

DAFTAR ISI

Halaman Judul	
i	
Daftar Isi	
ii	
Daftar Tabel	
iii	
BAB I PENDAHULUAN.....	1
A. Tujuan Evaluasi	1
B. Definisi Evaluasi	2
C. Pengukuran dan Penilaian.....	2
D. Prinsip-prinsip Evaluasi.....	3
E. Sifat dan Prinsip Pengukuran	5
F. Syarat Alat Ukur	9
G. Sistem	
Penilaian	10
H. Aspek-aspek Evaluasi Pembelajaran Seni	
Musik	11
I.	
Rangkuman	12
J. Pertanyaan dan	
Tugas	13

BAB II ASPEK-ASPEK PENILAIAN DALAM PEMBELAJARAN MUSIK ..14

A. Aspek Kognitif..... .15

B. Aspek

Psikomotor 20

C. Aspek

Afektif 22

D.

Rangkuman 24

E. Pertanyaan dan

Tugas 24

BAB III MENYUSUN SOAL

26 A. Jenis Tes..... 26

B. Prosedur Membuat Soal..... 29

BAB IV VALIDITAS DAN RELIABILITAS TES..... 37

A. Validitas..... 37

B. Reliabilitas.....

49

C. Pertanyaan dan Tugas

55

BAB V SISTEM PENILAIAN

56

A. Skoring Soal

Hal

Tabel 1. Contoh pembobotan materi tes apresiasi musik	32
Tabel 2. Contoh Kisi-kisi Tes Apresiasi Musik	33
Tabel 3. Contoh Rubrik Penilaian Materi Pelajaran Direksi	33
Tabel 4. Data Uji Coba Instrumen Pengukuran Prestasi Belajar Teori Musik..	42
Tabel 5. Data Urutan Sekor dari Sekor Rendah ke Sekor Tinggi	43
Tabel 6. Data Jawaban Siswa Kelompok Atas	44
Tabel 7. Data Jawaban Siswa Kelompok Bawah	44
Tabel 8. Data Soal nomor 4	46
Tabel 9. Distribusi Sekor Hasil Uji Coba Tes	51
Tabel 10. Data Perhitungan Reliabilitas	52
Tabel 11. Hasil Rating 3 Rater terhadap 10 Subyek.....	53

PENDAHULUAN

A. Tujuan Evaluasi

Dalam kehidupan sehari-hari kita tidak dapat lepas dari kegiatan evaluasi. Pada saat kita menyatakan bahwa suatu barang mahal atau murah, seseorang itu tampan atau cantik, sifat seseorang itu baik atau buruk, suatu tindakan itu sopan atau tidak, dan bahkan pada saat kita menentukan bahwa suatu lagu atau suara itu indah atau tidak, semua kegiatan tersebut merupakan kegiatan evaluasi.

Evaluasi merupakan rangkaian kegiatan dari suatu program yang bertujuan untuk menentukan keberhasilan suatu program. Worthen & Sanders (1981 : 19) mengungkapkan bahwa evaluasi adalah kegiatan untuk menentukan nilai sesuatu, didalamnya terkandung pemerolehan informasi yang digunakan untuk menentukan baik buruknya suatu program, produk, prosedur, tujuan, atau rancangan pendekatan untuk mencapai tujuan-tujuan tertentu. Isaac & Michael (1981 : 2) mengutip pendapat Stufflebeam mengemukakan bahwa tujuan evaluasi adalah untuk memperbaiki, bukan untuk membuktikan. Berdasarkan pendapat tersebut Isaac dan Michael berargumentasi bahwa karena tujuan evaluasi adalah untuk memperbaiki maka kegiatan evaluasi adalah kegiatan untuk menentukan efektivitas, kelayakan, atau nilai dari suatu proses atau produk berdasarkan persepsi evaluator. Menurut Stufflebeam yang kemudian dikembangkan oleh Alkin, evaluasi bertujuan untuk menyediakan pengetahuan dan nilai-nilai yang akan menjadi dasar suatu keputusan. Sedangkan Scriven berpendapat bahwa tujuan evaluasi adalah untuk menentukan nilai dari barang dan

jasa.

B. Definisi Evaluasi

Menurut Stufflebeam & Shinkfield (1985 : 159) evaluasi didefinisikan sebagai proses penggambaran, pemerolehan, dan penyediaan informasi keputusan tentang baik buruknya tujuan, rancangan, implementasi, dan dampak, yang dapat digunakan oleh pengambil kebijakan sebagai dasar pertanggungjawaban dalam memahami suatu fenomena. Fernandes (1984 : 8) mengemukakan bahwa dalam evaluasi mengandung dua kegiatan, yaitu mendeskripsikan dan memberi keputusan.

Berdasarkan pendapat-pendapat tersebut maka evaluasi dapat didefinisikan sebagai kegiatan menentukan baik buruknya suatu obyek berdasarkan deskripsi data-data yang diperoleh dengan prosedur yang dapat dipertanggungjawabkan. Kegiatan menentukan baik atau buruknya suatu obyek disebut penilaian, sedangkan kegiatan mendeskripsikan obyek disebut pengukuran.

C. Pengukuran dan Penilaian

Pengukuran adalah kegiatan yang dilakukan sebelum penilaian. Pengukuran adalah kegiatan pengumpulan data yang dideskripsikan baik dalam bentuk kuantitatif maupun kualitatif. Data yang diperoleh dari hasil pengukuran kemudian dibandingkan dengan suatu standar atau suatu norma. Perbandingan hasil pengukuran dengan standar atau norma kemudian diinterpretasikan yang merupakan hasil dari kegiatan penilaian.

Sebagai contoh seorang wanita mengukur tinggi badan dengan sebuah alat

meteran untuk mengetahui tinggi badannya. Kegiatan mengukur tinggi badan merupakan kegiatan pengukuran. Seandainya dalam kegiatan pengukuran tersebut diperoleh data bahwa tinggi badannya adalah 160 cm, maka dapat dinyatakan bahwa wanita tersebut tinggi, karena rata-rata tinggi badan wanita Indonesia adalah 150 cm. Namun apabila wanita tersebut ingin mendaftarkan diri sebagai pramugari atau sebagai model, maka wanita tersebut termasuk dalam kategori pendek, karena syarat minimal tinggi badan seorang pramugari atau seorang model adalah 170 cm. Kegiatan menentukan bahwa wanita tersebut tinggi atau pendek merupakan kegiatan penilaian.

Dengan demikian dapat dikatakan bahwa pengukuran adalah tindakan mengidentifikasi besar kecilnya gejala untuk menetapkan dengan pasti suatu luas, dimensi, dan kuantitas dari sesuatu, caranya membandingkan dengan ukuran tertentu, hasilnya dapat berupa angka atau pernyataan. Sedangkan penilaian adalah proses memberikan interpretasi terhadap hasil pengukuran dengan menggunakan norma tertentu untuk mengetahui tinggi rendahnya atau baik buruknya aspek tertentu.

D. Prinsip-prinsip Evaluasi

Kegiatan evaluasi dapat dipertanggungjawabkan apabila memenuhi prinsip-prinsip evaluasi. Joesmani (1988 : 22) mengemukakan prinsip-prinsip evaluasi antara lain : kepastian dan kejelasan, teknik, komprehensif, kesadaran adanya kesalahan pengukuran.

1. Kepastian dan Kejelasan

Evaluasi tidak dapat dilaksanakan apabila tujuan evaluasi tidak dirumuskan secara jelas dalam definisi operasional. Apabila kita akan mengevaluasi kemajuan belajar siswa maka kita harus mengidentifikasi dan mendefinisikan tujuan pembelajaran. Sebagai contoh apabila kita ingin mengetahui pemahaman siswa tentang pembuatan tangganda baru, kita perlu menjelaskan sampai berapa kuis atau mol.

Pada umumnya alat evaluasi dalam pendidikan adalah tes, baik tes praktik maupun tes tertulis. Tes tersebut harus mencerminkan karakteristik yang diukur. Sebagai contoh apabila kita akan mengevaluasi kemampuan siswa dalam menyanyi, maka komponen-komponen menyanyi harus dirumuskan secara jelas, misalnya kita akan melihat aspek kualitas suara, teknik, dan interpretasi dan ekspresi.

2. Teknik Evaluasi

Teknik evaluasi harus dipilih yang paling sesuai dengan tujuan evaluasi. Apabila kita akan mengevaluasi kemampuan siswa dalam bermain piano, maka teknik yang tepat adalah teknik observasi, yaitu *tester*(penguji) mengamati penampilan *testee*(orang yang diuji) berdasarkan lembar observasi berisi komponen-komponen yang dinilai. Apabila kita akan mengevaluasi pemahaman siswa tentang teori musik maka kita perlu membuat tes yang berisi pertanyaan-pertanyaan tentang teori musik.

3. Komprehensif

Evaluasi dalam pembelajaran adalah proses kegiatan yang bertujuan

untuk mengetahui prestasi belajar yang dicapai siswa. Agar kita dapat memperoleh informasi yang tepat maka kita perlu melakukan tes dengan berbagai teknik agar kita dapat mengetahui kemampuan siswa yang sebenarnya.

4. Kesadaran adanya kesalahan pengukuran

Evaluator harus menyadari keterbatasan dan kelemahan dalam teknik evaluasi yang digunakan. Atas dasar kesadaran ini maka evaluator harus lebih hati-hati dalam mengambil kebijakan, karena pengukuran yang dilaksanakan hanya berupa sampel, atau sebagian kecil dari kemampuannya.

E. Sifat dan Prinsip Pengukuran

Joesmani (1988 : 4) mengemukakan tiga sifat-sifat pengukuran, yaitu langsung dan tidak langsung, pengukuran adalah suatu proses, dan ketepatan pengukuran. Pengukuran dapat dilakukan secara langsung pada suatu obyek karena alat ukur dengan obyek begitu erat hubungannya. Sebagai contoh adalah pengukuran tinggi badan dan berat benda yang dapat diukur langsung dengan meteran atau timbangan. Adapun pengukuran tidak langsung contohnya adalah pengukuran kemampuan seseorang dalam memahami teori musik. Seseorang perlu membuat seperangkat tes untuk mengetahui penguasaan seseorang untuk mengetahui kaidah musik.

Pengukuran bukan merupakan kegiatan final karena hasil pengukuran perlu dibandingkan dengan kriteria atau standar tertentu untuk mengetahui kualitas hasil pengukuran. sebagai contoh seseorang siswa dapat menjawab 30 soal teori musik yang diberikan, untuk mengetahui apakah anak tersebut termasuk

pandai atau tidak kita harus mengetahui jumlah soal seluruhnya, apabila jumlah soal seluruhnya 50, berarti siswa tersebut termasuk dalam kategori sedang, namun apabila jumlah soal seluruhnya 100, berarti termasuk kurang. Namun skor 30 dapat berarti sangat baik, seandainya dalam suatu kelas tidak ada yang dapat mencapai lebih dari 30.

Ketepatan pengukuran sifatnya bervariasi. Pengukuran di bidang eksakta pada umumnya mempunyai derajat ketepatan yang lebih tinggi dari pada bidang sosial. Sebagai contoh pengukuran terhadap temperatur ruang atau suhu badan dapat diperoleh hasil yang akurat dari pada pengukuran terhadap tingkat musikalitas, karena untuk mengukur musikalitas kita harus mendeskripsikan atribut-atribut, setelah itu menjabarkan atribut tersebut dalam soal, dan hasilnya dapat berbeda jika pengukuran dilakukan oleh orang yang berbeda.

Pengukuran harus dilakukan berdasarkan prinsip-prinsip pengukuran. Adapun prinsip-prinsip pengukuran yang harus dipenuhi adalah menyeluruh, obyektif, terbuka, representatif, sasaran yang jelas, dan adanya kontrol, aturan skoring, dan keseksamaan.

Dalam pembelajaran, aspek-aspek yang dinilai harus menyeluruh, baik dalam segi materi pelajaran maupun aspek-aspek yang dinilai. Pengukuran harus obyektif artinya harus benar-benar didasarkan pada kemampuan yang dicapai siswa. Agar kita dapat mengukur kemampuan siswa sebenarnya, maka pengukuran harus terbuka, artinya siswa harus mengetahui apa saja yang hendak diukur sehingga mereka dapat mempersiapkan diri secara optimal. Siswa juga perlu mengetahui standar penilaian yang digunakan, sebagai contoh nilai C akan

dicapai siswa jika dapat menyanyikan lagu dengan lancar, nilai B jika siswa dapat menyanyikan lagu dengan penjiwaan yang baik.

Pengukuran harus didasarkan pada prinsip obyektivitas, artinya unsur-unsur subyektif *tester* (penguji) antara lain sikap simpati dan rasa benci harus benar-benar dihilangkan. Pusat perhatian pengukuran harus benar-benar tertuju pada atribut-atribut atau ciri-ciri yang melekat pada obyek yang diukur. Sebagai contoh seseorang yang ingin mengukur kemampuan menyanyi, maka pusat perhatian pengukuran harus pada suara *testee* (orang yang diuji), tidak terpengaruh oleh ketampanan atau kecantikannya.

