

Unit 4

MENGEMBANGKAN TES SEBAGAI INSTRUMEN EVALUASI

Endang Poerwanti

Masduki

PENDAHULUAN

Saudara, Anda seharusnya sudah tahu bahwa pelaksanaan Kurikulum Berbasis Kompetensi mengharuskan semua guru sebagai pendidik untuk pengembangan silabus dan merubah pula sistem penilaian yang digunakan dengan menerapkan sistem penilaian berbasis kompetensi. Sistem penilaian berbasis kompetensi lebih mengarah pada penilaian kelas, yaitu penilaian yang dilakukan secara menyeluruh dan berkesinambungan. Penilaian ini tidak hanya menitikberatkan pada kemampuan kognitif tetapi juga mencakup ranah psikomotor, dan afektif.

Anda pasti juga masih ingat bahwa silabus merupakan acuan untuk merencanakan dan melaksanakan program pembelajaran, sedang sistem penilaian berbasis kompetensi mencakup jenis tagihan, dan bentuk soal. Jenis tagihan adalah berbagai bentuk ulangan dan tugas-tugas, tagihan, seperti ulangan atau tugas-tugas yang harus dikerjakan peserta didik. Bentuk soal terkait dengan jawaban yang harus dilakukan oleh peserta didik, seperti bentuk isian singkat, pilihan ganda, uraian, objektif, uraian non objektif, dan sebagainya. Untuk itu sebaiknya dirancang secara tertulis dan rapi sistem penilaian yang akan dilakukan selama satu semester. Rancangan penilaian bersifat terbuka bagi siswa, guru lain, dan kepala sekolah.

Dalam merancang penilaian, pendidik dapat melakukannya dengan cara, yakni: (1) mencermati silabus dan sistem penilaian yang sudah ada, (2) menyusun sistem penilaian dengan KBK berdasarkan silabus dan sistem penilaian yang telah disusun, (3) menentukan bobot masing-masing jenis tagihan, dan (4) menyusun rancangan sistem penilaian dengan KBK. Rancangan penilaian ini diinformasikan kepada siswa pada awal pertemuan (awal semester). Dengan demikian sistem penilaian yang dilakukan guru semakin sempurna atau semakin memenuhi prinsip-prinsip penilaian.

Dalam melakukan asesmen atau penilaian Anda dapat menggunakan berbagai jenis pendekatan dan instrumen untuk dapat memperoleh data yang akurat tentang kemajuan belajar peserta didik, secara garis besar teknik tersebut dibagi menjadi dua,

yaitu teknik tes dan nontes. Dalam penilaian klasik, teknik tes merupakan teknik yang paling banyak digunakan, namun dalam penilaian KBK kedua teknik harus digunakan sesuai dengan kebutuhan di kelas. Anda sebagai pendidiklah yang tahu persis kapan harus menggunakan teknik tes dan kapan pula harus menggunakan teknik nontes, sesuai dengan kompetensi dasar dan indikator keberhasilan belajar peserta didik yang akan diukur, karena tes merupakan alat ukur untuk mengetahui tingkat pencapaian siswa terhadap kompetensi yang dipersyaratkan. Dalam kaitan dengan pembelajaran, aspek tersebut adalah indikator pencapaian kompetensi. Pada Unit 4 ini Anda akan dapat mencermati bagaimana mengembangkan tes sebagai instrumen asesmen, sedangkan di Unit 5 akan dapat Anda pelajari secara tuntas bagaimana mengembangkan asesmen nontes. Jadi, setelah mempelajari Unit 4 ini diharapkan Anda dapat:

1. memahami pengertian dan jenis tes;
2. langkah-langkah menyusun tes;
3. kriteria tes yang baik;
4. mengembangkan Tes sebagai instrumen asesmen.

Anda akan dapat memahami unit ini secara optimal dengan cara membaca dengan cermat semua uraian yang ada, mendiskusikan dengan teman-teman berdasarkan pengalaman di kelas, mengerjakan tugas-tugas latihan serta mengerjakan tes yang ada pada setiap subunit dengan bersungguh-sungguh. Selamat belajar, kesuksesan akan menyertai orang yang bersungguh-sungguh, kita akan mencapai kesuksesan tersebut karena kita bersungguh-sungguh.

Subunit 1

Pengertian dan Jenis Tes Sebagai Instrumen Asesmen

Pengantar

Saudara, setelah kita mengetahui hakekat tes, yaitu sebagai alat ukur, perlu kiranya kita membahas jenis-jenis tes. Perlu kita bersama mengetahui bahwa para ahli dalam bidang tes tidak semuanya seragam dalam mengklasifikasikan tes. Heaton (1988), misalnya, membagi jenis tes menjadi 4 bagian utama, yaitu: (1) tes hasil belajar (*achievement test*), (2) tes penguasaan (*proficiency test*), (3) tes bakat (*aptitude test*), dan (4) tes diagnostik (*diagnostic test*). Sementara itu, Brown (2004) melengkapi satu lagi jenis tes terhadap penggolongan yang telah dilakukan oleh Heaton, yaitu tes penempatan (*placement test*). Nampaknya, penggolongan jenis tes tersebut hanya mengacu kepada satu kriteria saja yaitu tujuan penyelenggaraan tes.

Saudara diajak untuk mencermati pembagian jenis-jenis tes. Pembagian yang ia tawarkan nampak lebih luas dan rinci dengan mengacu pada sejumlah kriteria. Kriteria yang dapat digunakan untuk membedakan jenis tes meliputi: 1) tujuan penyelenggaraan, 2) tahapan/waktu penyelenggaraan, 3) cara mengerjakan, 4) cara menyusun, 5) bentuk jawaban, 6) cara penilaian, dan 7) acuan penilaian. Uraian tentang jenis tes berikut dilengkapi dengan uraian mengenai arti, cakupan, ciri-ciri serta contohnya

1. Pengertian Tes

Tes secara sederhana dapat diartikan sebagai himpunan pertanyaan yang harus dijawab, pernyataan-pernyataan yang harus dipilih/ditanggapi, atau tugas-tugas yang harus dilakukan oleh peserta tes dengan tujuan untuk mengukur *suatu aspek tertentu* dari peserta tes. Dalam kaitan dengan pembelajaran aspek tersebut adalah indikator pencapaian kompetensi. Tes berasal dari bahasa Perancis yaitu “*testum*” yang berarti piring untuk menyisihkan logam mulia dari material lain seperti pasir, batu, tanah, dan sebagainya. Kemudian diadopsi dalam psikologi dan pendidikan untuk menjelaskan sebuah instrumen yang dikembangkan untuk dapat melihat dan mengukur dan menemukan peserta Tes yang memenuhi kriteria tertentu. Cronbach (dalam Azwar, 2005) mendefinisikan tes sebagai “*a systematic procedure for*

observing a person's behavior and describing it with the aid of a numerical scale or category system". Menurut Ebster's Collegiate (dalam Arikunto, 1995), tes adalah serangkaian pertanyaan atau latihan atau alat lain yang digunakan untuk mengukur keterampilan, pengetahuan, intelegensia, kemampuan atau bakat yang dimiliki oleh individu atau kelompok. Dari dua definisi tersebut dan uraian lebih jauh tentang itu dapat ditarik pengertian bahwa: (1) tes adalah prosedur pengukuran yang sengaja dirancang secara sistematis, untuk mengukur atribut tertentu, dilakukan dengan prosedur administrasi dan pemberian angka yang jelas dan spesifik, sehingga hasilnya relatif ajeg bila dilakukan dalam kondisi yang relatif sama; (2) tes pada umumnya berisi sampel perilaku, cakupan butir tes yang bisa dibuat dari suatu materi tidak terhingga jumlahnya, yang secara keseluruhan mungkin mustahil dapat tercakup dalam tes, sehingga tes harus dapat mewakili kawasan (domain) perilaku yang diukur, untuk itu perlu pembatasan yang jelas; (3) tes menghendaki subjek agar menunjukkan apa yang diketahui atau apa yang dipelajari dengan cara menjawab atau mengerjakan tugas dalam tes. Respon subjek atas tes merupakan perilaku yang ingin diketahui dari penyelenggaraan tes, karena tes memang mengukur perilaku, sebagai manifestasi atribut psikologis yang mau diukur.

Tes pada dasarnya adalah alat ukur atribut psikologis yang objektif atas sampel perilaku tertentu. Dalam psikologi, tes dapat diklasifikasikan menjadi empat, yaitu: (1) tes yang mengukur intelegensia umum yang dirancang untuk mengukur kemampuan umum seseorang dalam suatu tugas; (2) tes yang mengukur kemampuan khusus atau tes bakat yang dibuat untuk mengungkap kemampuan potensial dalam bidang tertentu; (3) tes yang ditujukan untuk mengukur prestasi yang digunakan untuk mengungkapkan kemampuan aktual sebagai hasil belajar; (4) tes yang mengungkap aspek kepribadian (*personality assesment*) yang bertujuan mengungkap karakteristik individual subjek dalam aspek yang diukur. Dengan melihat penggolongan di atas, tes dalam pembelajaran di kelas yang menjadi pembahasan ini adalah tes prestasi atau hasil belajar. Tes sebagai alat ukur dapat menyediakan informasi-informasi obyektif yang dapat digunakan sebagai pertimbangan dalam penentuan keputusan yang harus diambil pendidik terhadap proses dan hasil belajar yang dilakukan siswa dapat dibagi menjadi tiga kelompok besar yaitu:

a. Keputusan yang diambil pada pemulaan proses pembelajaran

Penggunaan tes sebagai dasar pengambilan keputusan pada permulaan proses pembelajaran bermuara pada dua pertanyaan yang harus dijawab oleh pendidik sebelum memulai proses pembelajaran yaitu; (1) sejauhmanakah pengetahuan, keterampilan dan kemampuan yang harus dimiliki oleh siswa sebelum mengikuti

proses pembelajaran yang berupa kemampuan awal yang diperlukan untuk mengikuti proses pembelajaran, (2) sejauhmanakah kemampuan dan keterampilan yang telah dicapai peserta didik terhadap pembelajaran yang direncanakan. Keduanya akan menentukan keputusan guru dalam merancang materi dan metode pembelajaran yang direncanakan.

b. Keputusan selama proses pembelajaran

Tes dapat pula digunakan selama proses pembelajaran (tes formatif). Tes formatif dapat diberikan baik dalam bentuk tes tulis maupun tes lisan, baik dengan jawaban uraian maupun tes obyektif.

c. Keputusan-keputusan pada akhir pembelajaran

Tes formatif yang diberikan guru pada akhir pembelajaran ditujukan untuk mengetahui apakah kompetensi dasar yang dirumuskan dalam program pembelajaran (satuan pembelajaran) telah tercapai atau belum. Jadi, fungsi tes pada akhir pembelajaran adalah untuk mengukur daya serap siswa pada materi pembelajaran. Sehingga guru dapat merencanakan tindak lanjut terhadap rencana, proses, media, metode, dan suasana pembelajaran. Seperti penilaian selama proses keputusan akhir pembelajaran dapat berasal dari informasi tes obyektif atau tes subyektif.

2. Jenis-jenis Tes

Bila kita membahas jenis-jenis tes, Anda akan dapat mencermati dalam lima jenis atau cara pembagian yaitu:

- a. Pembagian jenis tes berdasarkan tujuan penyelenggaraan.
- b. Jenis tes berdasarkan waktu penyelenggaraan.
- c. Pembagian jenis tes berdasarkan cara mengerjakan.
- d. Pembagian jenis tes berdasarkan cara penyusunan.
- e. Pembagian jenis tes berdasarkan bentuk jawaban.

Uraian selengkapnya adalah sebagai berikut.

1) Jenis Tes Berdasarkan Tujuan Penyelenggaraan

Saudara, untuk mengawali pembahasan tentang jenis-jenis tes, Anda akan diminta untuk menjawab satu pertanyaan di bawah ini.

Untuk apakah Anda menyelenggarakan tes?

Saudara, cobalah berpikir sejenak untuk merenungkan jawaban atas pertanyaan tadi. Anda diminta untuk tidak tergesa-gesa melanjutkan membaca

penjelasan berikut ini sebelum Anda menentukan jawaban. Jika Anda sudah menemukan jawaban, barangkali jawaban Anda akan lebih dari satu jawaban, seperti yang tertera di bawah ini, tes diselenggarakan dengan tujuan:

- a) untuk keperluan seleksi,
- b) untuk menempatkan orang pada kelas-kelas tertentu,
- c) untuk mengetahui hasil belajar,
- d) untuk keperluan diagnostik, dan
- e) untuk keperluan uji coba

Penjelasan secara rinci adalah sebagai berikut:

a) Tes Seleksi (*Selection Test*)

Saudara, Anda bisa memahami hakekat dari tes seleksi ini dari arti kata “seleksi” itu sendiri, yaitu memilih. sederhana bukan? Jadi, tes seleksi diselenggarakan untuk memilih peserta guna diikutsertakan dalam kegiatan yang menuntut kemampuan tertentu. Penentuan jenis kemampuan dan tingkat penguasaan pada tes seleksi, sepenuhnya tergantung pada kebutuhan akan kemampuan yang dibutuhkan untuk dapat mengikuti kegiatan. Dengan demikian, berdasarkan hasil tes seleksi, seseorang dapat dinyatakan diterima atau berhasil dan tidak diterima atau tidak lolos untuk mengikuti program kegiatan yang direncanakan. Sebagai contoh, jika kita menyelenggarakan tes seleksi untuk pemandu wisata, maka akan lebih baik menitikberatkan kemampuan berbicara daripada kemampuan menulis.

