



EVALUASI

PENDIDIKAN JASMANI

OLAHRAGA

**DAN
KESEHATAN**

Drs. Ngatman, M.Pd



PENERBIT CV. SARNU UNTUNG

Penulis:

Drs. Ngatman, M.Pd

ISBN : 978-602-61658-5-5

Desain sampul dan tata letak:
Yahya Abdullah

Ukuran Buku 23 x 15,5 cm

Penerbit:

CV. Sarwa Untung

Redaksi:

Jalan R. Supripto, Gg. Pringgondani, RT 07, RW 21,

Purwodadi, Grobogan, Jawa Tengah, 58111

No. HP 085726280111

Email: nlceng87@yahoo.co.id

Anggota IKAPI (No. 146/JTE/2015)

Cetakan pertama, Mei 2017

Hak cipta dilindungi undang-undang

Dilarang memperbanyak karya tulis ini dalam bentuk dan

dengan cara

Apapun tanpa ijin tertulis dari penerbit

KATA PENGANTAR

Kegiatan evaluasi atau penilaian merupakan hal yang tidak dapat dipisahkan dalam pembelajaran pendidikan jasmani di sekolah. Kegiatan tersebut bermakna ganda baik bagi guru maupun bagi peserta didik. Bagi guru, hasil evaluasi akan menjadi sumber informasi mengenai kemampuan awal dan kemampuan akhir peserta didik, serta peningkatan kemampuan peserta didik yang pada akhirnya menjadi bahan bagi guru dalam mengevaluasi program pembelajaran yang telah dilaksanakannya. Sedangkan bagi peserta didik sendiri, hasil evaluasi menyediakan informasi bagi peserta didik mengenai tingkat pengetahuan, kecakapan/keterampilannya, dan posisi keterampilannya dibandingkan dengan teman sebayanya sehingga dapat menjadi bahan motivasi dan evaluasi diri (*self evaluator*).

Evaluasi merupakan salah satu elemen yang esensial dalam pendidikan jasmani. Para calon guru dan guru pendidikan jasmani perlu memahami dengan seksama mengenai prosedur penyusunan alat evaluasi yang baik, bagaimana cara mendesain, menyusun, dan melakukan penilaian terhadap ranah-ranah PJOK, serta menganalisis nilai hasil belajar siswa. Dengan bekal pengetahuan dan pemahaman terhadap hal-hal tersebut diharapkan penilaian yang dilakukan oleh guru PJOK menjadi lebih berkualitas sehingga guru PJOK dapat menilai



hasil belajar siswa dilakukan dengan akurat (valid),
ajeg (reliabel), dan objektif.

Dengan diselesaikannya buku evaluasi pendidikan jasmani olahraga dan kesehatan yang terdiri dari 11 bab ini diharapkan dapat memperkaya literasi dalam evaluasi pendidikan jasmani dan olahraga. Sajian substansi yang disusun secara lengkap, detail, dan cermat ditunjukkan supaya buku ini dapat membawa manfaat optimal bagi penggunaannya. Di sekolah, nama buku mata pelajaran pendidikan jasmani saat ini yaitu Pendidikan Jasmani, Olahraga, dan Kesehatan (PJOK). Mempertimbangkan kemudahan dan keefektifan dalam penulisan dan tanpa bermaksud mengurangi makna, maka istilah PJOK dalam buku ini ditulis sebagai pendidikan jasmani olahraga dan kesehatan atau penjasorkes. Akhir kata, semoga penulisan buku ini membawa manfaat nyata dalam dunia pendidikan jasmani olahraga dan kesehatan tanah air.

Yogyakarta, 15 April 2017

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
KATA PENGANTAR.....	iii
DAFTAR ISI.....	v
BAB I. SIFAT DAN RUANG LINGKUP PROGRAM PENGUKURAN, EVALUASI PJOK.....	1
BAB II. KARAKTERISTIK ALAT EVALUASI RANAH PSIKOMOTOR DALAM PJOK.....	16
BAB III. KRITERIA-KRITERIA ALAT EVALUASI PJOK.....	44
BAB IV. PROSEDUR PENYUSUNAN TES PJOK.....	61
BAB V. PENDEKATAN ACUAN PENILAIAN DALAM PJOK.....	70
BAB VI. SISTEM EVALUASI DALAM PJOK.....	81
BAB VII. PENYUSUNAN TES BENTUK OBJEKTIF UNTUK EVALUASI RANAH KOGNITIF DALAM PJOK.....	106
BAB VIII. PENYUSUNAN TES BENTUK ESAI (URAIAN) UNTUK EVALUASI RANAH KOGNITIF DALAM PJOK.....	130
BAB IX. PENGUKURAN DAN PENILAIAN RANAH AFEKTIF.....	138
BAB X. PENILAIAN ALTERNATIF (ALTERNATIVE ASSESSMENT) DALAM PJOK.....	173
BAB XI. PENILAIAN OTENTIK (AUTHENTIC ASSESSMENT) DALAM PJOK.....	203
DAFTAR PUSTAKA.....	222



BAB I

SIFAT DAN RUANG LINGKUP PROGRAM PENGUKURAN, EVALUASI PENDIDIKAN JASMANI OLAHRAGA DAN KESEHATAN (PJOK)

A. ARTI PENGUKURAN DAN PENILAIAN

Dua istilah ini mempunyai hubungan yang sangat erat sekali sehingga penggunaannya sering dicampur adukkan. Namun demikian, antara keduanya sebenarnya terdapat perbedaan-perbedaan. Dalam pendidikan termasuk juga pendidikan jasmani olahraga dan kesehatan (PJOK) kepada anak-didik sering diberikan bermacam-macam tes oleh guru. Semua itu dilakukan dengan tujuan untuk mengetahui seberapa jauh hasil dan proses pembelajaran PJOK yang telah direncanakan oleh guru secara saksama telah dicapai.

Evaluasi dapat diberi batasan sebagai satu proses menilai keefektifan tujuan-tujuan pendidikan telah dicapai. Evaluasi dalam pendidikan dapat dilakukan terhadap macam-macam aspek pendidikan, dan bukan hanya terhadap anak-anak saja. Menurut Bovard, Cozens dan Hagma proses evaluasi mencakup tiga langkah. **Langkah pertama** adalah menentukan dan menilai (*appraise*) tujuan-tujuan. Evaluasi mempunyai dasar anggapan bahwa tujuan-tujuan mengenai pendidikan telah dipahami sungguh-sungguh, karena prinsip dasar dari evaluasi adalah dilakukan terhadap tujuan-tujuan yang ingin dicapai. Selanjutnya evaluasi beranggapan bahwa tujuan-tujuan itu harus bernilai. Hanya mempunyai tujuan-tujuan saja belumlah cukup, nilai dari tujuan itu perlu diyakini terlebih dahulu.

Langkah kedua adalah mengumpulkan data atau informasi. Dalam proses evaluasi semua prosedur dipergunakan baik prosedur kuantitatif maupun kualitatif. Prosedur tersebut bermanfaat untuk mengumpulkan data yang diperlukan untuk menilai seberapa jauh tujuan-tujuan pendidikan tersebut dicapai. **Langkah ketiga** adalah mempertimbangkan arti dari pendidikan ditinjau dari sudut tujuan-tujuan yang ingin dicapai dan dari keterangan-keterangan dan data yang dikumpulkan.

Dari uraian di atas, pengukuran dapat dibatasi dari sudut langkah kedua dari proses evaluasi, yaitu prosedur evaluasi untuk mengumpulkan data. Pengukuran itu berkenaan dengan prosedur-prosedur evaluasi yang bersifat tepat, objektif, kuantitatif dan

hasil-hasilnya dapat diolah secara statistik. Hasil pengukuran yang bersifat kuantitatif itu belumlah mempunyai arti apa-apa. Hasil pengukuran baru mempunyai arti hanya setelah diadakan interpretasi dari semua data yang dikumpulkan. Pengukuran menggunakan tes-tes standard, skala prestasi, skala *rating*, kartu-kartu skor dan alat-alat pengukur lainnya yang menghasilkan data yang dapat dicatat dalam detik, jarak, angka atau simbol-simbol kuantitatif lainnya.

Prosedur evaluasi kualitatif dapat berbentuk observasi dan *rating*, wawancara, daftar *check* dan teknik-teknik lainnya dengan mempergunakan pertimbangan subjektif. Prosedur evaluasi ini bertujuan untuk mengumpulkan keterangan mengenai faset-faset tingkah laku manusia atau organisasi sosial, yang tidak dapat dianalisa secara objektif.

Jadi semua prosedur evaluasi yang bertujuan untuk menentukan jumlah atau kuantitas adalah pengukuran. Sedangkan semua prosedur yang bertujuan untuk menentukan mutu atau kualitas adalah termasuk evaluasi atau penilaian.

Larson dan Yocum membedakan pengukuran dan penilaian sebagai berikut; **Pengukuran** menentukan sifat-sifat dan kemampuan dari individu atau kelompok. Sifat-sifat dan kemampuan itu telah ditetapkan dalam tujuan-tujuan pendidikan jasmani olahraga dan kesehatan (PJOK). Karena itu pengukuran itu berkenaan dengan hasil (*product*) dari aktivitas pendidikan. **Evaluasi** menentukan bagaimana tujuan-tujuan telah dicapai. Informasi ini meliputi sifat kepemilikan, materi program, fasilitas, alat-alat, pemberian nilai, penelitian, waktu untuk program, partisipasi dan administrasi. Karena itu evaluasi berkenaan dengan proses aktivitas-aktivitas pendidikan.

B. PENTINGNYA EVALUASI

Pada hakikatnya pelaksanaan pendidikan sama dengan pelaksanaan perusahaan yang cukup besar, yang mempergunakan sejumlah besar pegawai. Apabila perusahaan itu ingin mengetahui maju mundurnya perusahaan diperlukan penilaian terhadap hasil-hasil yang telah dicapai. Langkah yang dikembangkan oleh perusahaan pada prinsipnya sama dengan langkah yang ditempuh oleh instansi lain terlebih lagi dalam dunia pendidikan.

Bagaimana seorang pendidik dapat mengetahui apakah pelajaran yang diberikan kepada siswa-siswinya itu sudah dipahami, kalau tidak mengadakan penilaian terhadap hasil-hasil belajar mereka. Bagaimana seorang pendidik dapat mengetahui, apakah si A atau si B itu maju dalam belajarnya, kalau ia tidak mengadakan pencatatan-pencatatan

serta penilaian terhadap hasil-hasil belajar si A dan si B. Demikian pula bagaimana seorang guru dapat mengetahui bagian-bagian pelajaran manakah yang dianggap sukar oleh para anak-didik, jika guru tersebut tidak mengadakan pencatatan-pencatatan dan penilaian secara teliti terhadap hasil-hasil yang dicapai oleh mereka ?

Bagaimana seorang kepala sekolah dapat mengetahui apakah Rencana Pendidikan sudah dapat dilaksanakan dengan baik, jika ia tidak mengadakan pemeriksaan serta penilaian terhadap hal-hal yang telah dapat dilakukan? Bagaimana ia dapat mengetahui hal-hal apakah yang menjadi penghambat atau penghalang, dan bagaimana kiranya kemungkinan-kemungkinan untuk mengatasinya, apabila ia tidak menilai dan mempelajari usaha-usaha serta hasil-hasil yang telah dicapai? Untuk itu semua diperlukan penilaian atau evaluasi dalam pendidikan.

C. OBJEK-OBJEK EVALUASI

Ruang lingkup pendidikan cukup luas, meliputi banyak segi dan menyangkut banyak pihak. Segi-segi pendidikan di antaranya adalah :

1. Segi tujuan pendidikan, yaitu ke arah mana anak-didik itu dibawa
2. Segi materi atau bahan, yang disusun menjadi kurikulum bagi satu jenis dan tingkatan sekolah tertentu.
3. Segi metodologi, yang diperinci dalam metodologi umum dan metodologi khusus bagi mata-mata pelajaran tertentu.
4. Segi teknik administratif, yang menyangkut hal-hal ketatausahaan, misalnya; mengenai surat menyurat, gaji guru dan karyawan, kenaikan tingkat, dsb
5. Segi teknik edukatif, yang menyangkut penentuan policy dalam pendidikan serta pengawasan terhadap pelaksanaannya.

Sedangkan pihak-pihak yang bersangkutan dalam pelaksanaan pendidikan ialah:

1. Pihak anak-didik, yaitu pihak yang menjadi sasaran atau objek utama dalam pendidikan
2. Pihak pendidik, yaitu pihak yang menjadi subjek dalam pendidikan. Pihak yang secara langsung berusaha untuk membawa anak-didik ke arah tujuan pendidikan.
3. Pihak orang-tua, yaitu pihak yang mempercayakan anak-nya untuk dididik oleh para pendidik.

4. Pihak masyarakat, yaitu pihak yang turut serta berpengaruh terhadap berhasilnya pendidikan.
5. Pihak tokoh-tokoh pendidikan, yang secara tidak langsung berhubungan dengan anak-didik namun mereka mempunyai pengaruh terhadap pelaksanaan kegiatan-kegiatan pendidikan.

Terhadap semua segi dan pihak diatas seharusnya diadakan evaluasi, misalnya;

1. Apakah tujuan pendidikan itu sudah sesuai dengan cita-cita orang tua, bangsa dan negara ?
2. Apakah kurikulum sekolah itu sudah sesuai dengan kebutuhan masyarakat dalam masa pembangunan ini ?
3. Apakah metode yang dipergunakan adalah metode mengajar yang *up to date* ?
4. Apakah pekerjaan-pekerjaan administrasi telah dikerjakan dengan sebaik-baiknya ?
5. Apakah para pemimpin pendidikan telah melaksanakan kewajiban mereka dengan penuh pengabdian?
6. Bagaimanakah hasil-hasil pendidikan yang telah dicapai oleh anak-didik ?
7. Bagaimana partisipasi orang tua terhadap usaha-usaha memajukan pendidikan ?
8. Bagaimanakah partisipasi masyarakat dalam usaha-usaha memajukan pendidikan untuk membina moral masyarakat yang baik ?

Uraian selanjutnya tidak akan membahas semua masalah yang dikemukakan di atas. Pembicaraan akan dibatasi hanya pada masalah penilaian terhadap anak-didik sebagai objek utama dalam pendidikan. Penilaian terhadap hasil belajarnya yang di arahkan kepada tujuan-tujuan pendidikan olahraga dan kesehatan (PJOK).

D. TUJUAN-TUJUAN EVALUASI

Dalam pembicaraan mengenai pentingnya evaluasi sebenarnya telah disinggung pula tentang apa tujuan mengadakan evaluasi. Kiranya tidak ada salahnya apabila masalah tujuan ini dibicarakan dalam bagian tersendiri. Seperti telah diuraikan di depan, seseorang melakukan suatu pekerjaan ia mesti mempunyai tujuan. terlebih lagi dalam melaksanakan penilaian. Dalam mengadakan penilaian terhadap anak-didik pendidik mempunyai banyak tujuan baik tujuan itu bersifat langsung maupun tidak langsung.

