yang direntang mulai dari skala 1 sampai 4 tersebut.*

BAB VII
PENGOLAHAN SKOR

A. CARA PEMBERIAN SKOR

Sebelum skor diolah menjadi nilai, terlebih dulu harus ditentukan bagaimana cara pemberian skornya. Apakah betul mengerjakan suatu butir soal akan mendapat skor 1, sehingga betul 57 butir soal mendapat skor 57, atau tidak demikian ?

Cara pemberian skor dapat dibedakan atas dasar bentuk butir soal yang digunakan.

1. Cara Pemberian Skor Hasil Ujian Bentuk Uraian

Pemberian skor terhadap hasil pekerjaan suatu butir soal bentuk uraian didasarkan pada beberapa aspek yang harus dimunculkan dalam jawaban, kemudian bagaimana bobot tiap aspek.

Contoh :

Jelaskan cara-cara menggunakan termometer badan untuk mengukur suhu pasien !

Misalkan setelah dibuat kuncinya diperoleh kunci sebagai berikut :

Kunci : Langkah penggunaan termometer badan	Bobot
a. Termometer dokeluarkan dari tempatnya dengan memegang bagian ujung yang tidak terdapat air raksa	1
b. Posisi air raksa diturunkan pada skala terendah dengan cara mengayunkan termometer dengan cepat dari atas ke bawah, dan posisi tangan saat memegang tetap pada ujung yang tidak terdapat air raksanya	1
c. Melakukan pengecekan apakah air raksa sudah berada pada ujung penam-pungnya.	1
d. Minta ijin kepada pasien untuk mengukur suhu tubuhnya	1
e. Memasang termometer pada bagian tubuh yang peka paling sedikit 3 menit	1
f. Mengambil termometer dari tubuh pasien dengan tetap memegang bagian ujung yang tidak ada air raksanya	1
g. Memabaca tinggi air raksa dengan posisi mata sejajar dengan permukaan cairan air raksa.	2

Setelah dicocokkan dengan kunci dapat dihitung berapa skor yang diperoleh oleh peserta uji. Selanjutnya skor yang diperoleh baru dikonversi ke dalam nilai.

Jika butir-butir soal yang digunakan memiliki bobot yang berbeda-beda, para peserta ujian dapat diberi informasi berapa bobot tiap butir soal. Dengan mengetahui bobot masing-masing butir soal, peserta ujian dapat memilih butir soal yang mana yang akan diselesaikan terlebih dulu. Misal, butir soal pertama dicantumkan bobotnya sebesar 10, butir soal kedua 5, butir soal ketiga 6, dan seterusnya.

2. Cara Pemberian Skor Hasil Ujian Bentuk Obyektif.

Pemberian skor soal bentuk obyektif, ada dua cara :

a. Setiap jawaban yang benar dari suatu butir soal diberi skor satu, sehingga skor total akan sama dengan jumlah seluruh jawaban yang benar.
b. Memperhatikan adanya peluang terjadinya tebakan. Dengan demikian, skor yang diperoleh sama dengan jumlah jawaban yang benar yang dikoreksi dengan besarnya tebakan.

B. CARA MENGUBAH/MENGOLAH/MENGKONVERSI SKOR MENJADI NILAI

Ada dua cara mengubah atau mengolah skor menjadi nilai. Pertama, cara pengolahan nilai dengan menggunakan acuan patokan atau kriteria, yang biasa dikenal dengan nama **Penilaian Acuan Patokan** (*Criterion Referenced Evaluation*). Karena dalam menentukan nilai akhir terhadap seorang peserta ujian, dilakukan dengan cara membandingkan skor mentah dengan patokan atau kriteria yang telah ditetapkan telebih dahulu, maka sifatnya absolut, sehingga sering disebut **Penilaian Absolut**. Kedua, cara pengolahan nilai dengan menggunakan acuan norma, dan biasa disebut dengan **Penilaian Acuan Norma** (*Norm Referenced Evaluation*). Pada acara ini, penilaian akhir terhadap seorang peserta ujian, dibandingkan dengan prestasi seluruh peserta ujian dengan kaidah distribusi normal. Dengan cara ini akan dapat dilihat kedudukan siswa di dalam kelompoknya.