Latihan

Sekarang cobalah Anda tentukan, kemampuan manakah yang lebih Anda pentingkan dalam tes seleksi untuk seorang yang akan melakukan tugas redaksional dalam sebuah media cetak? Kemampuan menulis, kemampuan berbicara atau kemampuan menyimak? Mengapa kemampuan yang Saudara tentukan itu penting?

b) Tes Penempatan (*Placement Test*)

Saudara, adalah suatu keniscayaan bahwa kemampuan seseorang tidaklah bisa sama. Sekelompok orang barangkali memiliki kemampuan lebih tinggi dari pada kelompok lainnya. Permasalahan yang muncul adalah, bagaimanakah jika kemampuan siswa dalam satu kelas relatif beragam? Hal ini akan bisa mempersulit jalannya proses pengajaran yang Anda lakukan. Untuk itu perlu dilakukan tes penempatan. Tes penempatan umumnya diselenggarakan menjelang dimulainya suatu program pengajaran, dengan maksud untuk menempatkan seseorang pada kelompok yang sesuai dengan tingkat kemampuan yang dimilikinya.

c) Tes Hasil Belajar (*Achievement Test*)

Tes hasil belajar tentu tidak lagi asing bagi Saudara. Brown (2004) memberikan pengertian tes hasil belajar merupakan “a test to see how far students achieve materials addressed in a curriculum within a particular time frame”. Hasil belajar yang diungkap lewat tes hasil belajar dapat mengacu pada hasil pengajaran secara keseluruhan pada akhir penyelenggaraan atau pada kurun waktu tertentu.

Sebagai tes yang memfokuskan pada hasil yang telah dapat dicapai oleh suatu bentuk pengajaran, tes hasil belajar memiliki kaitan yang erat dengan apa yang telah diajarkan (kurikulum). Kaitan itu terutama dalam hal isi tes. Isi tes harus secara jelas mencerminkan isi pengajaran yang secara nyata telah diselenggarakan.

d) Tes Diagnostik (*Diagnostic Test*)

Secara etimologis, diagnostik diambil dari bahasa Inggris “diagnostic”. Bentuk kata kerjanya adalah “to diagnose”, yang artinya “to determine the nature of disease from observation of symptoms”. Mendiagnosis berarti melakukan observasi terhadap penyakit tertentu, sebagai dasar menentukan macam atau jenis penyakitnya. Jadi, tes diagnostik sengaja dirancang sebagai alat untuk menemukan kesulitan belajar yang sedang dihadapi siswa. Hasil tes diagnostik dapat digunakan sebagai dasar penyelenggaraan pengajaran yang lebih sesuai dengan kemampuan siswa sebenarnya, termasuk kesulitan-kesulitan belajarnya. Tes ini dilakukan apabila diperoleh informasi bahwa sebagian besar peserta didik gagal dalam mengikuti proses pembelajaran pada mata pelajaran tertentu. Hasil tes diagnostik memberikan

informasi tentang konsep-konsep yang belum dipahami dan yang telah dipahami. Oleh karenanya, tes ini berisi materi yang dirasa sulit oleh siswa, namun tingkat kesulitan tes ini cenderung rendah.

e) Tes Uji Coba

Apabila Saudara sebagai seorang guru pasti pernah mengembangkan tes. Tes yang dikembangkan belum tentu memenuhi kualifikasi sebagai tes yang “baik” dalam arti luas. Untuk mengetahui apakah tes yang dikembangkan bagus, perlu serangkaian uji coba, untuk memperoleh informasi, tidak hanya tentang ciri-ciri tes yang penting, seperti validitas, reliabilitas, tingkat kesulitan, dan tingkat pembeda, melainkan juga segi-segi lain, seperti kecukupan waktu, kejelasan tulisan maupun perintah tes, dan lain sebagainya.

2) Jenis Tes Berdasarkan Tahapan/Waktu Penyelenggaraan

Saudara, selanjutnya Anda diajak untuk memperhatikan jenis tes berdasar waktu penyelenggaraan tes, yang terbagi menjadi 4 yaitu:

a) Tes Masuk (*Entrance Test*)

Tes masuk diselenggarakan sebelum dan menjelang suatu program pengajaran dimulai. Sama dengan tes seleksi, tes masuk diselenggarakan untuk menentukan apakah seorang calon dapat diterima sebagai peserta program pengajaran karena ia memiliki jenis dan kemampuan yang dipersyaratkan. Tes masuk dirancang secara khusus dan disesuaikan dengan tujuan program pengajaran. Semakin sesuai isi tes masuk itu dengan tujuan pokok program pengajaran, maka akan semakin tinggi tingkat relevansi serta efektivitas dari tes masuk tersebut.

b) Tes Formatif (*Formative Test*)

Tes formatif dilakukan pada saat program pengajaran sedang berlangsung (*progress*), tujuannya untuk memperoleh informasi tentang jalannya pengajaran sampai tahap tertentu. Informasi tersebut penting untuk mengetahui apakah program pengajaran berjalan sesuai dengan format yang ditentukan sehingga dipertahankan atau program pembelajaran memerlukan perubahan atau penyesuaian, hasilnya berguna untuk memperbaiki strategi mengajar. Tes ini dilakukan secara periodik sepanjang rentang proses pembelajaran, materi tes dipilih berdasarkan tujuan pembelajaran tiap pokok bahasan atau sub pokok materi. Jadi tes untuk

menentukan keberhasilan belajar dan untuk mengetahui keberhasilan proses pembelajaran.

c) Tes Sumatif (*Summative Test*)

Kata dari “sumatif” adalah “sum” yang berarti “*total obtained by adding together items, numbers or amounts*”. Dengan demikian, tes sumatif diselenggarakan untuk mengetahui hasil pengajaran secara *keseluruhan* (total). Konsekuensi dari tes yang menekankan hasil pengajaran secara keseluruhan, maka item tes sumatif atau bahan cakupannya meliputi seluruh materi yang telah disampaikan. Tes sumatif diberikan di akhir suatu pelajaran, atau akhir semester. Hasilnya untuk menentukan keberhasilan belajar peserta didik. Tingkat keberhasilan dinyatakan dengan skor atau nilai, pemberian sertifikat, dan sejenisnya.

d) Pra-tes dan *Post-test*

Untuk mengetahui kemampuan yang dimiliki seorang siswa di awal program pengajaran, kadang-kadang diselenggarakan pra-tes. Hasil pra-tes digunakan untuk mengetahui tingkat kemampuan siswa pada awal program pengajaran. Tingkat kemampuan awal ini penting untuk menentukan sejauhmana kemajuan seorang siswa. Kemajuan yang dicapai bisa dilihat dari perbandingan hasil pra-tes dengan hasil tes yang diselenggarakan di akhir program pengajaran (*post-test*).

3) Jenis Tes Berdasarkan Cara Mengerjakan

Saudara, secara umum, tes dapat dikerjakan secara tertulis dan secara lisan. Selanjutnya, Saudara dapat mencermati pembahasan berikut ini.

a) Tes Tertulis

Tes tertulis adalah tes yang dilakukan secara tertulis baik dalam hal soal maupun jawabannya, namun tes yang disampaikan secara lisan dan dikerjakan secara tertulis masih digolongkan ke dalam jenis tes tertulis. Sebaliknya, tes yang soalnya diberikan dalam bentuk tulisan sedangkan jawabannya berbentuk lisan tidak dapat dikategorikan ke dalam bentuk tes tertulis.

b) Tes Lisan

Pada tes lisan, baik pertanyaan maupun jawaban (*response*) semuanya dalam bentuk lisan. Karenanya, tes lisan relatif tidak memiliki rambu-rambu penyelenggaraan tes yang baku, karena itu, hasil dari tes lisan biasanya tidak menjadi informasi pokok tetapi pelengkap dari instrumen asesmen yang lain.

c) Tes Unjuk Kerja

Pada Tes ini peserta didik diminta untuk melakukan sesuatu sebagai indikator pencapaian kompetensi yang berupa kemampuan psikomotor.

4) Jenis Tes Berdasarkan Cara Penyusunan

Berdasarkan kriteria ini, tes dapat dibedakan menjadi dua, yaitu (1) tes buatan guru dan (2) tes terstandar.

a) Tes Buatan Guru (*Teacher-made Test*)

Saudara tentu mengetahui tugas-tugas utama yang harus diemban oleh seorang guru. Untuk melakukan tugas evaluasi itu, seorang guru harus mengembangkan alat ukur, salah satunya tes. Tes yang dikembangkan sendiri oleh guru disebut tes buatan guru (*teacher-made test*). Jadi tes buatan guru adalah tes yang dirancang dan dipersiapkan oleh guru, tetap dengan mengacu pada karakteristik tes yang baik dan dilakukan secara cermat, untuk tetap menjamin validitas maupun reliabilitasnya.

b) Tes Terstandar (*Standardized Test*)

Saudara, dari istilah yang digunakan saja, barangkali Anda sudah bisa memperkirakan apa yang dimaksud dengan tes terstandar. Benar, tes terstandar adalah tes yang dikembangkan dengan mengikuti prosedur serta prinsip pengembangan tes secara ketat. Semua prosedur pengembangan tes diikuti sehingga ciri-ciri tes sebagai alat ukur yang baik senantiasa dapat dipenuhi. Dengan demikian, tingkat validitas, reliabilitas, kepraktisan, maupun daya beda sudah bukan menjadi masalah lagi. Bagaimana cara mengembangkan tes sebagai alat ukur yang baik, Saudara bisa membaca bagian lain dari Bahan Ajar ini.

5) Jenis Tes Berdasarkan Bentuk Jawaban

Saudara, jika kita melihat bentuk jawaban yang diberikan oleh peserta tes, kita dapat membedakan tiga jenis tes, yaitu; (a) tes esei, (b) tes jawaban pendek, dan (c) tes obyektif. Untuk lebih jelasnya, cobalah perhatikan bahasan berikut ini.

a) Tes Esei (*Essay-type Test*)

Tes bentuk uraian adalah tes yang menuntut siswa mengorganisasikan gagasan-gagasan tentang apa yang telah dipelajarinya dengan cara mengemukakannya dalam bentuk tulisan. Keunggulan tes uraian, guru dapat mengukur kemampuan siswa dalam hal mengorganisasikan pikirannya, mengemukakan pendapatnya, dan mengekspresikan gagasan dengan menggunakan kata-kata atau kalimat sendiri. Sedang keterbatasannya adalah cakupan materi pelajaran yang terbatas, waktu pemeriksaan jawaban yang lama, penskorannya cenderung subyektif dan umumnya kurang handal dalam pengukuran.

b) Tes Jawaban Pendek

Saudara, tes dapat digolongkan menjadi tes jawaban pendek jika peserta tes diminta menuangkan jawabannya bukan dalam bentuk esei, tetapi memberikan jawaban-jawaban pendek, dalam bentuk rangkaian kata-kata pendek, kata-kata lepas, maupun angka-angka. Termasuk ke dalam tes jenis ini adalah tes yang mewajibkan siswa untuk mengisi bagian yang kosong dari sebuah kalimat atau teks. Sehingga diharapkan dapat memberikan jawabannya sesingkat mungkin.

c) Tes Objektif

Tes objektif adalah tes yang keseluruhan informasi yang diperlukan untuk menjawab tes telah tersedia. Oleh karenanya sering pula disebut dengan istilah tes pilihan jawaban (*selected response test*). Butir soal telah mengandung kemungkinan

jawaban yang harus dipilih atau dikerjakan oleh peserta tes. Menurut Subino (1987) perbedaan yang khas bentuk soal objektif dibanding dengan soal esei adalah tugas peserta tes (*testee*) dalam merespons tes. Pada tes objektif, tugas *testee* adalah memanipulasikan data yang telah ada dalam butir soal. Oleh karenanya, tes objektif adalah tes yang dalam pemeriksaannya dapat dilakukan secara objektif. Karena sifatnya yang objektif maka penskorannya dapat dilakukan dengan bantuan mesin. Soal ini tidak memberi peluang untuk memberikan penilaian yang bergradasi karena dia hanya mengenal benar dan salah. Soal tes objektif sangat bermanfaat untuk mengukur hasil belajar kognitif tingkat rendah. Hasil-hasil belajar kompleks seperti menciptakan dan mengorganisasikan gagasan kurang cocok diukur menggunakan soal bentuk ini. Soal objektif sangat bervariasi bentuknya. Variasi yang bisa dibuat dari soal objektif adalah benar-salah, pilihan ganda, menjodohkan, melengkapi dan jawaban singkat.

Rangkuman

Tes secara sederhana dapat diartikan sebagai himpunan pertanyaan yang harus dijawab, pernyataan-pernyataan yang harus dipilih/ditanggapi, atau tugas-tugas yang harus dilakukan oleh peserta tes dengan tujuan untuk mengukur suatu aspek tertentu dari peserta tes. Jenis-jenis tes dapat dikelompokkan menjadi beberapa model klasifikasi yaitu:

Pembagian jenis Tes berdasarkan tujuan penyelenggaraan.

Jenis Tes berdasarkan waktu penyelenggaraan.

Pembagian jenis tes berdasarkan cara mengerjakan.

Pembagian jenis Tes berdasarkan cara Penyusunan.

Pembagian jenis Tes berdasarkan bentuk jawaban.

Tes Formatif 1

Di bawah ini dicantumkan tes formatif yang bertujuan untuk mengukur pemahaman Anda mengenai uraian, contoh, dan rangkuman yang tercantum dalam subunit 1. Jawablah pertanyaan berikut sesuai dengan permintaan!