Adapun tujuan-tujuan evaluasi terhadap anak-didik yang penting di antaranya adalah:

1. Untuk mengetahui sampai dimana potensi anak-didik itu berada. Apakah ia memperoleh kemajuan dalam berolahraga atau tidak. Jika ada kemajuannya sampai dimanakah atau berapakah kemajuan tersebut. Sebaliknya jika terdapat kemunduran, haruslah dicari apakah yang menyebabkannya. Selanjutnya diusahakan untuk menghilangkan atau setidaknya-tidaknya mengurangi hambatan-hambatan tersebut.
2. Untuk mengadakan seleksi. Misalnya seleksi untuk menentukan siapa saja yang akan dipilih sebagai anggota tim bola voli sekolah. Seleksi terhadap calon-calon siswa atau mahasiswa untuk masuk satu sekolah atau perguruan tinggi.
3. Untuk mengetahui apa yang telah dicapai anak-didik dalam pelajaran olahraga. Jadi untuk mengetahui prestasi atau pengetahuan peserta didik.
4. Untuk mengetahui kelemahan-kelemahan atau kesulitan-kesulitan yang dialami anak-didik. Kesulitan itu baik yang bersifat umum maupun yang bersifat perseorangan. Dengan diketahui letak kesulitan-kesulitan ini, maka dapat diberikan bantuan yang lebih tepat, baik bersifat klasikal maupun yang bersifat perseorangan.
5. Untuk memberi bantuan dalam pengelompokan anak-didik untuk tujuan-tujuan tertentu. Misalnya pengelompokan diadakan untuk bermain bola voli agar kedua regu yang bertanding kira-kira sama kuatnya.
6. Dapat memberi dorongan atau motivasi bagi anak-didik dalam berolahraga. Dengan mengetahui hasil-hasil yang dicapainya, hal ini dapat menjadi cambuk baginya untuk berusaha lebih giat lagi. Yang mendapat hasil kurang memuaskan berusaha untuk memperbaikinya, dan yang berhasil baik, berusaha untuk mempertahankannya.
7. Dapat memberikan bantuan dalam bimbingan ke arah pemilihan yang sesuai dengan bakat dan kemampuan anak-didik. Pemilihan yang tepat akan lebih menguntungkan bagi anak-didik terutama dalam memilih satu profesi.
8. Memberikan data bukti untuk dilaporkan kepada orang tua dan juga kepada masyarakat yaitu pihak-pihak yang memerlukan keterangan-keterangan tentang seorang anak-didik. Laporan itu dapat berbentuk surat-surat keterangan, sertifikat, raport, tanda tamat belajar, ijazah dan lain-lain.
9. Dapat memberikan data-data untuk keperluan penelitian atau riset.

E. PRINSIP-PRINSIP EVALUASI

Dalam pelaksanaan evaluasi dibedakan dua macam prinsip, yaitu:

1. **Prinsip-prinsip pelaksanaan**, yaitu prinsip-prinsip yang perlu diikuti dalam melakukan atau mengadakan evaluasi
2. **Prinsip-prinsip dasar**, yaitu prinsip-prinsip yang harus dipegang sebagai pedoman kerja dalam pelaksanaan evaluasi

Adapun **prinsip-prinsip pelaksanaan** yang penting ialah :

- a. Evaluasi harus dilakukan secara **objektif**. Yang dimaksud secara objektif disini ialah bahwa evaluasi itu harus berdasarkan bukti-bukti yang nyata. Tidak dapat dibenarkan memberikan evaluasi hanya berdasarkan ingatan saja. Evaluasi yang diberikan harus berdasarkan testing-testing yang diadakan. Juga berdasarkan catatan-catatan yang secara cermat dilakukan, yang merupakan hasil-hasil pengamatan dalam jangka waktu yang cukup. Berdasarkan daftar *check* atau skala rating yang diberikan dan sebagainya. Memang dalam praktik sering diadakan “pertimbangan” terhadap hasil-hasil penilaian yang dilakukan. Tetapi ini tidak berarti, bahwa membolehkan bekerja dengan semauanya saja. Pertimbangan-pertimbangan boleh diadakan, tetapi itupun harus berdasarkan pada bukti-bukti yang jelas dan tegas.
- b. Evaluasi harus diadakan secara **kontinu**. Artinya harus diadakan evaluasi terhadap anak-didik secara terus menerus, tanpa putus-putusnya. Tidak perlu menunggu-nunggu sampai satu bulan, dua bulan atau lebih untuk mengadakan evaluasi. Sejak anak diserahkan untuk menjadi anak-didik, sejak itu pula penilaian diadakan, sampai saat anak itu sudah tidak lagi menjadi anak-didik. Evaluasi itu dapat berupa aktivitas-aktivitas sehari-hari, ulangan-ulangan, baik ujian-ujian kwartalan atau ujian kenaikan tingkat atau ujian akhir. Untuk hal-hal yang berhubungan dengan kepribadian (kelakuan, kemauan, kerajinan dll, nya) harus diadakan terus-menerus pengamatan atau observasi yang disertai dengan catatan-catatan yang teliti dan teratur.
- c. Evaluasi hendaknya dilakukan secara **komprensif**. Artinya bahwa evaluasi itu sejauh mungkin harus menilai semua aspek dari keseluruhan kepribadian anak-didik. Evaluasi terhadap anak-didik harus mencakup hal-hal yang berhubungan

dengan kemampuan dan pengetahuan dalam pelajaran olahraga, kelakuannya, kerajinannya, sikap-sikap sosialnya, kegiatannya dalam mengikuti pelajaran, kegiatannya dalam mengikuti aktivitas sosial baik dalam maupun diluar sekolah dan lain-lain lagi.

Prinsip-prinsip dasar yang penting sebagai pedoman kerja dalam melaksanakan evaluasi ialah :

- a. Evaluasi adalah alat komunikasi. Komunikasi antar sekolah dengan sekolah, antara sekolah dengan orang tua, antara sekolah dengan masyarakat.

Komunikasi antar sekolah dengan sekolah, misalnya untuk keperluan-keperluan melanjutkan sekolah. Untuk pindah dari sekolah yang satu kesekolah yang lain. Untuk ini laporan penilaian dari sekolah yang lama sangat diperlukan oleh sekolah yang akan menerima anak-didik tersebut.

Komunikasi antara sekolah dengan orang tua. Ini adalah laporan tanda bukti dari sekolah kepada orang tua anak-didik, bahwa sekolah benar-benar telah menjalankan tugasnya untuk mendidik anak yang telah diserahkan kepadanya.

Komunikasi antara sekolah dengan masyarakat, misalnya pada waktu anak-didik mencari pekerjaan. Pihak-pihak dalam masyarakat yang akan mempergunakan tenaga-tenaga pasti membutuhkan keterangan-keterangan yang menyangkut pendidikannya. Oleh karena evaluasi itu merupakan alat komunikasi untuk berbagai pihak maka evaluasi itu harus “mempunyai arti (*meaningfull*)” bagi semua pihak yang bersangkutan . oleh karena itu laporan hasil penilaian harus lebih terperinci dan meliputi seluruh segi.

- b. Maksud utama dari evaluasi terhadap anak-didik adalah membantu anak-didik untuk mencapai perkembangan potensinya semaksimal mungkin. Ini tidak berarti bahwa diinginkan membuat anak-didik menjadi pandai semua, dan pandai dalam segala hal, tanpa mengingat batas-batas pembawaan kemampuan setiap anak-didik. Diakui adanya batas-batas pembawaan kemampuan bagi setiap anak-didik. Ada anak yang berpembawaan kemampuan lemah, dan adapula yang kuat atau pembawaan kemampuannya sedikit, sedang dan banyak. Dalam batas-batas kemampuannya dan dalam bidang-bidang sesuai dengan bata inilaberusaha anak-didik agar dapat mencapai batas-batas tersebut dengan baik. Jadi untuk semua jenis

tingkatan batas-batas kemampuan itu pendidik harus berusaha agar anak-didik dapat mencapai batas-batas perkembangannya.

- c. Dalam evaluasi anak-didik terutama harus dibimbing dengan dirinya sendiri, jangan hanya dibandingkan dengan anak didik lainnya. Dengan dibandingkan dengan dirinya sendiri pendidikan akan mendapatkan gambaran yang jelas tentang kemajuan anak-didik atau kemundurannya. Misalnya si A pada suatu saat mendapat skor 80 untuk kesegaran jasmaninya. Jika dibandingkan dengan teman-temannya dalam kelas, si A tetap menduduki tempat tertinggi, sebab tidak ada di antara teman-temannya yang mendapatkan skor 80. Jika dibandingkan dengan dirinya sendiri sebenarnya ia telah mengalami kemunduran. Dalam tes kesegaran jasmani terakhir ia mendapat skor 80.
- d. Dalam mengadakan evaluasi hendaknya dimanfaatkan berbagai macam alat atau teknik evaluasi. Dengan demikian diperoleh gambaran atau kesimpulan yang lebih dapat dipercaya. Misalnya dalam evaluasi hasil-hasil belajar anak-didik untuk berbagai macam pelajaran. Dengan cara ini pendidik dapat mengadakan penilaian melalui pekerjaan-pekerjaan harian, tes-tes atau ulangan-ulangan baik secara tertulis, lisan maupun perasaan. Baik ulangan-ulangan atau tes-tes dengan pemberitahuan sebelumnya maupun tanpa pemberitahuan. Dengan demikian hasil-hasil itu dapat melukiskan kemampuan yang sebenarnya dari seorang siswa.
- e. Evaluasi setidaknya menyarankan langkah-langkah atau tindakan lanjut yang perlu diambil. Misalnya saja saran-saran didalam pemilihan jurusan pendidikan.

F. KEMUNGKINAN-KEMUNGKINAN KESALAHAN DALAM EVALUASI PJOK

Karena evaluasi penting sekali dalam pendidikan perlu di usahakan agar kemungkinan-kemungkinan kesalahan dalam evaluasi sekecil mungkin. Adapun hal-hal yang memungkinkan terjadinya kesalahan-kesalahan itu di antaranya ialah :

1. Kesalahan dalam pengamatan atau observasi. Misalnya saja, observasi itu kurang dilakukan secara teliti. Observasi itu kurang menyeluruh terhadap aspek-aspeknya. Baru satu atau dua kali mengadakan observasi sudah mengambil kesimpulan. Cara atau teknik melakukan observasi tidak tepat.
2. Kesalahan-kesalahan pada alat pengukur, misalnya tes yang dipergunakan itu tidak memenuhi syarat-syarat pembuatan tes yang baik (mengenai reliabilitasnya,

validitasnya, dan sebagainya). Soal-soal yang dibuat oleh pendidik kurang memperhatikan petunjuk-petunjuk dalam penyusunannya.

3. Kesalahan-kesalahan dalam prose analisis data. Misalnya kesalahan-kesalahan dalam perhitungan-perhitungannya seperti kesalahan menjumlah, mengalikan dan sebagainya.
4. Pengaruh dari pekerjaan-pekerjaan yang mendahului. Misalnya kalau pekerjaan yang dulu baik, maka pendidik dipengaruhi untuk memberikan angka yang baik pula. Begitu pula sebaliknya.
5. Kecendrungan dari seorang pendidik untuk menilai lebih rendah atau lebih tinggi. Misalnya pendidik A mempunyai kecenderungan untuk memberikan angka-angka yang rendah seperti 4, 5, 6, 7 jarang sekali memberikan angka 8. Begitu pula sebaliknya.
6. Pengaruh dari kesan-kesan luar. Misalnya pakian yang teratur, bersih dan serasi. Tingkah laku yang sopan, wajah yang manis/ganteng dan sebagainya. Ini semuanya mempengaruhi seorang pendidik untuk menilai lebih baik. Demikian pula pakian yang bersih, teratur, mudah dibaca dan sebagainya. Kesan-kesan luar yang tidak menyenangkan kepada seorang pendidik mempengaruhi pendidik itu untuk menilai tidak baik.

BAB II

KARAKTERISTIK ALAT EVALUASI RANAH PSIKOMOTOR DALAM PJOK

A. TUJUAN DAN ALAT PENGUKUR PJOK

Dalam uraian mengenai evaluasi dalam Bab I dijelaskan bahwa langkah pertama dalam melakukan evaluasi adalah menentukan tujuan. Tujuan-tujuan itu dapat bersifat umum, misalnya tujuan-tujuan pendidikan jasmani olahraga dan kesehatan (PJOK). Dapat bersifat lebih khusus seperti tujuan-tujuan dalam memberi pelajaran permainan bola voli. Penentuan dan perumusan tujuan-tujuan itu penting sekali dalam proses penilaian karena tujuan-tujuan itu bukan saja menentukan inti atau bahan pelajaran dan metode mengajarnya tetapi juga alat-alat pengukur yang akan dipergunakan ditentukan oleh tujuan-tujuan tersebut. Salah satu tujuan dari pelajaran bola voli ialah mengetahui sejarah dan peraturan bermain bola voli. Untuk mengetahui seberapa jauh siswa dalam mengetahui sejarah dan aturan bermain bola volley itu dipergunakan tes pengetahuan yang dapat berbentuk tes jawaban singkat atau tes objektif atau mengarang atau uraian. Tes pengetahuan itu dapat diberikan dalam bentuk lisan maupun tertulis. Apabila tujuan lain dari pengajaran permainan bola voli itu ialah agar para siswa mempunyai kecakapan bermain bola voli haruslah berbentuk tes kecakapan pula. Jadi macam-macam alat pengukur yang akan dipilih atau di buat dan kemudian dipergunakan dalam pelajaran olahraga tergantung pada tujuan-tujuan yang ditentukan dan dirumuskan untuk pelajaran tersebut.

Mengenai tujuan-tujuan pendidikan jasmani olahraga dan kesehatan (PJOK) para ahli pendidikan PJOK telah banyak memikirkannya. Pada umumnya mereka sependapat bahwa yang ingin dicapai dengan pendidikan olahraga dan kesehatan di lembaga pendidikan adalah :

1. Perkembangan kemampuan organik (*physical domain*)
2. Perkembangan keterampilan atau *neuromuscular* (*psicomotor psikomotor*)
3. Perkembangan kemampuan intelektual (*kognitive domain*)
4. Perkembangan penyesuaian diri atau sosial (*afektif domain*)

Macam-macam alat pengukur telah mereka buat dan masih dibuat untuk mengukur seberapa jauh tujuan-tujuan itu tercapai oleh para siswa.

Ranah psikomotor sering tidak mendapat perhatian khusus dan posisinya sering diabaikan oleh beberapa pendidik mata pelajaran tertentu. Seharusnya ranah psikomotor merupakan kecakapan pamungkas setelah ranah kognitif dan ranah afektif dikuasai. Hal ini terjadi barangkali ada anggapan tidak semua mata pelajaran dapat dinilai ranah psikomotornya (sesuai dengan tuntutan kompetensi dasar dan kompetensi inti yang harus dicapai oleh siswa) menyebabkan ranah ini kurang mendapat perhatian. Biasanya mata pelajaran-mata pelajaran yang memerlukan banyak kecakapan aplikasi atau praktik adalah: Pendidikan Agama, Pendidikan Seni, Praktik IPA dan Bahasa, serta Pendidikan Jasmani Olahraga dan Kesehatan (PJOK) yang menempatkan ranah psikomotor ini sebagai domain evaluasi yang penting.