Keterbukaan pengukuran sangat dituntut agar siswa dapat mengetahui tingkat kemajuan yang dicapai sehingga dapat menambah motivasi untuk meningkatkan prestasi belajar. Siswa yang diuji selain harus mengetahui aspek apa saja yang hendak diukur, dia juga harus mengetahui skor yang dicapai dalam pengukuran sehingga dia memahami bagian pelajaran yang belum dicapai dan kedudukan peringkatnya dalam suatu kelas. Hal ini akan memberikan informasi pada siswa tentang apa yang harus mereka pelajari dan kepada siapa mereka dapat belajar.

Pengukuran juga harus representatif, artinya bahwa atribut yang diukur harus benar-benar mewakili atau menggambarkan kemampuan siswa yang sebenarnya. Sebagai contoh kita akan mengukur kemampuan siswa dalam bernyanyi, maka atribut-atribut yang melekat dalam menyanyi harus dipertimbangkan dalam pengukuran. atribut-atribut tersebut antara lain : kualitas suara, teknik vokal, interpretasi dan ekspresi, serta penampilan.

Penentuan tujuan pengukuran harus mempunyai sasaran yang jelas sehingga memberikan panduan yang tepat bagi *tester* untuk melakukan pengukuran. Sebagai contoh seseorang ingin mengukur pengetahuan siswa tentang sejarah musik. Tester harus menetapkan terlebih dahulu ruang lingkup sejarah musik yang hendak diujikan. Dengan demikian perlu disusun *blue print* yang memberikan arah pengukuran yang kemudian dijabarkan dalam kisi-kisi sebelum soal disusun. Hal ini berlaku baik untuk pengukuran yang bersifat teori maupun praktek. Sebagai contoh seseorang ingin mengetahui tingkat keterampilan bermain piano, maka dia harus mempersiapkan rubrik(kriteria penilaian keseluruhan) yang berisi komponen-komponen keterampilan piano, bukan menyiapkan soal-soal tentang sejarah piano.

Kegiatan pengukuran harus dapat mengontrol pengaruh luar yang dapat mempengaruhi hasil pengukuran. Misalnya seseorang ingin mengetahui kemampuan *solfegio* para siswanya, maka ruangan yang digunakan harus terhindar dari suara gaduh.

Aturan skoring dalam pengukuran adalah kriteria tingkat skor. Aturan skoring dapat berupa daftar kemampuan beserta sekornya, atau ketentuan-ketentuan pemberian skor. Aturan skoring ini dapat juga berupa *benchmark* yaitu standar penilaian yang dapat digunakan untuk mengetahui tingkat prestasi seseorang atau suatu kelompok. Berikut ini salah satu contoh aturan skoring: Dalam pelajaran ansambel guru menetapkan bahwa nilai C apabila siswa dapat memainkan lagu yang menjadi bagiannya dengan lancar, nilai B apabila lagu dapat dimainkan secara lancar dan kompak, dan nilai A apabila siswa

mampu menyanyikan lagu dengan lancar, kompak, dan penuh ekspresi.

Pengukuran dapat menyangkut nasib seseorang, apakah seseorang dapat di terima di suatu sekolah atau perguruan tinggi, apakah seseorang dapat lulus atau tidak, bahkan apakah seseorang tersebut dapat diterima kerja atau tidak, hal itu banyak tergantung pada hasil pengukuran. Oleh karena itu pengukuran harus dilakukan dengan seksama, karena kegiatan pengukuran dapat berpengaruh besar pada hidup seseorang.

F. Syarat Alat Ukur

Kegiatan pengukuran dapat dilakukan apabila tersedia alat ukur yang disebut instrumen pengukuran. Kebenaran hasil pengukuran tergantung pada alat ukur yang digunakan. Oleh karena itu alat ukur atau instrumen pengukuran yang digunakan harus dapat dipertanggungjawabkan kualitasnya agar dapat memberikan informasi obyek yang diukur secara tepat. Adapun syarat yang harus dipenuhi oleh suatu instrumen pengukuran ditinjau dari dua hal, yaitu tingkat validitas dan reliabilitasnya.

Validitas adalah ketepatan, sedangkan reliabilitas adalah ketepatan. Suatu alat ukur dikatakan valid apabila dapat mengukur atribut suatu obyek secara tepat. Sebagai contoh seseorang ingin mengetahui tinggi badan, dia mengukur tinggi badan menggunakan timbangan, atau sebaliknya dia ingin mengetahui berat badan dengan menggunakan meteran. Tentu saja kedua alat tersebut bukan merupakan alat ukur yang tepat, karena dengan timbangan tidak mungkin akan didapat informasi tinggi badan seseorang, demikian juga sebaliknya dengan

meteran tidak mungkin didapat informasi tentang berat badan. Dengan demikian timbangan bukan merupakan alat ukur yang valid untuk mengukur tinggi badan, demikian juga meteran bukan merupakan alat ukur yang valid untuk mengukur berat badan.

Reliabilitas adalah sesuatu yang berkaitan dengan ketetapan atau keajegan. Suatu alat ukur dikatakan reliabel apabila dapat menghasilkan suatu hasil yang relatif sama walaupun dilakukan oleh orang yang berbeda dan ditempat serta waktu yang berbeda. Sebagai contoh seseorang ingin mengukur lebar suatu ruangan dengan menggunakan hasta. Alat ukur tersebut dikatakan valid karena dengan tangan bagian hasta dapat dihasilkan berapa hasta lebar suatu ruangan. Namun hasta bukan merupakan alat yang reliabel karena panjang hasta setiap orang berbeda, sehingga masing-masing orang dapat memberikan informasi yang berbeda untuk mengukur lebar suatu ruang, tergantung pada panjang hasta masing-masing.

G. Sistem Penilaian

Secara garis besar terdapat dua macam sistem penilaian, yaitu sistem Penilaian Acuan Patokan, dan sistem Penilaian Acuan Norma. Sistem penilaian acuan patokan adalah pemberian makna terhadap hasil pengukuran dengan didasarkan pada kriteria yang sudah ditentukan terlebih dahulu. Penilaian Acuan Norma adalah pemaknaan hasil pengukuran dengan membandingkan antara hasil pengukuran dan kedudukan peringkat dalam suatu kelompok atau kelas.

Sebagai contoh seorang siswa dapat menjawab 90 soal dari 100 soal,

maka secara otomatis siswa tersebut mendapat nilai A karena dalam kriteria penilaian yang sudah ditetapkan skor 80 – 100 adalah A, 66 – 79 adalah B, 56 – 65 adalah C, 40 – 56 adalah C, dan 39 ke bawah adalah E. Keputusan tersebut didasarkan pada sistem Penilaian Acuan Patokan karena pemberian makna hasil pengukuran didasarkan pada acuan yang sudah ditetapkan. Namun seorang siswa yang dapat menjawab 60 soal dari 100 soal belum tentu mendapat nilai C apabila menggunakan sistem Penilaian Acuan Norma. Siswa tersebut mungkin saja akan mendapat nilai A atau B apabila dalam suatu kelas skor 60 termasuk skor tinggi karena tidak ada siswa yang mencapai skor 100.

H. Aspek- aspek Evaluasi Pembelajaran Seni Musik

Tujuan diadakannya sekolah adalah untuk membentuk kepribadian pelajar melalui materi-materi pelajaran (Mursell, 1982: 5). Dengan pemberian materi pelajaran tersebut diharapkan siswa tidak sekedar dapat memahami suatu konsep dan menguasai suatu keterampilan, namun nilai-nilai yang terkandung dalam materi pelajaran tersebut juga dapat berpengaruh dalam kehidupan mereka. Sebagai contoh, seseorang dapat hidup hemat karena mendapat pelajaran ekonomi, seseorang dapat mengantisipasi permasalahan-permasalahan yang mungkin timbul dimasa yang akan datang karena mendapat pelajaran sejarah, seseorang dapat berpikir secara logis karena mendapat pelajaran matematika, seseorang dapat hidup secara lebih praktis karena mendapat pelajaran fisika, dan seseorang dapat merasakan keindahan karena mendapat pelajaran kesenian, serta seseorang dapat merasakan kebahagiaan karena mendapat pelajaran agama. Dengan demikian aspek-aspek yang dievaluasi dalam pembelajaran bukan sekedar

penguasaan materi pelajaran, tetapi juga dampaknya dalam kehidupan pelajar.

Prestasi belajar yang harus dievaluasi dalam proses pembelajaran meliputi tiga aspek, yaitu aspek kognitif, aspek psikomotor, dan aspek afektif. Aspek kognitif adalah kemampuan yang berkaitan dengan penalaran, aspek psikomotor adalah kemampuan dibidang keterampilan, sedangkan aspek afektif adalah aspek yang berkaitan dengan sikap.

I. Rangkuman

Evaluasi adalah kegiatan yang bertujuan untuk mengetahui keberhasilan suatu program. Kegiatan evaluasi tersebut meliputi dua tahap, yaitu tahap pengukuran dan tahap penilaian. Pengukuran adalah kegiatan mendiskripsikan obyek baik secara kuantitatif maupun kualitatif, sedangkan penilaian adalah kegiatan pemaknaan atau interpretasi dari hasil pengukuran.

Evaluasi harus dilakukan berdasarkan prinsip-prinsip kepastian dan kejelasan, teknik evaluasi, komprehensif, dan kesadaran adanya kesalahan pengukuran. adapun prinsip-prinsip yang harus dipenuhi oleh kegiatan pengukuran adalah menyeluruh, obyektif, terbuka, representatif, sasaran yang jelas, dan adanya kontrol, aturan skoring, dan keseksamaan.

Kegiatan pengukuran dapat dilakukan apabila tersedia alat ukur yang disebut instrumen pengukuran. Kebenaran hasil pengukuran tergantung pada alat ukur yang digunakan. Adapun syarat yang harus dipenuhi oleh suatu instrumen pengukuran ditinjau dari dua hal, yaitu tingkat validitas dan reliabilitasnya.

Secara garis besar terdapat dua macam sistem penilaian, yaitu sistem Penilaian Acuan Patokan, dan sistem Penilaian Acuan Norma. Prestasi belajar yang harus dievaluasi dalam proses pembelajaran meliputi tiga aspek, yaitu aspek kognitif, aspek psikomotor, dan aspek afektif.

J. Pertanyaan dan Tugas

1. Apakah yang dimaksud evaluasi ?
2. Bagaimanakah hubungan antara pengukuran, penilaian, dan evaluasi ?
3. Jelaskan mengapa dalam melakukan kegiatan evaluasi harus memenuhi prinsip kepastian dan kejelasan, serta komprehensif ?
4. Mengapa dalam melakukan pengukuran harus obyektif, representatif, dan terbuka ?
5. Buatlah contoh aturan skoring dalam kegiatan pengukuran musik ?
6. Apakah yang dimaksud validitas dan reliabilitas ?
7. Seorang pemburu ingin menembak burung yang hinggap di atas bunga, ternyata pemburu tersebut berhasil menembak bunga, apabila dianalogikan dengan alat ukur, maka pemburu tersebut tidak memenuhi kriteria ? jelaskan !
8. Menurut anda kapan saat yang tepat menggunakan sistem PAN ? Jelaskan !
9. Berikan contoh penggunaan sistem PAP !
10. Mengapa dalam mengukur prestasi belajar siswa perlu memperhatikan aspek kognitif, psikomotor, dan afektif ?

BAB II

ASPEK-ASPEK PENILAIAN DALAM PEMBELAJARAN SENI MUSIK

Dalam kurikulum berbasis kompetensi 2004 (Depdiknas, 2003 : 1) disebutkan bahwa Pendidikan seni di sekolah umum pada dasarnya diarahkan untuk menumbuhkan kepekaan rasa estetik dan artistik sehingga terbentuk sikap kritis, apresiatif dan kreatif pada diri siswa secara menyeluruh. Sikap ini hanya mungkin tumbuh jika dilakukan serangkaian proses kegiatan pada siswa yang meliputi kegiatan pengamatan, penilaian, serta penumbuhan rasa memiliki melalui keterlibatan siswa dalam segala aktivitas seni di dalam kelas dan atau di luar kelas. Dengan demikian aspek-aspek yang dinilai dalam pembelajaran musik bukan hanya meliputi keterampilan bermain musik atau bernyanyi, tetapi juga tentang wawasan musik dan sikapnya terhadap seni musik. Namun demikian karena pembelajaran musik bersifat praktik maka aspek psikomotor mendapat persentase yang lebih banyak.

Pada saat ini klasifikasi prestasi hasil belajar yang lebih banyak digunakan adalah perpaduan klasifikasi Bloom, Kratwohl dan Simpson. Bloom, Madaus,& Hasting mengklasifikasikan aspek kognitif dan afektif, sedangkan E. Simpson dan A. Harrow berhasil mengembangkan klasifikasi hasil belajar ranah (aspek) kognitif (Kun S. Astuti, 2001 : 40).

Berikut ini akan dijabarkan klasifikasi aspek kognitif menurut Bloom, Madaus, dan Hasting (1981 : 331) secara lebih terperinci.