Sistem PAP dapat diterapkan secara luas, baik untuk memberikan penilaian prasarat, penilaian penempatan, penilaian formatif maupun penilaian diagnostik.. Pertimbangan yang digunakan yaitu bahwa seseorang harus memiliki kemampuan prasarat minimal agar ia dapat mengikuti proses belajar-mengajar selanjutnya, demikian pula seseoranghanya dapat mengikuti proses belajar mengajar apabila ia berada pada posisi yang benar-benar sesuai dengan kemampuan yang dimiliki. Seseorang harus diremidiasi,

jika ternyata tidak dapat menampilkan perubahan perilaku sebagaimana yang ditargetkan dalam TIK. Sebaliknya sitem PAN akan lebih tepat jika digunakan dalam penilaian subsumatif dan sumatf, dengan pertimbangan, bahwa tujuan penilaian subsumatif dan sumatif adalah untuk melihat bagaimana posisi seseorang setelah ia mengikuti suatu program. Namun demikian, dalam kurikulum yang digunakan di Indonesia, seperti kurikulum 1984 SMA ataupun kurikulum untuk sekolah menengah lainya, ditegaskan bahwa sistem penilaian yang digunakan adalah berdasar sistem PAP atau sistem mutlak (*absolut*). Yang perlu diperhatikan adalah sepanjang instrumen yang digunakan memenuhi syarat agar data yang diperoleh dapat diolah dengan sistem PAP, kiranya bukan menjadi masalah.

BAB VIII
TINDAK LANJUT HASIL PENILAIAN

Seorang guru wajib melakukan tindak lanjut setelah selesai melakukan penilaian karena gurulah yang menyusun dan melaksanakan program pembelajaran. Ada dua macam tindak lanjut, yakni menggunakan prinsip *assessment for learning* dan *assessment of learning.*

a. A*ssessment for learning*

Dalam konteks *assessment for learning* maka guru wajib menggunakan hasil penilaian untuk meninjau ulang program pembelajarannya. Dalam konteks formatif berarti hasil penilaian digunakan untuk meninjau ulang apakah strategi, media, metode, sumber belajar, dan teknik asesmen yang digunakan sudah mendukung pencapaian KD yang ditargetkan.

Dalam konteks penempatan (*pleacement*), tidak lanjut yang dilakukan guru adalah menempatkan peserta didik sesuai dengan hasil pengukuran kemampuan yang diperoleh. Dalam hal ini, guru akan dapat mendudukkan mana peserta didik yang memiliki kemampuan tinggi dan mana pula pendidik yang memiliki kemampuan rendah.

Dalam konteks penguasaan prerekuisit, guru harus menindaklanjuti dengan memberikan layanan pembelajaran bagi peserta didik yang tidak menguasai prerekuisit yang ditetapkan, agar peserta didik tidak mengalami kesulitan untuk mengikuti proses pembelajaran.

Dalam konteks sumatif, guru harus menindaklanjuti dengan memberikan remediasi bila peserta didik gagal menguasai KD yang bersangkutan. Dalam hal ini diperlukan kajian yang mendalam sebelum guru memberikan layanan remediasi.

b. A*ssessment of learning*

Dalam konteks *assessment of learning* maka evaluasi dilakukan terhadap keberhasilan peserta didik setelah seleai mengikuti suatu program pembelajaran yang diselenggarakan. Dalam hal ini harus diartikan bahwa guru adalah penyusun program, pelaksana program dan dengan demikian harus melakukan evaluasi program. Dalam hal

ini, evaluasi program dapat dilihat dari tingkat efektif dan efisiensinya strategi/metode, sumber belajar, dan teknik penilaian yang dirancang.