Kecakapan psikomotor ini secara khusus menjadi domain evaluasi penting dalam mata pelajaran PJOK. Keterampilan psikomotorik adalah merupakan tipe hasil belajar yang berbentuk keterampilan dan kecakapan untuk bergerak. Kecakapan ini secara teoritis menurut I Made Sriundi Mahardika (2014: 37) terdiri atas:

- 1) Kecakapan yang berkaitan dengan aspek gerakan refleks,
- 2) Kecakapan yang berkaitan dengan aspek keterampilan gerak dasar,
- 3) Kecakapan yang berkaitan dengan kemampuan *perceptual*,
- 4) Kecakapan yang berkaitan dengan keharmonisan dan ketepatan gerak,
- 5) Kecakapan yang berupa keterampilan kompleks dan aspek gerakan ekspresif serta gerakan interpretatif.

Kecakapan-kecakapan tersebut dapat dilakukan melalui kegiatan meniru (*perception*), menyusun (*manipulating*), melakukan dengan prosedur (*precision*), melakukan dengan baik dan tepat (*articulation*), maupun melakukan tindakan secara alamiah (*naturalization*)

Uraian selanjutnya akan dibatasi pada alat evaluasi untuk mengukur kecakapan atau keterampilan pendidikan jasmani olahraga dan kesehatan khususnya ranah psikomotor yang diambil dari pendapat Scott and French

B. KRITERIA BAGI TES KETERAMPILAN PJOK

Kriteria yang dipergunakan untuk tes telah banyak dijelaskan oleh banyak penganalisis-penganalisis buku tes dan pengukuran. Mereka pada umumnya berpendapat bahwa kriteria yang dipakai meliputi ukuran-ukuran statistik yang dilengkapi dengan pertimbangan-pertimbangan praktis dan pendapat objektif. Biasanya diberikan bukti-bukti mengenai reliabilitas, validitas, objektivitas, petunjuk-petunjuk yang distandar/baku, cara menilai atau mensekor, kemudahan

dalam pelaksanaan, ekonomis dalam waktu, adanya norma adalah beberapa pertimbangan praktis. Semuanya ini dan kriteria lainnya yang berhubungan dengan evaluasi pemilihan satu tes kecakapan akan dibicarakan selanjutnya.

1. TES HARUS MENGUKUR KEMAMPUAN-KEMAMPUAN YANG PENTING

Tes-tes kecakapan olahraga biasanya diberikan untuk pembentukan kecakapan umum dalam satu cabang olahraga ataupun kecakapan khusus untuk satu teknik yang penting untuk berhasilnya satu pelaksanaan (*performance*). Berartinya satu kecakapan khusus sebagai satu ukuran sukses dalam memainkan satu permainan tergantung pada relatif pentingnya kecakapan khusus itu bagi semua teknik yang diperlukan dalam permainan tersebut. Jumlah keseluruhan dari semua teknik itu biasanya besar dan tidaklah selalu mudah untuk menentukan pentingnya secara relatif. Juga tidak mudah dalam menentukan yang mana saja yang cukup penting untuk dimasukkan kedalam tes. Dengan kata-kata lain, tidak mudah menentukan standard minimal dari sudut pentingnya. Bila tujuannya untuk menilai kecakapan umum bermain, kelirulah bila mengetes kecakapan-kecakapan yang tidak penting. Bila tujuannya adalah untuk mengadakan diagonosa dari kemampuan-kemampuan untuk keperluan mengajar atau untuk memilih anggota team, maka standard pentingnya satu teknik adalah lebih rendah.

Bukti statistik dari realibilita dan validita suatu tes belumlah cukup. Misalnya, tes “lemparan jauh” secara keliru telah diberikan atau dipergunakan untuk mengukur kemampuan bermain bola basket. Sebagian besar kesalahan disebabkan karena ada bukti-bukti statistik yang dikemukakan yang menunjukkan bahwa korelasi tes itu tinggi dengan kemampuan bermain bola basket. Seseorang harus dapat menduga bahwa skor-skor tes melempar jauh itu akan mempunyai korelasi yang tinggi dengan kemampuan umum dalam hampir semua cabang olahraga, karena melempar biasanya dimasukkan kedalam hampir semua baterai kemampuan gerak umum (*general motor ability*).

Dalam permainan bola basket, lemparan yang jauh relatif tidak begitu penting dan terbatas penggunaannya. Dibandingkan dengan kecakapan seperti menembak, menggiring, *pivoting*, kemampuan melempar jauh relatif tidak penting. Jadi pertimbangan-pertimbangan logis disamping statistik harus dipergunakan untuk memilih tes kecakapan. Apabila memilih tes-tes kecakapan untuk olahraga, guru harus membuat daftar dari kecakapan yang penting yang berhubungan dengan cabang olahraga yang diajarkan dan memilih tes-tes

mengenai kecakapan-kecakapan itu. Misalnya, permainan *hockey* terutama memerlukan kecakapan memainkan bola dan kemampuan pemain untuk bergerak dengan cepat dan pengontrolan badan yang baik. Kemampuan memainkan bola dapat berupa *passing*, menggiring, menembak ke gawang, menghindarkan diri dan mentekel. Kemampuan dapat bergerak dengan cepat dan pengontrolan badan yang baik dapat dibagi kedalam lari dengan bola (menggiring) mengubah arah, kerja kaki untuk *passing*, menembak, mengelak, dan mentekel. Pengontrolan berat badan tidak dapat dipisahkan dari kepandaian memainkan bola dengan tongkat dan menggerakkan badan kedalam posisi untuk menerima *passing*. Mentekel lawan, atau untuk mencegat *passing*. Penelitian-penelitian tentang permainan *hockey* menunjukkan bahwa kebanyakan dari *passing* itu tidak jauh. Dan biasanya dilakukan selagi pemain sedang bergerak dan dalam menembak ke gawang hampir selalu bola dalam keadaan bergerak dilakukan agar jauh jaraknya dari gawang. Dengan analisa seperti ini, guru akan mempunyai sedikit alasan untuk memilih tes *roll-in* atau tes menembak bola ke gawang dalam keadaan bola dan badan berhenti bergerak, untuk menentukan kecakapan bermain *hockey*.

2. TES HARUS MENYERUPAI SITUASI BERMAIN SESUNGGUHNYA

Tes harus sedapat mungkin menyerupai situasi bermain sesungguhnya. Tujuannya seorang pemain tenis lapangan mungkin saja menjatuhkan bola dan memukulnya ke ujung terjauh lapangan tenis. Tes dapat dibuat dalam bentuk memukul bola secara berturut-turut ke suatu sasaran tertentu dan akan diperoleh skor-skor mengenai ketepatan memukul ke sasaran tersebut. Namun pukulan-pukulan atau pengembalian bola dalam permainan tenis dilakukan terhadap bola yang datang mendekati yang biasanya memerlukan posisi kaki dan timing. Jarang sekali reli dalam permainan berlangsung lambat atau sengaja diperlambat. Bola dipukul bolak-balik secara kontinu. Untuk mengukur kecakapan yang sama seperti dalam permainan tes itu harus dilaksanakan secara kontinu. Satu kombinasi dari kecakapan-kecakapan belumlah membuat tes menyerupai situasi permainan sesungguhnya, bahkan mungkin akan mengakibatkan pelanggaran peraturan permainan atau menghasilkan bentuk permainan yang tidak baik. Tes yang dibuat demikian menghasilkan alat mengajar yang tidak bermutu. Contoh-contoh yang dapat dijumpai dalam tes baterai bola basket. “Pivot dan *shooting*” dan “pantulan dan menembak/*shooting*” diberi skor sebanyak jumlah tembakan yang berhasil. Kedua jenis tes itu apabila ditinjau sekilas akan mempunyai

kualitas menyerupai permainan. Namun kurangnya tekanan untuk bertindak cepat sebagai satu ciri dalam permainan bola basket banyak mengurangi nilai tes tersebut. Tidak ada sesuatu apapun yang menghalangi pemain untuk berhenti sebelum melakukan tembakan, yang melenyapkan tujuan seseorang pemain melakukan *pivot* atau pantulan sebelum menembak. Selanjutnya, pelanggaran peraturan berjalan dengan bola dapat terjadi dalam pelaksanaan kedua macam tes tersebut. Apabila pelanggaran-pelanggaran itu dipertimbangkan dalam memberikan skor, diperlukan tambahan seorang pembantu untuk melaksanakan tes itu dan keputusan dari seorang itu sifatnya subjektif.

Faktor waktu merupakan variabel yang penting dalam tes PJOK dan biasanya hal ini disadari. Dalam membuat satu tes ketepatan dalam permainan *softball* dapat dilihat bahwa melempar bola ke sasaran yang tidak bergerak secara berturut-turut memungkinkan seorang pemain lebih banyak mempergunakan waktu untuk mengarahkan lemparan daripada waktu yang tersedia dalam situasi permainan sebagainya. Tes itu disempurnakan, pemain menerima bola dari lontaran sebuah ketapel dan kemudian dengan cepat melemparkan bola ke sasaran yang tidak bergerak. Tes ini menimbulkan kesulitan karena penggunaan ketapel untuk mengetes satu kelompok tidak praktis. Pemain masih dapat untuk tetap melempar setelah menangkap bola, sehingga situasi hampir sama dengan melempar dalam keadaan tidak bergerak. Langkah berikut dalam pembuatan tes ini adalah mengubah urutan dan adanya pengontrolan waktu. Jadi pemain melemparkan bola ke dinding, menangkap bola *softball* setelah memantul dari dinding, memutar badan dan lari cepat ke suatu titik di lapangan yang diumpamakan sebuah base, dan kemudian melemparkan bola ke satu sasaran di dinding yang diumpamakan seorang pemain lainnya. Stop watch dijalankan pada aba-aba saat melakukan lemparan pertama dan dihentikan ketika bola mengenai dinding pada lemparan kedua. Stopwach dari lemparan kedua dicatat. Waktu yang sesungguhnya diperlukan untuk melakukan sepuluh percobaan ternyata relatif tidak penting, tapi dilakukan untuk meniadakan keragu-raguan (kelambatan) melempar bola segera ke sasaran.

Pelaksanaan tes ini lebih banyak mendekati situasi permainan sesungguhnya daripada kedua jenis tes terdahulu. Perbedaan antara kedua tes tersebut dapat dibandingkan sebagai perbedaan antara kedua tes tersebut dapat dibandingkan sebagai perbedaan antara

melempar-lemparkan bola ke *base* untuk latihan dengan usaha sungguh-sungguh mematikan pelari yang menuju ke *base* dalam situasi permainan sebenarnya.

3. TES HARUS MENDORONG BENTUK GERAKAN YANG BAIK

Pengukuran bentuk gerakan di hampir semua cabang olahraga menimbulkan satu masalah. Seorang pemain tenis mungkin dapat membuat skor yang tinggi dalam ketepatan dengan cara memukul bola asal masuk yang seharusnya bola itu dipukul ke daerah tertentu (*baseline*) dari lapangan tenis. Bola yang dipukul asal masuk jalannya mungkin lambat, melengkung dan diakhiri dengan pantulan yang tinggi. Seorang pemain lain mungkin memukul bola itu ke daerah yang sama tetapi jalannya bola itu cepat, mendatar dan menghasilkan satu pukulan rendah. Untuk mengembalikan bola yang dipukul seperti ini lebih sulit dari pada mengembalikan bola dari pemain pertama (bola asal masuk). Beberapa tes gagal untuk membedakan kemampuan kedua pemain itu dan sudah jelas pemain kedua telah melakukan permainan tenis yang jauh lebih baik.

Berbagai usaha telah dilakukan untuk mendorong pemain melakukan bentuk gerakan yang baik. *Rating* secara subjektif mengenai bentuk gerak telah dipergunakan sebagai tambahan sebuah tes. Usaha ini akan efektif untuk mendorong pelaksanaan bentuk gerak yang baik, apabila rating itu dikerjakan ketika tes dilakukan. Penggunaan batas waktu juga telah dicobakan. Bila waktu keseluruhannya itu pendek penting sekali adanya ketelitian dari *stop watch* dan keajegan waktu reaksi dari pencatat waktu. Dimasukkannya unsur waktu dalam tes mungkin lebih bernilai untuk memperoleh situasi permainan sebenarnya daripada untuk mendapatkan bentuk gerak yang baik. Masalah pengukuran secara objektif dari bentuk gerak belum dapat dipecahkan untuk cabang olahraga atau untuk aktivitas menari. Dalam memilih tes harus diusahakan untuk menghindari tes dimana pemain-pemain yang menggunakan bentuk gerak yang kurang baik memperoleh sekor tinggi dari pada pemain-pemain yang menggunakan bentuk gerak yang baik.

4. TES HARUS DILAKUKAN OLEH HANYA SEORANG SAJA

Dalam uraian di atas mungkin telah disarankan menggunakan dua pemain atau lebih dalam situasi tes, karena dalam situasi bermain seorang pemain mesti selalu memperhitungkan pemain lain dari siapa ia menerima bola, kepada siapa ia melemparkan/memberikan bola, serta perilaku dari lawan-lawannya. Penggunaan beberapa pemain dalam satu tes akan memenuhi kriteria menyerupai situasi permainan sebenarnya.

Namun kerjasama atau kompetisi untuk semua pemain harus sama untuk situasi tes. Orang-orang yang di tes harus mendapatkan bola-bola yang enak atau mudah diumpangkan, ukuran sasaran yang sama, dan semua peralatan lainnya harus sama untuk setiap orang yang di tes. Seorang pemain tidak boleh di tes dengan seorang teman yang tidak dapat melempar dengan baik, sedangkan pemainnya mempunyai teman yang pandai. Jelaslah, bahwa tidak ada satu kriteria dipertimbangkan terpisah dari kriteria-kriteria lainnya.

Tes harus dilakukan oleh hanya seorang saja yang sedang dites. Mengingat kriteria itu banyak, maka telah dihasilkan suatu ketentuan tes bahwa tes dengan menggunakan sebuah dinding/tembok sebagai pengganti teman bermain dari yang di tes. Satu kelemahan terhadap jenis tes jika menggunakan tembok sebagai pengganti teman bermain ialah situasi tidak menyerupai permainan yang sebenarnya. Tes memantulkan bola ke tembok merupakan kesepakatan antara kriteria menyerupai situasi permainan yang sebenarnya dan dilakukan oleh hanya seorang saja. Tes yang demikian membantu menghasilkan permainan yang sebenarnya, karena pemain dapat melempar dan menerima bola dan permainan dapat berlangsung terus-menerus tanpa dipengaruhi oleh orang lain. Tuntutan untuk bertindak dapat tidak berbeda dari situasi permainan yang sebenarnya. Lagi pula pemain itu sendiri bertanggungjawab penuh terhadap hasilnya.

5. TES HARUS MENARIK DAN MEMPUNYAI ARTI

Bila guru menginginkan usaha dari siswa sebaik-baiknya maka tes tersebut harus menarik bagi mereka. Tes-tes biasanya mempunyai kualitas tertentu yang menarik bagi siswa atau yang tidak disukai oleh para siswa. Salah satu sifat kualitas tes PJOK adalah memiliki daya tarik bagi siswa adalah tes itu menyerupai situasi permainan yang sebenarnya. Macam-macam usaha untuk motivasi yang dapat dipergunakan guru adalah kepada setiap siswa diberikan kartu skor untuk mencatat hasil-hasilnya secara kontinu, menempelkan dan juga mengumumkan skor-skor dan nama-nama dari pemain-pemain terbaik atau dari siswa-siswa yang memperoleh kemajuan yang besar, membandingkan skor-skor dengan kelas-kelas sebelumnya atau dengan kelompok-kelompok lain, atau dengan standard/kriteria tertentu.