A. Aspek Kognitif

Aspek kognitif adalah kemampuan seseorang dalam menggunakan daya penalarannya. Bloom dkk. Mengklasifikasikan domain (aspek) kognitif dalam enam tingkatan yang bersifat hierarki, artinya tingkat kemampuan yang lebih rendah merupakan syarat dicapainya kemampuan yang berada di atasnya. Adapun ke-enam tingkatan tersebut adalah sebagai berikut :

1. Pengetahuan

Tingkatan ini sering disingkat dengan C_1 (Kognitif tingkat 1). Aspek pengetahuan adalah kemampuan yang berhubungan dengan kemampuan mengingat sesuatu, baik berupa kemampuan mengingat hal-hal yang bersifat spesifik maupun pengetahuan yang bersifat umum, mengingat kembali suatu metode atau proses, atau mengingat kembali suatu tahapan, metode, struktur atau seting. Tingkat ini dirinci lagi menjadi tiga kemampuan, yaitu pengetahuan yang bersifat spesifik, pengetahuan tentang cara-cara dan makna yang spesifik, serta pengetahuan umum dan abstraksinya dalam suatu bidang.

a. Pengetahuan Spesifik

Kemampuan ini merupakan kemampuan yang paling rendah, sebagai contoh adalah mengingat simbol-simbol spesifik. Dalam bidang musik dapat berupa kemampuan mengingat simbol-simbol musik, seperti bentuk dan nilai not, tanda-tanda ekspresi, tanda ulang, dan lain sebagainya.

b. Pengetahuan tentang Cara dan Makna Spesifik

Kemampuan ini berupa pengetahuan tentang cara menyusun, mempelajari, memutuskan, dan mengkritisi. Sebagai contoh adalah pengetahuan tentang cara

memegang biola, posisi yang benar dalam bermain gitar, cara menentukan posisi duduk yang tepat dalam bermain piano.

c. Abstraksi Pengetahuan Umum

Abstraksi pengetahuan umum adalah pengetahuan tentang skema dan tahap-tahap umum dari suatu fenomena atau penyusunan ide-ide. Hal itu dapat berupa pengetahuan tentang teori-teori umum. Sebagai contoh adalah pengetahuan tentang hukum-hukum ilmu harmoni.

2. Pemahaman

Pemahaman adalah kemampuan menjelaskan sesuatu tanpa harus dihubungkan dengan materi yang lain. Kemampuan ini meliputi kemampuan menterjemahkan (*translation*), menginterpretasikan, dan mengekstrapolasi.

a. Menterjemahkan

Kemampuan menterjemahkan bukan berarti menterjemahkan suatu kalimat dari suatu bahasa dalam bahasa lain, tetapi berupa kemampuan menjelaskan suatu obyek dengan kalimat sendiri. Sebagai contoh seorang siswa mampu menjelaskan arti dinamik *crescendo* dengan kalimat sendiri.

b. Interpretasi

Interpretasi adalah kemampuan merangkum kembali suatu komunikasi atau suatu bacaan. Apabila kemampuan menterjemahkan hanya mengungkap kembali suatu materi dengan kalimat sendiri, maka kemampuan interpretasi mengandung kemampuan menyusun pandangan baru. Sebagai contoh seseorang dapat menjelaskan makna suatu lagu berdasarkan penjelasan informasi dari suatu artikel, namun dia juga mampu mengungkapkan pandangannya sendiri tentang lagu

tersebut.

c. Ekstrapolasi

Ekstrapolasi adalah kemampuan menjelaskan kecenderungan dan menentukan implikasi, konsekuensi, atau dampak suatu fenomena. Sebagai contoh seseorang mampu memprediksikan suasana lagu yang banyak menggunakan nada-nada *disonan*.

3. Aplikasi

Aplikasi adalah kemampuan menggunakan kemampuan abstrak dalam situasi konkrit. Abstraksi tersebut dapat berupa ide-ide yang bersifat umum, aturan prosedur, atau generalisasi suatu metode. Abstraksi juga dapat berupa prinsip-prinsip suatu teknik, ide, atau teori yang harus diingat dan diterapkan. Sebagai contoh seorang siswa mampu menyusun tangga nada baru setelah diberi aturan jarak tangga nada mayor.

4. Analisis

Analisis adalah kemampuan untuk memecah atau memerinci suatu dalam elemen-elemen atau unsur-unsur yang lebih kecil, seperti memperjelas suatu ide rumit, menggambarkan proses terjadinya dampak dari suatu sebab. Kemampuan analisis ini sering disingkat C4 (kognitif 4), dan terdiri dari tiga jenis, yaitu menganalisis unsur-unsur, menganalisis hubungan, dan menganalisis prinsip-prinsip yang telah tersusun.

a. Menganalisis Unsur-unsur

Kemampuan ini meliputi kemampuan mengidentifikasi atau memerinci suatu obyek dalam bagian-bagian yang lebih kecil. Sebagai contoh seorang siswa mampu memerinci susunan nada dalam suatu akord, memerinci bagian-bagian dari suatu lagu.

b. Menganalisis Hubungan

Kemampuan menganalisis hubungan adalah kemampuan seseorang untuk menjelaskan hubungan dan interaksi antar unsur. Sebagai contoh seorang siswa mampu menjelaskan hubungan antara unsur ritme dengan nada, dan harmoni dalam pokok bahasan unsur-unsur musik.

c. Menganalisis Prinsip-prinsip

Kemampuan analisis jenis ini adalah kemampuan seseorang dalam menganalisis prinsip-prinsip yang meliputi dasar-dasar, hubungan-hubungan, dan mekanisasinya. Sebagai contoh seseorang mampu menjelaskan elemen-elemen yang menyusun karakteristik style atau gaya musik dangdut, keroncong, jazz, pop, dan campur sari.

5. Evaluasi

Pada buku Bloom edisi lama jenjang evaluasi merupakan jenjang tertinggi di atas jenjang sintesa, namun pada buku Bloom edisi terakhir posisi tersebut dibalik, sehingga kemampuan sintesa merupakan kemampuan yang paling tinggi di atas kemampuan evaluasi.

Kemampuan evaluasi adalah kemampuan untuk menilai suatu metode atau suatu materi baik atau tidak, benar atau salah, dan bagus atau tidak. Penilaian tersebut dapat bersifat kuantitatif maupun kualitatif didasarkan pada

kriteria. Kriteria tersebut dapat berupa standar yang sudah ada atau didasarkan pada pendapat siswa sendiri.

a. Keputusan Evaluasi Didasarkan pada Evidensi Internal

Ketepatan evaluasi didasarkan pada logika, konsistensi, dan kriteria internal yang lain. Sebagai contoh seorang siswa menilai kualitas lagu didasarkan pada pendapat logis mereka sendiri.

b. Keputusan Didasarkan pada Kriteria Eksternal

Evaluasi didasarkan pada ketentuan kriteria yang telah ada. Sebagai contoh seorang siswa menilai kualitas karya aransemen berdasarkan kaidah-kaidah harmoni yang sudah ada.

6. Sintesa

Jenjang C6 ini merupakan tingkat tertinggi dari aspek kognitif. Sintesa adalah kemampuan seseorang dalam menyusun elemen-elemen atau unsur-unsur dan bagian dalam bentuk satu kesatuan menjadi sesuatu yang unik atau baru. Produk kemampuan sintesa terdiri dari tiga bentuk, yaitu dapat berupa sesuatu yang bersifat unik, rencana suatu program, dan derivasi dari hubungan yang bersifat abstrak.

a. Menyusun Sesuatu yang Bersifat Unik

Kemampuan ini dapat berupa kemampuan seseorang dalam mengembangkan ide atau mengungkapkan perasaan kepada orang lain dengan media yang mereka kuasai, sebagai contoh adalah seorang seniman membuat suatu karya lagu, lukisan, atau seorang penulis menyusun sebuah buku atau novel.

b. Menyusun Proposal atau Rencana Kerja

Hasil aspek kognitif jenjang sintesa jenis ini dapat berupa pengembangan rencana kerja atau proposal rencana kerja. Rencana tersebut dapat berupa inisiatif pengembangan dari ide siswa sendiri atau tugas yang diberikan guru pada siswa. Sebagai contoh rencana program pertunjukan yang akan digelar siswa.

c. Derivasi dari Seperangkat Hubungan yang Bersifat Abstrak

Kemampuan ini berupa kemampuan menderivasikan atau menurunkan suatu konsep atau prinsip yang bersifat abstrak atau menjelaskan suatu fenomena atau data, atau menjelaskan proposisi secara deduktif dari seperangkat proposisi dasar atau simbol-simbol. Sebagai contoh seorang siswa mampu mendiskripsikan suatu karya musik dengan ulasan yang didasarkan pada prinsip-prinsip seni.

B. Aspek Psikomotor

Aspek psikomotor adalah kemampuan yang banyak berhubungan dengan hal-hal yang bersifat motorik, dapat berupa *skill* atau keterampilan. Simpson (Sirait, 1989 : 104) mengklasifikasi kemampuan psikomotor dalam tujuh tingkatan, yaitu set, imitasi, respon terbimbing, mekanisme, respon kompleks, adaptasi dan orisinalitas.

1. Persepsi

Set adalah kemampuan motorik yang berhubungan dengan kepekaan penginderaan, baik indra penglihatan, pendengaran, maupun indera perasa. Sebagai contoh seseorang mampu membedakan nada yang fals dan nada yang tepat, seseorang dapat merasakan bahwa suatu masakan lezat atau tidak, atau seseorang mampu merasakan bahwa suatu perpaduan warna serasi atau tidak.

2. Set

Set adalah kemampuan untuk menunjukkan jenis aksi tertentu. Kemampuan ini meliputi mental set (kesediaan mental untuk beraksi), fisik set (kesediaan fisik untuk beraksi), dan emosional set (kemauan beraksi). Contoh kemampuan ini adalah seorang siswa yang mampu memperagakan cara memegang gitar yang benar, posisi duduk yang benar dalam bermain piano, atau posisi penjarian yang benar dalam bermain biola.

3. Responsi Terbimbing

Responsi terbimbing merupakan tahap permulaan dalam belajar keterampilan yang lebih kompleks. Selain diperlukan kemampuan peniruan, responsi terbimbing ini juga mengandung *trial and error*. Sebagai contoh seorang siswa mampu menirukan teknik-teknik penjarian yang diajarkan oleh guru, seorang siswa mampu memainkan teknik *staccato* pada biola sebagaimana dicontohkan instruktur.

4. Mekanisme

Mekanisme berkenaan dengan gerakan memadukan berbagai teknik yang sudah menjadi kebiasaan. Sebagai contoh seorang anak mampu memainkan piano dengan menampilkan teknik legato, *staccato*, dan penjarian secara benar.

5. Respon Nyata yang Kompleks

Kemampuan ini berkenaan dengan penampilan pola gerakan-gerakan kompleks secara mahir. Kemahiran dapat dilihat dari kecepatan, keluwesan, atau ketepatan, sehingga suatu ketrampilan dapat ditunjukkan secara otomatis atau spontan. Sebagai contoh seorang pemain piano dapat memainkan lagu *Turkish*

March dengan lancar dan penuh ekspresi tanpa membaca partitur.

6. Adaptasi

Adaptasi atau penyesuaian berkenaan dengan keterampilan yang telah dikembangkan dengan amat baik sehingga siswa dapat mengolah gerakan sesuai dengan situasi yang dituntut. Sebagai contoh seorang penyanyi mampu menyanyikan lagu dengan gaya campur sari, pop, dan dangdut secara baik, sesuai dengan permintaan penonton.

7. Orisinalitas

Kemampuan ini berkenaan dengan penciptaan pola gerakan yang baru. Hasil belajar pada tingkatan ini menekankan kreativitas yang dikembangkan berdasarkan keterampilan tingkat tinggi, yang dapat dijadikan pemilikinya sebagai ciri identitas dirinya. Sebagai contoh Idris Sardi mampu memainkan lagu dengan *tone colour* yang sangat khas dan indah, sehingga tanpa melihat permainannya seseorang dapat mengenal bahwa itu permainan biola Idris Sardi. Contoh lain adalah Yuni Sara yang mampu menyanyikan lagu Mutiara yang Hilang dengan gaya yang lebih hidup, Kris Dayanti mampu menciptakan pola improvisasi baru, Kenny G mampu memainkan Saxophone dengan *tone colour*, teknik, dan gaya yang sangat bagus dan khas, yang belum pernah ada sebelumnya.

C. Aspek Afektif

Aspek afektif adalah aspek yang banyak berhubungan dengan sikap. Krathwohl (Sirait, 1989 : 103) mengklasifikasikan aspek afektif pada jenjang menerima, menjawab, menilai, dan mengorganisasi.

1. Menerima

Menerima adalah kemauan siswa atau seseorang dalam mengikuti respon tertentu. Sebagai contoh kemauan siswa untuk mengikuti peraturan sekolah, kemauan siswa dalam memperhatikan penjelasan guru dengan tenang pada saat proses pembelajaran, dan kemauan siswa untuk tidak merokok.

2. Menjawab

Menjawab adalah aspek afektif yang berhubungan dengan partisipasi aktif siswa dalam merespon sesuatu. Hasil belajar dalam bidang ini dapat berupa persetujuan bisu (membaca tugas bacaan), kemauan untuk menjawab (membaca buku secara suka rela di luar hal yang ditugaskan, dan kepuasan untuk menjawab (membaca untuk kesenangan). Contoh dalam bidang musik adalah kemauan siswa untuk berlatih memainkan lagu-lagu sesuai yang ditugaskan dalam pelajaran piano, seorang siswa dengan suka rela meluangkan waktu untuk berlatih memainkan lagu-lagu diluar yang ditugaskan, atau seorang siswa yang sangat kecanduan berlatih untuk mencapai tingkat kemampuan yang setinggi-tingginya.