Karena peserta program sudah selesai dalam mengikut suatu program, maka tinjauan dalam konteks *assessment of learning* difokuskan kepada peninjauan program pembelajaran yang telah disusun secara keseluruhan, baik dalam bentuk silabus meupun RPP. Dengan demikian, ada kemungkinan guru memperbaiki silabus dan/atau RPP, dan ada pula kemungkinan guru tidak memperbaiki silabus dan/atau RPP bila hasil akhir dari penilaian menunjukkan hasil yang menggembirakan.

Namun demikian harus diingat, karena Standar yang ditetapkan dimungkinkan untuk ditingkatkan, maka guru dapat membenahi silabut dengan meningkatkan kualitas SK dan/atau KD dan/atau indikator.

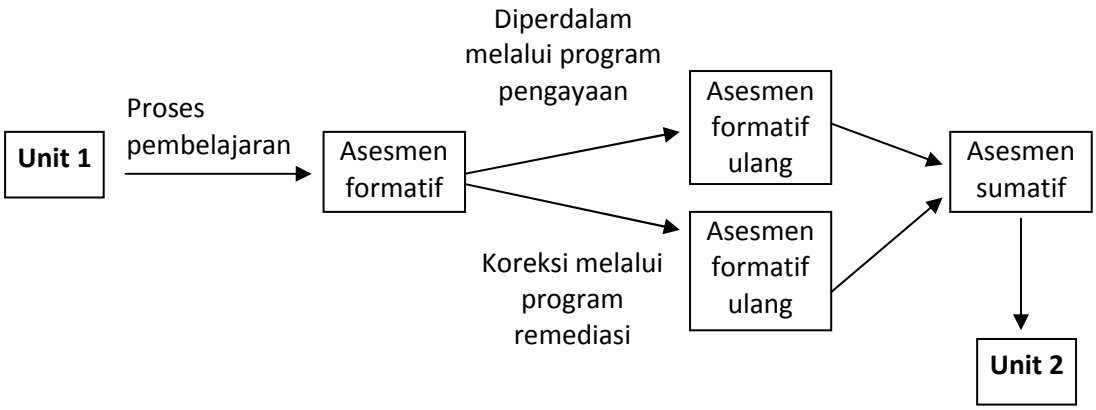
BAB IX
PROGRAM REMEDIASI DAN PENGAYAAN

B. Prinsip Program Remediasi dan Pengayaan

Suatu pembelajaran harus dipantau tinkat keterlaksanaannya dengan melakukan evaluasi formatif. Dasar evaluasi formatif adalah penilaian/asesmen formatif. Bila hasil asesmen formatif menunjukkan setiap anak mengalami kegagalan maka dapat disimpulkan bahwa hasil evaluasi formatif menunjukkan seluruh subjek belajar mengalami kesulitan mengikuti proses pembelajaran. Hal ini tentu saja dapat dihindari bila rancangan pembelajaran (silabus maupun RPP) sudah dievaluasi kelayakannya.

Ada beberapa kemungkinan yang terjadi dalam suatu proses pembelajaran. Kemungkinan yang terjadi adalah semua subjek belajar berhasil mencapai hasil belajar sebagaimana yang ditargetkan, dapat pula sebagian besar saja yang berhasil, sebagian kecil saja, atau semuanya gagal.

Program remediasi merupakan suatu rancangan pembelajaran ulang yang dikenakan bagi peserta didik yang gagal menguasai suatu KD yang ditargetkan. Namun demikian, bukan sekedar mengulang proses pembelajaran seperti yang sudah dilakukan sebelumnya. Proses pembelajaran ulang hanya akan berhasil bila disesuaikan dengan penyebab kegagalannya. Program pengayaan adalah program yang diberikan kepada subjek belajar yang telah berhasil menguasai hasil belajar sesuai yang ditargetkan. Bila dibuat bagan adalah sebagai berikut.