Sistem penskoran harus berarti bagi para siswa, misalnya; penskoran dari beberapa tes dapat didasarkan pada perkenaan yang dapat dilakukan dalam satu periode waktu tertentu atau pada waktu yang diperlukan untuk membuat sejumlah perkenaan. Seorang siswa

dengan segera mengetahui perkenaan yang baik pada sasaran dan jumlah skor keseluruhan dihubungkan dengan jumlah lemparan yang dilakukannya. Namun pengertian siswa mengenai hasil tes nya akan kurang jelas apabila hasil-hasil itu dinyatakan dalam interval waktu, jumlah detik. Walaupun siswa tidak begitu memahami arti dari sekornya, siswa akan membuat perbandingan dengan sekor-sekor teman-temannya, tetapi ia tidak mempunyai pengertian tentang standard dari satu hasil yang optimum atau ideal. Demikian juga, satu skor janganlah dinyatakan dalam bentuk standard deviasi atau cara-cara lain yang tidak dipahami oleh para siswa.

Para siswa akan tetap tertarik kepada tes hanya apabila skor-skor dari tes itu dipergunakan. Skor-skor dari tes harus diumumkan kepada para siswa secepat mungkin, barangkali salah satu kepopuleran tes menembak ke keranjang dalam permainan bola basket adalah hasil tes menembak tersebut akan segera dapat diketahui oleh siswa. Perhatian akan cepat berkurang apabila tes terlalu banyak memakan waktu sehingga tidak ada kesempatan lagi bagi siswa untuk melakukan aktivitas olahraga atau jika tes tersebut menjadi tidak disenangi siswa jika menyebabkan otot-otot menjadi sakit. Kebanyakan tes-tes keterampilan olahraga tidak mengakibatkan hal-hal yang tidak disenangi walaupun kemungkinan ada beberapa tes yang tidak disenangi siswa, seperti tes daya tahan (*endurance*). Persiapan yang cukup sebelum tes tes daya tahan diberikan, seperti melakukan latihan kondisioning akan mencegah otot-otot menjadi sakit atau kaku.

Waktu untuk tes harus sesuai dengan waktu yang biasanya dipergunakan untuk pelajaran latihan-latihan tehnik. Para siswa akan memandang tes itu sebagai bentuk lebih objektif dari latihan dan mereka belajar membuat semua latihan-latihan sebagai pengalaman mengetes diri sendiri.

Dalam memilih tes PJOK pertimbangan harus difokuskan pada tes yang disenangi dan dipahami oleh para siswa. Ini bukanlah berarti bahwa ada perlunya mereka diberi hiburan. Sungguhpun demikian usaha maksimal diperlukan dari siswa dalam melakukan tes. Kegembiraan dan pengertian akan membuat siswa berusaha semaksimal mungkin.

6. TES HARUS CUKUP SUKAR

Tingkat kesulitan dari tes harus disesuaikan dengan kemampuan kelompok yang diberikan tes. Skor-skor harus menunjukkan distribusi yang baik tanpa ada terkumpulnya skor-skor pada suatu angka tertentu. Jika tes terlalu mudah kebanyakan dari skor-skor itu

akan tinggi dan tes itu tidak akan membedakan antara pemain-pemain yang mempunyai perbedaan kemampuan yang kecil. Tes yang mudah tidak dapat dipakai sebagai motivasi untuk selalu berusaha keras. Namun sebaliknya, apabila tes terlalu sulit kebanyakan skor akan mendekati nilai terendah. Satu tes yang sulit akan mengecilkan arti para siswa dan tidak akan berfungsi untuk membedakan pemain-pemain yang memiliki tingkat kemampuan yang berbeda-beda dari kemampuan yang diukur.

Tes dapat diubah atau ditangguhkan sampai para siswa telah lebih berkembang kemampuan mereka atau bentuk tes lain lain dapat dipilih sebagai alternative tes pengganti. Apabila diadakan perubahan, tes itu janganlah dipakai untuk diagnosis atau memberi nilai kecuali jika tes yang diubah itu telah memberikan bukti-bukti statistik tentang bernilai atau tidaknya tes tersebut. Tes yang direvisi harus juga mempunyai norma yang baru atau ditentukan skala prestasinya karena tes yang lama sudah tidak dapat dipergunakan lagi.

7. TES HARUS DAPAT MEMBEDAKAN TINGKAT KEMAMPUAN

Hampir semua kegiatan dalam pendidikan didasarkan pada pengakuan adanya perbedaan-perbedaan dari para siswa. Sudah barang tentu semua dasar prosedur-prosedur pengukuran dan evaluasi berlandaskan pada pengakuan adanya perbedaan-perbedaan tertentu, dan pengakuan adanya perubahan status dari setiap siswa, karena disebabkan oleh: waktu, pengalaman dan latihan mengubah tingkah lakunya. Ini berarti bahwa skor-skor tes harus mencerminkan perbedaan-perbedaan tersebut. Apabila sebuah kelompok sangat berbeda kemampuannya, skor-skor tes akan menunjukkan perbedaan itu. Jika kelompok tersebut kemampuannya homogen masih ada perbedaan-perbedaan antara siswa dalam kelompok itu yang perlu ditentukan.

Perbedaan-perbedaan yang menyolok tidak begitu sulit untuk diketahui guru ketika latihan, pada waktu bermain, atau dari kebanyakan tes yang diberikan. Namun demikian, akan lebih sulit untuk membedakan kemampuan antara sebagian besar dari para siswa yang tergolong sedang kemampuan mereka di dalam situasi pelajaran olahraga biasa. Tes harus dapat membedakan kemampuan mereka yang tergolong sedang ini selain dapat membedakan kemampuan dari yang tergolong berkemampuan tinggi dan rendah. Maksud dari tes adalah untuk mendapatkan satu ranking atau tingkatan kemampuan dari siswa sesuai dengan kemampuan masing-masing.

Tes-tes akan mencapai tingkatan diskriminasi yang tinggi hanya apabila tugas yang diberikan kepada para siswa disesuaikan dengan tingkatan kemampuan mereka. Sungguhpun demikian, kriteria tentang kesukaran ini tidak menjamin diskriminasi dan sifat-sifat yang lain dari tes harus pula dipertimbangkan.

8. TES HARUS DILENGKAPI DENGAN CARA MENS KOR YANG TELITI

Objektivitas dari suatu tes banyak tergantung pada ketelitian dalam menskor. Beberapa tes dapat disekor dengan tingkat ketelitian yang cukup tinggi dengan menggunakan pita pengukur atau *stop watch*. Tes-tes senam yang disekor dengan cara “berhasil” atau “gagal” ditentukan berdasarkan pertimbangan subjektif dan beberapa nomor pertandingan senam sulit dapat disekor secara teliti. Keranjang dengan jala sudah berlubang-lubang akan sedikit menimbulkan kesangsian untuk menentukan apakah bola tersebut masuk atau tidak melalui keranjang. Panah yang mengenai sasaran dapat disekor dengan teliti, sebuah bola yang di lempar kesuatu sasaran sulit untuk disekor.

Ketelitian menskor lemparan-lemparan ke satu sasaran yang terdiri dari lingkaran-lingkaran dapat mempermudah dengan membuat pusat lingkaran dari sengk dan dipakukan pada sasaran. Lingkaran kedua, ketiga, keempat dan seterusnya dibuat dari sengk juga dan dipakukan pada sasaran. Suara akan membantu penglihatan. Sasaran-sasaran dapat diberi cat dengan warna yang berlainan untuk memudahkan menskor, misalnya sasaran untuk cabang olahrag panahan. Apabila waktu dari perlombaan lari ditentukan dengan menggunakan zona, kesalahan-kesalahan dalam penentuan waktunya kadang-kadang terjadi. Jika perlombaan itu dilakukan dalam ruang yang berdekatan dengan sebuah dinding, ketelitian menskor dapat ditingkatkan apabila benar-benar dari zona dituliskan pada dinding dan bukan di lantai. Tes-tes yang tidak dapat disekor secara teliti biasanya mempunyai realibilitas dan objektivitas rendah dan hal ini menyebabkan ketidakpuasan bagi yang siswa yang di tes.

9. TES HARUS MEMPUNYAI CUKUP JUMLAH PERCOBAAN/TRIAL

Jumlah percobaan harus cukup untuk meniadakan terjadinya faktor kebetulan. Kebanyakan tes yang mengharuskan usaha maksimal dapat diukur satu sampai tiga kali percobaan, misalnya tes melempar sejauh mungkin, tes kekuatan, tes kecepatan dimana unsur pengontrolan bukan merupakan faktor yang pokok. Apabila tingkat ketelitian yang tinggi diperlukan untuk pelaksanaan gerak yang baik maka jumlah percobaan harus lebih

banyak. Kecepatan lari untuk jarak pendek tidak berbeda banyak dari hari-kehari, kemampuan untuk melakukan lemparan bebas hasilnya menunjukkan perbedaan yang sangat besar dari hari kehari.

Hanya dengan jalan eksperimen dapat ditentukan jumlah percobaan yang diperlukan untuk mendapatkan skor yang betul-betul representatif. Jumlah percobaan harus ditambah apabila unsur kebetulan bertambah pula. Cara mensekor mempengaruhi jumlah percobaan yang diperlukan. Apabila tiap percobaan setiap tes keterampilan diberi skor atas dasar keberhasilan dan gagal, seperti pada menembak bola ke keranjang, jumlah percobaan perlu lebih banyak dari pada apabila setiap percobaan diskor atas dasar ketelitian secara relatif, seperti dalam satu tes yang menggunakan sasaran dengan lingkaran-lingkaran konsentris. Jumlah percobaan yang diperlukan berhubungan erat dengan kemampuan para siswa. Jumlah yang diperlukan untuk kelompok dari siswa-siswa yang pandai biasanya akan lebih sedikit dari pada untuk siswa-siswa yang kurang pandai untuk tes yang sama.

Bilamana tes itu dipergunakan untuk alat latihan atau motivasi, jumlah percobaan dapat lebih kurang jika dibandingkan tes untuk penggolongan atau digunakan untuk sebagian dasar menentukan nilai. Jumlah waktu untuk memberikan tes biasanya merupakan salah satu factor yang dipertimbangkan dalam memilih tes. Misalnya; 15 atau 20 percobaan mungkin memerlukan waktu yang lebih banyak daripada yang tersedia. Waktu untuk tes tidak dapat diperkirakan hanya dengan jumlah percobaan saja. Tes yang diberikan secara perorangan, seperti tes kekuatan dengan menggunakan *dynamometer*, mungkin lebih banyak memakan waktu dari pada tes yang diberikan kepada sejumlah siswa pada waktu yang sama walaupun diperlukan lebih banyak percobaan.

Untuk mengukur kemampuan siswa secara cermat maka jumlah percobaan/*triall* harus cukup untuk meniadakan terjadinya faktor kebetulan. Daripada mengurangi jumlah percobaan untuk menghemat waktu dan mengorbankan faktor ketelitian tes maka akan lebih baik membatasi pemilihan tes yang berisi kemampuan-kemampuan penting. Kesungguhan guru PJOK dalam merencanakan pelaksanaan tes bertujuan agar tidak ada waktu yang terbuang dengan percuma. Persiapan tambahan akan semakin meningkatkan ketelitian dalam mengumpulkan data sehingga akan menghemat waktu yang dipergunakan pendidik.

10. TES HARUS SEBAGIAN DIPERTIMBANGKAN DENGAN BUKTI-BUKTI STATISTIK

Daftar kriteria untuk menilai satu tes yang biasanya dijumpai dalam *teks book* adalah reliabilitas, validitas, objektivitas, ekonomis, dan adanya norma. Reliabilita atau keajegan pengukuran dari tingkatan kemampuan yang sama hanya dimungkinkan apabila:

- a. Si pelaku cukup tertarik untuk berusaha semaksimal mungkin
- b. Hanya kecakapannya sendiri yang diukur
- c. Alat dan kondisi tes itu seragam
- d. Jumlah percobaannya cukup untuk meniadakan faktor kebetulan
- e. Cara mensekor yang objektif.

Validitas suatu instrumen penilaian dapat diartikan seberapa jauh instrument penilaian tersebut mengukur apa yang harus diukurnya. Pertimbangan logika maupun statistik harus dipergunakan dalam mempelajari validitas. Istilah "*face validity*" kadang-kadang dipakai. Ini adalah suatu analisa rasional dari tugas tes atau satu ketentuan mengenai apa yang harus diukur oleh tes. Akan menjadi sia-sia untuk mencoba menentukan nilai dari satu tes kecuali apabila orang merasa bahwa secara rasional dapat dianggap bahwa tes itu mengukur kemampuan tertentu. Ini adalah satu dasar untuk membuat item-item dalam tes pengetahuan atau item tes dalam tes keterampilan. Namun demikian, akan lebih diinginkan untuk mengecek atau mengesahkan pemikiran ini sehingga perkiraan validitas dapat dibuat untuk satu kelompok siswa atau lebih.

Ekonomi adalah terutama hasil dari pemilihan yang hati-hati dari beberapa tes yang tinggi validitasnya dan dilaksanakan secara efisien. Perhitungan validitas dari satu baterai tes atau kombinasi dari tes dapat membantu dalam pemilihan ini. Pengakuan kenyataan bahwa setiap kecakapan dalam satu aktivitas tidak dapat dan tidak perlu diukur akan membantu kemudahan dalam mengukur.

Norma atau standard prestasi tidak disangsikan lagi akan menambah kegunaan suatu tes, tetapi tanpa adanya norma tidaklah menjadikan tes itu tidak bernilai. Dari data-data yang dikumpulkan oleh guru PJOK yang mempunyai beberapa kelas dengan para siswa yang sama umur dan pengalamannya, guru dapat membuat satu skala. Skala-T terbukti telah memuaskan dan dapat dibuat dari sebanyak lima puluh buah data lebih.

Skala buatan guru mungkin lebih disukai dari pada penggunaan skala yang diterbitkan, apabila skala yang terakhir ini tidak cocok untuk kelompok yang di ukur. Seringkali skala yang diterbitkan dibuat dari sejumlah kecil orang coba dari suatu tempat tertentu. Nilai norma nasional masih sering dipertanyakan. Agar norma itu dapat dipergunakan, norma tersebut perlu berisikan standard untuk pemain-pemain wanita dan pria dan yang memiliki pengalaman-pengalaman berbeda-beda misalnya satu tes dari norma untuk tes permainan bola basket belumlah memadai. Ukuran badan, umur, jenis kelamin dan kematangan-kematangan pemain-pemain perlu dipertimbangkan dalam membuat norma, terutama untuk tingkatan-tingkatan umur tertentu.