3. Menilai

Menilai berkenaan dengan nilai yang diberikan siswa pada suatu obyek. Menilai didasarkan pada internalisasi dari sejumlah nilai yang dinyatakan dalam tingkah laku. Hasil belajar dalam tahap ini berkenaan dengan tingkah laku yang bersifat tetap. Sebagai contoh kebiasaan datang tepat waktu, kebiasaan menepati janji, dan kebiasaan mementingkan memilih kegiatan yang positif dan mendukung studi dari pada kegiatan hura-hura.

4. Organisasi

Organisasi berkenaan dengan mempersatukan nilai-nilai yang berbeda, memecahkan pertentangan, dan mulai membangun suatu nilai batiniah yang bersifat tetap. Jadi penekanannya adalah dalam membandingkan, menghubungkan, dan mensintesis nilai-nilai. Sebagai contoh adalah mengembangkan kegiatan seperti membentuk grup Paduan Suara, band, atau kelompok belajar.

D. Rangkuman

Secara garis nesar terdapat tiga aspek yang perlu diperhatikan dalam menilai prestasi belajar siswa, yaitu aspek kognitif, psikomotor, dan afektif. Aspek kognitif adalah kemampuan yang berhubungan dengan penalaran, aspek psikomotor berkenaan dengan kemampuan motorik, dan aspek afektif adalah kemampuan yang berhubungan dengan sikap atau perasaan. Aspek kognitif diklasifikasikan secara hierarki dalam enam jenjang, yaitu C1(pengetahuan), C2 (pemahaman), C3(aplikasi), C4(analisa), C5(evaluasi), dan C6(sintesa). Aspek psikomotor terbagi dalam jenjang P1(Persepsi), P2(Set), P3(Respon terbimbing), P4(Mekanisme), P5(Respon kompleks), P6(Adaptasi), dan P7 (Orisinalitas). Aspek afektif meliputi empat jenjang yaitu A1(Menerima), A2(Menjawab), A3 (Menilai), dan A4 (Mengorganisasi).

E. Pertanyaan dan Tugas

1. Berikan dua contoh kemampuan aspek kognitif jenjang pengetahuan !
2. Berikan dua contoh kemampuan aspek kognitif jenjang pemahaman !
3. Berikan dua contoh kemampuan aspek kognitif jenjang aplikasi !

4. Seorang anak mampu menyusun tangga nada baru setelah diberi teori tentang pembentukan tangga nada. Kemampuan tersebut termasuk aspek kognitif, psikomotor. Atau afektif ? dan termasuk jenjang apa ? jelaskan alasannya !
5. Seorang anak mampu memainkan lagu sebagaimana telah dicontohkan oleh instruktur. Berarti anak tersebut mampu menunjukkan kemampuan jenis aspek apa ? jenjang yang mana ?
6. Berikan contoh kemampuan aspek kognitif jenjang sintesa ?
7. berikan contoh kemampuan aspek kognitif jenjang analisis ?
8. Berikan contoh kemampuan aspek psikomotor aspek adaptasi !
9. Berikan contoh kemampuan aspek psikomotor aspek orisinalitas !
10. Seorang siswa mampu mengkoordinir teman-temannya untuk membentuk suatu grup band dan mementaskannya. Hal itu termasuk kemampuan jenis aspek apa ? jenjang apa ? Jelaskan alasannya !

BAB III

MENYUSUN SOAL

A. Jenis Tes

Setelah kita mengetahui aspek-aspek apa saja yang perlu diperhatikan dalam menilai prestasi belajar musik, langkah selanjutnya adalah menyusun soal. Jenis soal atau tes, ditinjau dari fungsinya terdapat berbagai macam, antara lain adalah tes seleksi, tes penempatan, tes diagnostik, dan tes prestasi. Cronbach (Saifuddin Azwar, 1996 : 5) mengklasifikasikan tes menjadi dua golongan yaitu *maximum performance* (power tes) dan performansi tipikal. Ditinjau dari prosedurnya tes dibagi dalam empat jenis, yaitu pre-tes, post tes, formatif tes, dan sumatif tes. Apabila ditinjau dari bentuk jawaban, tes dibagi menjadi tes obyektif dan tes subyektif. Dan ditinjau dari sifatnya, tes dapat dibagi menjadi dua yaitu tes teori dan tes praktik.

Tes seleksi adalah tes yang bertujuan untuk memilih calon, baik calon siswa, mahasiswa, atau pegawai. Tes penempatan adalah tes yang diselenggarakan untuk menempatkan seseorang dalam kelompok tertentu. Sebagai contoh setelah diterima sebagai siswa, seorang anak apakah akan dikelompokkan pada kelas akselerasi atau kelas reguler. Contoh lain, apakah seorang siswa lebih cocok masuk jurusan eksakta, atau sosial. Tes diagnostik adalah tes yang bertujuan untuk mengetahui kelemahan atau kesulitan siswa. Sebagai contoh seorang guru menemukan bahwa terdapat siswa yang selalu mendapat nilai rendah, kemudian guru tersebut menyusun soal khusus, untuk mengetahui sebab-sebab kesulitan siswa, oleh karena itu tes diagnostik ini berisi

soal-soal mempunyai taraf kesukaran yang rendah.

Maximum performance test (tes performansi maksimal) adalah tes yang dirancang untuk mengungkap apa yang mampu dilakukan oleh seseorang dan seberapa baik ia melakukan. Tes yang termasuk jenis ini antara lain tes inteligensi, tes bakat, dan tes prestasi. Dalam tes jenis ini peserta selalu didorong untuk mencapai nilai tinggi, dan mereka diberi informasi tentang kriteria penilaian. *Typical performance* (tes performansi tipikal) dirancang untuk mengungkap kecenderungan reaksi atau perilaku individu ketika berada dalam situasi-situasi tertentu. Jadi yang diukur bukan apa yang mampu dilakukan siswa melainkan kecenderungan perilaku.

Pre-tes adalah tes yang diberikan sebelum pelajaran dimulai. Tes ini dimaksudkan untuk mengetahui kemampuan awal siswa, sehingga guru dapat mengawali pelajaran dengan materi yang tepat. Post-tes adalah tes yang diberikan setelah pelajaran selesai, tujuannya untuk mengetahui seberapa jauh siswa menyerap materi pelajaran yang telah diajarkan. Formatif tes adalah tes yang dilakukan setelah satu pokok bahasan selesai diajarkan. Formatif tes dapat dilakukan setelah dua atau tiga kali tatap muka, tergantung pada banyak dan sedikitnya atau tingkat kesulitan materi pelajaran. Tes formatif ini dapat juga dianggap sebagai ulangan harian. Tes sumatif adalah tes yang dilakukan setelah seluruh pelajaran selesai disampaikan, misalnya setelah satu semester atau satu tahun. Tujuan pre-tes, post-tes, dan formatif tes adalah sama, yaitu untuk memberi motivasi siswa untuk mencapai prestasi sebaik-baiknya, sedangkan tes-sumatif tujuannya adalah untuk menentukan jenjang seseorang. Sebagai contoh apakah

seorang siswa naik kelas atau tidak, atau seorang siswa lulus atau tidak.

Tes obyektif adalah tes yang kebenaran jawabannya sudah disediakan sehingga siswa tidak mendapat kesempatan untuk mengemukakan pendapat secara leluasan. Atau dengan kata lain hanya terdapat satu jawaban yang benar. Tes subyektif adalah tes yang memberi keleluasaan kepada siswa untuk mengemukakan pendapat, sehingga kemungkinan tidak hanya terdapat satu jawaban yang benar. Secara tradisional tes obyektif pada umumnya berbentuk soal benar-salah, soal pilihan ganda dengan empat alternatif jawaban, soal menjodohkan, dan jawaban singkat. Sedangkan tes subyektif pada umumnya adalah soal berbentuk essay.

Asmawi Zainul (2001, 11) mengemukakan berbagai jenis tes alternatif yang diharapkan mampu mengukur dan menilai prestasi belajar siswa secara lebih akurat. Jenis tes tersebut antara lain : tes pilihan ganda yang diperluas, *extended-response* atau *open ended question*, *group performance assessment*, *individual performance assesment*, *interview*, *nontraditional test items*, *observasi*, *portofolio*, *project*, *exhibition*, atau *demonstration*, dan *short answer*, *open ended*.

Tes pilihan ganda diperluas yaitu bentuk tes obyektif yang menuntut siswa tidak hanya menjawab benar, tetapi juga alasan mengapa milih jawaban tersebut. Sebagai contoh adalah soal obyektif yang berisi hubungan sebab-akibat, atau pilihan ganda dengan studi kasus. Soal-soal tersebut banyak terdapat pada tes Seleksi Mahasiswa Baru. *Extended-response* atau *open ended question* adalah tes yang tidak menuntut satu jawaban benar dengan satu pola. *Group performance assesment* adalah tugas-tugas yang menuntut kerjasama kelompok, *individual*

performance assesment adalah tugas-tugas yang bersifat individu. *Interview* adalah tes yang dilakukan asesor secara lisan, sehingga siswa harus menjawab pertanyaan-pertanyaan asesor dengan cara lisan. *Nontraditional test items* adalah tes yang tidak bersifat obyektif, tetapi merupakan seperangkat respon yang mengharuskan siswa memilih berdasarkan kriteria yang ditetapkan. *Observasi* adalah tes yang menuntut siswa melakukan tugas. Selama melakukan tugas tersebut, siswa diobservasi baik secara terbuka maupun tertutup. *Portofolio* adalah kumpulan hasil karya siswa yang disusun berdasarkan urutan waktu maupun urutan kategori kegiatan. *Project, exhibition, atau demonstration* adalah penyelesaian tugas-tugas yang kompleks dalam suatu jangka waktu tertentu yang dapat memperlihatkan penguasaan kemampuan sampai pada tingkatan tertentu pula. dan *Short answer, open ended* adalah tes yang menuntut jawaban singkat dari siswa, tetapi bukan memilih jawaban dari sederet kemungkinan jawaban yang telah disediakan.

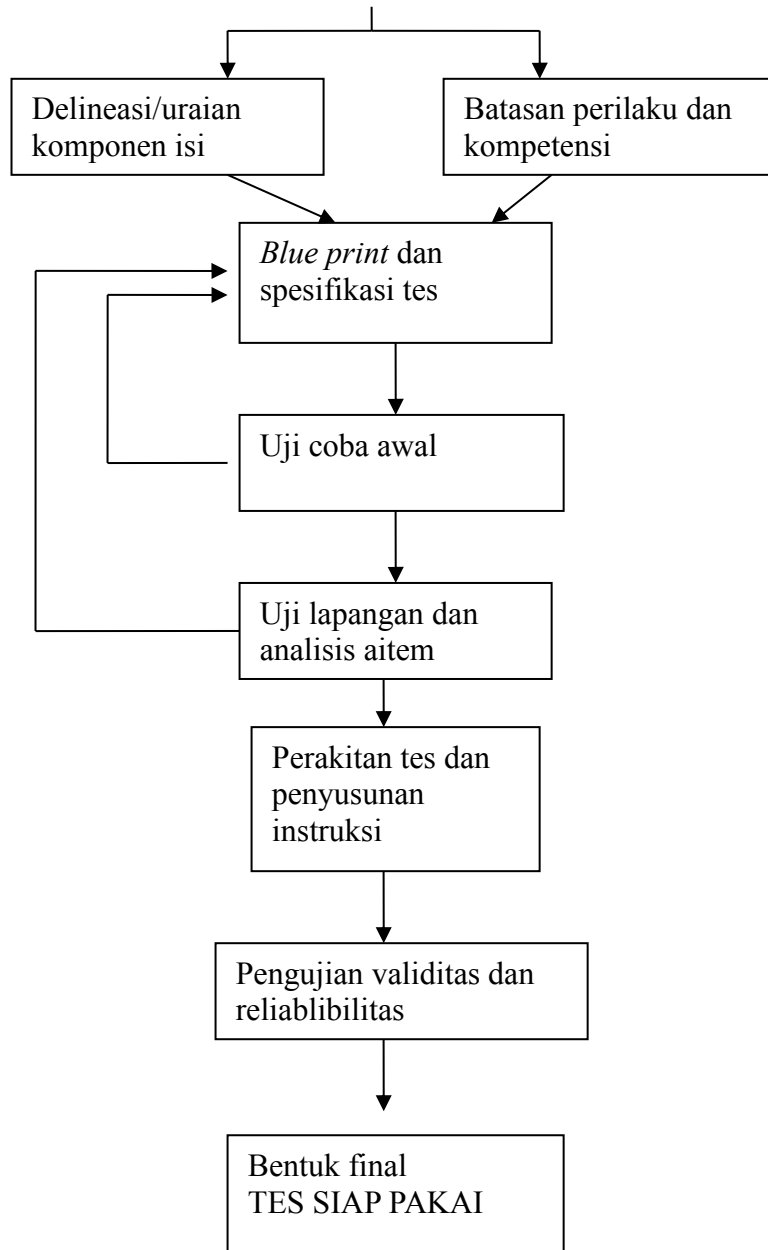
B. Prosedur Membuat Soal

Prosedur membuat soal meliputi langkah-langkah : identifikasi tujuan dan kawasan ukur, delineasi/ uraian komponen isi, batasan perilaku dan kompetensi, membuat *blue print* dan spesifikasi tes, penulisan aitem/soal dan review aitem, uji coba awal, uji lapangan dan analisis aitem, perakitan tes dan penyusunan instruksi, pengujian validitas dan reliabilitas, dan bentuk final, artinya tes siap digunakan. Berikut ini skema urutan pembuatan soal.