1. Jelaskan pembagian jenis tes dilihat dari tujuan penyelenggaraannya!
2. Jelaskan pembagian jenis tes berdasarkan bentuk jawabannya!
3. Jelaskan dan berikan contoh tes yang melandasi pengambilan keputusan guru diawal proses pembelajaran!
4. Jelaskan dan berikan contoh jenis tes yang melandasi pengambilan keputusan guru selama proses pembelajaran berlangsung!
5. Jelaskan dengan contoh perbedaan antara tes obyektif dengan tes esei!

Umpan Balik dan Tindak Lanjut

Jawablah pertanyaan tes formatif di atas, setelah selesai baru cocokkan dengan kunci jawabannya. Diskusikan dengan teman bila jawaban belum sesuai atau Anda belum merasa masih ada hal-hal yang meragukan. Hal ini sangat diperlukan karena kesepahaman tentang pengertian ini akan mendasari dan mempengaruhi langkah dan kegiatan dalam menyelesaikan mata kuliah ini.

Subunit 2

Langkah-Langkah Menyusun Tes

Pengantar

Telah Anda pahami dari penjelasan-penjelasan sebelumnya bahwa proses pengukuran merupakan proses kuantifikasi terhadap atribut, benda atau gejala tertentu. Proses pengukuran diharapkan dapat menghasilkan data yang valid dan akurat sehingga harus dilakukan secara terencana dan sistematis. Pengukuran berbagai atribut yang berupa benda ataupun aspek-aspek fisik seperti mengukur tinggi bangunan, mengukur tinggi bangunan imbang beras, mengukur tinggi badan, berat badan, luas tanah, suhu udara, ataupun kecepatan motor sangat mungkin dapat dilakukan dengan tepat karenanya dapat diterima secara universal karena validitasnya sangat mudah dibuktikan. Tinggi suatu bangunan dengan mudah dapat diukur dengan centimeter, meter, berat beras dengan cepat dapat diukur dengan timbangan dan sebagainya, dimana ketepatan (validitas) maupun keajegan hasil pengukurannya (reliabilitas) serta obyektivitas hasil pengukurannya tidak lagi perlu diragukan, karena dengan mudah akan dapat dilakukan pengukuran ulang dengan hasil yang sama persis.

Bagi kita sebagai pendidik, yang menjadi persoalan kemudian adalah pengukuran hasil belajar yang termasuk bidang non fisik atau aspek yang bersifat abstrak. Dalam hal ini pendidik harus paham bahwa aspek yang bersifat abstrak seperti hasil belajar ini dalam melakukan pengukuran memerlukan perencanaan dan pelaksanaan yang sistematis. Alat yang biasa digunakan sebagai alat ukur dari hasil belajar adalah tes. Sehingga dapat dikatakan bahwa tes merupakan salah satu alat ukur dalam melakukan asesmen proses dan hasil pembelajaran. Seperti halnya atribut psikologis yang lain ketika melakukan pengukuran terhadap hasil belajar, tes sebagai alat ukur mungkin tidak akan pernah dapat menggambarkan hasil dengan validitas dan reliabilitas ataupun obyektivitas yang sempurna. Untuk itu dalam menyusun tes sebagai alat ukur hasil belajar perlu dipertimbangkan beberapa permasalahan yang merupakan keterbatasan dari tes sebagai alat ukur psikologis (Saifuddin, 2005):

1. Atribut psikologis termasuk hasil belajar bersifat abstrak dan laten sehingga apa yang diukur adalah suatu konstruk yang memang tidak dapat diukur secara langsung, sehingga dilakukan berdasarkan indikator perilaku yang mungkin belum tentu mewakili domain yang tepat, karena batasan dari konstruk tersebut tidak mungkin dapat dijabarkan dalam akurasi yang sempurna.
2. Dalam atribut yang bersifat laten atau abstrak sering kali didasari pada indikator yang jumlahnya terbatas, keterbatasan dalam menjabarkan indikator perilaku ini menyebabkan hasil pengukuran menjadi kurang komprehensif, di samping itu penjabaran indikator perilaku tersebut masih mungkin terjadi tumpang tindih dengan indikator dari atribut psikologis yang lain.
3. Respon yang diberikan oleh siswa sebagai subyek sedikit banyak akan dipengaruhi oleh faktor lain yang tidak relevan, baik yang bersumber dari dirinya sendiri maupun dari variabel dari luar dirinya, seperti misalnya suasana hati, sakit, kondisi dan situasi sekitar, cetakan yang tidak jelas, pengawasan waktu pengerjaan, sistem administrasi dan sebagainya.
4. Atribut psikologis termasuk hasil belajar yang terdapat pada diri siswa, sering kali bersifat tidak stabil dan mudah sekali berubah, seiring dengan perubahan situasi dan kondisi sesaat hingga interpretasi terhadap hasil tes sebagai alat ukur hanya dapat dilakukan secara normatif, dalam pengertian banyak sekali sumber bias yang harus diperhitungkan.

Anda juga harus memahami bahwa keterbatasan-keterbatasan dalam melakukan pengukuran hasil belajar dengan menggunakan tes ini menjadikan kita sebagai pendidik yang juga penyusun tes harus mempersiapkan semuanya secara lebih teliti, karena prosedur konstruksi psikologis lebih rumit sehingga harus dilakukan dengan perencanaan yang sangat teliti dan mengikuti langkah-langkah yang sistematis untuk meminimalkan berbagai sumber kesalahan yang mungkin terjadi.

1. Langkah Pokok Mengembangkan Tes

Anda sebagai pendidik seringkali kurang menyadari bahwa, mengembangkan tes sebagai instrumen asesmen proses dan hasil belajar adalah menyusun alat ukur suatu gejala yang bersifat abstrak yaitu pemahaman dan penguasaan anak terhadap materi yang berupa seperangkat kompetensi dipersyaratkan, dan dalam kenyataan di lapangan sebagian besar tenaga pengajar memang menggunakan teknik tes sebagai upaya untuk mengukur hasil belajar tersebut. Karena demikian seringkali pengajar menyusun tes hasil belajar, justru sering menimbulkan kecerobohan, karena menganggap hal ini sebagai hal yang sudah biasa/umum dilakukan, dan kurang perlu

mempersiapkannya secara cermat. Padahal penyusunan tes, sangat besar pengaruhnya terhadap siswa yang akan mengikuti tes, untuk mengurangi kesalahan dalam pengukuran maka tes harus direncanakan secara cermat. Secara umum ada lima langkah pokok yang harus dilewati yaitu:

a) Perencanaan Tes

Dalam langkah perencanaan tes ada beberapa kegiatan yang harus dilakukan guru sebagai pendidik yaitu:

- (1) **Menentukan cakupan materi yang akan diukur** yang menyangkut penetapan cakupan materi dan aspek (ranah) kemampuan yang akan diukur. Penetapan ini penting mengingat bahwa kemampuan belajar merupakan proses yg kompleks dan menyangkut pemahaman yang bersifat abstrak, sehingga harus jelas pada bagian mana cakupan materi yang akan diukur dan dikembangkan dalam soal tes, langkah ini biasanya dilakukan dengan menyusun kisi-kisi soal yaitu daftar spesifikasi. Ada tiga langkah dalam mengembangkan kisi-kisi tes dalam sistem penilaian berbasis kompetensi dasar, yaitu; (1) Menulis kompetensi dasar, (2) Menulis materi pokok, (3) Menentukan indikator, dan (4) Menentukan jumlah soal.
- (2) **Bentuk Tes:** Pemilihan bentuk tes akan dapat dilakukan dengan tepat bila didasarkan pada tujuan tes, jumlah peserta tes, waktu yang tersedia untuk memeriksa lembar jawaban tes, cakupan materi tes, dan karakteristik mata pelajaran yang diujikan. Misalnya, bentuk tes objektif pilihan ganda dan bentuk tes benar salah cocok digunakan bila jumlah peserta tes banyak, waktu koreksi singkat, dan cakupan materi yang diujikan banyak. Bentuk tes objektif lebih cocok digunakan pada mata pelajaran yang batasnya jelas, misalnya mata pelajaran Matematika, Biologi, dan sebagainya. Dalam memilih teknik tes mana yang akan digunakan Pendidik juga harus mempertimbangkan ciri indikator, contoh, apabila tuntutan indikator melakukan sesuatu, maka teknik penilaiannya adalah tes unjuk kerja (*performance*), sedang bila tuntutan indikator berkaitan dengan pemahaman konsep, maka teknik penilaiannya adalah tes tertulis. Tingkat berpikir yang digunakan dalam mengerjakan tes harus mencakup mulai yang rendah sampai yang tinggi, dengan proporsi yang sebanding sesuai dengan jenjang pendidikan.
- (3) **Menetapkan panjang Tes:** langkah menetapkan panjang tes, meliputi berapa waktu yang tersedia untuk melakukan tes, hal ini terkait erat dengan penetapan

jumlah item-item tes yang akan dikembangkan. Apabila oleh pendidik ada materi yang dinilai lebih penting dan mempunyai tingkat kesulitan yang lebih tinggi, guru bisa memberikan pembobotan yang berbeda dari setiap soal yang disusun. Ada tiga hal yang harus dipertimbangkan dalam menentukan jumlah soal, yaitu bobot masing-masing bagian yang telah ditentukan dalam kisi-kisi, keandalan yang diinginkan, dan waktu yang tersedia.

b) Menulis Butir Pertanyaan

Setelah selesai mencermati dan menjabarkan setiap indikator menjadi diskriptor-diskriptor, dan telah ditetapkan ukurannya, maka pendidik mulai dapat mengembangkan atau menulis butir pertanyaan sesuai dengan kisi-kisi yang telah ditetapkan. Ada 3 kegiatan pokok dalam menulis butir soal yaitu:

- (1) **Menulis draft soal:** Menulis soal bagi Anda pasti sudah menjadi pekerjaan rutin sebelum ulangan, tetapi seharusnya Anda perlu mencermatinya karena langkah ini juga memerlukan kecermatan dalam memilih kalimat-kalimat yang mudah dimengerti dan tidak menimbulkan interpretasi ganda. Ada dua hal yang perlu mendapat perhatian dalam penulisan butir pertanyaan yaitu format pertanyaan dan alternatif jawaban. Dalam hal ini perlu diperhatikan beberapa hal yaitu, (1) apakah pertanyaan mudah dimengerti? (2) apakah sudah sesuai dengan indikator (3) apakah tata letak keseluruhan baik? (4) apakah perlu pembobotan (5) apakah kunci jawaban sudah benar?
- (2) **Memantapkan Validitas Isi (*Content Validity*):** *Content validity* atau validitas isi pada dasarnya merupakan koefisien yang menunjukkan kesesuaian antara draft tes yang telah disusun dengan isi dari konsep dan kisi-kisi yang telah disusun, apakah semua materi telah terjabar dalam item, dan apakah soal yang disusun telah pula sesuai ranah atau kawasan yang akan diukur. Langkah ini dapat dilakukan dengan beberapa cara misalnya diskusi dengan sesama pendidik ataupun dengan cara mencermati kembali substansi dari konsep yang akan diukur.
- (3) **Melakukan Uji Coba (*try out*):** Mungkin Anda mengira bahwa *try out* hanya digunakan untuk tes standard dan tidak perlu dilakukan untuk tes buatan guru. Anggapan itu kurang benar karena uji coba tetap diperlukan dalam penyusunan tes buatan guru, *try out* tidak harus dilakukan secara formal dan dalam skala besar, yang perlu Anda perhatikan adalah bahwa *try out* dapat dilakukan untuk berbagai kepentingan diantaranya adalah untuk; (1) analisis item, (2) bagaimana

rencana pelaksanaan, (3) memperkirakan penggunaan waktu pengerjaan, (4) kejelasan format tes, (5) kejelasan petunjuk pengisian, dan (6) pemahaman bahasa yang digunakan dsbnya.

- (4) **Revisi soal:** Hasil dari uji coba kemudian dilakukan analisis untuk mencari tingkat kesulitan soal dan penggunaan bahasa yang kurang komunikatif, untuk kemudian dilakukan revisi sesuai dengan kebutuhan. Misalnya revisi dilakukan untuk; (1) eliminasi butir-butir yang jelek, (2) menambah butir-butir baru, (3) memperjelas petunjuk, dan (4) memodifikasi format dan urutan, dsbnya.

c) Melakukan pengukuran dengan tes

Sekarang tes sebagai instrumen sudah selesai disusun! Apa saja yang harus Anda lakukan selanjutnya? Ada beberapa langkah yang harus diperhatikan pada saat menyelenggarakan tes untuk siswa yaitu:

- (1) **Menjaga obyektivitas pelaksanaan tes:** Kegiatan pengukuran yang berupa penyelenggaraan tes juga sudah menjadi kegiatan Anda sehari-hari, meskipun demikian pendidik tetap harus menjaga obyektifitas, baik dalam pengawasan, menjaga kerahasiaan soal, dan berbagai kode etik penyelenggaraan tes yang lain. Setelah ujian dilaksanakan maka langkah berikut adalah koreksi, dan interpretasi dari hasil ujian tersebut, untuk kemudian berdasar data hasil analisis tersebut akan diambil keputusan dalam berbagai kepentingan.
- (2) **Memberikan skor pada hasil tes:** Yaitu memeriksa hasil jawaban dari para siswa, untuk memberikan skor/angka sebagai penghargaan terhadap setiap poin soal yang dapat dikerjakan, hasilnya berupa angka yang disebut skor mentah, angka yang menunjukkan berapa soal yang bisa dijawab benar oleh siswa. Penentuan jumlah soal yang bisa dijawab benar ini tidak menjadi masalah untuk tes obyektif. Namun untuk bentuk soal tes uraian masalah ini akan menjadi persoalan, karena setiap siswa akan mengemukakan argumentasi yang berbeda-beda untuk menjawab soal dan permasalahan tes. Sehingga dalam melakukan langkah ini harus pula dijaga obyektivitas dengan selalu menggunakan kunci jawaban dan indikator keberhasilan.