Bila kegagalan terjadi pada semua atau sebagian besar subjek didik maka program remediasi dilaksanakan secara klasikal. Bila kegagalan terjadi pada sebagian besar subjek didik dan program remediasi dilaksanakan secara klasikal maka kepada sebagian kecil

siswa yang telah berhasil menguasai target pembelajaran diberi program pengayaan. Demikian pula Sebaliknya, bila kegagalan terjadi hanya pada sebagian kecil subjek didik maka program remediasi hanya dikenalkan pada mereka yang gagal, sedangkan kepada sebagian besar subjek didik diberikan program pengayaan.

Baik program remediasi maupun program pengayaan sebaiknya sudah dirancang disesuaikan dengan dengan jenis kegagalan yang terjadi. Program remediasi ataupun program pengayaan memerlukan pertimbangan apakah akan dkenakan dalam bentuk klasikal, atau dalam bentuk kelompok, ataukah dalam bentuk individual. Jika direncanakan secara klasikal (yang tentunya diharapkan kecil kemngkinannya akan terjadi bila rancangan awal sudah baik), maka diperlukan strategi baru yang dapat membantu subjek didik untuk mengatasi kegagalannya. Bila dilakukan terhadap kelompok, maka strategi pembelajaran kelompok termasuk strategi *peer-group* dengan memanfaatkan siswa yang benar-benar tuntas, ataupun melalui *study group*.

B. Identifikasi kegagalan siswa:

Untuk mengidentifikasi adanya kegagalan atau kesulitan belajar yang dialami siswa dapat dilakukan melalui

a) Prestasi akademis selama siswa menempuh pembelajarn
   1) interaksi kelas: Anak yang mengalami kesulitan belajar sering memberikan jawaban salah atau memberikan jawaban dengan kalimat yang kacau ketika ditanya atau bahkan tidak memberikan jawaban sama sekali.
   2) Tugas rumah: Anak yang mengalami kesulitan belajar tidak dapat mengerjakan tugas dengan benar, bahkan hanya mengkopi pekerjaan teman.
   3) Hasil tes: Anak yang mengalami kesulitan belajar tidak mampu mengerjakan tes dengan baik. Apabila diberi tes bentuk uraian tidak mampu menyusun kalimat dengan baik meskipun sudah berkala-kali mengganti jawaban, atau bahkan tidak menjawab sama sekali. Bila ada kesempatan ia akan mencontoh pekerjaan pasangan duduknya.

b) Aspek perilaku siswa selama proses pembelajaran
   1) Sikap anak yang mengalami kesulitan belajar tidak menunjukkan minat belajar, lesu, menghindari pertanyaan guru (dengan menunduk misalnya), pasif dalam diskusi atau kerja kelompok.
   2) Menghindar untuk hadir di kelas dengan memberikan berbagai alasan.

3) Bermain-main dengan buku catatannya, seperti dengan menggambar atau membuat coretan-coretan selama pelajaran berlangsung.

C. Jenis-jenis kegagalan

Jenis-jenis kegagalan yang dapat didiagnosis disesuaikan dengan karakteristik kompetensi yang dikembangkan. Beberapa contoh kesulitan yang dapat dihadapi siswa antara lain

1. Sulit menguasai konsep atau rumus
2. Tidak memiliki keterampilan komputasi
3. Tidak menguasai prosedur pemecahan masalah sehingga sering memberikan jawaban yang salah jika diminta memecahkan suatu masalah
4. Tidak memiliki kemampuan mengaplikasikan konsep pada situasi baru, tidak mampu memilih rumus yang tepat untuk menyelesaikan masalah, tidak mampu mengidentifikasi variabel, tidak mampu menginterpretasi makna suatu kalimat.

D. Penyebab kegagalan siswa dalam belajar

Ada banyak faktor yang menyebabkan siswa gagal dalam belajar, baik faktor yang bersifat instrinsik maupun ekstrinsik. Berikut adalah faktor-faktor yang memungkinkan siswa menalamai kesulitan belajar.