Skema 1. Prosedur Pembuatan Soal (dikembangkan dari saiffuddin Azwar

Identifikasi tujuan dan Kawasan ukur

1997 : 54)



1. Identifikasi Tujuan dan Kawasan Ukur

Identifikasi tujuan merupakan penegasan tujuan pengukuran yang akan dicapai dan pembatasan kawasan ukur. Isi dari identifikasi tujuan tes tergantung pada fungsi tes. Sebagai contoh bagi tes prestasi yang dijadikan dasar penempatan, tujuan tes harus mencakup sampel perilaku yang luas yang mewakili indikator kemampuan yang disyaratkan. Jenis soal pre-tes, kawasannya cenderung mengacu pada silabi yang akan diajarkan pada suatu mata pelajaran tertentu. Sedangkan jenis formatif yang tujuannya untuk memberi umpan balik, cakupan soalnya meliputi semua unit pembelajaran yang telah diajarkan.

2. Penguraian Isi

Penguraian isi perlu dilakukan untuk menghindari soal keluar dari lingkup materi yang telah diajarkan, selain itu penguraian isi juga berfungsi untuk menghindari materi penting yang terlewatkan. Penguraian isi dapat didasarkan pada bab-bab yang ada dalam buku pelajaran, atau didasarkan pada standar kompetensi dan kompetensi dasar sebagaimana dirancang dalam silabi.

Setelah pengelompokan materi selesai ditetapkan, dilakukan pembobotan untuk setiap sub bab, agar bagian-bagian yang dianggap lebih penting mendapat proporsi yang lebih banyak. Sebagai contoh seorang guru ingin menguji tingkat apresiasi musik siswanya dengan pembobotan masing-masing sub pokok bahasan sebagai berikut :

Tabel 1. Contoh pembobotan materi tes apresiasi musik

Uraian	Isi
bobot (%)	
Pendahuluan	
10%	
Karakter bunyi dan	Musik
30 %	
Medium	Musikal
25%	
Elemen	Musik
35%	
Total	
100%	

3. Batasan Perilaku dan Kompetensi

Batasan perilaku merupakan operasionalisasi tujuan pembelajaran atau kompetensi dasar yang hendak dicapai dalam proses pembelajaran. Indikator perilaku harus dapat diukur. Dasar penjabaran perilaku adalah aspek-aspek penilaian prestasi belajar yang meliputi aspek kognitif, psikomotor, dan afektif, sebagaimana dijelaskan pada Bab II. Pembuat soal atau guru dapat menentukan aspek dan jenjang apa saja yang hendak dinilai.

4. Tabel Spesifikasi Tes

Tabel spesifikasi tes adalah tabel yang memuat uraian isi tes, dan tingkat kompetensi yang akan diungkap. Tabel yang demikian disebut *blue print*, atau

kisi-kisi. Berikut ini contoh *Blue Print* tes dalam bidang apresiasi musik.

Tabel 2. Contoh Kisi-kisi Tes Apresiasi Musik

Komponen Uraian Isi	Komponen Perilaku						Jumlah soal
	C1	C2	C3	C4	C5	C6	
Pendahuluan	5	5	-	-	-	-	10
Karakter Bunyi dan Musik	5	5	5	10	5	-	30
Medium musikal	5	5	5	5	5	-	25
Elemen Musik	5	5	5	5	5	10	35
Jumlah soal	20	20	15	20	15	10	100

Kriteria banyaknya soal tiap jenjang tergantung pada perkembangan mental dan keterampilan pelajar. Bagi siswa jenjang sekolah dasar awal sebaiknya soal tes lebih banyak mengukur kemampuan C1 dan C2. Bagi siswa SD kelas 4,5, dan 6 lebih ditekankan pada soal jenjang C2, dan C3. Siswa SMP sebaiknya sudah dikenalkan dengan soal C4, dan siswa SMA diarahkan pada soal-soal C4, dan C5. Untuk tingkat mahasiswa soal-soal yang tepat adalah C5 dan C6. Namun kriteria tersebut bersifat fleksibel, artinya soal-soal dapat disusun tidak harus berdasarkan kriteria jenjang tersebut, tetapi lebih didasarkan pada kebutuhan.

Soal aspek psikomotor dikembangkan berdasarkan spesifikasi dalam bentuk rubrik, yaitu kriteria komponen penilaian. Berikut ini contoh rubrik penilaian materi pelajaran direksi.

Tabel 3. Contoh Rubrik Penilaian Materi Pelajaran Direksi

No.	Indikator	SK	K	C	B	SB	Bobot	Keterangan
		1	2	3	4	5		
1.	Sikap							
2.	Attack							
3.	Pola aba-aba							
4.	Tempo							
5.	Isyarat							
6.	Komunikasi							
7.	Ekspresi							
8.	Release							

Keterangan :

SK = sangat kurang

K = kurang

C = cukup

B = baik

SB = sangat baik

Spesifikasi untuk aspek afektif berupa penjabaran nilai dalam komponen-komponen. Bloom, Madaus, dan Hasting (1981 : 301) memberikan contoh spesifikasi kecenderungan siswa dalam suatu aspek sebagai berikut :

No.	Indikator	SK	K	C	B	SB	Bobot	Keterangan
		1	2	3	4	5		
1.	Kepuasan							
2.	Keinginan untuk meningkatkan apresiasi							
3.	Keinginan untuk mengetahui lebih banyak							
4.	Inisiatif untuk berkreasi							
5.	Keinginan untuk mengevaluasi							

5. Memilih Tes dan Menulis Soal

Penentuan bentuk soal disesuaikan dengan tujuan pembelajaran dan situasi serta kondisi yang ada. Tes yang paling ideal adalah tes yang memungkinkan bagi guru untuk mengetahui prestasi belajar siswa secara menyeluruh. Namun pada umumnya guru tidak mempunyai waktu yang cukup untuk mengetahui prestasi siswa, sehingga guru hanya mendapat kesempatan untuk memilih soal-soal yang cukup mewakili suatu kemampuan. Bagi guru yang tidak mempunyai waktu banyak untuk mengoreksi hasil belajar siswa, sebaiknya menggunakan tes obyektif, namun apabila guru mempunyai kesempatan untuk mengoreksi dalam waktu yang relatif lama dianjurkan untuk memilih tes subyektif atau essay untuk aspek kognitif. Hanya kelemahan tes subyektif adalah guru tidak mempunyai kesempatan untuk mengetahui kemampuan siswa secara komprehensif karena jumlah soal subyektif terbatas. Aspek psikomotor dapat diukur dengan tes observasi, yaitu siswa diminta untuk menampilkan keterampilannya, sedangkan guru mengamati kemampuan siswa dan memberikan skor sesuai dengan rubrik penilaian yang sudah tersedia.

Guru dapat juga memilih portofolio sebagai salah satu alternatif penilaian, untuk mengetahui perkembangan siswa. Portofolio ini juga dapat mendorong siswa untuk mengerjakan tugas yang diberikan guru secara lebih serius karena proses pembelajaran juga diperhitungkan dalam pemberian nilai.

5. Uji Coba Awal

Soal yang baik adalah soal yang mudah dipahami maksudnya oleh siswa dan tidak menimbulkan penafsiran yang salah. Untuk itu soal harus ditulis dengan petunjuk dan jelas dengan bahasa yang baik dan benar. Oleh karena itu

sebelum soal diberikan kepada siswa sebaiknya diujicobakan terlebih dahulu sehingga diketahui soal-soal yang tidak memenuhi syarat dan harus dibuang.

6. Uji Coba Lapangan

Setelah soal diuji coba awal, soal di uji coba di lapangan, untuk mengetahui indeks daya beda, indeks kesukaran dan analisis butir soalnya dalam hubungannya dengan skor total. Dengan diketahui indeks kesukaran, daya beda dan koefisien korelasinya, suatu soal dapat diketahui apakah layak digunakan sebagai alat ukur atau tidak. Uji coba lapangan ini sekaligus dapat dilakukan signifikansi validitas dan reliabilitas suatu tes.

C. Pertanyaan dan Tugas

1. Buatlah kisi-kisi tes kognitif berdasarkan tujuan yang telah diidentifikasi !
2. Buatlah 50 soal tes obyektif berdasarkan kisi-kisi yang dibuat pada nomor 1 !
2. Buatlah rubrik penilaian aspek psikomotor sesuai dengan Praktik Instrumen Mayor anda termasuk pembobotannya !

BAB IV

VALIDITAS DAN RELIABILITAS TES

A. Validitas

Sebagaimana telah dikemukakan pada bagian pendahuluan, bahwa alat ukur atau instrumen pengukuran yang digunakan untuk mengukur prestasi belajar harus valid dan variabel, dengan demikian kita perlu mempunyai seperangkat soal test atau alat ukur yang mempunyai tingkat validitas dan reliabilitas yang memadai agar kita memperoleh data yang tepat tentang prestasi belajar siswa.

Validitas berasal dari kata *validity* yang mempunyai arti sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya (Saifuddin Azwar,1977:5). Suatu tes atau insrtumen pengukuran mempunyai validitas yang tinggi apabila dapat memberikan hasil yang sesuai dengan maksud pengukuran tersebut. Sebagai contoh seorang guru ingin mengetahui kemampuan siswanya dalam membaca notasi angka. Guru tersebut memberikan soal-soal yang diberikan berisi melodi yang belum dikenal siswanya, maka soal tersebut dapat dikatakan valid, karena benar-benar mengukur kemampuan siswa membaca notasi angka, bukan karena mereka sudah mengenal dan hafal melodi yang diujikan.

Valid juga berarti cermat, artinya alat ukur tersebut bukan hanya tepat, tetapi juga cermat, artinya alat ukur tersebut mampu memberikan gambaran yang cermat terhadap suatu data. Sebagai contoh untuk mengukur berat emas akan lebih tepat apabila menggunakan neraca dari pada timbangan, karena dengan neraca dapat mengukur berat emas sampai berat yang sekecil-kecilnya, bahkan sampai ukuran

miligram.

Tipe validitas pada umumnya digolongkan pada tiga kategori, yaitu *content validity* (validitas isi), *construct validity* (validitas konstruk), dan *criterion-related validity* (validitas kriteria) (Saifuddin Anwar, 1997:45). Validitas konstruk dan validitas kriteria termasuk validitas empiri, karena tingkat validitasnya hanya dapat diketahui setelah dilakukan uji coba di lapangan. Validitas isi sangat tergantung pada penilaian subyektif individu, karena estimasi validitas ini hanya berdasarkan analisis rasional. Penilaian validitas isi dilakukan oleh ahli, sehingga keputusan valid atau tidaknya suatu tes atau alat ukur tergantung pada ahli yang bersangkutan, hal ini disebut *professional judgment*.

Fokus validitas adalah untuk menentukan apakah item-item dalam tes atau alat ukur sudah mencakup keseluruhan kawasan obyek yang hendak diukur atau mencerminkan ciri atribut yang hendak diukur. Dengan demikian tes atau alat ukur selain komprehensif juga harus relevan, tidak keluar dari batasan tujuan pengukuran.

Validitas isi terbagi dalam 2 tipe, yaitu *face validity* (validitas muka) dan *logical validity* (validitas logik) (Saifuddin Anwar, 1997:46). Validitas muka hanya dilihat berdasarkan format penampilan, sehingga signifikansinya rendah. Namun demikian format penampilan juga penting karena dapat mempengaruhi keseriusan *testee* (peserta tes) untuk menjawab pertanyaan atau merespon tugas.

Validitas logik adalah validitas yang didasarkan pada representasi ciri-ciri yang hendak diukur. Artinya validitas logik menunjuk pada tingkat terwakilinya ciri-ciri atribut yang hendak diukur. Sebagai contoh dalam soal teori musik pokok

bahasan unsur-unsur dasar musik, butir-butir soal harus mencakup unsur ritme, melodi, dan harmoni. Dalam praktik menyanyi, komponen yang diukur harus menggambarkan semua atribut menyanyi, yaitu kualitas suara, teknik, dan pembawaan, apabila penampilan dianggap penting, maka komponen penampilan juga dijadikan salah satu kriteria pengukuran.

Validitas logik dalam pembelajaran dapat pula dilihat berdasarkan kompetensi dasar dan standar kompetensi. Dengan demikian, pokok-pokok bahasan atau atribut-atribut yang diukur dijabarkan berdasarkan kurikulum berbasis kompetensi. Validitas konstruk adalah tipe validitas yang menunjukkan sejauh mana tes mengungkap suatu trait atau konstruk teoritik yang hendak diukur (Siafuddin Anwar, 1997:48). Untuk mengetahui koefisien validitas konstruk dapat menggunakan analisis faktor dengan bantuan program komputer SPSS.