(3) Melakukan Analisis Hasil Tes

Setelah semua pekerjaan siswa dikoreksi langkah berikutnya adalah melakukan analisis terhadap skor hasil tes. Materi tentang ini akan secara khusus dibahas pada **UNIT 6**.

2. Mengembangkan Tes Sebagai Instrumen Asesmen di SD

Anda telah memahami langkah-langkah pokok yang seharusnya dilakukan dalam pelaksanaan tes. Dengan tetap mengacu pada langkah-langkah pokok tersebut, berikut ini akan dikemukakan langkah-langkah detail yang diharapkan dapat menuntun Anda mengembangkan tes sebagai instrumen asesmen di kelas.

a. Menjabarkan Kompetensi Dasar ke dalam Indikator Pencapaian Hasil Belajar

Kegiatan ini, dalam langkah kegiatan umum masuk dalam langkah “menentukan cakupan materi yang akan diukur”. Indikator merupakan ukuran, karakteristik, ciri-ciri, pembuatan atau proses yang berkontribusi/menunjukkan ketercapaian suatu kompetensi dasar. Indikator dirumuskan dengan menggunakan kata kerja operasional yang dapat diukur, seperti: menyebutkan, memberikan contoh, mengidentifikasi, menghitung, membedakan, menyimpulkan, mempraktekkan, mendemonstrasikan.

Sesuai dengan Kurikulum Tingkat Satuan Pendidikan (KTSP), maka indikator pencapaian hasil belajar dikembangkan oleh pendidik dengan memperhatikan perkembangan dan kemampuan setiap peserta didik, keluasan dan kedalaman kompetensi dasar, dan daya dukung sekolah, misalnya kemampuan guru dan sarana atau perasarana penunjang. Setiap kompetensi dasar dapat dikembangkan menjadi beberapa indikator pencapaian hasil belajar. Indikator-indikator pencapaian hasil belajar dari setiap kompetensi dasar merupakan acuan yang digunakan untuk menyusun butir tes.

Contoh 1

Mata pelajaran : Pendidikan Jasmani, Olah Raga dan Kesehatan

Kelas/Semester : IV/1

Standar Kompetensi	Kompetensi Dasar	Indikator*
Mempraktikkan gerak dasar ke dalam permainan sederhana dan olahraga serta nilai-	Mempraktikkan gerak dasar dalam permainan bola kecil sederhana dengan peraturan yang	✓ Melakukan berbagai teknik dasar permainan kasti.
		✓ Menerapkan kerjasama team dalam permainan kasti.

Standar Kompetensi	Kompetensi Dasar	Indikator*
nilai yang terkandung didalamnya	dimodifikasi, serta nilai kerjasama tim, sportivitas, dan kejujuran**)	✓ Menyebutkan manfaat permainan kasti terhadap kesehatan tubuh.

Contoh 2

Mata pelajaran : IPS

Kelas/Semester : I/1

Standar Kompetensi	Kompetensi Dasar	Indikator*
Memahami identitas diri dan keluarga, serta sikap saling menghormati dalam kemajemukan keluarga.	1.1.Mengidentifikasi identitas diri, keluarga, dan kerabat. 1.2.Menceriterakan pengalaman diri.	✓ Siswa dapat menyebutkan identitas diri secara lisan di depan teman-temannya. ✓ Siswa dapat menceritakan pengalamannya dalam bentuk karangan sederhana

Contoh 3

Mata pelajaran : Bahasa Indonesia

Kelas / Semester : III/2

Standar Kompetensi	Kompetensi Dasar	Indikator*
Menulis Mengungkapkan pikiran, perasaan dan informasi dalam karangan sederhana dan puisi.	Menulis puisi berdasarkan gambar dengan pilihan kata yang menarik	✓ Siswa dapat menyebutkan ciri-ciri kalimat dalam puisi. ✓ Siswa dapat menulis puisi dengan benar .

Indikator*: dikembangkan oleh guru sekolah sesuai dengan kondisi daerah dan sekolah masing-masing. Satu KD dapat dikembangkan menjadi satu atau lebih indikator .

Sumber: Pedoman Penilaian SD (Depdiknas, 2006).

d. Menetapkan Jenis Tes dan Penulisan Butir Soal

Setelah Anda menjabarkan standar kompetensi, kompetensi dasar, dan indikator keberhasilannya, maka Anda mulai dapat menetapkan indikator yang menunjukkan tingkat pencapaian kompetensi tersebut sebaiknya dapat diukur dengan menggunakan alat ukur apa, bila ditetapkan tes, akan pula dapat ditetapkan jenis tes yang mana. Di samping itu pemilihan bentuk tes akan dapat dilakukan dengan tepat bila didasarkan pada tujuan tes, cakupan materi tes, karakteristik mata pelajaran yang diukur pencapaiannya, jumlah peserta tes, termasuk waktu yang tersedia untuk memeriksa lembar jawaban tes. Dalam menyusun instrumen penilaian tertulis perlu

dipertimbangkan; (1) **materi**, misalnya kesesuaian soal dengan kompetensi dasar dan indikator pencapaian pada kurikulum tingkat satuan pendidikan, (2) **konstruksi**, misalnya rumusan soal atau pertanyaan harus jelas dan tegas, (3) **bahasa**, misalnya rumusan soal tidak menggunakan kata/kalimat yang menimbulkan penafsiran ganda, dan (4) **kaidah penulisan**, harus berpedoman pada kaidah penulisan soal yang baku dari berbagai bentuk soal penilaian. Rancangan penilaian ini diinformasikan kepada siswa pada awal pertemuan (awal semester). Dengan demikian sistem penilaian yang dilakukan guru semakin sempurna atau semakin memenuhi prinsip-prinsip penilaian.

Dalam pembelajaran di SD Anda dapat melihat beberapa contoh di bawah ini yang secara jelas memberikan paparan tentang keterkaitan hubungan antara pemetaan Standar Kompetensi, Kompetensi Dasar, Indikator dan Teknik Penilaian yang bersumber dari Model Penilaian Kelas SD (Depdiknas 2006).

Contoh 1

Mata Pelajaran: Matematika
Kelas/Semester: II/1

No.	Standar Kompetensi	Kompetensi Dasar	Indikator	Aspek	Tehnik penilaian
1	Menggunakan pengukuran waktu, panjang, dan berat dalam pemecahan masalah.	Menggunakan alat ukur tidak baku dan baku (cm, m) yang sering digunakan	<ul style="list-style-type: none"> • Siswa menyebutkan macam-macam alat ukur panjang tidak baku dalam kehidupan sehari-hari (jengkal, depa, langkah kaki, dll). • Siswa dapat menggunakan alat ukur tidak baku (jengkal, depa, pecak (panjang telapak kaki, langkah kaki, dll) • Siswa menyebutkan alat ukur baku (cm, m) yang biasa digunakan dalam kehidupan sehari-hari. • Siswa dapat menggunakan alat ukur baku . • Siswa dapat menarik kesimpulan bahwa pengukuran dengan alat ukur tidak baku hasilnya berbeda. 	Geometri dan pengukuran	Penilaian Kinerja Test tertulis Contoh alat penilaian terlampir

Untuk standar kompetensi, kompetensi dasar dan indikator materi di atas dan dengan didasarkan kemampuan penalaran anak kelas 2 SD maka ditetapkan tes pilihan ganda dan isian berikut ini.

Bentuk Pilihan Ganda

Berilah tanda silang pada huruf di depan jawaban yang paling tepat!

Skor: Setiap jawaban benar diberi nilai 1 .

1. Yang termasuk alat ukur tidak baku yaitu
 - a. meter
 - b. centimeter
 - c. jengkal
2. Yang termasuk alat ukur baku ialah
 - a. cm
 - b. depa
 - c. langkah kaki

Bentuk Isian

Isilah titik-titik di bawah ini dengan jawaban yang singkat dan tepat !

Skor: Setiap jawaban benar diberi nilai 2.

1. Satuan panjang Centimeter dan Meter adalah contoh alat ukur
2. Satuan panjang langkah kaki, depa dan jengkal termasuk alat ukur
3. Karena menggunakan alat ukur tidak baku, maka hasil pengukurannya

Pemberian Skor:

$$\text{Nilai} = \frac{\text{Banyak jawaban benar}}{\text{Banyak soal}} \times 100$$

Contoh 2

Mata Pelajaran : Pendidikan Kewarganegaraan
Kelas/Semester : IV / 2

No.	Standar Kompetensi	Kompetensi Dasar	Indikator	Aspek	Tehnik penilaian
1	Menunjukkan sikap terhadap globalisasi di lingkungannya.	Memberi contoh sederhana pengaruh globalisasi di lingkungannya.	<ul style="list-style-type: none"> • Siswa dapat menjelaskan pengertian globalisasi • Siswa dapat memberikan salah satu contoh pengaruh positif globalisasi bidang komunikasi • Siswa dapat 	Globalisasi	<ul style="list-style-type: none"> ✓ Test tertulis ✓ Pengamatan sikap Contoh alat penilaian terlampir

No.	Standar Kompetensi	Kompetensi Dasar	Indikator	Aspek	Tehnik penilaian
			memberikan salah satu contoh pengaruh negatif globalisasi bidang kebudayaan		

Untuk Standar kompetensi , kompetensi dasar dan indikator materi diatas dan dengan didasarkan kemampuan penalaran anak kelas 4 SD maka ditetapkan tes isian dan jawaban uraian sebagai berikut:

Bentuk tes Isian

Isilah titik–titik pada soal di bawah ini dengan jawaban singkat dan tepat!

1. Pengaruh positif globalisasi di bidang komunikasi di lingkungan masyarakat misalnya
2. Kecenderungan masyarakat menyukai jenis musik jaz termasuk pengaruh negatif globalisasi bidang
3. Terhadap kebudayaan asing yang masuk ke Indonesia sikap kita seharusnya

Bentuk Soal tes Uraian

Jawablah pertanyaan di bawah ini!

1. Jelaskan yang dimaksud dengan istilah globalisasi!
2. Berikan tanggapan dan alasan terhadap pernyataan di bawah ini?
 - a. Dengan globalisasi kita semakin mudah menikmati siaran televisi luar negeri.
 - b. Karena pengaruh globalisasi masyarakat cenderung bersikap konsumtif.

Pemberian Skor:

$$\text{Nilai} = \frac{\text{Banyak jawaban benar}}{\text{Banyak soal}} \times 100$$

Contoh 3

Mata Pelajaran : Ilmu Pengetahuan Alam
Kelas/Semester : IV / 1

No.	Standar Kompetensi	Kompetensi Dasar	Indikator	Aspek	Tehnik penilaian
1	Memahami daur hidup beragam jenis makhluk hidup	Mendeskripsikan daur hidup beberapa hewan di lingkungan sekitar, misalnya kecoa, nyamuk, kupu-kupu, kucing.	<ul style="list-style-type: none">• Mendeskripsikan urutan daur hidup hewan, misalnya kupu-kupu, nyamuk dan kecoa secara sederhana.• Menyimpulkan berdasarkan pengamatan bahwa tidak semua hewan berubah bentuk dengan cara yang sama.• Menyimpulkan bahwa berubahnya bentuk pada hewan menunjukkan adanya pertumbuhan.• Menyimpulkan hasil pengamatan daur hidup hewan yang dipeliharanya*)	Mahluk hidup dan proses kehidupan	<ul style="list-style-type: none">• Jenis: ulangan• Bentuk: tes tertulis, penugasan

Depdiknas 2006

Latihan

1. Buatlah contoh tes tertulis tentang penguasaan konsep dan tes unjuk kerja dari Standar kompetensi, kompetensi dasar dan indikator materi di atas dan dengan didasarkan kemampuan penalaran anak kelas 4 SD!
2. Setelah soal tersusun diskusikan dengan teman Anda, sebagai bentuk pemantapan internal validity dan uji-coba terbatas!

4. Mengembangkan tes pada Kawasan Kognitif, Afektif, dan Psikomotor

Mungkin masih ada anggapan termasuk mungkin juga anggapan Anda bahwa tes tertulis khususnya dalam bentuk tes obyektif hanya cocok untuk mengukur pencapaian hasil belajar pada kawasan kognitif saja. Anggapan itu tidak bisa dibenarkan karena dengan pemahaman yang tinggi terhadap cakupan materi maupun teknik evaluasi, pendidik akan dapat mengembangkan tes tertulis yang dapat meliputi dua kawasan yang lain yaitu afektif maupun psikomotor. Marilah bersama-sama kita cermati penjelasan berikut.

a. Mengembangkan Tes pada Domain Kognitif

Pada dasarnya akan sangat mudah mengembangkan tes untuk mengukur indikator pencapaian hasil belajar pencapaian kawasan (domain) kognitif, hampir semua jenis tes dengan berbagai bentuk soal dapat digunakan untuk mengukur kawasan ini seperti misalnya:

1) Tes Lisan

Pertanyaan secara lisan masih sering digunakan untuk mengukur daya serap peserta didik pada kawasan kognitif. Yang perlu Anda ingat tes lisan harus disampaikan dengan jelas, dan semua peserta didik harus diberi kesempatan yang sama. Beberapa prinsip yang harus dipedomani adalah memberi waktu untuk berpikir, baru menunjuk peserta untuk menjawab pertanyaan. Tingkat berpikir untuk pertanyaan lisan di kelas cenderung rendah, seperti pengetahuan dan pemahaman. Jawaban salah satu siswa harus dikembalikan ke forum kelas untuk ditanggapi siswa yang lain.