1. **Memori/daya mengingat**: Kapasitas mengingat sangat terbatas sehingga mudah lupa.
2. **Pemahaman**: Tidak mampu memahami apa yang dibaca/dipelajari.
3. **Presentasi**: Sukar menyampaikan atau menyatakan pendapat atau tidak memiliki kosa kata yang memadai untuk merangkai kalimat yang akan diucapkan sehingga tergagap-gagap.
4. **Kesenjangan pengetahuan**: Tidak menguasai unit-unit yang dibahas/dipelajari karena sering tidak masuk kelas.
5. **Latar belakang orang tua**: status sosial-ekonomi dan pendidikan
6. **Sikap orang tua**: Sikap orang tua yang acuh tak acuh terhadap anak atau menuntut/berharap tertalu tinggi sehingga anak menjadi tertekan.
7. **Faktor sekolah**: Ketiadaan lingkungan dan peralatan yang kurang memadai atau suasana kelas yang terlalu bising/ramai karena ukuran kelas terlalu besar.

8. **Bahasa pengantar**: siswa tidak menguasai bahasa pengantar yang digunakan di kelas.
9. **Faktor fisik**: Penglihatan dan atau pendengaran yang lemah, sakit-sakitan, mengalamipenyakit yang kronis.
10. **Faktor individual**: tidak siap/malas belajar, tidak disiplin mengerjakan tugas, tidakmau belajar secara teratur; tidak berminat dan/atau tidak mampu berkonsentrasi dengan program yang ditempuh; memiliki sifat kurang percaya diri, rendah diri, takut salah, mudah cemas; takut gagal; berusaha menghindari kelas; ketidakstabilan emosional.
11. **Faktor guru**: Tiada keyakinan diri guru; ketiadaan waktu yang disediakan guru; metode pembelajaran yang tidak tepat, ketidakmampuan guru memotivasi siswa; tugas rumah tidak mendukung apa yang dipelajari/terlalu sedikit tugas rumah yang diberikan guru; cara koreksi pekerjaan rumah yang tidak memadai; kurangnya waktu atau tidak tepatnya bimbingan yang diberikan kepada siswa; tidak mampu menerangkan atau memahamkan siswa terhadap konsep yang sulit dipelajari siswa; ada kesalahan konsep; kurang memberi perhatian atau kurang memberikan bantuan kepada siswa yang memerlukan.

E. Cara melakukan remediasi

Beberapa hal yang perlu dipertimbangkan untuk melakukan remediasi adalah sebagai berikut.

a. Tidak bijaksana kalau peserta remediasi dalam satu kelas lebih dari 5 – 10 siswa.\
b. Memberi perhatian secara individual dan personal kepada peserta remedial.
c. Jangan sampai mencemooh siswa yang mengalami kegagalan.
d. Untuk meningkatkan pemahaman siswa maka proses belajar dapat dilakukan dengan cara “drill” yakni Membaca-Membaca ulang-Menulis-Menulis ulang-Menulis dengan kata-katanya sendiri
e. Memberikan asesmen secara khusus kepada peserta program remedial yang penyebab kegagalannya disebabkan lamban dalam berpikir. Asesmen secara khusus perlu dilakukan secara bertingkat mulai asesmen denan item yang berkategori sangat mudah, kemudian meningkat kepada asesmen dengan item kategori yang lebih sukar, dan diakhiri dengan asesmen yang itemnya kompleks.

f. Model belajar kelompok atau belajar melalui kelompok belajar akan sangat membantu.
g. Berlatih membuat catatan-cacatan kecil untuk mempermudah belajar.
h. Pembelajaran remedial hanya difokuskan pada target dalam silabus.

DAFTAR PUSTAKA

Allen, M.J. & Yen, W.M. (1979*). Introduction to measurement theory*. California: Broaks/Cole Publisihing Company.

Ary, D. & Jacobe, L.Ch., & Razavieh, A. (1985). *Introduction to research in education*, 3-rd ed. New York: Holt, Rinehart and Winston.