1. Validitas Instrumen Pengukuran Aspek Kognitif

Instrumen pengukuran aspek kognitif pada umumnya berupa tes teori. Valid atau tidaknya suatu tes teori secara empiris juga dapat diketahui dengan menemukan indeks daya beda, taraf kesukaran, dan analisis butir soal. Indeks daya beda adalah kemampuan soal membedakan kemampuan siswa, indeks kesukaran berkaitan dengan tingkat kemudahan atau kesulitan soal bagi siswa, dan analisis butir soal adalah koefisien korelasi butir soal dengan skor total.

Suatu soal dikatakan layak digunakan apabila mempunyai daya beda positif, indeks kesukaran sedang (0,3 – 0,7), dan koefisien korelasi butir soal dengan skor total paling sedikit 0,4. Pada soal jenis pilihan ganda dengan beberapa alternatif jawaban, terdapat satu hal lagi yang perlu ditambahkan sebagai syarat

soal yang layak digunakan, yaitu jumlah *ommit* (soal yang tidak dijawab) tidak boleh lebih dari 5%, selain itu setiap alternatif jawaban harus dipilih siswa paling sedikit 2 %.

a. Penghitungan Indeks Daya Beda

Cara memperoleh indeks daya beda menggunakan rumus sebagai berikut :

Rumus 4.1

$$D = PH - PL$$

D adalah indeks daya beda

PH adalah prosentase jawaban betul pada siswa kelompok atas.

PL adalah prosentase jawaban betul pada siswa kelompok bawah.

b. Penghitungan Indeks Kesukaran

Adapun rumus untuk memperoleh indeks kesukaran adalah sebagai berikut :

Rumus 4.2

$$P = (PH + PL)/2$$

P adalah indeks kesukaran

PH adalah prosentase jawaban betul pada kelompok atas.

PL adalah prosentase jawaban betul pada kelompok bawah.

c. Penghitungan Analisis Butir Soal

Koefisien korelasi butir soal dengan skor total diperoleh dengan menggunakan rumus korelasi *Product Moment*

Rumus 4.3

$$r_{xy} = (N \sum xy - \sum x \sum y) / \sqrt { N \sum x^2 - (\sum x)^2 } \times \sqrt { N \sum y^2 - (\sum y)^2 }$$

d. Penghitungan *Ommit*

Adapun untuk memperoleh persentase *ommit* dilakukan dengan cara menghitung persentase jumlah soal yang tidak dijawab oleh siswa dengan rumus :

Rumus 4.4

$$(O : N) \times 100\%$$

O adalah jumlah *ommit* (siswa yang tidak menjawab soal).

N adalah jumlah siswa.

Sebagaimana telah dikemukakan pada bagian terdahulu bahwa suatu soal dikatakan baik apabila jumlah *ommit* tidak lebih dari 5 %.

Cara yang sama dilakukan untuk menentukan persentase jumlah *option* (alternatif jawaban pada soal pilihan ganda) dengan menjumlahkan siswa yang memilih alternatif jawaban a, b, c, atau d, dibagi dengan seluruh siswa dalam hal ini sampel yang mengikuti uji coba, kemudian dikali 100 %, dengan rumus sebagai berikut :

Rumus 4.5

$$(Op : N) \times 100 \%$$

Op adalah jumlah siswa yang memilih suatu alternatif jawaban

N adalah jumlah siswa

Suatu alternatif jawaban layak digunakan apabila paling sedikit terdapat 2 % siswa memilih alternatif jawaban tersebut.

Berikut ini contoh penghitungan indeks daya beda, indeks kesukaran, dan analisis butir soal.

Berdasarkan uji coba instrumen pengukuran prestasi belajar teori musik yang terdiri dari 10 soal terhadap 15 siswa diperoleh data sebagai berikut :

Tabel 4. Data Uji Coba Instrumen Pengukuran Prestasi Belajar Teori Musik

Soal N	1	2	3	4	5	6	7	8	9	10	Total
1.	1	1	1	0	0	0	0	1	0	0	4
2.	1	1	1	0	1	0	0	1	0	0	5
3.	1	1	1	1	0	1	1	1	1	1	9
4.	1	0	0	1	1	0	1	1	1	1	7
5.	1	1	1	0	1	0	1	1	0	0	6
6.	1	1	1	1	0	0	1	1	1	0	7
7.	1	1	1	1	1	0	1	1	0	0	7
8.	0	1	1	1	1	0	1	1	1	1	8
9.	1	1	1	0	0	0	1	1	1	1	7
10.	1	1	1	0	0	0	0	1	0	0	4
11.	0	1	1	1	1	0	1	1	1	0	7
12.	1	1	1	1	1	1	1	1	1	0	9
13.	0	0	0	1	1	1	1	0	1	1	6
14.	1	1	1	0	0	0	1	1	1	0	6
15.	1	1	1	1	1	0	1	1	1	1	9
Σ	12	13	13	9	9	3	12	14	10	6	101

Keterangan :

N adalah siswa

0 adalah jawaban salah

1 adalah jawaban betul

Σ adalah jumlah

e. Menghitung Indeks Kesukaran dan Indeks Daya Beda

Sebelum kita mengoperasikan Rumus 4.1 dan 4.2 terlebih dahulu kita mengelompokkan siswa pada kelompok atas dan kelompok bawah. Apabila jumlah sampel (dalam hal ini siswa) relatif banyak, yaitu 30 atau lebih, maka kelompok atas dapat diambil dari 30 % nilai siswa teratas, dan kelompok bawah

adalah 30 % siswa yang mendapat nilai terendah. Pada jumlah sampel yang jumlahnya relatif sedikit kita dapat membagi sampel menjadi dua kelompok, yaitu kelompok atas dan kelompok bawah berdasarkan urutan ranking nilai yang diperoleh.

Berikut ini adalah contoh pengelompokan kelompok atas dan kelompok bawah. Untuk memudahkan contoh perhitungan sampel 15 **dianggap** banyak. Siswa yang termasuk kelompok atas adalah 30 % dari 15 siswa yang memperoleh nilai tertinggi, yaitu berjumlah 5 orang, sedangkan siswa yang termasuk kelompok bawah adalah 5 orang yang mendapat nilai terendah.

Berdasarkan Tabel 4, diperoleh data nilai sebagai berikut :

Tabel 5. Data Urutan Sekor dari Sekor Rendah ke Sekor Tinggi

Siswa	1	10	2	5	13	14	4	6	7	9	11	8	3	12	15
Sekor	4	4	5	6	6	6	7	7	7	7	7	8	9	9	9

Berdasarkan tabel no 5, maka siswa yang termasuk kelompok bawah adalah siswa nomor 1, 10,2,5, dan 13. siswa no 13 dan 14 mendapat nilai yang sama, yaitu 6, tetapi karena kita hanya memerlukan lima sampel, maka kita boleh memilih antara siswa no 13 dan 14. pada perhitungan ini kita memutuskan nomor 13 sebagai sampel kelompok bawah. Hal yang sama dilakukan pada penentuan sampel siswa pada kelompok atas. Siswa yang tergolong kelompok atas adalah nomor 11, 8, 3, 12, dan 15.

Pada perhitungan selanjutnya kita hanya memperhatikan data siswa yang menjadi sampel. Untuk sementara data siswa nomor 14, 4, 6, 7, dan 9 kita

abaikan.

1) Penghitungan Indeks Daya Beda dan Indeks Kesukaran Soal Nomor 1

Indeks daya beda dan indeks kesukaran dapat diperoleh dengan menghitung PH dan PL terlebih dahulu. PH adalah persentase jawaban benar pada kelompok atas, sedangkan PL adalah persentase jawaban benar pada kelompok bawah. Berdasarkan Tabel 4, diperoleh data sebagai berikut :

Tabel 6. Data Jawaban Siswa Kelompok Atas

Soal N	1	2	3	4	5	6	7	8	9	10	Total
3.	1	1	1	1	0	1	1	1	1	1	9
8.	0	1	1	1	1	0	1	1	1	1	8
11.	0	1	1	1	1	0	1	1	1	0	7
12.	1	1	1	1	1	1	1	1	1	0	9
15.	1	1	1	1	1	0	1	1	1	1	9

Tabel 7. Data Jawaban Siswa Kelompok Bawah

Soal N	1	2	3	4	5	6	7	8	9	10	Total
1.	1	1	1	0	0	0	0	1	0	0	4
2.	1	1	1	0	1	0	0	1	0	0	5
5.	1	1	1	0	1	0	1	1	0	0	6
10.	1	1	1	0	0	0	0	1	0	0	4
13.	0	0	0	1	1	1	1	0	1	1	6

Berdasarkan data pada Tabel 6 diperoleh informasi bahwa jumlah siswa yang menjawab benar soal nomor satu pada kelompok atas adalah 3 orang, dengan demikian $PH = (3/5) \times 100\% = 60\%$ atau 0,6, sedangkan pada kelompok bawah terdapat empat siswa yang menjawab benar, dengan demikian $PL = (4/5) \times 100\% = 80\%$ atau 0,8.

Dengan demikian indeks daya beda soal nomor 1 adalah

$$D = PH - PL = 0,6 - 0,8 = -0,2$$

Indeks kesukaran soal nomor 1 adalah

$$P = (PH + PL) / 2 = (0,6 + 0,8) / 2 = 0,7$$

Berdasarkan perhitungan tersebut maka ditinjau dari indeks kesukaran soal tersebut termasuk baik karena indeks kesukaran berada antara $0,3 \leq$ dan $0,7 \geq$, namun ditinjau dari indeks daya beda soal tersebut tidak dapat membedakan anak yang pandai dan yang tidak, karena indeks daya beda negatif, berarti anak-anak yang pandai justru tidak dapat menjawab soal, sedangkan anak-anak yang tidak pandai dapat menjawab soal.

2) Penghitungan Indeks Daya Beda dan Indeks Kesukaran Soal nomor 4

Sekarang kita lakukan analisis soal nomor 4. Berdasarkan Tabel 6 dan Tabel 7 diperoleh informasi bahwa semua siswa kelompok atas dapat menjawab soal nomor 4, dengan demikian PH soal nomor 4 adalah $(5/5) \times 100\% = 100\%$ atau 1, sedangkan pada kelompok bawah hanya satu orang yang menjawab benar, dengan demikian PL adalah $(1/5) \times 100\% = 20\%$ atau 0,2. Dengan demikian indeks daya beda soal nomor 4 adalah

$$D = PH - PL = 1 - 0,2 = 0,8$$

$$P = (PH + PL) / 2 = (1 + 0,2) / 2 = 0,6$$

Ditinjau dari indeks daya beda, soal tersebut dapat membedakan siswa yang pandai dan yang tidak karena hasilnya positif, selain itu soal tersebut juga berada dalam kategori indeks kesukaran sedang karena berada di antara $0,3 \leq$ dan $0,7 \geq$, dengan demikian soal tersebut dinyatakan layak digunakan bila ditinjau dari indeks daya beda dan indeks kesukaran.

Hal yang sama dilakukan pada seluruh soal. Setelah diperoleh soal-soal yang layak digunakan, soal-soal tersebut perlu dianalisis lagi dengan mengkorelasikan skor butir soal dengan skor total.

3) Menghitung Koefisien Analisis Butir Soal

Kali ini kita akan menghitung koefisien korelasi butir soal nomor 4, karena berdasarkan perhitungan indeks daya beda dan indeks kesukaran soal tersebut memenuhi syarat. Langkah pertama menghitung x^2 , y^2 , dan xy dengan membuat tabel sebagai berikut :

Tabel 8. Data Soal nomor 4

N	x	y	x^2	y^2	Xy
1.	0	4	0	16	0
2.	0	5	0	25	0
3.	1	9	1	81	9
4.	1	7	1	49	7
5.	0	6	0	36	0
6.	1	7	1	49	7
7.	1	7	1	49	7
8.	1	8	1	64	8
9.	0	7	0	49	0
10.	0	4	0	16	0
11.	1	7	1	49	7
12.	1	9	1	81	9
13.	1	6	1	36	6
14.	0	6	0	36	0
15.	1	9	1	81	9
Σ	9	101	9	717	69

Berdasarkan tabel tersebut diperoleh data sebagai berikut :

$$N = 15$$

$$\Sigma x = 9 \quad \Sigma y = 101$$

$$\sum x^2 = 9 \sum y^2 = 717 \quad \sum xy = 69$$

Berdasarkan rumus 4.3, maka komputasi koefisien korelasi butir soal nomor 4 dengan skor total adalah sebagai berikut :

$$\begin{aligned} r_{xy} &= (N \sum xy - \sum x \sum y) / \sqrt{\{N \sum x^2 - (\sum x)^2\} \times \sqrt{\{N \sum y^2 - (\sum y)^2\}}} \\ r_{xy} &= (15 \times 69) - (9 \times 101) / \sqrt{\{15 \times 9 - (9)^2\} \times \sqrt{\{15 \times 717 - (101)^2\}}} \\ &= 1035 - 909 / \sqrt{(135 - 81) \times \sqrt{(10755 - 10201)}} \\ &= 126 / (7,3485 \times 23,5372) \\ &= 126 / 172,9631 = 0,729 \end{aligned}$$

Ternyata koefisien korelasi butir soal nomor 4 dengan skor total adalah 0,73, karena lebih dari 0,4, maka soal nomor 4 dinyatakan memenuhi syarat. Dengan demikian berdasarkan D, P, dan analisis butir soal, soal nomor empat dapat digunakan.