2) Tes Pilihan Ganda

Ketika Anda mengembangkan tes pilihan ganda hendaknya memperhatikan sepuluh pedoman penulisannya yaitu: (1) soal harus jelas, (2) isi pilihan jawaban homogen dalam arti isi, (3) panjang kalimat pilihan jawaban relatif sama, (4) tidak ada petunjuk jawaban benar, (4) hindari menggunakan pilihan jawaban “semua benar “ atau “semua salah”, (6) pilihan jawaban angka diurutkan, (7) pilihan jawaban logis dan tidak menggunakan negatif ganda, (8) kalimat yang digunakan sesuai dengan tingkat perkembangan peserta tes, (9) menggunakan bahasa Indonesia yang baik dan baku, dan (10) letak pilihan jawaban benar ditentukan secara acak.

3) Bentuk Tes uraian Obyektif

Bentuk ini tepat digunakan untuk bidang Matematika dan IPA, karena kunci jawabannya hanya satu. Pengerjaan soal ini melalui suatu prosedur atau langkah-langkah tertentu. Setiap langkah ada skornya. Objektif disini dalam arti apabila diperiksa oleh beberapa guru dalam bidang studi tersebut hasil penskorannya akan sama. Pertanyaan pada bentuk soal ini di antaranya adalah: hitunglah, tafsirkan, buat kesimpulan dsbnya.

4) Bentuk Tes Uraian

Tes ini menuntut siswa menyampaikan, memilih, menyusun, dan memadukan gagasan dan ide-idenya dengan menggunakan kata-katanya sendiri. Keunggulan bentuk tes ini dapat mengukur tingkat berpikir dari yang rendah sampai yang tinggi, yaitu mulai dari hapalan sampai dengan evaluasi. Kelemahan bentuk tes ini adalah: (1) penskoran sering dipengaruhi oleh subjektivitas penilai, (2) memerlukan waktu yang lama untuk melakukan koreksi, (3) cakupan materi yang diujikan sangat terbatas, (4) dan adanya efek *bluffing*. Untuk menghindari kelemahan tersebut cara yang ditempuh adalah: (a) jawaban tidak panjang, sehingga bisa mencakup materi yang banyak, (b) tidak melihat nama peserta ujian, (c) memeriksa tiap butir secara keseluruhan, dan (d) menyiapkan pedoman penskoran.

5) Bentuk Tes jawaban Singkat

Tes ini mengharuskan siswa menuliskan jawaban singkatnya sesuai dengan petunjuk. Ada tiga jenis soal bentuk ini, yaitu: jenis pertanyaan, jenis melengkapi atau isian, dan jenis identifikasi atau asosiasi. Ketika Anda menyusun tes bentuk ini perhatikan keharusannya yaitu; (1) soal mengacu pada indikator, (2) rumusan kalimat soal harus komunikatif, dan (3) tidak menimbulkan interpretasi ganda.

6) Bentuk Tes Menjodohkan

Pengerjaan tes ini dilakukan dengan menjodohkan atau memasangkan suatu premis dengan daftar kemungkinan jawaban, dan suatu petunjuk untuk menjodohkan masing-masing premis itu dengan satu kemungkinan jawaban. Bila Anda menuliskan soal bentuk ini perhatikan bahwa: (1) soal harus sesuai dengan indikator, (2) jumlah alternatif jawaban lebih banyak dari pada premis, (3) alternatif jawaban berhubungan secara logis dengan premisnya, (4) rumusan kalimat soal harus komunikatif, dan (5) butir soal menggunakan Bahasa Indonesia yang baik dan benar.

7) Bentuk Tes Unjuk Kerja (*Performance*)

Tes bentuk ini sering pula diklasifikasikan dalam bentuk penilaian autentik atau penilaian alternatif yang bertujuan untuk mengetahui tingkat kemampuan peserta didik dalam menyelesaikan masalah-masalah di kehidupan nyata. Yang dapat Anda cermati uraiannya pada Unit 5.

b. Mengembangkan Tes pada Domain Afektif

Anda dapat mengembangkan tes pada domain afektif ini, untuk beberapa fokus sikap diantaranya adalah:

1) Sikap terhadap mata pelajaran

Tes sikap terhadap mata pelajaran dapat diberikan pada awal atau akhir program agar siswa memiliki sikap yang lebih baik pada suatu mata pelajaran. Perlu dilakukan tindakan bila sebagian besar siswa bersikap negatif pada mata pelajaran tertentu

2) Sikap positif terhadap belajar

Siswa diharapkan memiliki sikap yang baik terhadap belajar. Siswa yang memiliki sikap positif terhadap belajar cenderung menjadi pembelajar pada masa depan.

3) Sikap terhadap diri sendiri

Meskipun harga diri siswa dipengaruhi oleh keluarga dan kejadian di luar sekolah, hal-hal yang terjadi di kelas diharapkan dapat meningkatkan harga diri siswa.

4) Sikap positif terhadap perbedaan

Siswa perlu mengembangkan sikap yang lebih toleran dan menerima perbedaan seperti etnik, jender, kebangsaan dan keagamaan.

5) Sikap terhadap permasalahan faktual yang ada di sekitarnya

Penilaian afektif juga dapat melihat fokus nilai semacam kejujuran, integritas, keadilan, dan nilai kebebasan. Fokus penilaian afektif dapat dikenakan terhadap permasalahan-permasalahan aktual di sekitar siswa.

Pertanyaan yang berikutnya muncul adalah “**Bagaimanakah tes pada domain afektif dilaksanakan?**” Penilaian sikap dapat dilakukan dengan beberapa cara atau teknik antara lain: observasi perilaku, pertanyaan langsung, dan laporan pribadi. Hasil observasi perilaku dapat dijadikan sebagai umpan balik dalam pembinaan. Perilaku adalah kecenderungan seseorang dalam sesuatu hal.

Pada tes ini biasanya digunakan dengan memanfaatkan skala likert. Langkah-langkah dalam menyusun skala likert antara lain adalah: (1) Memilih variabel afektif yang akan diukur; (2) Membuat beberapa pernyataan tentang variabel afektif yang

dimaksudkan; (3) Mengklasifikasikan pernyataan positif atau negatif; (4) Menentukan jumlah gradual dan frase atau angka yang dapat menjadi alternatif pilihan; (5) Menyusun pernyataan dan pilihan jawaban menjadi sebuah alat penilaian; (6) Melakukan ujicoba; (7) Membuang butir-butir pernyataan yang kurang baik; dan (8) Melaksanakan penilaian.

Di bawah ini adalah satu contoh tes afektif yang mengases sikap siswa terhadap pelajaran sains.

Sikap terhadap Pelajaran Sains

Petunjuk:

1. Pengisian skala ini tidak ada hubungannya dengan prestasi belajar, Anda tidak perlu mencantumkan nama dan nomor absen!
2. Pilihlah dengan melingkari jawaban yang paling sesuai dengan pernyataan di bawah ini!

No.	Pernyataan	Sangat Setuju	Setuju	Tidak Setuju	Sangat Tidak Setuju
1	Saya mempersiapkan diri untuk menerima pelajaran sains di kelas	SS	S	TS	STS
2	Saya berperan aktif dalam pembelajaran sains	SS	S	TS	STS
3	Saya suka melakukan percobaan sains	SS	S	TS	STS
4	Saya tertarik artikel yang berhubungan dengan sains	SS	S	TS	STS
5	Saya memperkaya keterangan guru sains dengan membaca buku-buku penunjang	SS	S	TS	STS
6	Saya mengulang pelajaran sains di rumah	SS	S	TS	STS
7	SS	S	TS	STS

Depdiknas 2006

Latihan

Cobalah menyusun tes domain afektif yang berupa sikap siswa pada peristiwa factual yang ada di sekitar lingkungan sekolah!

c. Mengembangkan Tes pada Domain Psikomotor

Pada umumnya pelajaran yang termasuk kelompok psikomotor adalah mata pelajaran yang indikator keberhasilan yang lebih beorientasi pada gerakan dan menekankan pada reaksi-reaksi fisik atau keterampilan tangan. Hasil belajar psikomotor dapat diklasifikasikan menjadi tiga, yaitu: (1) *specific responding*, siswa baru mampu merespons hal-hal yang sifatnya fisik, yang dapat didengar, dilihat, atau

diraba, misalnya memegang raket, memegang bed untuk tenis meja dsb. dan (2) *motor chaining*, siswa sudah mampu menggabungkan lebih dari dua keterampilan dasar menjadi satu keterampilan gabungan, misal memukul bola, menggerakkan, menggunakan jangka sorong. Pada tingkat *rule using* siswa sudah dapat menggunakan hukum-hukum dan atau pengalaman-pengalaman untuk melakukan keterampilan yang kompleks, misal bagaimana memukul bola yang tepat agar dengan tenaga yang sama namun hasilnya lebih keras. Gagne (1977) berpendapat bahwa ada 2 kondisi yang dapat mengoptimalkan hasil belajar keterampilan yaitu kondisi internal dan eksternal. Untuk kondisi internal dapat dilakukan dengan cara, yakni (a) mengingatkan kembali sub-sub keterampilan yang sudah dipelajari dan (b) mengingatkan prosedur-prosedur atau langkah-langkah gerakan yang telah dikuasainya. Untuk kondisi eksternal dapat dilakukan dengan: (a) instruksi verbal, (b) gambar, (c) demonstrasi, (d) praktik, dan (e) umpan balik.

Soal untuk ranah psikomotor juga harus mengacu pada standar kompetensi yang sudah dijabarkan menjadi kompetensi dasar. Setiap butir standar kompetensi dijabarkan menjadi 3 sampai dengan 6 butir kompetensi dasar. Selanjutnya setiap butir kompetensi dasar dapat dijabarkan menjadi 3 sampai dengan 6 indikator dan setiap indikator harus dapat dibuat lebih dari satu butir soal. Namun, ada kalanya satu butir soal ranah psikomotor terdiri dari beberapa indikator. Instrumen psikomotor ini terdiri dari dua macam, yaitu (1) soal dan (2) lembar yang digunakan untuk mengamati dan menilai jawaban siswa terhadap soal tersebut.

1) Menyusun Soal

Menyusun soal dapat diawali dengan mencermati kisi-kisi instrumen psikomotor yang telah dibuat. Soal harus dijabarkan dari indikator dengan memperhatikan materi pokok dan pengalaman belajar. Namun adakalanya soal ranah psikomotor untuk ujian blok yang biasanya sudah mencapai tingkat psikomotor manipulasi, mencakup beberapa indikator.

2) Menyusun Lembar Observasi dan Lembar Penilaian

Lembar observasi dan lembar penilaian harus mengacu pada soal. Soal atau lembar tugas atau perintah kerja inilah yang selanjutnya dijabarkan menjadi aspek-aspek keterampilan. Lembar observasi pada tes unjuk kerja dapat Anda cermati juga pada UNIT 5.

Teknik asesmen, pendekatan, dan metode pembelajaran serta hasil belajar pada semua ranah merupakan hal yang tak terpisahkan satu dengan yang lain karena semua di desain untuk mencapai kompetensi yang dipersyaratkan. Pertanyaan yang

kemudian muncul adalah: Sejauhmana pola pembelajaran mampu mengembangkan kemampuan kognitif, afektif, dan psikomotor siswa. Pedoman Penilaian Depdiknas (2006) memvisualkan gambaran tersebut dalam Tabel berikut untuk mempermudah Anda mencermati keterkaitan ranah Kognitif, Afektif, dan Psikomotor dalam Penilaian.

Tabel: Keterkaitan Ranah Kognitif, Afektif, dan Psikomotor dalam Penilaian

No	TINGKATAN DOMAIN	Pola mengajar		Cara penilaian						
		Tradisional	Belajar aktif	Tulis Objektif	Tulis Subjektif	Lisan	Unjuk kerja	Produk	Portofolio	Tingkah laku
KOGNITIF										
VI.	Evaluasi	-	v	-	v	-	v	-	v	v
V.	Sintesis	-	v	-	v	-	v	-	v	v
IV.	Analisis	-	v	-	v	-	v	-	v	v
III.	Aplikasi	-	v	-	v	v	v	v	v	v
II.	Pemahaman	v	v	v	v	v	v	v	v	v
I.	Pengetahuan	v	v	v	v	v	v	v	v	v
AFEKTIF										
V.	Karakterisasi	-	v	-	-	-	-	-	v	-
IV.	Organisasi	-	v	-	-	-	v	-	v	-
III.	Acuan nilai	-	v	-	-	-	v	v	v	v
II.	Responsi	v	v	-	-	-	v	v	v	v
I.	Penerimaan	v	v	-	-	-	v	v	v	v
PSIKOMOTOR										
VI.	Gerakan indah dan kreatif	-	v	-	-	-	v	v	-	-
V.	Gerakan terampil	-	v	-	-	-	v	v	-	-
IV.	Gerakan kemampuan fisik	-	v	-	-	-	v	v	-	-
III.	Gerakan persepsi	-	v	-	-	-	v	v	v	-
II.	Gerakan dasar	v	v	-	-	-	v	v	v	-
I.	Gerakan refleks	v	v	-	-	-	v	v	v	v
Jumlah		6	17	2	8	3	16	12	14	10
Persentase		35%	100%	12%	47%	18%	94%	71%	82%	59%

(Depdiknas 2006)

Rangkuman

Mengembangkan Tes sebagai instrumen asesmen proses dan hasil belajar adalah menyusun alat ukur suatu gejala yang bersifat abstrak yaitu pemahaman dan penguasaan anak terhadap materi yang berupa seperangkat kompetensi dipersyaratkan. Untuk dapat mengembangkan tes yang baik perlu diperhatikan langkah pokok mengembangkan Tes yang meliputi:

1. Perencanaan Tes

- a. Menentukan cakupan materi yang akan diukur
- b. Memilih bentuk tes
- c. Menetapkan panjang tes

2. Menulis Butir Pertanyaan

- a. Menulis draft soal
- b. Memantapkan validitas isi (Content Validity)
- c. Melakukan uji-coba (try out)
- d. Revisi soal

3. Melakukan pengukuran dengan tes

- a. Menjaga obyektivitas pelaksanaan
- b. Memberikan skor pada hasil tes
- c. Melakukan analisis hasil tes

Tes Formatif 2

Di bawah ini dicantumkan tes formatif yang bertujuan untuk mengukur pemahaman Anda mengenai uraian, contoh, dan rangkuman yang tercantum dalam subunit 2. Jawablah pertanyaan berikut sesuai dengan permintaan!