Apabila seseorang akan memanfaatkan hasil kerja pembuatan tes dari orang lain, penting adanya beberapa konsep statistik tertentu. *Range* Dari kemampuan atau distribusi dari sekor adalah dasar untuk kebanyakan konsep-konsep ini. Dalam setiap kelompok yang besar untuk satu kemampuan tertentu akan dijumpai variasi sekor yang besar dalam tingkatan kemampuan tersebut. Dengan alat pengukur yang baik, *range* dari kemampuan akan terlihat pada sekor-sekor tes. Tidak akan dijumpai jumlah individu yang sama setiap titik dari *range* itu. Ke arah pusat *range* (di sekitar nilai rata-ratanya) akan dijumpai banyak sekor dan sedikit sekor yang jauh dari pusat *range*. Dapat diduga bahwa dalam setiap kelompok besar banyak orang yang mendapat skor mendekati nilai rata-rata (*mean*), sedikit orang yang akan memperoleh skor tinggi dan begitu pula untuk skor rendah.

Skala prestasi menunjukkan distribusi kemampuan dari kelompok tertentu dalam mencapai suatu tes tertentu. Distribusi-distribusi berbeda untuk kelompok-kelompok yang berbeda. Laporan skala prestasi yang disusun harus memuat dari populasi mana skala tes tersebut dibuat. Di samping itu juga harus memuat umur atau tingkatan kelas, kelamin, jumlah orang coba, tingkatan pengalaman mereka, dan sifat-sifat khusus lain yang akan mempengaruhi hasil-hasil.

Reliabilitas dan validitas dari tes dinyatakan dengan koefisien korelasi. Koefisien-koefisien itu adalah angka-angka yang menunjukkan tingkatan hubungan. Koefisien reliabilitas menunjukkan antara pemberian tes dua kali berurutan-turut dan menyatakan ketetapan (keajegan) hasil tes apabila tes itu mengukur apa yang di ukurnya. Seringkali tidak praktis untuk melaksanakan tes yang sama dua kali untuk mencari reliabilita satu

tes. Sebagai penggantinya jika tes itu terdiri beberapa kali percobaan, misalnya: enam kali, maka hasil percobaan yang genap dikorelasikan dengan hasil percobaan yang ganjil. Jadi skor-skor dari percobaan kesatu, ketiga, kelima, ketujuh, kesembilan dijumlahkan dan dikorelasikan dengan jumlah skor yang diperoleh dari percobaan kedua, keempat dan keenam, kedelapan, dan kesepuluh. Hal yang sama dikerjakan untuk semua individu yang di tes. Koefisien korelasi yang dihasilkan menunjukkan keajegan hasil dari semua percobaan. Koefisien yang diperoleh biasanya akan lebih rendah daripada apabila tes keseluruhannya dilaksanakan dua kali. Ini disebabkan karena korelasi itu sebenarnya diperoleh dari setengah percobaan. Perbedaan ini mungkin dapat dikoreksi dan biasanya dikerjakan. Metode untuk mengoreksi yang biasanya dipakai adalah dengan rumus *Spearman-Brown Prophecy*. Apabila memilih tes hendaknya selalu dicermati apakah telah dilakukan koreksi terhadap hal ini agar dapat melakukan interpretasi yang baik mengenai koefisien reliabilitanya.

Koefisien validitas menunjukkan tingkatan hubungan antar *criterion* dan tes. Semakin tinggi tingkatan hubungannya akan semakin nyata bahwa tes tersebut mengukur kemampuan yang diselidiki, atau hampir sempurna ia membenarkan anggapan dasar dari *face validity*. *Criterion* yang dikorelasikan dengan skor-skor tes dapat berupa tes lain yang telah terbukti baik, atau dapat berupa skor-skor yang ditentukan secara subjektif seperti hasil dari *subjektif rating*. Karena *criterion* adalah alat pengukur yang dipergunakan untuk menilai tes maka penting sekali *criterion* yang dipergunakan itu harus baik. Jika *criterion* tidak baik penyusun tes dapat membuat kekeliruan karena ia membuat tes yang mungkin sekali lebih bernilai pada *criterionnya* tetapi terpaksa dibuang atau tidak dipakai. Karena *criterion* kurang baik, tes yang dibuat kelihatannya juga menjadi kurang baik.

Pengguna tes mempunyai hak untuk memperoleh kejelasan tentang *criterion* yang dipakai agar supaya pengguna tes dapat menginterpretasikan koefisien validitas. Apabila *subjektif rating* telah dipergunakan sebagai *criterion* prosedur yang dilakukan harus dapat diterima. Koefisien validitas yang rendah akan diperoleh apabila *criterion* berasal dari skor-skor yang diperoleh melalui *subjektif rating*. Hal ini akan berbeda apabila *criterion* yang dipergunakan oleh pengguna tes berupa tes yang sudah memenuhi persyaratan

sebagai tes yang baik. Dengan mengetahui kualitas tes tersebut pengguna tes akan lebih dapat menginterpretasikan koefisien validitas.

Seringkali diajukan pertanyaan tentang berapa besarnya koefisien korelasi. Tidak dapat ditentukan besar minimalnya secara sembarangan, akan tetapi secara umum dapat dibuat. Koefisien reliabilitas dapat diharapkan lebih besar dari pada koefisien validitas. Koefisien-koefisien yang rendah dapat diharapkan pada tes-tes PJOK daripada tes-tes kemampuan kognitif. Hal ini mungkin disebabkan karena lebih banyak fluktuasi dalam pelaksanaan tes yang pertama. Tes-tes yang diberikan kepada siswa puteri biasanya menghasilkan reliabilitas yang lebih rendah dari pada apabila tes-tes itu diberikan kepada siswa putera. Kenyataan ini dapat terjadi karena siswa puteri lebih sukses untuk dimotivasi agar mereka berusaha sekeras mungkin. Hasil prestasi dari pemain-pemain yang kurang berpengalaman biasanya kurang reliabel dari pada hasil-hasil pemain-pemain yang lebih berpengalaman. Apabila satu tes dilaporkan mempunyai koefisien reliabilita sebesar 0,78 misalnya, ini bukan berarti bahwa kelompok siswa manapun yang memakai tes yang sama akan menghasilkan koefisien reliabilitas yang sama apabila dihitung dari data-data kelompok tersebut. Reliabilitas tes itu berlaku khusus bagi kelompok yang di tes. Koefisien reliabilitas yang sama dapat diharapkan hanya untuk kelompok yang serupa dalam kondisi testing yang serupa.

Tes *Braddy Volley Ball* menghasilkan koefisien reliabilitas sebesar 0,73 untuk 47 orang siswa putri SLTA. Dengan menggunakan sebagian skor tes lima percobaan terbaik dari 10 kali percobaan. Tes yang sama memberikan koefisien reliabilita sebesar 0,92 untuk 75 mahasiswa puteri dengan menggunakan sebagian skor tes dari 10 kali percobaan. Perbedaan ini tidak seluruhnya disebabkan karena pelaksanaan tes karena tes bola voli tersebut diberikan oleh siswa putri yang sama dalam kondisi test yang sama. Jumlah percobaan yang digunakan sebagai sekor akhir dari tes mungkin yang menyebabkan perbedaan hal tersebut. Telah diketahui bahwa sampai batas jumlah tertentu reliabilitas dari satu pengukuran meningkat dengan ditambahkan jumlah percobaan. Sebagian besar dari percobaan kedua koefisien reliabilitas tersebut mungkin disebabkan karena perbedaan pengalaman dari siswa-siswi yang di tes. Jumlah orang coba dapat pula mempengaruhi koefisien realibilitas tes. Koefisien reliabilitas tes akan meningkat apabila jumlah orang coba yang dipakai tes jumlahnya semakin diperbanyak.

Dengan semua penjelasan sebagaimana diuraikan di atas, kesimpulan dapat dibuat mengenai interpretasi kualitatif dari angka-angka koefisien korelasi. Koefisien di atas 85 dianggap sangat baik. Koefisien validitas tes jarang sekali di atas 0,85, koefisien realibilita sering lebih besar 0,85. Koefisien 0,75 sampai 0,85 dianggap sudah cukup layak dipergunakan untuk berbagai tujuan tes. Apabila koefisien reliabilita lebih kecil dari 0,75 menunjukkan ketidak ajegan dan sebagai cerminan alat ukur yang tidak baik. Koefisien dibawah 60 sampai 65 menunjukkan nilai peramal yang kurang baik tetapi tes itu dapat dipakai untuk tujuan-tujuan yang lain. Koefisien reliabilitas yang rendah menunjukkan ketidak ajegan dalam pengukuran dan hal ini dapat menyarankan perlunya tambahan percobaan atau perbaikan dalam petunjuk pelaksanaan dan kondisi pada saat tes itu dilaksanakan. Koefisien validitas yang rendah menunjukkan bahwa tes itu tidak bernilai untuk meramalkan kemampuan yang seharusnya di ukur. Tetapi jika ternyata secara subjektif mempunyai nilai serta mempunyai reliabilitas yang tinggi tes tersebut dapat menjadi tes yang baik untuk latihan, serta dapat pula dipakai untuk menjamin adanya perhatian dan usaha dari para siswa.

Kadang-kadang beberapa tes dipelajari karena berhubungan dengan kemampuan umum dalam pendidikan olahraga dan kesehatan (PJOK). Setiap tes barangkali mempunyai koefisien validitas yang tinggi apabila masing-masing di korelasikan dengan kriteria. Mungkin sekali beberapa tes tersebut barangkali mengukur kemampuan yang sama. Tes-tes dikorelasikan dengan masing-masing tes untuk menentukan apakah terdapat beberapa tes mengukur hal yang sama. Jika koefisien korelasi antara dua tes itu tinggi salah satu tes mungkin dapat dibuang dan dipilih tes yang bervaliditas tinggi. Korelasi antara dua tes dikenal dengan istilah interkorelasi. Dalam menggabungkan tes untuk membuat tangkai tes (*battery test*) tes-tes yang nilai interkorelasinya rendah harus dipilih karena tes tersebut mengukur aspek yang berbeda dari kemampuan umum.

Validitas dari rangkaian tes keseluruhannya dihitung dengan menggunakan analisis statistik korelasi berganda (*multiple correlation*). Untuk mendapatkan validitas yang tinggi dari salah satu kombinasi tes penting sekali bahwa setiap tes itu reliabel dan cukup valid. Kadang-kadang beberapa tes disimpan untuk digunakan di masa mendatang karena ada sumbangannya terhadap tes baterai meskipun tes-tes itu tidak mempunyai validitas yang dapat diterima. Interkorelasi antar item tes yang dipergunakan untuk baterai tes

harus cukup rendah. Koefisien korelasi berganda dinyatakan dalam bentuk yang sama dengan korelasi sederhana (korelasi *pearson product moment*) tetapi koefisien korelasinya selalu positif. Nilai korelasi berganda agak lebih tinggi daripada yang diharapkan dari satu tes tunggal karena lebih banyak pengukuran-pengukuran yang dipertimbangkan.

Korelasi berganda ($R_{Y X_1, X_2, \dots, X_n}$) sesungguhnya adalah koefisien validitas dari baterai tes. Dalam mengerjakan daftar tes eksperimen yang panjang beberapa kombinasi dapat dicobakan dan R korelasi-korelasi berganda itu akan menunjukkan validitas dari tiap-tiap kombinasi itu. Orang coba setelah itu dapat memilih satu baterai tes hanya berdasarkan nilai R yang tinggi atau memilih beberapa tes yang paling valid dengan mempertimbangkan tes tersebut dapat dilaksanakan dan biayanya terjangkau, peralatan khusus, dan hubungan dari tes-tes itu ditinjau dari apa yang telah diajarkan kepada para siswa, serta dari apa yang diharapkan dari mereka dan beberapa pertimbangan lain yang serupa. Harus diingat bahwa perbedaan angka dalam nilai koefisien validitas baterai tes ($R_{Y X_1, X_2, \dots, X_n}$) tidak selalu berarti, dan pertimbangan lain itu sungguh-sungguh penting sekali dalam pelaksanaan dan penggunaan dari tes.

BAB III

KRITERIA-KRITERIA ALAT EVALUASI PJOK

Menyusun tes prestasi belajar yang baik merupakan salah satu tugas dari guru pendidikan jasmani paling menantang. Dikatakan demikian sebab tes yang berkualitas tidak dengan sendirinya terjadi, melainkan perlu dikerjakan dengan sungguh-sungguh dan perlu dipersiapkan secara matang. Tidak mengherankan apabila Fernandes (1984) menyatakan, "*The skill in writing items tes is crucial*". Untuk menjadi penyusun tes yang baik, pembuat tes perlu memiliki bekal pengetahuan dan pemahaman tentang tes dan evaluasi secara baik pula. Prinsip-prinsip penyusun kisi-kisi atau *blue print*, penulisan butir soal, penyusunan format yang tepat, sistem penskoran, serta analisis butir perlu dikuasai, selain isi materi sebagai bahan penyusunan tes.

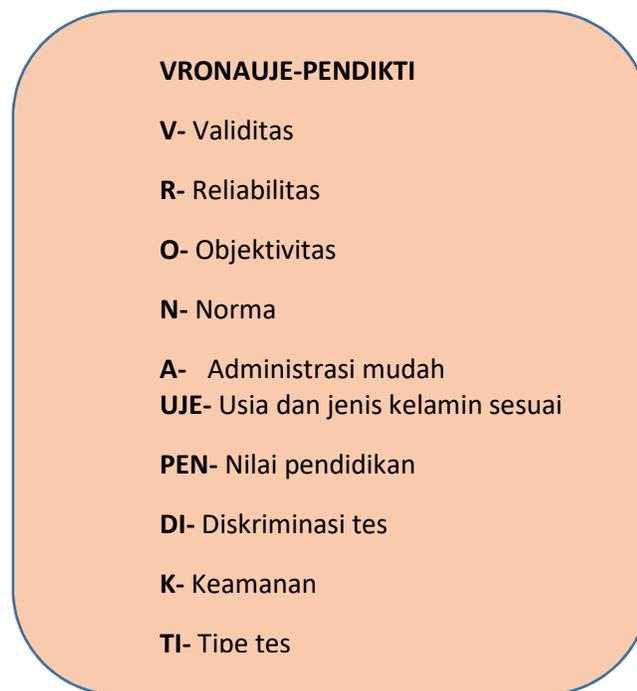
Pengetahuan tentang penyusunan tes dan evaluasi saja tidaklah cukup untuk menghasilkan tes yang baik. Ternyata masih diperlukan praktik menyusun tes serta menganalisisnya secara berulang-ulang untuk memperoleh apa yang diinginkan. Guru yang berpengalaman menyusun/menulis soal akan terbiasa menghadapi liku-liku permasalahan tes, yang pada akhirnya akan menghantarkan guru tersebut pada keberhasilan penyusunan soal. Hopkins (1981: 26) mengemukakan bahwa sebenarnya penyusunan tes adalah "**lebih pada seni daripada ilmu**" dan seni menyusun tes dapat dipelajari lewat petunjuk-petunjuk yang jelas, praktik penyusunan yang terus menerus, serta umpan balik (*feed back*) dari soal yang disusunnya.

Tes yang baik perlu direncanakan dengan hati-hati dan teliti. Petunjuk yang biasa diberikan untuk itu adalah:

1. Sesuaian tes yang disusun dengan tujuan kurikulum, bukan pada apa yang tertulis melainkan pada apa yang diajarkan.
2. Berikan proporsi isi materi yang menurut pentingnya dan tekanannya dalam pembelajaran.
3. Perhatikan tujuan diselenggarakannya tes tersebut, seperti untuk melihat perbedaan individu atau untuk mengetahui penguasaan kelas antara materi yang diajarkan.
4. Sesuaikan tes dengan tingkat kemampuan siswa.