Cara yang sama dilakukan pada semua soal. Soal yang memenuhi syarat dinyatakan valid dan dapat digunakan sebagai alat untuk mengukur prestasi belajar.

2. Validitas Instrumen Pengukuran Aspek Psikomotor

Alat ukur atau instrumen pengukuran aspek psikomotor pada umumnya berupa lembar observasi yang berisi komponen-komponen yang akan diukur. Validitas konstruk alat ukur tersebut secara empiris dapat diketahui dengan menggunakan analisis faktor yang dapat dihitung dengan bantuan program SPSS.

Suatu atribut dikatakan valid apabila mempunyai korelasi tinggi karena berarti mempunyai sifat-sifat yang sama dengan komponen pengukuran secara keseluruhan. Dengan demikian suatu atribut yang mempunyai korelasi rendah

atau bahkan negatif tidak termasuk dalam atribut komponen yang diukur karena tidak mempunyai sifat-sifat yang sesuai dengan komponen secara keseluruhan, sehingga komponen tersebut harus digugurkan.

3. Validitas Instrumen Pengukuran Aspek Afektif

Pada umumnya pengukuran aspek afektif menggunakan instrumen berupa angket dalam bentuk skala Likert, yaitu soal yang berisi pertanyaan-pertanyaan dengan lima *option* (alternatif jawaban), sangat setuju (SS), Setuju (S), tidak tahu (TT), tidak setuju (TS), dan, sangat tidak setuju (STS). Angket tersebut tidak mengukur benar salah, tetapi mengukur kecenderungan sikap seseorang, dengan demikian hasil pengukurannya berupa sikap positif atau negatif. Karena tidak menyangkut benar salah, maka aspek afektif tersebut biasanya digunakan sebagai syarat diperbolehkannya seorang siswa mengikuti ujian akhir. Sebagai contoh seorang mahasiswa diperbolehkan mengikuti ujian dengan syarat telah mengikuti perkuliahan minimal 75% selama satu semester, seorang siswa diperbolehkan mengikuti pelajaran apabila mengenakan baju rapi sesuai dengan peraturan sekolah.

Validitas aspek afektif dapat diketahui dengan cara mengkorelasikan butir soal dengan skor total, dengan menggunakan **rumus 4.3** analisis butir soal pada aspek kognitif.

Seperti halnya pada aspek kognitif, validitas alat ukur aspek afektif signifikansinya ditentukan oleh koefisien korelasi. Suatu butir soal dikatakan valid apabila mempunyai koefisien korelasi 0,4 atau lebih.

B. Reliabilitas

Reliabilitas merupakan penerjemahan dari kata *reliability* yang mempunyai asal kata *rely* dan *ability* (Saifuddin Azwar, 1997 : 4). *Rely* berarti terpercaya dan *ability* berarti kemampuan. Instrumen yang reliabel berarti instrumen atau alat ukur tersebut terpercaya. Reliabel artinya ketetapan, keajegan, kestabilan, konsistensi, kepercayaan, atau keterandalan. Seberapa besar reliabilitas suatu tes, tergantung pada kemampuannya untuk mengukur obyek yang sama pada saat dan tempat yang berbeda dengan hasil yang sama.

1. Reliabilitas Pengukuran Aspek Kognitif

Signifikansi reliabilitas pada soal kognitif dapat diketahui dengan tiga pendekatan, yaitu tes ulang, bentuk paralel, dan konsistensi internal.

a. Pendekatan Test Ulang

Pendekatan tes ulang dilakukan dengan melakukan tes dua kali pada siswa yang sama. Jarak waktu antara tes pertama dan tes kedua sebaiknya tidak terlalu lama, tetapi juga tidak terlalu singkat. Apabila tenggang waktu antara tes pertama dan kedua terlalu lama, karena siswa sudah mengalami perkembangan karena pembelajaran atau kematangan, terdapat kemungkinan hasilnya berubah. Namun sebaliknya apabila tenggang waktu antara test pertama dan test kedua terlalu singkat, juga tidak dapat memberikan kondisi yang relatif sama, karena terdapat kemungkinan siswa dapat merespon secara lebih baik karena soal-soal tes masih segar dalam ingatan siswa, tetapi dapat juga terjadi sebaliknya, siswa menjadi tidak serius dalam mengerjakan soal, karena pernah diuji dengan soal yang sama. Penghitungan koefisien reliabilitas dapat menggunakan rumus korelasi *product moment* sebagaimana pada rumus 4.3.

b. Pendekatan Bentuk Pararel

Pendekatan bentuk paralel dengan memberikan sekaligus dua bentuk tes yang paralel satu sama lain, kepada sekelompok subjek (Saifuddin Anwar,1977;59). Kedua tes tersebut digabungkan terlebih dahulu sebelum diberikan pada siswa, setelah selesai dijawab item pada masing-masing tes dipisah kembali menjadi dua tes, sehingga diperoleh dua distribusi skor. Kedua distribusi tersebut kemudian dikorelasikan dengan komputasi korelasi *product moment*. Penghitungan koefisien *korelasi product moment* dapat dilakukan dengan bantuan program SPSS.

c. Pendekatan Konsistensi Internal

Pendekatan Konsistensi Internal prosedurnya hanya memerlukan satu kali penguasaan tes. Hal ini dimaksudkan untuk mengatasi kelemahan pendekatan tes-ulang (*test-retest*) dan bentuk paralel. Penguasaan tes hanya satu kali akan menghasilkan satu distribusi skor tes. Analisis reliabilitasnya diarahkan pada analisis butir soal, atau kelompok soal, dengan cara melakukan pembelahan soal. Prosedur pembelahan tergantung pada sifat dan fungsi tes.

Tes yang mengandung soal-soal tes yang bersifat homogen relatif lebih mudah prosedur pembelahannya karena tidak harus memasangkan tes yang mempunyai bobot yang sama sebagaimana pada *power test* (test yang mempunyai tujuan untuk mengetahui kemampuan optimal siswa).

Pembelahan dapat dilakukan dengan beberapa cara antara lain : pembelahan cara random, pembelahan gasal-genap, dan pembelahan *Matched Random Subsets* (Saifuddin Anwar, 1997:67). Pembelahan cara random dapat

dilakukan dengan pengundian. Sebagai contoh terdapat seperangkat tes berisi 50 soal, maka untuk menentukan apakah suatu soal termasuk belahan kedua atau pertama dilakukan dengan cara undian. Prosedur pembelahan gasal-genap dilakukan dengan mengelompokkan soal-soal bernomor ganjil pada belahan pertama, dan soal-soal bernomor genap pada belahan kedua. Pembelalahan *Matched Random subsets*, digunakan apabila indeks daya beda dan indeks kesukaran dijadikan pertimbangan dalam pembelahan soal.

Pada kajian ini juga akan disajikan cara penghitungan reliabilitas dengan pendekatan belah dua gasal genap.

Tabel 9. Distribusi Sekor Hasil Uji Coba Tes

Siswa	Nomor Item												To Tal	0	E	OE
	1	2	3	4	5	6	7	8	9	10	11	12				
A	1	1	1	1	1	1	1	1	1	1	1	1	12	6	6	36
B	1	1	1	1	1	1	1	1	0	1	0	1	10	4	6	24
C	1	1	1	1	1	1	1	0	1	1	1	0	10	6	4	24
D	1	1	1	1	1	1	0	0	0	0	0	0	6	3	3	9
E	1	1	1	1	1	1	1	1	1	0	0	0	9	5	4	20
F	1	1	1	1	1	0	0	0	0	0	0	0	5	3	2	6
G	1	1	1	1	0	0	1	0	1	0	0	0	6	4	2	8
H	1	1	1	0	0	0	0	0	0	0	0	0	3	2	1	1
I	1	1	0	0	0	0	0	0	0	0	0	0	2	1	1	1
J	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
To Tal													64	35	24	129

Keterangan :

1 adalah jawaban benar

0 adalah jawaban salah

O (odd) adalah jawaban benar pada nomor genap

E(Even) adalah Jawaban Benar pada nomor gasal

Berdasarkan tabel no 9, dikembangkan tabel 10 yang berisi data perhitungan reliabilitas.

Tabel 10. Data Perhitungan Reliabilitas

N	X	O	E	O²	E²	OE
A	12	6	6	36	36	36
B	10	4	6	16	36	24
C	10	6	4	36	16	24
D	6	3	3	9	9	9
E	9	5	4	25	16	20
F	5	3	2	9	4	6
G	6	4	2	16	4	8
H	3	2	1	4	1	1
I	2	1	1	1	1	1
J	1	1	0	1	0	0
	64	35	24	153	123	129

Langkah selanjutnya adalah menghitung koefisien korelasi dengan rumus 4.3. Dengan ketentuan simbol O kemudian dilambangkan dengan X dan E dilambangkan dengan Y. Berdasarkan perhitungan korelasi diperoleh koefisien korelasi 0,827. langkah selanjutnya koefisien korelasi tersebut diberi simbol r_{xx} , dan kita perlu menghitung koefisien estimasi dengan rumus :

Rumus 4.6

$$r_{xy} = (2 X r_{xx}) / 1 + r_{xx}$$

$$= (2 X 0,827) / 1 + 0,827 = 0,905$$

2. Reliabilitas Aspek Psikomotor

Koefisien reliabilitas untuk aspek psikomotor, biasanya menggunakan instrumen pengukuran yang berupa lembar observasi. Setelah syarat validitas terpenuhi, perlu dilakukan analisis signifikansi reliabilitasnya. Cara untuk mengetahui reliabilitas suatu test psikomotor terdapat beberapa cara, antara lain

dengan rumus cohen-kappa dan interater reliabilitas. Berikut akan disajikan contoh penggunaan *interater reliability*.

Rumus 4.7 Interater reliability {Saifuddin Azwar,1997;108 }

$$r_{xx} = (S_s^2 - S_e^2) / S_s$$

$$S_e^2 = \{ \sum i^2 - (\sum R^2) / n - (\sum T^2) / k + (\sum i^2) / nk \} : (n - 1) (k - 1)$$

$$S_s^2 = \{ (\sum T^2) / k - (\sum i^2) / nk \} : n - 1$$

Keterangan

N adalah jumlah siswa

k adalah jumlah rater (penguji)

Tabel 11. Hasil Rating 3 Rater terhadap 10 Subyek

Subyek	Rater			T	T ²
	I	II	III		
A	3	3	5	11	121
B	4	5	6	15	225
C	9	7	11	27	729
D	2	2	3	7	49
E	7	4	10	21	441
F	10	9	12	31	961
G	6	6	7	19	361
H	1	1	4	6	36
I	8	10	9	27	729
J	5	8	8	21	441
R	55	55	75		
R ²	3025	3025	5625		

Berdasarkan tabel 11 diperoleh data sebagai berikut :

$$N = 10 \quad ; k = 3$$

$$\sum R = \sum T = \sum I = 55 + 55 + 75 = 185$$

$$\sum R^2 = 3025 + 3025 + 5625 = 11675$$

$$\sum T^2 = 121 + 225 + 729 \dots\dots\dots + 441 = 4093$$

$$\sum i^2 = 3^2 + 4^2 + 9^2 + \dots + 8^2 = 1415$$

Berdasarkan data tersebut diperoleh

$$Se^2 = \{1415 - 11675/10 - 4093/3 + 185^2/(10)(3)\} : (10 - 1)(3 - 1)$$

$$Se^2 = 1,333$$

$$Ss^2 = \{4093/3 - 185^2/(10)(3)\} : \{10 - 1\} = 24,833$$

$$r_{xx} = (24,833 - 1,333)/24,833 = 0,946$$

Contoh penghitungan interrater reliability dengan program software SPSS dapat dibaca pada lampiran 1.

3. Kriteria reliabilitas

Kriteria reliabel atau tidaknya suatu tes atau instrumen pengukuran tergantung pada tingkat koefisien korelasi. Pada umumnya kriterianya adalah sebagai berikut :

0	-	0,2	sangat rendah
0,2	-	0,4	rendah
0,4	-	0,6	cukup
0,6	-	0,8	tinggi
0,8	-	1	sangat tinggi

Dengan demikian instrumen pengukuran pada sedangkan instrumen psikomotor tersebut adalah sangat tinggi.

C. Pertanyaan dan Tugas

1. Uji cobakan soal aspek kognitif dan psikomotor yang telah dibuat pada soal bab III pada sampel yang sesuai dengan soal yang anda buat !
2. Hitunglah Indeks Daya Beda, Indeks Kesukaran, Analisis butir soal pada soal-soal aspek kognitif.
3. Berdasarkan perhitungan pada soal no 2, pilihlah soal-soal yang layak digunakan.
4. Hitunglah reliabilitas soal aspek kognitif pada soal nomor 1.
5. Uji cobakan lembar observasi aspek psikomotor.
6. Bagaimanakah reliabilitas instrumen pengukuran pada soal nomor 5 ?

BAB V

SISTEM PENILAIAN

Sebagaimana telah di kemukakan pada bagian pendahuluan, evaluasi merupakan kegiatan yang meliputi dua tahap yaitu pengukuran dan penilaian. Ketepatan hasil Pengukuran sangat tergantung pada validitas dan reliabilitas alat ukur atau instrumen pengukuran.