1. Jelaskan langkah-langkah pokok penyusunan tes!
2. Buatlah matrik keterkaitan yang menggambarkan tata hubungan antara standar kompetensi, kompetensi dasar, indikator dan penetapan jenis tes serta contoh soal yang dikembangkan!
3. Jelaskan dengan contoh bagaimana sebaiknya try out dilakukan untuk tes buatan guru!
4. Bagaimanakah langkah-langkah mengembangkan tes untuk mengukur domain afektif!
5. Bagaimanakah langkah-langkah mengembangkan tes untuk mengukur domain psikomotor!

Umpan Balik dan Tindak Lanjut

Jawablah pertanyaan tes formatif di atas, setelah selesai baru cocokkan dengan kunci jawabannya. Diskusikan dengan teman bila jawaban belum sesuai atau Anda belum merasa masih ada hal-hal yang meragukan. Hal ini sangat diperlukan karena kesepahaman tentang pengertian ini akan mendasari dan mempengaruhi langkah dan kegiatan dalam menyelesaikan mata kuliah ini.

Subunit 3

Kriteria Tes Yang Baik

Pengantar

Dalam proses pembelajaran, tes merupakan alat ukur dalam proses asesmen maupun evaluasi yang memiliki peranan sangat penting untuk mengetahui keberhasilan proses belajar-mengajar di sekolah. Dalam hal ini, tes memiliki fungsi ganda, yaitu mengukur tingkat pencapaian siswa pada kompetensi yang dipersyaratkan, yang terjabar dalam indikator pencapaian, dan mengukur keberhasilan program pengajaran sekaligus *kualitas* pendidik dalam mengelola proses pembelajaran. Untuk bisa memberikan data yang akurat, sesuai dengan fungsinya maka ada beberapa persyaratan yang harus dipenuhi, untuk dapat dikatakan sebagai tes yang baik. Secara umum tes yang baik memiliki syarat-syarat antara lain (1) hanya mengukur satu aspek saja. Tes yang baik memiliki sebuah aspek saja yang akan di ukur, jadi tes matematika misalnya hanya menguji kemampuan matematika seseorang, (2) handal dalam pengukuran; kehandalan ini meliputi ketepatan hasil pengukuran dan keajegan hasil pengukuran.

Dengan memahami betapa pentingnya tes dalam kegiatan asesmen pembelajaran di sekolah, pastilah Anda sebagai pendidik menjadi penasaran untuk lebih dalam mempelajari apa sebenarnya persyaratan atau kriteria dari tes yang baik. Rasa penasaran Anda akan terlunasi dengan mencermati uraian pada subunit ini. Pemahaman itu akan semakin sempurna bila Anda mengerjakan semua soal latihan dan tes formatif yang ada di akhir subunit ini. Anda juga disarankan untuk membaca referensi lain serta mendiskusikannya dengan teman-teman Anda .

1. Kriteria Tes yang Baik

Tes atau soal ujian merupakan alat ukur yang memiliki fungsi ganda yaitu untuk mengukur efektivitas belajar dan mengukur efektivitas guru dalam mengajar. Untuk dapat menjadi alat ukur yang baik dan dapat memberikan informasi yang akurat maka setiap soal sebagai bagian dari konstruksi tes harus dijaga kualitasnya. Ada beberapa kriteria yang dapat dipakai untuk menyusun butir-butir tes yang berkualitas yaitu:

a. **Valid**

Soal dikatakan valid bila dapat mengukur apa yang seharusnya diukur, validitas soal

dapat dilihat dari kesesuaian soal dengan tujuan instruksional khusus dan tujuan pengukuran yang telah ditetapkan. Validitas dapat pula dilihat dari kemampuannya memprediksi prestasi di masa yang akan datang,

b. **Relevan**

Tes yang relevan mengandung soal-soal yang dapat mengukur kemampuan belajar sesuai dengan tingkat kemampuan yang ditetapkan dalam indikator pencapaian hasil belajar (Ranah kognitif, afektif dan psikomotor). Bila kompetensi dasar dan indikator bertujuan mengungkap ranah afektif, pertanyaan soal harus pula mengarah ke sikap dan seterusnya.

c. **Spesifik**

Soal harus direncanakan sedemikian rupa agar jawabannya pasti dan tidak menimbulkan ambivalensi atau spakulasi dalam memberikan jawaban. Kesulitan soal tidak saja kesulitan materi juga bisa ditambah kesulitan dalam memahami soal bila soal tidak disusun secara spesifik.

d. **Representatif**

Soal tes sebaiknya dikembangkan dari satuan materi yang jelas cakupannya, dan bersifat komprehensif dalam pengertian materi tes harus mencakup seluruh materi pengajaran, untuk itu seluruh pokok bahasan (sub pokok bahasan) idealnya harus terwakili dalam soal tes. Syarat ini akan dapat mengurangi error terhadap hasil pengukuran.

e. **Seimbang**

Dalam proses pengajaran dosen akan tahu persis, bahwa setiap pokok bahasan memiliki tingkat kesulitan yang berbeda, soal tes dikatakan seimbang bila pokok bahasan yang terpenting mendapat porsi terbanyak dalam soal. Kalau dalam keadaan terpaksa hal tersebut tidak dapat dilakukan maka keseimbangan dapat dicapai dengan memberikan bobot yang berbeda pada pokok bahasan yang memiliki tingkat kesulitan yang berbeda.

f. Sensitif

Syarat ini berkait erat dengan taraf kesukaran soal, butir tes yang baik harus memiliki sensitivitas untuk membedakan siswa yang benar-benar menguasai materi dengan yang tidak, hal ini tidak akan tercapai bila soal terlalu sulit sehingga semua siswa tidak dapat mengerjakan, atau soal yang terlalu gampang sehingga semua siswa dapat mengerjakan dengan benar.

g. Fair

Tes hasil ujian hendaklah bersifat terbuka dalam pengertian tidak mengandung jebakan, jelas cakupan materinya, kejelasan norma yang dipakai serta kriteria keberhasilannya. Dalam pelaksanaannya obyektif, tidak merugikan kelompok tertentu.

h. Praktis

Dalam pengertian bahwa tes tidak sulit untuk dilaksanakan dilihat dari segi pembiayaan maupun pelaksanaannya. Tes yang baik harus efisien dan mudah untuk dilaksanakan.

Kriteria yang dikemukakan di atas, tidak dimaksudkan untuk memberikan belenggu pada guru dalam menyelesaikan tugasnya di kelas khususnya dalam mengembangkan tes, tetapi lebih diarahkan pada pengenalan kondisi ideal yang seharusnya dipenuhi oleh soal-soal yang disusun oleh pendidik, atau paling tidak memberikan arah kepada perbaikan Anda dalam memperbaiki sistem penilaian yang telah Anda lakukan selama ini.

2. Validitas dan Reliabilitas

Kualitas instrumen sebagai alat ukur ataupun alat pengumpul data diukur dari kemampuan alat ukur tersebut untuk dapat mengungkapkan dengan secermat mungkin fenomena-fenomena ataupun gejala yang diukur. Kualitas yang menunjuk pada tingkat kejelasan, kemantapan serta konsistensi dari data yang diperoleh itulah yang disebut dengan validitas dan reliabilitas.

a. Validitas

Validitas alat ukur menunjukkan kualitas kesahihan suatu instrumen atau alat pengumpul data dapat dikatakan valid atau sah apabila alat ukur tersebut mampu mengukur apa yang seharusnya diukur/diinginkan, sehingga alat ukur dikatakan sah apabila dapat mengungkap secara cermat dan tepat data dari variabel yang diteliti. Tinggi rendahnya tingkat validitas instrumen menunjukkan sejauhmana data dari variabel yang terkumpul tidak menyimpang dari gambaran tentang variabel yang dimaksud. Kerlinger (1986) menyatakan bahwa validitas alat ukur tidak cukup ditentukan oleh derajat ketepatan alat ukur dapat mengukur apa yang seharusnya diukur, tetapi perlu pula dilihat dari tiga kriteria yang lain yaitu *Appropriatness*, *Meaningfulness* dan *Usefulness*. Bila dikaitkan dengan pengukuran aspek perilaku sebagai hasil belajar, penjelasan ketiga kriteria tersebut secara bebas dapat diterjemahkan sebagai berikut: (1) ***Appropriatness***: Kriteria ini menunjuk pada kelayakan dari tes sebagai alat ukur tersebut, yaitu seberapa jauh alat ukur dapat menjangkau keragaman aspek perilaku tertentu; (2) ***Meaningfulness***: Adalah kriteria yang didasarkan pada kemampuan alat ukur untuk dapat memberikan keseimbangan item-item pengukurannya berdasar tingkat kepentingan/urgensi dari setiap bagian gejala; dan (3) ***Usefulness to inferences***: yakni kriteria ini menunjuk pada sensitif tidaknya alat ukur untuk dapat menangkap gejala perilaku, dan tingkat ketelitian yang ditunjukkan dalam pembuatan kesimpulan.

Jenis-jenis validitas yang dapat dipakai sebagai kriterium, dalam menetapkan tingkat kehandalan tes, diantaranya adalah:

- 1) **Validitas Permukaan (*Face Validity*)**: Validitas ini sering pula disebut sebagai validitas tampak. Validitas jenis ini menggunakan kriterium yang paling sederhana karena yang menjadi kriterianya hanya tampak atau penampakan dari instrumen itu sendiri. Apabila tes sebagai instrumen pengukuran, berdasar pengamatan sepintas telah dapat mengungkap fenomena yang akan dicari, bila secara sepintas sudah dianggap baik, maka alat tersebut sudah dapat dianggap memenuhi kriteria *face validity*, sehingga tidak diperlukan adanya pertimbangan mendalam.
- 2) **Validitas konsep (*Construct Validity*)**: Validitas ini disebut juga sebagai validitas konstruksi teori. Dalam hal ini alat ukur dikatakan valid apabila item sebagai alat ukur telah mencerminkan konsep perilaku yang diukur, dan memiliki tingkat kesesuaian dengan konstruksi teoritiknya. Validitas konstruksi ini sering

pula disebut sebagai **logical Validity**. Penggunaan validitas logis terutama dalam pengukuran-pengukuran gejala perilaku yang abstrak misalnya ukuran tentang kesetiakawanan, kematangan emosi, sikap terhadap KB, motivasi dan sebagainya.

- 3) **Validitas Isi (Content Validity)**: Sesuai dengan namanya validitas ini disebut pula sebagai validitas isi, pada validitas ini yang menjadi kriterium untuk menetapkan valid atau tidaknya alat ukur adalah isi/substansi dari variabel yang akan diukur, sehingga pada umumnya validitas ini hanya digunakan untuk mengukur variabel dengan cakupan materi yang jelas, misalnya saja dalam tes hasil belajar, alat ukur digunakan untuk dapat mengukur penguasaan siswa terhadap kompetensi bidang studi yang dipersyaratkan. Derajat validitas menunjuk pada kemampuan tes dalam menggambarkan topik-topik dan ruang lingkup cakupan materi yang akan diukur. Apabila alat ukur yang dikembangkan telah representatif, dalam arti mewakili semua cakupan materi, maka alat ukur tersebut telah memenuhi syarat *content validity*. Karena secara umum cakupan materi bidang studi biasanya berpedoman pada kurikulum yang telah ditetapkan maka *content validity* sering pula disebut sebagai "**Curriculair Validity**".
- 4) **Concurrent Validity**: Validitas ini dikenal pula dengan nama validitas bandingan, karena dalam menetapkan tingkat validitas alat ukur diperlukan kriterium luar yang berupa alat ukur lain yang serupa dan sudah dibakukan validitasnya. Apabila hasil pengukuran yang dilakukan dengan alat ukur baru, mempunyai tingkat kesesuaian dengan hasil yang pengukuran yang diperoleh dari alat ukur yang sudah dibakukan, maka tes sebagai alat ukur ini dianggap memenuhi *concurrent validity*.
- 5) **Factorial Validity**: Dalam kegiatan penelitian, tidak jarang terjadi sebuah skala pengukuran variabel terdiri dari beberapa faktor. Faktor-faktor tersebut diperoleh berdasar dimensi/indikator dari variabel/gejala yang diukur, sesuai yang terungkap dalam konstruksi teoritisnya. Meskipun variabel terdiri dari beberapa faktor, prinsip homogenitas untuk keseluruhan faktor harus tetap dipertahankan. Disamping perlu dicegah adanya overlap antara satu faktor dengan faktor yang lain. Sehingga kriterium yang digunakan dalam *factorial validity* ini dapat dilihat dengan menghitung homogenitas skor setiap faktor dengan total skor, serta homogenitas antara skor dari faktor yang satu dengan skor dari faktor yang lain.