Agar mampu membuat/menyusun tes prestasi belajar yang berkualitas, Suryabrata (1997: 27) berpendapat bahwa, kemampuan-kemampuan yang harus dimiliki oleh seorang guru diantaranya adalah: (1) menguasai materi yang akan diujikan, (2) memahami tata nilai yang mendasari pendidikan, (3) memahami karakter siswa (4) mampu menggunakan bahasa yang efektif, dan (5) menguasai teknik penulisan soal tes. Banyaknya persyaratan dalam membuat dan menyusun tes prestasi belajar menunjukkan bahwa pembuatan tes prestasi belajar yang baik bukanlah merupakan proses yang sederhana. Oleh karena itu, agar mampu membuat tes prestasi belajar yang berkualitas, maka guru dituntut untuk memiliki sejumlah pengetahuan dan keterampilan dalam membuat soal yang baik.

Menyadari bahwa kualitas informasi yang diperoleh saat pengukuran tergantung pada kualitas tes yang dipakai, maka guru perlu memahami kriteria untuk memilih tes yang telah tersedia atau untuk menyusun tes baru untuk kepentingan pembelajaran pendidikan jasmani. Kriteria itu mencakup: (1) validitas, (2) reliabilitas, (3) objektivitas, (4) norma, (5) kemudahan administrasi tes, (6) kesesuaian usia dan jenis kelamin, (7) nilai pendidikan, (8) diskriminasi tes, (9) keamanan, dan (10) tipe tes yang akan dijelaskan berikut ini¹.



Gambar 2. Kriteria tes yang baik

¹ Strand, Bradford N. & Wilson, R. (1993). *Assessing Sport Skills halm 9-16*. Champaign: Human Kinetics Publishers.

A. Validitas

Validitas merupakan salah satu faktor yang sangat penting untuk mengevaluasi tes yang baik. Validitas merupakan derajat kemampuan tes untuk mengukur yang seharusnya diukur (Johnson & Nelson, 1986). Misalnya, suatu tes yang didesain untuk mengukur akurasi penempatan servis tenis seharusnya tidak dapat digunakan untuk mengukur kecepatan atau power servis. Contoh yang lain, misalnya Markus, siswa kelas 7, yang mampu mengalahkan semua teman sekelasnya dalam bermain tenis namun melakukan salah satu ites tes tenis dengan buruk. Jika tes keterampilan tenis yang diberikan pada Mark memiliki koefisien korelasi validitas 0,40; maka dipertanyakan apakah tes tersebut betul-betul mengukur apa yang seharusnya diukur. Dengan hasil tersebut, maka tester harus mencari tes lain dengan validitas yang lebih tinggi.

Korelasi menunjukkan hubungan antara 2 variabel (Hastad & Lacy, 1989). Derajat hubungan antara 2 variabel (sesuatu yang dapat diskor, misal berat badan, skor tes, kecepatan) disebut sebagai koefisien korelasi. Hubungan sempurna antara 2 variabel ditunjukkan jika koefisien korelasi di antara keduanya adalah 1,00. Rentang koefisien korelasi adalah dari +1,0 sampai -1,0. Misalnya 0,96; 0,80; -0,45.

Meskipun tidak ada satupun tes, skala ataupun instrumen yang dapat dinilai 100% valid (Barrow et al., 1989), anda harus mencari sebisa mungkin tes yang validitasnya mendekati 1,00. Selanjutnya, meskipun para ahli telah melaporkan berbagai level koefisien korelasi yang dapat diterima, penerimaan suatu koefisien korelasi tergantung pada kesesuaian dari kriteria, tujuan tes, dan tujuan penggunaan tes. Safrit (1990) menyatakan bahwa penerimaan koefisien korelasi tergantung pada tipe validitas yang ditentukan. Misalnya, validitas konstruk harus memiliki koefisien korelasi minimal 0,8. Sementara untuk validitas prediktif hanya perlu memiliki koefisien korelasi 0,5. Barrow & McGee (1979) menyatakan standar untuk mengevaluasi koefisien korelasi validitas yang dapat dilihat pada tabel 1. Ketika menginvestigasi validitas suatu tes, maka anda mengevaluasi kewajaran dan kesesuaian tes untuk suatu tujuan tertentu dan kelompok tertentu.

Tabel 1. Standard untuk Menginterpretasikan Koefisien Korelasi Validitas

Koefisien Validitas	Kategori
0.85 – 0.99	<i>Excellent</i> (sangat bagus sekali)
0.80 – 0.84	<i>Very good</i> (sangat bagus)
0.70 – 0.79	<i>Acceptable</i> (dapat diterima)
0.60 – 0.69	<i>Poor</i> (buruk)

Sumber: A Practical Approach to Measurement in Physical Education (p.42) by H.M. Barrow and R. McGee, 1979, Malvern, PA: Lea&Febiger.

B. Reliabilitas

Reliabilitas adalah konsistensi suatu tes dalam mengukur apa yang seharusnya diukur (Bosco & Gustafson, 1983). Tes yang reliabel harus menunjukkan hasil yang mirip/serupa kapanpun diberikan. Reliabilitas tinggi mensyaratkan minimalisasi kesalahan pengukuran. Teknik pengukuran, kondisi pengetesan, administrasi tes, persiapan peserta tes, dan kondisi lingkungan adalah variabel-variabel yang harus distandarkan sedapat mungkin.

Reliabilitas dan validitas memiliki hubungan secara langsung (Johnson & Nelson, 1986). Ketika reliabilitas suatu tes tinggi, skor tes tidak berbeda jauh ketika diulang-ulang; jika skor tes berbeda jauh maka reliabilitas tes tersebut rendah. Misalnya, skor Jefry 40 poin pada tes keterampilan *wall-volley* tenis pada hari Senin dan 41 pada tes yang sama 1 minggu sesudahnya, tes tersebut dianggap memiliki derajat reliabilitas tinggi. Sebaliknya, jika skor Jefry 40 pada hari Senin dan menjadi 20 satu minggu sesudahnya, maka disimpulkan tes tersebut memiliki derajat reliabilitas rendah.

Meskipun reliabilitas tinggi tidak menjamin validitas tinggi, namun reliabilitas rendah menjamin rendahnya validitas suatu tes. Ketika skor Jefry 40 dan 41, reliabilitas tes tampak baik karena ditunjukkan skor yang mirip pada pengulangan tes yang sama. Tetapi, meskipun item tes memiliki koefisien reliabilitas tinggi, kita harus memastikan bahwa item tes sesuai keterampilan yang akan diukur. Jika siswa yang sama mencetak angka baik pada tes *wall-volley* namun tampil buruk saat turnamen, maka validitasnya dipertanyakan.

Ketika menentukan reliabilitas tes, anda mengevaluasi konsistensi tes pada pengukuran berulang. Barrow & McGee (1979) menyatakan standard untuk mengevaluasi koefisien korelasi reliabilitas yang dapat dilihat pada tabel 2.

Tabel 2. Standard untuk Menginterpretasikan Koefisien Reliabilitas

Koefisien Reliabilitas	Kategori
0.95 – 0.99	<i>Excellent</i> (sangat bagus sekali)
0.90 – 0.94	<i>Very good</i> (sangat bagus)
0.80 – 0.89	<i>Acceptable</i> (dapat diterima)
0.70 – 0.79	<i>Poor</i> (buruk)
0.60 – 0.79	<i>Questionable</i> (dipertanyakan)

Sumber: A Practical Approach to Measurement in Physical Education (p.42) by H.M. Barrow and R. McGee, 1979, Malvern, PA: Lea&Febiger.

C. Objektivitas

Objektivitas adalah suatu bentuk reliabilitas yang disebut sebagai reliabilitas *rater* (penilai), terjadi ketika 2 orang atau lebih mengadministrasi tes yang sama pada orang yang sama dan memperoleh hasil yang sama (John & Nelson, 1986). Koefisien korelasi diperoleh dengan membandingkan hasil tes yang menunjukkan derajat hubungan di antara *judges* mengenai nilai pengukuran. Objektivitas sangat tergantung dari kejelasan dan kelengkapan petunjuk tes yang diberikan dan seberapa cermat prosedur tes diikuti.

Ketika *judges* memberi nilai berbeda ketika mengevaluasi subjek yang sama, maka objektivitas tidak ada dan tidak ada satu skor pun yang dapat dianggap reliabel atau valid. Perbedaan hasil terjadi ketika *judges* menggunakan teknik skor atau penilaian berbeda dan ketika prosedur tes tidak diikuti. Misalnya, dua orang guru (Ferina dan Andre) menilai Jihan ketika dia lari melewati rangkaian kun. Andre menambahkan 1 detik pada skor final tiap kali Jihan membuat kesalahan seperti menabrak kun. Ferina mencatat waktu tanpa mempertimbangkan kesalahan. Secara jelas, skor final akan sangat berbeda. Untuk memastikan objektivitas *judges* harus mengevaluasi subjek dengan cara serupa dan mengikuti prosedur tes semirip mungkin.

D. Norma

Norma adalah nilai-nilai yang berasal dari skor kasar yang diperoleh dari populasi tertentu pada suatu tes (Johnson & Nelson, 1986). Secara umum, guru dan siswa menggunakan norma untuk membandingkan skor siswa. Hal yang perlu diperhatikan adalah bahwa tabel norma hanya dapat diterapkan pada kelompok tertentu dimana norma tersebut

dihimpun. Misalnya norma tes keterampilan untuk siswa sekolah menengah tentu tidak dapat diterapkan untuk mahasiswa.

Penggunaan norma nasional mungkin tidak praktis karena banyak program lokal yang memiliki keterbatasan waktu, guru, ruang, dan peralatan. Pada kasus ini, sebaiknya norma ditetapkan berdasarkan sekolah atau area tertentu sehingga lebih bermakna. Sebelum menggunakan norma, evaluasi kecukupan kondisi dengan mempertimbangkan pertanyaan-pertanyaan berikut (Hastad & Lacy, 1989; Johnson & Nelson, 1986).

1. Seberapa besar database normatif? Semakin besar sampel, maka semakin merepresentasikan populasi keseluruhan. Jika tabel norma ditetapkan kurang dari beberapa ratus skor untuk tiap usia dan jenis kelamin, maka gunakan norma tersebut dengan hati-hati.
2. Dari kelompok mana norma tersebut disusun? Administrator tes tidak boleh mengevaluasi siswanya berdasarkan norma yang diperoleh dan didesain untuk populasi yang berbeda. Misalnya, norma yang diperoleh dari mahasiswa tidak boleh digunakan untuk mengevaluasi siswa sekolah menengah.
3. Dari lokasi geografis mana norma tersebut dihimpun? Variasi penampilan siswa seringkali ditemukan pada lokasi geografis yang berbeda karena perbedaan iklim, level sosio-ekonomi, pengaruh budaya, penekanan program, dan kondisi lingkungan lainnya. Norma loka seringkali lebih membantu daripada norma nasional.
4. Apakah petunjuk administrasi dan penskoran tes cukup jelas untuk diikuti secara eksplisit? Jika tidak, norma menjadi tidak bermakna dan tidak boleh digunakan untuk tujuan perbandingan.
5. Kapan norma tersebut disusun? Norma lawas harus digunakan secara hati-hati karena ciri-ciri, karakteristik, dan kemampuan siswa saat ini banyak berbeda dari masa lampau. Lihat norma secara temporer dan revisi atau update norma secara reguler.

E. Administrasi Tes Mudah

Guru yang bekerja dalam batas waktu yang ketat seringkali tidak melakukan tes keterampilan karena mereka terlalu kesulitan untuk mengeset, yang dapat terjadi karena petunjuk tes kurang jelas atau karena hanya sedikit siswa yang dapat dites pada 1 waktu. Berikut ini adalah faktor-faktor yang mempengaruhi kemudahan administrasi tes:

1. Peralatan dan Personel

Tes yang memerlukan banyak peralatan seringkali tidak praktis. Akan menjadi terlalu lama untuk mengeset tes dan peralatan yang digunakan bisa jadi terlalu mahal. Tes yang berkualitas menyediakan petunjuk yang mengidentifikasi tipe dan jumlah peralatan dan bahan yang diperlukan untuk pelaksanaan tes.

Selanjutnya, Anda dapat menilai siswa secara lebih efisien jika terdapat dukungan bantuan. Merupakan hal yang sangat penting bahwa pengadministrasi tes dilatih untuk menjamin kesuksesan administrasi tes. Tes yang memerlukan banyak administrator akan menyusahkan dalam pelaksanaannya.

2. Ruang yang Diperlukan

Menata pos-pos tes memerlukan perencanaan yang matang untuk menyesuaikan dengan ruang yang tersedia berdasarkan ukuran kelas. Pengeluaran energi pada tiap pos, ketersediaan peralatan dan fasilitas adalah faktor-faktor yang harus dipertimbangkan saat merencanakan pos. Ruang yang diperlukan pada tes-tes tertentu akan membuat anda untuk mengetes pada hari berbeda.

3. Waktu Persiapan dan Administrasi Tes

Usahakan waktu persiapan tes seminimal mungkin. Susun dan organisasikan sebelum administrasi tes karena waktu dan disiplin bisa hilang jika persiapan terjadi pada waktu kelas sudah dimulai. Gunakan bantuan personel untuk menyiapkan dan melaksanakan tes. Tes yang memerlukan waktu persiapan dan administrasi lama menjadi kurang efektif dalam pengadministrasiannya.

Pertimbangkan pertanyaan-pertanyaan berikut ketika mengevaluasi kemudahan administrasi tes; meskipun jawabannya tidak baik sehingga perlu penyesuaian tanpa mempengaruhi validitas dan reliabilitas tes:

- a. Dapatkah tes mengakomodasi sejumlah besar siswa?
- b. Apakah terdapat petunjuk standard?
- c. Apakah petunjuknya jelas dan sederhana?
- d. Berapa banyak ruang yang dibutuhkan?
- e. Apakah tes tersebut wajar terkait kebutuhan fisik yang diperlukan siswa? (siswa tidak boleh terlalu lelah atau terlalu sakit setelah pelaksanaan tes)
- f. Berapa banyak waktu persiapan yang diperlukan?

- g. Apakah peralatan yang diperlukan dalam tes wajar?
- h. Apakah asisten dibutuhkan untuk mengadministrasi tes?
- i. Apakah tes memerlukan minimal praktik/latihan?
- j. Apakah keterampilan yang diteskan sesuai dengan keterampilan yang diperlukan siswa dalam situasi permainan?

F. Kesesuaian Usia dan Jenis Kelamin

Tes keterampilan harus spesifik pada usia, jenis kelamin, level keterampilan, kekuatan, dan variabel-variabel lain pada siswa anda. Tes harus mempertimbangkan perbedaan antara laki-laki dan perempuan dan tidak bias antara satu dan yang lain. Tes yang bagus untuk siswa sekolah menengah pertama tidak dapat diasumsikan dapat bekerja dengan baik untuk siswa sekolah menengah atas.