Setelah kita membahas aspek-aspek pengukuran, bentuk dan jenis tes, serta validitas dan reliabilitas instrumen pengukuran, berikut ini akan dikemukakan tentang cara menilai prestasi dengan sistem penilaian acuan patokan (PAP) dan Sistem penilaian norma (PAN).

Soal tes atau alat ukur yang valid atau reliabel merupakan alat yang dipercaya mampu mengukur prestasi belajar secara tepat. Data prestasi belajar yang diperoleh lewat soal test atau instrumen pengukuran perlu diinterpretasikan. Berikut ini akan dijelaskan prosedur penilaian setelah tersedia soal test atau instrumen pengukuran yang valid dan reliabel.

A. Skoring soal subyektif (esai)

Kelebihan soal subyektif atau *essay* adalah guru mampu menguji kemampuan optimal siswa, karena siswa dapat dengan leluasa mengemukakan pendapatnya. Namun kelemahannya, dengan soal esai sulit menentukan kriteria penilaian secara tepat. Untuk itu guru harus menentukan komponen jawaban apa saja yang perlu dinilai dalam tiap soal, sehingga setiap siswa mendapat jaminan bahwa mereka akan mendapat perlakuan penilaian yang sama sesuai dengan kualitas jawaban mereka.

Sebagai contoh pada soal tentang macam-macam birama, sebelum guru memberikan skor jawaban, guru harus terlebih dahulu memberi aturan skoring, misalnya apabila seorang siswa mampu menjelaskan :

4 macam atau lebih macam birama, sekornya 10

3 macam birama, maka sekornya 7,5

2 macam birama, maka sekornya 5

1 macam birama, maka sekornya 2,5

jawaban salah, maka sekornya 0

Adapun rumus pemberian Sekor total adalah :

Rumus 5.8

Sekor Total = (sekor yang diperoleh siswa : sekor maksimal) X 100 %

Sebagai contoh seorang siswa menjawab lima soal essay dengan skor tiap soal sebagai berikut :

Soal	1,	sekor	8
Soal	2,	“	5
Soal	3,	“	10
Soal	4,	“	7
Soal	5,	“	4

Maka skor totalnya adalah sebagai berikut :

Diketahui :

Sekor yang diperoleh siswa adalah $8 + 5 + 10 + 7 + 4 = 34$

Sekor maksimal adalah 50 (diperoleh dari 5 soal, kali 10)

Dengan demikian skor total adalah $(34 : 50) \times 100 \% = 68$

Nilai akhir diperoleh dengan rumus :

Rumus 5.9

$NA = (\text{sekor total} : 100) \times \text{nilai maksimal}$

NA = nilai akhir

Apabila nilai maksimal adalah 100, maka nilai siswa tersebut adalah

$$(68 : 100) \times 100 = 68$$

Apabila nilai maksimal adalah 10, maka nilai siswa tersebut adalah

$$(68 : 100) 10 = 6,8$$

B. Skoring Soal Test Obyektif

Kelebihan soal tes obyektif adalah mampu mengukur prestasi belajar dengan pokok bahasan yang lebih banyak dibanding tes obyektif, karena dengan tes obyektif memungkinkan siswa untuk menjawab 50 soal dalam satu kali tes, sehingga guru atau pembuat soal memungkinkan menanyakan berbagai hal, sedangkan dalam soal essay hal ini tidak mungkin dilakukan, karena dengan waktu 60 menit paling banyak siswa dapat menjawab 10 soal secara sempurna. Kelebihan lain dari test obyektif dibandingkan test subyektif adalah, guru dapat memberikan skoring secara lebih obyektif, karena perbedaan jawaban benar dan jawaban salah sangat jelas.

Kelemahan tes obyektif adalah terdapat kemungkinan siswa dapat menjawab soal karena terkaan. Untuk mengatasi jawaban benar karena kebetulan, maka aturan skoringnya adalah, jawaban salah mengurangi jawaban betul, dengan demikian maka siswa lebih berhati-bati dalam menjawab soal. Adapun skoring

soal obyektif menggunakan rumus sebagai berikut :

Rumus 5.10

$$\text{Sekor Total} = (B - S/a-1)$$

B adalah jumlah jawaban betul

S adalah jumlah jawaban salah

a adalah jumlah alternatif jawaban.

Contoh :

Seorang siswa menjawab semua soal pilihan ganda, dengan 4 alternatif jawaban sebanyak 50 soal. Siswa tersebut berhasil menjawab 41 soal dengan betul dan salah 9. dengan demikian sekor totalnya adalah :

$$41 - (9/4-1) = 32$$

Setelah diketahui sekor total yang diraih siswa, guru perlu mengkonversikannya dalam bentuk nilai dengan rumus 5.11 sebagai berikut :

Rumus 5.11

$$NA = (\text{sekor yang diperoleh/sekor maksimal}) \times \text{nilai maksimal}$$

NA adalah Nilai Akhir

Dengan demikian, apabila nilai maksimal 100 nilai yang diperoleh siswa adalah :

$$(32/50) \times 100 = 64$$

Skoring untuk bentuk soal obyektif Benar-Salah, menjodohkan, pilihan ganda dengan sejumlah alternatif jawaban yang lain juga mengikuti rumus 5.11 tersebut.

Sebagai contoh, dalam soal benar-salah berjumlah 20, seorang siswa

menjawab 18, betul 15, maka skor totalnya adalah :

$$B - S (a-1) = 15 - (3/2-1) = 15 - 3 = 12$$

Apabila nilai maksimal 10, maka nilainya dicari dengan rumus 5.11. menghasilkan nilai $(12/20) \times 10 = 6$

Adakalanya seorang pembuat tes atau guru membuat soal terdiri dari beberapa bentuk tes dengan bobot yang berbeda. Nilai akhir juga harus memperhitungkan masing-masing bobot bentuk soal. Sebagai contoh guru membuat tes terdiri dari tiga romawi dengan perincian, 10 soal benar-salah, 20 soal pilihan ganda, dan 10 soal menjodohkan dengan 13 pilihan jawaban. Bobot soal benar-salah 1, pilihan ganda 2, dan soal menjodohkan 2.

Seorang siswa mengerjakan semua soal benar-salah betul 8, mengerjakan 17 soal pilihan ganda betul 15, dan mengerjakan semua soal menjodohkan betul 7. apabila nilai maksimal 100, maka nilai akhir siswa tersebut sebagai berikut :

Tahap I Menentukan Skor tiap romawi.

$$\text{Skor romawi I} = 8 - (2/2-1) = 7$$

$$\text{Skor romawi II} = 15 - 3(4-1) = 15 - 9 = 6$$

$$\text{Skor Romawi III} = 7 - (3/13 -1) = 7 - 3/12 = 6,75$$

Tahap II Menentukan Nilai tiap romawi.

$$\text{Nilai romawi I} = 7/10 \times 100 = 70$$

$$\text{Skor romawi II} = 6/20 \times 100 = 30$$

$$\text{Skor Romawi III} = 6,75/10 \times 100 = 67,5$$

Tahap III Menentukan nilai akhir berdasarkan bobot

Perbandingan bobot nilai tiap romawi adalah 1 : 2 : 2, maka nilai tiap romawi dikalikan bobotnya masing-masing. Dengan demikian nilai masing-masing romawi adalah :

$$\text{Romawi I} = 70 \times 1 = 70$$

$$\text{Romawi II} = 70 \times 2 = 140$$

$$\text{Romawi III} = 67,5 \times 2 = 135$$

Nilai akhir siswa adalah :

Rumus 5.12

Nilai Akhir = Jumlah nilai/bobot

Dengan demikian nilai siswa tersebut adalah $345 / 5 = 69$.

C. Sistem Penilaian PAP

Pemberian nilai adalah proses penerjemahan hasil sekor tes. Sebagaimana dikemukakan pada bagian pendahuluan bahwa dalam menginterpretasikan hasil tes terdapat dua pendekatan yaitu pendekatan PAP (Penilaian Acuan Patokan) dan PAN (Penilaian Acuan Norma). Penilaian Acuan Patokan didasarkan pada kriteria yang sudah ditentukan. Pada umumnya setiap institusi sudah mempunyai patokan penilaian. Sebagai contoh :

Nilai	Konversi	Predikat /keputusan
85 – 100	A	Lulus/Dengan pujian
80 – 84	A-	Lulus/sangat memuaskan
75 – 79	B+	Lulus/memuaskan
70 – 74	B	Lulus/memuaskan
66 – 69	B-	Lulus/memuaskan
56 – 65	C	Lulus/cukup
0 - 55	D	Tidak lulus

Dengan demikian nilai 69 yang diperoleh siswa pada sub bab sebelum ini mempunyai makna bahwa siswa tersebut lulus dengan predikat memuaskan.

D. Sistem Penilaian Acuan Norma

Penilaian berdasarkan acuan norma bersifat relatif, artinya sekor yang diperoleh tidak secara otomatis dapat diketahui predikat yang diperoleh siswa karena tergantung pada nilai yang dicapai siswa dalam suatu kelompok. Sistem penilaian ini didasarkan pada kurva normal. Kriteria penentuan lulus dan tidaknya siswa ditentukan oleh nilai rata-rata dan standar deviasi. Adapun sistem penilaian didasarkan pada kriteria sebagai berikut :

- $M + 1,5 SD \leq x$ adalah A
 - $M + 0,5 SD \leq x < M + 1,5 SD$ adalah B
 - $M - 0,5 SD \leq x < M + 0,5 SD$ adalah C
 - $M - 1,5 SD \leq x < M - 0,5 SD$ adalah D
 - $M - 1,5 SD \geq x$ adalah E
- SD adalah standar deviasi.

Contoh, distribusi nilai hasil ujian teori musik adalah sebagai berikut :

Siswa	Sekor
A	50
B	60
C	70
D	55
E	35
F	40
G	30
H	45
I	55
J	50
K	45
L	35
M	30
N	60
O	60

Berdasarkan data tersebut diperoleh perhitungan :

Mean (rata-rata) = 48

SD = 12,218 (diperoleh dari perhitungan calculator 3600)

Dengan demikian kriteria penilaiannya adalah sebagai berikut :

66,34 \leq x adalah A

54,11 – 66,33 adalah B

41,89 – 54,1 adalah C

29,68 – 41,88 adalah D

0 – 29,67 adalah E

Berdasarkan kriteria tersebut maka siswa – siswa tersebut memperoleh nilai sebagai berikut :

Siswa	Sekor	Nilai
A	50	C
B	60	B
C	70	A
D	55	B
E	35	D
F	40	D
G	30	D
H	45	C
I	55	B
J	50	C
K	45	C
L	35	D
M	30	D
N	60	B
O	60	B

E. Pertanyaan dan Tugas

1. Berilah laporan tentang sekor yang diperoleh siswa berdasarkan soal yang sudah anda buat pada bab sebelumnya.
2. Berilah laporan nilai yang diperoleh siswa(pada soal nomor1) apabila

menggunakan PAP.

3. Berilah laporan nilai yang diperoleh siswa(pada soal nomor1) apabila menggunakan PAN.
4. Menurut anda kapan saat yang tepat anda menggunakan PAP, dan kapan anda memilih PAN adalah sistem penilaian anda ! Jelaskan.

DAFTAR PUSTAKA

- Asmawi Zainul. (2001). *Alternative Assesment*. Jakarta : Proyek Pengembangan Universitas Terbuka Direktorat Jenderal Pendidikan Tinggi Departemen Pendidikan Nasional.
- Depdiknas. (2003). *Kurikulum 2004 – Standar Kompetensi Mata Pelajaran Kerajinan Tangan dan Kesenian. Sekolah Dasar dan Madrasah Ibtidaiyah*. Jakarta : Depdiknas.
- Fernandes, HJX.(1984). *Evaluation of educational programs*. Jakarta : National Education Planning Evaluation and Curriculum Development.
- Isaac, Stephen & Michael, William B. (1981). *Hanbook in research and evaluation for education and the behavioral sciences*. California : Hange
- Joesmani. (1988). *Panduan Pengajar Buku Pengukuran dan Evaluasi dalam Pengukuran*. Jakarta : Depdikbud.
- Kun S.Astuti. (2001). Efektivitas pertunjukan sebagai fokus pembelajaran ansambel musik untuk menghasilkan prestasi belajar musik yang penuh makna. Tesis master, tidak diterbitkan, Pascasarjana UNY Yogyakarta.
- Mursell, James. (1982). *Pengajaran berhasil*. Jakarta : UI.
- Sirait, Bistok. (1989). *Bahan pengajaran untuk mata kuliah evaluasi hasil belajar siswa*. Jakarta : Depdikbud P2LPTK.
- Worthen, B.R.& Sanders,J.R.(1973) *Educational evaluation : Theory and practice*. Belmont, CA: Wadsworth Publishing Co.

Lampiran 1

Cara penghitungan interrater reliability dengan SPSS

Langkah 1 buka program SPSS

Langkah 2 masukkan data

Langkah 3 tekan Variabel view

Langkah 4 tuliskan sifat data, seperti rater 1,2, dan 3, type, numeric,

Langkah 4 tekan scale

Langkah 5 pilih reliability

Langkah 6 tekan okey/enter