Di samping pembagian validitas dengan jenis-jenis seperti telah diuraikan di atas, terdapat pula pembagian validitas yang hanya dikelompokkan menjadi dua kelompok besar yaitu validitas eksternal dan validitas internal.

b. Reliabilitas

Pengertian yang paling sederhana dari reliabilitas adalah kemantapan alat ukur dalam pengertian bahwa alat ukur tersebut dapat diandalkan atau memiliki keajegan hasil. Pada dasarnya hubungan antara validitas dan reliabilitas dapat dikemukakan bahwa alat ukur yang valid akan cenderung menghasilkan pengukuran yang reliabel, sebaliknya alat ukur yang reliabel sama sekali tidak menunjuk pada validitas alat ukur tersebut. Masalah validitas dan reliabilitas alat ukur nampak sangat jelas penggunaannya pada penelitian dengan pendekatan kauntitatif, karena penghitungan tingkat validitas dan reliabilitas pada umumnya juga menggunakan teknik statistik.

Kerlinger (1986: 443) mengemukakan bahwa reliabilitas dapat diukur dari tiga kriteria yaitu: (1) *Stability*, adalah kriteria yang menunjuk pada keajegan (konsistensi) hasil yang ditunjukkan alat ukur dalam mengukur gejala yang sama, pada waktu yang berbeda; (2) *Dependability*, yaitu kriteria yang mendasarkan diri pada kemantapan alat ukur atau seberapa jauh alat ukur dapat diandalkan; (3) *Predictability*, karena perilaku merupakan proses yang saling berkait dan berkesinambungan, maka kriteria ini mengidealkan alat ukur yang dapat diramalkan hasilnya dan meramalkan hasil pada pengukuran gejala selanjutnya.

Dengan mencermati pendapat di atas, maka batas reliabilitas atau keajegan dapat diartikan sebagai konsistensi skor yang diperoleh dari orang yang sama, pada gejala yang sama. Untuk itu ada kemungkinan skor pembandingan, mungkin berupa skor yang diperoleh dari alat ukur yang sama pada kesempatan yang berbeda, atau skor yang diperoleh dari alat ukur lain yang seimbang. Kerlinger menyatakan bahwa reliabilitas instrumen dikatakan baik bila alat tersebut dikenakan pada obyek yang sama, akan mendapatkan hasil yang sama pada beberapa kesempatan yang berbeda. Hal yang menjadi permasalahan dalam reliabilitas adalah kesalahan dalam penggunaan suatu alat ukur, semakin kecil kemungkinan kesalahan terjadi, maka akan semakin reliabel alat ukur tersebut. Dijelaskan lebih jauh bahwa reliabilitas alat ukur dapat ditingkatkan dengan cara memperbanyak butir item, dengan alasan bahwa secara statistik jumlah item yang banyak akan meningkatkan reliabilitas alat ukur. Meningkatkan reliabilitas alat ukur dapat pula dilakukan dengan menggunakan petunjuk pengerjaan yang jelas dan dengan menggunakan istilah-istilah yang jelas,

sesuai dengan tingkat pengetahuan dan bahasa responden, sehingga tidak menimbulkan keraguan atau kesalahpahaman dalam pengisian.

Pengukuran reliabilitas mendasarkan diri pada “*measurement error*” yaitu kesalahan yang bersumber dari proses pengukuran. Sehingga kesalahan dapat disebabkan oleh alat ukur ataupun dari perubahan-perubahan gejala yang diukur. Dalam penelitian sosial termasuk perilaku, sumber kesalahan pengukuran dapat ditengarai dari berbagai faktor diantaranya adalah (Kartono, 1996: 125): (1) hakekat dari gejala perilaku yang mudah sekali berubah, dan tidak dapat diulang dengan kondisi dan hasil yang sama, sebagai akibatnya hasil pengukuran perilaku juga akan selalu mengalami fluktuasi sejalan dengan perubahan waktu, dan kondisi-kondisi yang ada di sekitarnya; (2) kondisi pribadi yang ada pada diri seseorang bersifat tidak menetap, baik yang menyangkut tingkat kelelahan, suasana hati, dan sebagainya. Hal ini akan mempengaruhi perilaku, dan hasil pengukurannya; dan (3) ketidaktetapan hasil pengukuran juga dapat disebabkan oleh validitas alat ukur yang rendah, situasi pengukuran yang berubah-ubah, ketidaktetapan dalam pelaksanaan pengukuran maupun interpretasi terhadap hasil pengamatan serta kecermatan dalam pengadministrasian perlu mendapat perhatian.

Dengan mendasarkan diri pada keterbatasan penelitian sosial dan perilaku, maka dipahami bahwa angka yang diperoleh sebagai hasil pengukuran gejala sosial dan perilaku akan selalu berupa $True\ score + error$. *Error* yang terjadi bisa berarti skor yang diperoleh terlalu tinggi atau terlalu rendah. Sumber *error* (kesalahan atau penyimpangan) dapat berasal dari alat ukur, kondisi responden, pelaksanaan pengukuran ataupun interpretasi dan pengadministrasian. Langkah-langkah untuk menguji reliabilitas alat ukur pada dasarnya merupakan upaya untuk dapat mengetahui seberapa besar “salah ukur” dalam upaya mengukur gejala perilaku sebagai variabel penelitian. Hasil yang diperoleh disebut dengan “*Indeks Reliability*”. Koefisien reliabilitas selalu berada dalam rentangan 0 sampai dengan 1 yang menunjuk pada persentase varian *error* dengan sumber variasi yang berbeda. Misalnya koefisien reliabilitas menunjukkan 0.74 berarti 74 % varian skor yang bersumber pada keadaan yang diukur, sedang 26 % adalah kesalahan atau varian *error* yang bersumber dari keadaan di luar variabel yang diukur.

Cara mencari koefisien reliabilitas alat ukur, dapat dilakukan dengan menggunakan beberapa cara, yang masing-masing mempunyai kekurangan dan keunggulan. Berbagai pilihan tentang cara menetapkan tingkat reliabilitas alat ukur tersebut adalah:

1) Teknik Pengulangan (Test and Re Test Reliability)

Cara ini disebut sebagai teknik ulangan, karena dilakukan dengan memberikan dua kali pengukuran dengan rentang waktu tertentu dengan menggunakan alat ukur yang sama. Skor yang diperoleh pada pengukuran pertama dikorelasikan dengan skor dari hasil pengukuran pada pengukuran yang kedua. Koefisien yang diperoleh dengan cara ini menunjuk pada derajat stabilitas alat ukur. Pada umumnya sumber error pada teknik pengulangan ini dapat bersumber dari berbagai faktor yang menyebabkan seseorang mempunyai skor berbeda pada saat dua kali mengerjakan tes yang sama. Sangat mungkin perubahan skor yang terjadi bukan karena perubahan hal yang diukur, tetapi karena situasi yang berbeda atau pengalaman dari responden pada saat mengerjakan soal yang pertama, sehingga dalam pengerjaan tes kedua lebih hati-hati dan lebih baik hasilnya. Kebaikan utama dari cara ini adalah: karena subyek dan alat pengukuran yang digunakan sama, akan dapat memperkecil kemungkinan masuknya sumber error yang lain, tetapi perlu pula dipertimbangkan bahwa penggunaan subyek dan alat ukur yang sama dalam dua kali pengukuran, sekaligus juga merupakan kelemahan yang disebabkan karena adanya pengalaman mengerjakan akan mempengaruhi hasil pada pengukuran yang kedua.

2) Teknik Bentuk Paralel (*Alternate Form Reliability*)

Mencari reliabilitas dengan teknik bentuk paralel dilakukan dengan cara pengukuran pada subyek yang sama tetapi menggunakan alat ukur berbeda yang mempunyai tingkat kesamaan. Dengan cara ini peneliti perlu mempersiapkan dua set alat ukur yang berbeda dengan mempertimbangkan keseimbangan di antara kedua alat ukur tersebut. Keseimbangan diperlukan karena alat ukur ini ditujukan untuk mengukur gejala yang sama. Teknik ini sering juga disebut sebagai *Parallel Test Reliability*. Penggunaan dua set alat ukur dimaksudkan untuk mengurangi kemungkinan terjadinya pengaruh ingatan terhadap pengukuran yang pertama. Teknik ini dapat dilakukan dengan mengadakan pengukuran dengan alat ukur yang pertama berturut-turut waktunya dengan pengukuran dengan menggunakan alat ukur yang kedua pada subyek yang sama. Kemudian skor dari pengukuran alat ukur yang pertama dikorelasikan dengan skor hasil pengukuran yang kedua. Koefisien korelasi yang diperoleh akan mengungkap derajat ekuivalensi dan indeks stabilitas. Kemungkinan kesalahan pada cara ini dapat bersumber dari derajat keseimbangan antara dua alat ukur tersebut, serta kondisi yang mungkin berbeda pada saat pengukuran pertama dengan pengukuran kedua, meskipun dilakukan secara berturut-turut.

3) Teknik belah dua (*Split Half reliability*)

Teknik belah dua ini dikembangkan dengan menggunakan satu jenis alat ukur, dan hanya diberikan satu kali pada subyek, kemudian hasilnya diolah sedemikian rupa. Yaitu dengan cara mengelompokkan butir-butir itemnya menjadi dua bagian sama besar (belah dua). Pembagian item menjadi dua kelompok sama besar dapat dilakukan dengan cara acak atau pengelompokan berdasar nomor ganjil-genap, dapat pula dengan cara membagi menjadi separo kelompok bagian awal dan separo bagian akhir dalam jumlah yang sama. Setelah itu skor yang berasal dari belahan yang pertama dikorelasikan dengan skor pada belahan yang kedua. Koefisien korelasi yang diperoleh mencerminkan *derajat ekuivalensi* antara dua belahan tersebut. Teknik ini baru mencerminkan koefisien reliabilitas dari masing-masing belahan tersebut. Oleh karenanya untuk mendapatkan gambaran koefisien secara keseluruhan, koefisien antar belahan tersebut masih perlu dikoreksi dengan rumus sebagai berikut:

$$\text{Reliability} = \frac{N r_{x_1 x_2}}{1 + r_{x_1 x_2}}$$

Dimana x_1 adalah skor dari belahan satu, x_2 adalah skor dari belahan kedua, dan n adalah banyaknya subyek pada setia bagian (belahan). Rumus tersebut didasarkan pada asumsi bahwa kedua belahan mengukur hal yang sama, yang memiliki varian yang sama.

4) Kuder Richardson Reliability

Cara ini diberlakukan bila instrumen digunakan untuk mengukur satu gejala psikologis atau perilaku yang sama, artinya alat ukur tersebut dapat dikatakan reliabel bila terbukti ada konsistensi jawaban antar item yang satu dengan item yang lain. Sehingga apabila sifat dan tingkatan homogenitas antar item tidak terpenuhi, artinya alat tersebut dianggap mengukur lebih dari satu variabel. Bila dalam kenyataan dalam satu instrumen terdapat lebih dari satu skala pengukuran atau mengukur lebih dari satu variabel, dan setiap variabel memiliki beberapa dimensi, maka pengecekan reliabilitas dilakukan terhadap masing-masing skala pengukuran. Model Kuder Richardson Reliability ini menghasilkan *koefisien konsistensi internal* yang menunjuk pada derajat konsistensi antara item yang satu dengan item yang lain. Sehingga lebih cocok untuk alat ukur yang menggunakan item dua pilihan dengan salah satu jawaban benar.

6) Cronbach Alpha Reliability

Cara ini juga dikembangkan untuk mengujir konsistensi internal dari suatu alat ukur, perbedaan pokok dengan model Kuder Richardson adalah bahwa teknik ini tidak hanya untuk instrumen dengan dua pilihan tetapi tidak terikat pada dua pilihan saja, sehingga penerapannya lebih luas. Misalnya untuk menguji reliabilitas skala pengukuran sikap dengan 3, 5 atau 7 pilihan.

Satu hal yang tak kalah pentingnya adalah indeks sensitivitas, yang merupakan perbedaan kemampuan peserta didik antara setelah mengikuti proses pembelajaran dengan sebelum mengikuti proses pembelajaran. Indeks ini menyatakan tingkat keberhasilan peserta didik dalam mengikuti proses pembelajaran dan keberhasilan guru dalam melaksanakan proses pembelajaran. Besarnya indeks yang baik adalah positif dan besar.

Indeks ini sering dinyatakan dalam bentuk formula seperti berikut ini:

$$I_s = \frac{R_A - R_B}{T}$$

R_A = Jumlah peserta didik yang menjawab benar setelah mengikuti proses pembelajaran

R_B = Jumlah peserta didik yang menjawab benar sebelum mengikuti proses pembelajaran

T = Jumlah peserta didik yang mengikuti ujian

Rangkuman

Untuk dapat menjadi alat ukur yang baik dan dapat memberikan informasi yang akurat maka setiap soal sebagai bagian dari konstruksi tes harus dijaga kualitasnya. Ada beberapa kriteria yang dapat dipakai untuk menyusun butir-butir tes yang berkualitas yaitu; (1) valid, (2) relevan, (3) spesifik, (4) representatif, (5) seimbang (6) sensitif, (7) *fair*, dan (8) praktis

Validitas sebagai kriteria mutlak tes sebagai instrument terbagi menjadi 5 jenis yaitu; (1) validitas permukaan (*face validity*), (2) validitas konsep (*construct validity*), (3) validitas isi (*content validity*), (4) *concurrent validity*, dan (5) *factorial validity*.