G. Nilai Pendidikan

Pengetesan yang hanya sekedar pengetesan harus dihindari. Tes harus relevan dengan unit yang diajarkan, tujuan siswa/guru, dan pengalaman belajar yang dirancang untuk mencapai tujuan (Johnson & Nelson, 1986). Tes harus mensyaratkan siswa menggunakan bentuk dan teknik yang benar, mengikuti peraturan aktivitas, dan menampilkan keterampilan yang berhubungan dengan aktivitas. Melalui pengetesan, siswa harus belajar sesuatu mengenai dirinya dan variabel yang diteskan.

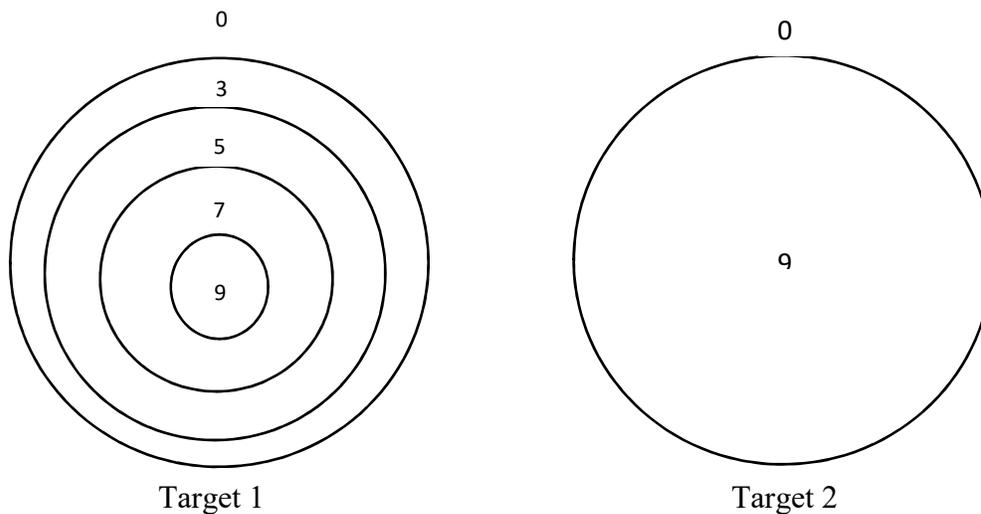
Atribut afektif yang didapat dari partisipasi pada pendidikan jasmani dan olahraga mencakup pengembangan kejujuran, sportivitas, harga diri, integritas, dan keadilan. Beberapa situasi tes membuat siswa mengetes dirinya sendiri atau menghitung temannya. Selama situasi tes khusus ini, guru harus menekankan dan mengembangkan perilaku positif pada siswanya. Meskipun kecurangan sangat mudah dilakukan dan sulit dideteksi, namun penekanan guru pada kejujuran dan integritas dapat mengurangi atau bahkan menghilangkan kecurangan.

H. Diskriminasi (Pembeda)

Diskriminasi dalam pengetesan adalah kemampuan tes untuk membedakan siswa dalam berbagai tingkatan kemampuan. Guru dapat memilih tes yang cukup sulit sehingga tidak ada siswa yang mendapat nilai sempurna, tetapi juga cukup mudah sehingga tidak ada yang mendapat nilai nol (Hastad & Lacy, 1989).

Idealnya, skor harus berada dari rentang tinggi ke rendah sehingga pemisahan level keterampilan siswa menjadi jelas. Salah satu cara untuk memastikan bahwa evaluasi

keterampilan berjalan adil adalah melalui ilustrasi nilai poin pada target memanah. Pada target 1, Joni dapat memanah pada skor 9 sebanyak 10 kali sehingga mendapat nilai 90. Sementara pada target 2, Reno memanah skor 9 sebanyak 10 kali sehingga mendapat poin 90 juga. Namun jika kita perhatikan, tingkat kesulitan dan diskriminasi pada target 1 dan target 2 jelas berbeda. Jika memanah pada target 1, belum tentu Reno dapat memperoleh nilai 90; bahkan mungkin hanya mendapat poin 50. Namun memang menemukan tes yang betul-betul mendiskriminasi penampilan seluruh siswa adalah hal yang sulit.



Gambar 3. Contoh diskriminasi penskoran

I. Keamanan (*safety*)

Sebelum menggunakan tes apapun, kajilah apakah tes terlalu memforsir siswa atau dapat mengakibatkan cedera. Sebelum memulai tes fisik, siswa harus diijinkan melakukan pemanasan yang cukup, latihan, dan mempraktikkan item tes. Anak yang lebih muda tidak dapat diharapkan tampil pada intensitas yang sama dengan anak yang lebih tua. Bahkan dalam kelompok umur yang sama, pertimbangkan mengenai tinggi badan dan berat badan siswa.

Sebelum mengetes, ceklah peralatan dan fasilitas yang rusak atau berbahaya. Peralatan yang rusak atau patah tidak boleh digunakan. Peralatan penanda haruslah lembut, tidak terpatahkan, mudah dilihat, dan sesuai untuk tes. Misalnya, jangan gunakan kursi sebagai pengganti cones pada tes dribble bolabasket.

J. Tipe Tes

Tes kecakapan psikomotor dapat diukur dengan 3 cara: simulasi permainan, penampilan saat game, dan *rating scale*. Dari ketiga item tersebut, banyak guru penjas yang memilih

rating scale dalam menilai penampilan siswanya. Namun, dari ketiga item tersebut, *rating scale* memiliki skala objektivitas yang paling rendah sehingga pertanyaan akan validitasnya semakin besar (Baumgartner & Jackson, 1982).

Tujuan utama pengetesan adalah untuk menentukan apakah tujuan pendidikan sudah tercapai. Untuk mencapai tujuan tersebut, tes haruslah paralel dengan tujuan pembelajaran. Merupakan hal yang umum bahwa guru mengembangkan tujuan pembelajaran sesuai dengan kebutuhan, sehingga guru seringkali harus mengembangkan alat sendiri untuk mengevaluasi apakah tujuan pembelajarannya tercapai.

Banyak guru yang komplain bahwa tes keterampilan olahraga lebih merepresentasikan kondisi simulasi permainan daripada situasi game sesungguhnya. Hal ini seringkali benar; akan tetapi, suatu tes tidaklah selalu harus identik dengan lingkungan permainan sepanjang aktivitas pada tes mirip dengan olahraga yang sesungguhnya (Baumgartner & Jackson, 1982). Misalnya, tes servis pada AAHPERD Tennis Skill Test mengkondisikan siswa untuk menempatkan 16 servis melewati net ke area servis yang sesuai. Tes tersebut tidak sama dengan kondisi permainan, namun penampilan pada item tes mirip dengan yang dibutuhkan saat permainan.

Tes keterampilan olahraga yang digunakan untuk mengevaluasi kecakapan terbagi menjadi 6 kategori: (1) tes waktu, (2) tes jarak, (3) tes akurasi, (4) tes power, (5) tes pergerakan tubuh, dan (6) *form test*.

1. Tes waktu

Tes ini menggunakan peralatan pengukur waktu untuk mengevaluasi durasi suatu keterampilan (misal tes sprint) atau reprints dari keterampilan yang sama pada periode waktu tertentu (misal wall-volley). Safrit (1990) percaya bahwa ketika suatu objek dipukul ke udara (misal pada wall-volley), penggunaan waktu sebagai ukuran kecakapan menjadi dipertanyakan, namun pengukuran waktu adalah hal yang tepat untuk olahraga ground atau floor.

2. Tes jarak

Tes jarak menggunakan pita pengukur untuk mengukur seberapa jauh seseorang melempar (misal lemparan softball, tendangan football) atau mendorong suatu objek atau tubuh (misal lompat jauh). Tipe tes ini secara umum reliabel karena jarak objek berpindah dapat diukur secara akurat (Baumgartner & Jackson, 1982). Namun, Safrit (1990)

memperingatkan bahwa melempar memerlukan usaha dan akurasi dan bahwa tes kemampuan melempar harus mengukur kecepatan dan akurasi.

3. Tes akurasi

Tes akurasi menggunakan target untuk mengevaluasi seberapa akurat seseorang melempar (misal pada tes free throw bolabasket), memukul (pada servis tenis), atau menendang (pada tendangan penalti sepakbola). Banyak sistem penskoran tes akurasi yang gagal membedakan berbagai level keterampilan. Untuk mengatasi masalah ini dan meningkatkan reliabilitas tes akurasi, anda harus menyediakan target dengan suatu rentang skor (misalnya target dengan rentang skor 1 – 5 akan membedakan kecakapan keterampilan lebih baik daripada target dengan rentang skor 0 – 2) dan jumlah percobaan yang cukup (Baumgartner & Jackson, 1982).

4. Tes power atau kecepatan

Tes power atau kecepatan mengukur komponen kecepatan, sudut proyeksi, dan jarak objek. Tes kecepatan digunakan untuk mengukur keterampilan yang digunakan untuk memproyeksikan suatu objek ke udara dan merefleksikan usaha yang digunakan. Tes kecepatan dapat menggunakan tali untuk mengukur tinggi lintasan, dapat menilai sudut vertikal dari proyeksi untuk menentukan kategori tinggi lintasan, dapat menggunakan stopwatch dan target dinding untuk mengukur daya, atau menggunakan peralatan elektronik canggih seperti *velocimeter* atau *radar gun* (Safrit, 1990).

5. Tes pergerakan tubuh

Tes ini mirip dengan tes waktu, mengharuskan subjek menyelesaikan tes standar secepat mungkin menggunakan karakteristik gerak yang sesuai dengan suatu cabang olahraga. Misalnya, AAHPERD *Basketball Dribble Test* mengukur kecepatan pemain dalam mendribble bola melalui alur yang ditentukan (Baumgartner & Jackson, 1982).

6. *Form Tests*

Form test biasanya berbentuk *rating scale* atau *checklist*, mengukur bagaimana suatu keterampilan dieksekusi. Tes ini sering membantu mengukur pemula dan membantu aktivitas yang diskor secara subjektif (senam, tari, renang indah, dan aerobik). Ketika atlet lebih ahli dari level pemula, form evaluasi haruslah menggunakan perlombaan; dalam hal ini banyak individu telah mengembangkan polanya sendiri untuk melakukan keterampilan (Safrit, 1990).

BAB IV
PROSEDUR PENYUSUNAN TES
PENDIDIKAN JASMANI OLAHRAGA DAN KESEHATAN (PJOK)

Banyak guru penjas yang tidak menggunakan tes yang telah disusun sebelumnya karena menemui bahwa tidak ada tes terpublish yang memenuhi kebutuhan spesifik mereka. Guru juga mengeluhkan bahwa tes yang tersedia tidak sesuai karena tidak mengukur kecakapan keterampilan selama kondisi permainan. Berdasarkan hal tersebut, bab ini menyediakan panduan untuk membantu anda dalam memilih dan mengkonstruksi tes kecakapan motorik yang sesuai, yang terdiri dari 10 langkah: (1) mereview kriteria tes yang baik, (2) menganalisis cabang olahraga, (3) mereview literatur, (4) memilih atau mengkonstruksi item tes, (5) menetapkan prosedur, (6) melakukan peer review, (7) melakukan pilot study, (8) menentukan validitas, reliabilitas, dan objektivitas, (9) membuat norma, dan (10) menyusun manual tes²

A. Mereview Kriteria Tes yang Baik

Sebelum mengkonstruksi tes, anda harus terlebih dahulu mengetahui kriteria tes yang baik. Paparan mengenai kriteria tes yang baik dapat dilihat pada bab III yang terdiri dari: (1) validitas, (2) reliabilitas, (3) objektivitas, (4) norma, (5) kemudahan administrasi tes, (6) kesesuaian usia dan jenis kelamin, (7) nilai pendidikan, (8) diskriminasi tes, (9) keamanan, dan (10) tipe tes.

B. Menganalisis Cabang Olahraga

Langkah kedua dalam menyeleksi dan mengkonstruksi tes adalah menganalisis olahraga atau aktivitas yang akan dites. Langkah analisis ini sangat penting baik untuk perencanaan pembelajaran maupun untuk penyusunan tes. Untuk memulainya, diperlukan daftar dan deskripsi keterampilan-keterampilan yang penting untuk bermain olahraga secara terampil (Phillips & Hornak, 1979; Safrit, 1990). Misalnya, permainan bolabasket melibatkan dribbling, jump shot, free throw, lay up, rebounding, passing, dan defense. Guru harus menekankan keterampilan-keterampilan yang sesuai dengan usia siswa. Setelah keterampilan-keterampilan yang sesuai dengan usia siswa diidentifikasi, reviewlah tujuan pembelajaran untuk menentukan mana yang memiliki bobot terbesar. Merupakan hal yang

² Strand, Bradford N. & Wilson, R. (1993). *Assessing Sport Skills* halm 9-22. Champaign: Human Kinetics Publishers.

tidak fair dan etis ketika guru mengharapkan siswa menampilkan suatu keterampilan dengan baik sementara teknik tersebut hanya diajarkan selama satu jam pelajaran saja.

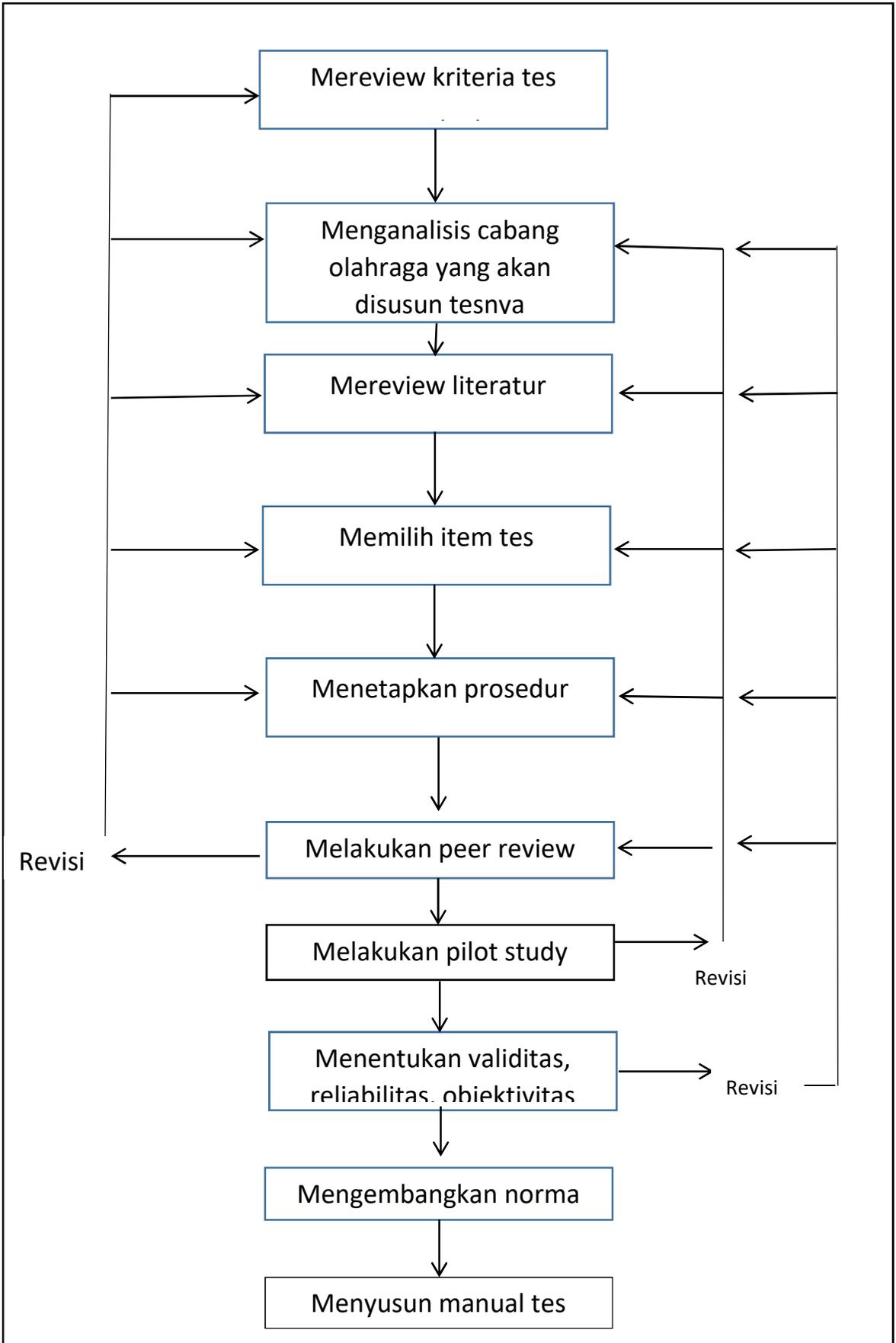
Jadi pada proses analisis cabang olahraga guna keperluan perencanaan pembelajaran dan penyusunan tes, lakukan langkah berikut:

1. Analisislah semua keterampilan yang diperlukan pada cabang olahraga tertentu.
2. Analisislah keterampilan-keterampilan yang sesuai dengan usia siswa.
3. Berdasarkan waktu pembelajaran yang ada, pilihlah keterampilan-keterampilan apa yang akan ditekankan. Misalnya alokasi waktu pembelajaran 3 kali pertemuan tidak mungkin untuk mengajarkan semua teknik. Maka anda perlu menyortir teknik-teknik dasar terpenting.
4. Penyusunan item tes dilakukan berdasarkan item teknik yang ditekankan pada poin 3.