Badan Standar Nasional Pendidikan (2007). *Panduan penilaian kelompok mata pelajaran ilmu pengetahuan dan teknologi.* Jakarta: Badan Standar Nasional Pembelajaran.

Brennan, R.L. (2006). *Educational measurement*, 4-th ed. Westport: American Council on Education and Praeger Publishers.

Bryce, T.G.K., McCall, J., MacGregor, J., Robertson, I.J., dan Weston, R.A.J. (1990). *Techniques for assessing process skills in practical science: Teacher's guide*. Oxford: Heinemann Instructional Books.

Carin, A.A. & Sund, R.B. (1989). *Teaching science through discovery* (*6-th ed*). Columbus: Merrill Publishing Company.

Dettmer, P. (2006). New Blooms in Established Fields: four domains of learning and doing [Versi elektronik]. *Roeper Review*, 28, 2, 70-78.

Djemari Mardapi. (2007). *Teknik penyusunan instrumen tes dan non tes*. Yogyakarta: Mitra Cendekia Press.

Drake, S.M. (2007). Creating *standards-based intregrated curriculum*: *Aligning curriculum, content, assessment, and Instructions*. Thousand Oaks: Corwin Press.

Ebel, R.L. & Frisbie, D.A. (1986). *Essentials of educational measurement*, 4-th ed. New Jersey: Prentice-Hall, Inc.

Glencoe. (t.t.). *Performance assessment in the science classroom*. Professional Glencoe Science series. New York: McGraw-Hill.

Gronlund, H.E. (1973). *Preparing criterion-referenced tests for classroom instruction.* New York: The Macmillan Company.Hart, D. (1994). *Authentic assessment: A handbook for educators*. California: Addison-Wiley Publishing Company.

Gronlund, N.E. (1998). *Assessment of student achievement (9-th ed).* Boston: Allyn and Bacon.

Hedges, W.D. (1969). *Testing and evaluation for the science*. Belmont, California: Wadsworth Publishing Company, Inc.

Hegge, Th.G., Kirk, S.A., & Kirk, W.D. (1965). *Remedial reading drills*. Novato, CA.: Academic Therapy Publications. Versi Elektronik. www.academictherapy.com. Diakses tanggal 2 Agustus 2010.

Hibbard, K.M. (t.t.). *Performance assessment in the science classroom*. New York: McGraw-Hill Companies.

Masri Singarimbun. (1982). Metode penelitian survei (Edt). Jakarta: LP3ES.

McMillan, J.H. (Ed). (2007). *Formative classroom assessment: Theory into practice.* New York: Teacher College, Columbia University.

Miller, P.W. (2008). Measurement and teaching. Munster: Patric W. Miller & Associates.

Popham, W.J. (2005). *Classroom assessment: What teachers need to know (4-thed*). Boston: Pearson Education, Inc.

Prabhat Marwaha. (2009). *Role of remedial teaching*. www.hicow.com/education/**teacher**/classroom-173494.html. Diakses tanggal 2 Agutus 2010.

Rezba, R.J., Sparague, C.S., Fiel, R.L., Funk, H.J., Okey, J.R., & Jaus, H.H. (1995). *Learning and assessing science process skills*. 3rd ed. Iowa: Kendall/Hunt Publishing Company.

Roid, G.H. & Haladyna, Th.M. (1982). *A technology for test-item writing*. Oriando: Academic Press, Inc.

Smith, J.K. (2003). Reconsidering reliability in classroom assessment and grading [Versi elektronik]. *Educational Measurement, Issues and Practice*, 22, 4, 26-33.

Stark, S., Chernyshenko, S., Chuah, D.,Wayne Lee, & Wilington, P. (2001). *IRT modeling lab: IRT tutorial* [Versi elektronik]. Urbana: University of Illinois.

Stark, S., Chernyshenko, S., Chuah, D.,Wayne Lee, & Wilington, P. (2001). *IRT modeling lab: Test Development Using Classical Test Theory* [Versi elektronik]. Urbana: University of Illinois.