Tes Formatif 3

Di bawah ini dicantumkan tes formatif yang bertujuan untuk mengukur pemahaman Anda mengenai uraian, contoh, dan rangkuman yang tercantum dalam subunit 3. Jawablah pertanyaan berikut sesuai dengan permintaan!

1. Jelaskan dengan contoh kriteria tes yang baik!
2. Jelaskan dengan contoh jenis-jenis validitas tes sebagai alat ukur!
3. Jelaskan mengapa *content validity* sering pula disebut sebagai validitas kurikulum!
4. Jelaskan jenis-jenis ukuran reliabilitas!
5. Jelaskan dengan contoh ketekaitan antara validitas dan reliabilitas!

Umpan Balik dan Tindak Lanjut

Jawablah pertanyaan tes formatif di atas, setelah selesai baru cocokkan dengan kunci jawabannya. Diskusikan dengan teman bila jawaban belum sesuai atau Anda belum merasa masih ada hal-hal yang meragukan. Hal ini sangat diperlukan karena kesepahaman tentang pengertian ini akan mendasari dan mempengaruhi langkah dan kegiatan dalam menyelesaikan mata kuliah ini.

Kunci Jawaban Tes Formatif

Tes Formatif 1

1. **Jenis tes dilihat dari tujuan penyelenggaraannya**, yaitu tes untuk keperluan seleksi, untuk menempatkan orang pada kelas-kelas tertentu, untuk mengetahui hasil belajar, untuk keperluan diagnostik, dan untuk keperluan uji coba.
2. **Jenis tes berdasar bentuk jawabannya adalah :**
 - a. **Tes Esei** : Tes yang jawabannya berbentuk uraian yang menuntut siswa mengorganisasikan gagasan-gagasan atau hal yang telah dipelajarinya dengan cara mengemukakan atau mengekspresikan gagasannya dalam bentuk tulisan.
 - b. Tes jawaban pendek peserta menuangkan jawabannya, dalam bentuk rangkaian kata-kata pendek, kata-kata lepas, maupun angka-angka.
 - c. Tes objektif adalah tes yang keseluruhan informasi yang diperlukan untuk menjawab tes telah tersedia.
3. **Tes melandasi pengambilan keputusan di awal proses pembelajaran, yaitu manfaat tes untuk mengetahui :**
 - a. Sejauhmanakah pengetahuan, keterampilan dan kemampuan yang dimiliki oleh siswa sebelum mengikuti proses pembelajaran hal ini diperlukan untuk menetapkan perencanaan pembelajaran.
 - b. Sejauhmanakah kemampuan dan keterampilan yang telah dicapai peserta didik terhadap pembelajaran yang direncanakan, untuk merancang materi dan metode pembelajaran yang direncanakan.
4. **Tes melandasi pengambilan keputusan selama proses pembelajaran.**

Tes dapat pula digunakan selama proses pembelajaran (tes formatif). Tes formatif dapat diberikan baik dalam bentuk tes tulis maupun tes lisan, baik dengan jawaban uraian maupun tes obyektif.
5. **Perbedaan antara tes obyektif dengan tes esei** : Tes objektif adalah tes yang keseluruhan informasi yang diperlukan untuk menjawab tes telah tersedia. Oleh karenanya sering pula disebut dengan istilah tes pilihan jawaban (*selected response test*). Butir soal telah mengandung kemungkinan jawaban yang harus dipilih atau dikerjakan oleh peserta tes. Menurut Subino (1987)

perbedaan yang khas bentuk soal objektif dibanding dengan soal esei adalah tugas peserta tes (*testee*) dalam merespons tes. Pada tes objektif, tugas *testee* adalah memanipulasikan data yang telah ada dalam butir soal. Oleh karenanya, tes objektif adalah tes yang dalam pemeriksaannya dapat dilakukan secara objektif.

Sedang **Tes Esei (*Essay-type Test*)** : Adalah tes bentuk uraian yang menghendaki siswa mengorganisasikan gagasan-gagasan atau hal yang telah dipelajarinya dengan cara mengemukakan atau mengekspresikan gagasannya dalam bentuk tulisan. Keunggulan tes uraian adalah dapat mengukur kemampuan siswa dalam hal mengorganisasikan pikirannya, mengemukakan pendapatnya, dan mengekspresikan gagasan-gagasan dengan menggunakan kata-kata atau kalimat sendiri. Sedang keterbatasannya adalah cakupan materi pelajaran yang terbatas, waktu pemeriksaan jawaban yang lama, penskorannya subyektif dan umumnya kurang handal dalam pengukuran.

Tes Formatif 2

1. Langkah-langkah pokok penyusunan tes :

- a. Langkah Perencanaan Tes: (1) Menentukan cakupan materi yang akan diukur, Menulis kompetensi dasar, Menulis materi pokok, Menentukan indikator, Menentukan jumlah soal, (2)Pemilihan bentuk tes (3)Menetapkan panjang tes.
 - b. Menulis Butir Pertanyaan meliputi: (1) Menulis draf soal (2) Memantapkan Validitas Isi (*Content Validity*), (3)Melakukan Uji Coba (*try out*), (4) Revisi soal.
2. Membuat matrik keterkaitan yang menggambarkan tata hubungan antara standar kompetensi, kompetensi dasar, indikator dan penetapan jenis tes serta contoh soal yang dikembangkan pada kelas saudara, dapat melihat contoh pada subunit 2 point 2.a.
 3. Try out dilakukan untuk tes buatan guru harus dilakukan dengan memperhatikan (1) apakah item-item soal cukup valid dan sesuai dengan kompetensi, (2) bagaimana rencana pelaksanaan tes untuk siswa, (3) berapa lama memperkirakan penggunaan waktu pengerjaan, (4) apakah kejelasan format tes, sudah cukup baik, (5) kejelasan petunjuk pengisian, dan (6) pemahaman bahasa yang digunakan dan sebagainya.
 4. Langkah mengembangkan tes untuk mengukur domain afektif. (1) Memilih variabel afektif yang akan diukur; (2) Membuat beberapa pernyataan tentang variabel afektif yang dimaksudkan; (3) Mengklasifikasikan pernyataan positif

atau negatif; (4) Menentukan jumlah gradual dan frase atau angka yang dapat menjadi alternatif pilihan; (5) Menyusun pernyataan dan pilihan jawaban menjadi sebuah alat penilaian; (6) Melakukan ujicoba; (7) Membuang butir-butir pernyataan yang kurang baik; dan (8) Melaksanakan penilaian.

5. Langkah mengembangkan tes psikomotor: (1) Menyusun soal dengan mencermati kisi-kisi instrumen psikomotor yang telah dibuat, menjabarkan indikator dengan memperhatikan materi pokok dan pengalaman belajar. (2) Menyusun Lembar Observasi dan Lembar Penilaian yang mengacu pada soal.

Tes Formatif 3

1. Kriteria tes yang baik :

- a. **Valid** : Soal dapat mengukur apa yang seharusnya diukur.
- b. **Relevan** : dapat mengukur kemampuan belajar sesuai dengan tingkat kemampuan yang ditetapkan dalam indikator pencapaian hasil belajar.
- c. **Spesifik** : tidak menimbulkan ambivalensi/spekulasi dalam memberikan jawaban.
- d. **Representatif** : dikembangkan dari satuan materi yang jelas cakupannya, dan bersifat komprehensif.
- e. **Seimbang** : pokok bahasan yang terpenting mendapat porsi terbanyak dalam soal.
- f. **Sensitif** : dapat membedakan siswa yang menguasai materi dengan yang tidak.
- g. **Fair** : terbuka tidak mengandung jebakan, jelas cakupan materinya, kejelasan norma yang dipakai serta kriteria keberhasilannya.
- h. **Praktis** : tidak sulit dilaksanakan dari segi pembiayaan maupun pelaksanaannya.

2. **Jenis-jenis validitas tes** : Validitas adalah kriteria mutlak tes sebagai instrumen.

Ada 5 jenis validitas yang dapat dipedomani yaitu yaitu; (1) validitas permukaan (*face validity*), (2) validitas konsep (*construct validity*), (3) validitas isi (*content validity*), (4) *concurrent validity*, dan (5) *factorial validity*.

3. **Content validity sering pula disebut sebagai validitas kurikulum.**

Derajat validitas isi (*content validity*) menunjuk pada kemampuan tes dalam menggambarkan topik-topik dan ruang lingkup cakupan materi yang akan diukur. Apabila alat ukur yang dikembangkan telah representatif, dalam arti mewakili

semua cakupan materi, maka alat ukur tersebut telah memenuhi syarat Content validity. Karena secara umum cakupan materi bidang studi biasanya berpedoman pada kurikulum yang telah ditetapkan maka content validity sering pula disebut sebagai “**Curriculair Validity**”.

4. Jenis-jenis ukuran Reliabilitas

- a. *Stability*, adalah kriteria yang menunjuk pada keajegan (konsistensi) hasil yang ditunjukkan alat ukur dalam mengukur gejala yang sama, pada waktu yang berbeda;
- b. *Dependability*, yaitu kriteria yang mendasarkan diri pada kemantapan alat ukur atau seberapa jauh alat ukur dapat diandalkan;
- c. *Predictability*, karena perilaku merupakan proses yang saling berkait dan berkesinambungan, maka kriteria ini mengidealkan alat ukur yang dapat diramalkan hasilnya dan meramalkan hasil pada pengukuran gejala selanjutnya.

5. Keterkaitan antara validitas dan reliabilitas.

Validitas menyangkut ketepatan tes dalam mengukur gejala yang diukur, sedang reliabilitas menunjuk pada konsistensi hasil pengukuran dari waktu ke waktu maupun antar bagian dari tes tersebut, sehingga tes yang valid yang dapat mengukur apa yang seharusnya dapat diukur pasti akan menunjukkan hasil yang konsisten atau reliabel tetapi hasil pengukuran yang konsisten tidak dapat menunjukkan dukungannya terhadap validitas, misalnya mengukur tingkat kecerdasan siswa dengan mengukur lingkaran kepala, hasilnya akan selalu konsisten dan reliabel, tetapi ukuran tersebut sama sekali tidak valid

Daftar Pustaka

- Arikunto, S. (2002). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Balitbang Depdiknas. (2006). *Panduan Penilaian Berbasis Kelas*. Jakarta: Depdiknas.
- Brookhart Susan M, Nitko J. Anthony. (2007). *Educational Assesment of Student*. Fifth edition. New Jersey: Meril Prentice Hall.
- Johson David, W & Johson, Roger T. (2002). *Meaningful Assessment*. Arlington Street Boston: Ally & Dacon A Pearson Education Company.
- Koufman, R. and Thomas, S. (1990). *Evaluations Without Fear*. New York: A Division of Franklin Watts.
- Poerwanti E. (2001). *Evaluasi Pembelajaran, Modul Akta mengajar*. UMM Press.
- Subino. (1987). *Konstruksi dan Analisis tes*. Jakarta: Dit-Jen Dikti.
- Silverius, S. (2001). *Evaluasi Hasil Belajar dan Umpan Balik*. Jakarta: Gramedia Widya Sarana.
- Stiggins, R. J. (1994). *Student Centered Classroom Assessment*. New York: Maxwell Macmillan International.
- Sudiyono, A. (1996). *Pengantar Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada.
- Syaifuddin, A. (2002). *Test Prestasi*. Yogyakarta.

Glosarium

- Indikator** adalah ciri-ciri atau tanda-tanda seseorang telah menguasai kompetensi standar.
- Kompetensi** adalah kemampuan bersikap, berpikir, dan bertindak secara konsisten sebagai perwujudan dari pengetahuan, sikap, dan keterampilan yang dimiliki oleh peserta didik.
- Kurikulum** adalah seperangkat rencana dan pengaturan mengenai tujuan, isi, dan bahan pelajaran serta cara yang digunakan sebagai pedoman penyelenggaraan kegiatan pembelajaran untuk mencapai tujuan pendidikan tertentu.
- KTSP** atau **Kurikulum tingkat satuan pendidikan** adalah kurikulum operasional yang disusun oleh dan dilaksanakan di masing-masing satuan pendidikan.
- Standar isi** adalah ruang lingkup materi dan tingkat kompetensi yang dituangkan dalam kriteria tentang kompetensi tamatan, kompetensi bahan kajian, kompetensi mata pelajaran, dan silabus pembelajaran yang harus dipenuhi oleh peserta didik pada jenjang dan jenis pendidikan tertentu.
- Standar Kompetensi Lulusan** adalah kualifikasi kemampuan lulusan yang mencakup sikap, pengetahuan, dan keterampilan.
- Standar Kompetensi Kelompok Mata Pelajaran** adalah kualifikasi kemampuan minimal peserta didik pada setiap kelompok mata pelajaran.
- Standar Kompetensi Mata Pelajaran** adalah kualifikasi kemampuan minimal peserta didik yang menggambarkan penguasaan sikap, pengetahuan, dan keterampilan yang diharapkan dicapai pada setiap tingkat dan/atau semester untuk mata pelajaran tertentu.
- Struktur kurikulum** adalah merupakan pola dan susunan mata pelajaran yang harus ditempuh oleh peserta didik pada satuan pendidikan dalam kegiatan pembelajaran.