C. MEREVIEW LITERATUR

Setelah guru memilih keterampilan-keterampilan yang sesuai dan menuliskan tujuan pembelajaran, ia harus mereview literatur tes-tes keterampilan yang mengukur keterampilan atau penampilan yang sama dengan apa yang akan ia lakukan (Miller,1988; Safrit, 1990). Jika tes terkait materi yang diajarkan guru tersedia dan valid, reliabel, dan mudah dilakukan, maka guru harus menggunakannya. Jika tes tidak tersedia, maka guru harus mengembangkannya.

Banyak guru mengeluh bahwa mereka tidak tahu kemana untuk menemukan tes-tes keterampilan. Secara umum, sumber pertama adalah buku teks pengukuran dan evaluasi. Sayangnya, di Indonesia buku-buku pengukuran dan evaluasi penjas masih terbatas. Anda dapat lebih mudah menemukan dalam versi bahasa Inggris. Demikian pula, banyak tes-tes yang telah disusun namun tersebar dalam berbagai jurnal yang sulit untuk diakses. Pun, sulit untuk mengaksesnya secara online. Oleh karena itu, pencarian sumber tes dapat diupayakan melalui perpustakaan kampus ataupun jurnal-jurnal internasional yang dilanggan oleh Universitas, misalnya jurnal berikut: *Research Quarterly for Exercise and Sport*; *The Physical Educator*; *Journal of Physical Education, Recreation, and Dance*; dan disertasi. AAHPERD menawar beberapa manual tes dalam bolavoli, bolabasket, tenis, *softball*, *football*, dan panahan. Ketika sudah *mereview literatur* dan guru tidak menemukan tes baku yang baik, maka guru harus menyusunnya.



Gambar 1. Sepuluh tahap (*flow-chart*) dalam menyusun tes PJOK

Sumber: Strand, Bradford N. & Wilson, R. (1993). *Assessing Sport Skills hal. 10*. Champaign: Human Kinetics Publishers.

D. MEMILIH ATAU MENGKONSTRUKSI ITEM TES

Panduan dalam mengkonstruksi tes adalah sebagai berikut:

1. Item tes harus mewakili keterampilan yang dibutuhkan nyata pada permainan (Kirkendall et al, 1987). Misalnya, servis adalah keterampilan tertutup dimana seseorang tampil pada lingkungan yang tidak berubah dengan performer yang melakukan gerak yang relatif sama setiap waktu. Di sisi lain, shooting adalah keterampilan terbuka dimana seseorang tampil pada lingkungan yang berubah-ubah sehingga performer perlu menyesuaikan diri sesuai situasi yang ada. Jadi, tes servis harus dilakukan dari posisi stasioner sementara tes shooting dilakukan dari posisi bergerak.
2. Item tes harus mudah dimengerti dan relatif mudah untuk ditampilkan (Miller, 1988). Jika item tes memerlukan detail yang rumit, maka kebanyakan siswa dapat melupakan beberapa aspek dan menimbulkan delay yang tidak penting pada tes. Selain itu, jika siswa tidak melakukan tes sebagaimana deskripsi tes, maka hal tersebut mempengaruhi validitas dan reliabilitas tes.
3. Untuk kepraktisan, item tes haruslah tidak mahal, mudah diadministrasi, mudah disiapkan, dan dibentuk dalam station (Miller, 1988). Waktu adalah hal yang paling utama bagi kebanyakan guru penjas dan tes yang membutuhkan banyak pengukuran dan penanda adalah hal yang tidak praktis. Terlalu banyak garis, tanda, dan zona skor pada lantai atau lapangan akan membingungkan administrator, tester, skorer, dan recorders.

E. MENETAPKAN PROSEDUR

Setelah item tes dipilih, tentukanlah lay out tes, zona skor, arah pelaksanaan, dan prosedur administrasi (Mileer, 1988). Kesimpulan dalam tiap aspek akan mengurangi waktu persiapan, waktu administrasi, waktu latihan, dan membuat siswa lebih baik dalam memahami tes. Pada titik ini, trial dan error terjadi karena guru mencoba untuk menentukan prosedur terbaik. Revisi dan adaptasi tidak terelakkan dan mambantu menjelaskan dan meningkatkan kualitas tes.

F. MELAKUKAN *REVIEW* TEMAN SEJAWAT

Suatu tes yang baru perlu mendapat review dari teman sejawat, kolega, ahli, atau siswa (Miller, 1988). Proses review ini ditujukan untuk memperoleh kritik yang konstruktif secara objektif guna memperhalus atau memperbaiki rancangan tes. Ketika seseorang mengembangkan sesuatu hal yang baru, subjektivitas personal seringkali mempengaruhi logika. Merupakan hal yang mudah untuk terjebak dalam prosedur teknis dan melupakan hal-hal yang simple dan sebaliknya.

Saran dari para reviewer tidak perlu secara eksplisit diikuti semua. Beberapa saran mungkin tidak layak dan saran yang lain mungkin menambah waktu yang diperlukan dalam melakukan tes. Saran yang lain, bisa jadi menunjukkan poin penting yang terlewatkan. Dengarkan semua saran secara terbuka, review saran tersebut secara kritis, dan putuskan saran apa dan bagaimana yang akan digunakan. Proses review ini mungkin akan membawa anda kembali pada tahap-tahap sebelumnya karena diperlukan proses perbaikan. Pada tahap manapun yang anda ulangi, buatlah perubahan berdasarkan rekomendasi reviewer, dan lakukan proses selanjutnya sesuai dengan tahapan-tahapan yang perlu dilakukan. Misalnya anda mengulangi dari tahap keempat, maka anda perlu melanjutkan tahap kelima, dan melakukan review teman sejawat lagi di tahap keenam dan seterusnya.

G. MELAKUKAN *PILOT STUDY*

Langkah ketujuh dalam proses penyusunan tes adalah melakukan pilot study, yang bertujuan untuk melihat apakah masih ada masalah-masalah pada tes dalam aspek administrasi, persiapan, penskoran, dan/atau petunjuk tes. Jika pilot study menunjukkan masalah pada tes, maka penyusun tes harus kembali pada tahap yang sesuai, melakukan revisi tambahan, dan mengulang kembali tahapan-tahapan selanjutnya. Pilot study dapat mengungkap misalnya, bahwa sistem penskoran terlalu sulit atau terlalu mudah, atau bahwa petunjuk tes terlalu rumit atau terlalu umum. Tanpa pilot study, kekurangan tes tidak akan nampak sampai tes diterapkan pada siswa dalam kelompok besar dengan hanya 1 guru dan tempat tes yang terbatas (Miller, 1988; Safrit, 1990).

H. MENENTUKAN VALIDITAS, RELIABILITAS, DAN OBJEKTIVITAS

Hal yang perlu diingat pada tahap ini adalah bahwa tes yang memiliki validitas tinggi selalu reliabel, namun tes yang reliabel belum tentu selalu valid. Hal ini karena reliabilitas

menunjukkan derajat suatu instrumen mengukur secara konsisten apa yang ingin diukur. Meskipun suatu instrumen reliabel, tes tersebut mungkin tidak mengukur apa yang seharusnya diukur.

Karena pengetesan memiliki tujuan-tujuan berbeda dan karena validitas dievaluasi terkait dengan tujuannya, terdapat beberapa tipe validitas. Secara umum, validitas terdiri dari: (1) validitas isi, (2) validitas muka, (3) validitas kriteria, dan (4) validitas konstruk. Masing-masing tipe validitas digunakan untuk menyediakan data yang menunjukkan bahwa tes mengukur apa yang seharusnya diukur.

Selanjutnya, sama halnya dengan validitas, reliabilitas dapat ditentukan dengan beberapa cara. Reliabilitas dikategorisasikan menjadi: (1) reliabilitas tes-retest, (2) reliabilitas bentuk sejajar/paralel, dan (3) reliabilitas belah dua atau ganjil-genap. Karena reliabilitas lebih mudah untuk dicapai daripada validitas, koefisien reliabilitas diekspektasikan lebih tinggi dari koefisien validitas pada tes keterampilan yang sama.

Terkait dengan objektivitas, penentuannya sering dilakukan bersamaan dengan waktu penentuan reliabilitas. Objektivitas diperoleh dengan menggunakan dua atau lebih testor yang mengadministrasi suatu tes pada grup siswa yang sama (Miller, 1988). Koefisien objektivitas akan ditunjukkan oleh jumlah kecocokan di antara para administrator tes. Penjelasan lebih lengkap mengenai jenis-jenis validitas, reliabilitas, dan objektivitas akan dibahas pada bab selanjutnya.

I. MENGEMBANGKAN NORMA

Ketika suatu tes mencapai validitas dan reliabilitas yang baik, merupakan hal yang umum untuk menetapkan norma atau standar yang berlaku secara lokal, regional, atau nasional. Norma-norma harus disusun untuk laki-laki dan perempuan dan untuk kelompok usia yang berbeda (Safrit, 1990). Untuk menyusun norma misalnya, anda meminta 20 guru penjas untuk melakukan pengetesan pada semua siswa di kelas yang mereka ajar selama 2 tahun. Setelah 2 tahun mentabulasi data, informasi dianalisis dan norma ditetapkan. T scores, percentiles, dan skala peringkar yang direkomendasikan diukur.

J. MENYUSUN MANUAL TES

Sebagai tahap final penyusunan tes, persiapkan manual tes yang mencakup informasi seperti *overview* terhadap cabang olahraga; sejarah pengetesan pada olahraga; instruksi dalam menggunakan manual tes; deskripsi tes dengan petunjuk tes; diagram, dan prosedur penskoran; *drill* untuk mengembangkan teknik yang benar; norma; formulir pencatat; dan referensi.

BAB V

PENDEKATAN ACUAN PENILAIAN DALAM PJOK

Dilihat dari perencanaan tes dan penafsiran hasil tes, menurut Mardapi (2004: 23) ada dua macam pendekatan acuan penilaian yang lazim dipakai dalam menentukan nilai hasil belajar pendidikan jasmani siswa di sekolah, yaitu: Pendekatan Acuan Norma (PAN) dan Pendekatan Acuan Kriteria (PAP). Kedua acuan ini menggunakan asumsi yang berbeda tentang kemampuan siswa. Bahkan ada pula yang menggunakan pendekatan penilaian yang mengkombinasikan kedua acuan penilaian tersebut dalam memberikan nilai akhir mata pelajaran pendidikan jasmani kepada para siswanya.

A. PENDEKATAN ACUAN PENILAIAN PJOK

(1). Penilaian Acuan Norma (*Norm-Referenced Measured*)

Penilaian acuan norma berasumsi bahwa kemampuan siswa itu berbeda dan dapat digambarkan menurut distribusi normal. Perbedaan itu harus ditunjukkan oleh hasil pengukuran, misalnya: setelah siswa mengikuti pelajaran satu semester, kemudian siswa diberikan tes. Hasil tes siswa ini kemudian dibandingkan dengan kelompoknya sehingga akan diketahui posisi siswa tersebut. Acuan ini digunakan terutama pada tes untuk seleksi.

Penilaian acuan norma memungkinkan penafsiran prestasi siswa dikaitkan dengan prestasi siswa lain yang juga menempuh tes yang sama. Satu cara sederhana untuk menggunakan penafsiran penilaian yang beracuan norma adalah membuat rangking skor dari skor tinggi ke skor yang rendah dan menentukan dimanakah skor individu berada.

Menurut Phillips (1979: 70-71) ada beberapa faktor penting yang perlu dipertimbangkan apabila guru pendidikan jasmani menggunakan dasar penilaian hasil belajar siswa yang beracuan norma. Adapun faktor tersebut di antaranya adalah: (1) sampel, (2) administrasi/*administration*, (3) bersifat mewakili/*representativeness*, (4) bersifat temporer/*temporariness*, (5) memiliki petunjuk/*presentation*, (6) dapat dibandingkan/*comparability*.

Sampel yang dipergunakan untuk menyusun norma harus mencakup jumlah kasus yang besar. Biasanya semakin banyak sampel akan semakin dapat mewakili populasinya. Prosedur pengambilan sampel harus didasarkan pada suatu distribusi populasi yang luas. Jika norma-norma nasional akan dikembangkan, distribusi geografi merupakan suatu faktor

penting yang perlu dijadikan bahan pertimbangan selain faktor usia, jenis kelamin, ras, tingkat pendidikan, status sosial ekonomi, serta teknik pengambilan sampel.

Administrasi tes dan sistem pencatatan hasil harus dibuat sedemikian rupa sehingga sudah baku. Petunjuk penyusunan norma harus dibuat dengan jelas sehingga dalam pelaksanaan tes tidak timbul keragu-raguan baik dari pihak guru maupun siswa. Petunjuk penilaian tes hasil belajar siswa harus tegas dan relatif sederhana.

Norma-norma harus dapat mewakili populasi untuk siapa tes itu dibuat. Untuk penyusunan norma yang berskala nasional prosedur teknik pengambilan sampel merupakan salah satu faktor yang sangat penting untuk dijadikan sebagai bahan pertimbangan. Dalam mengevaluasi norma-norma tes, harus diberikan penekanan pada teknik pengambilan sampel dan ukuran sampel. Makin banyak jumlah siswa yang memiliki karakteristik yang sama dengan karakteristik dimana norma tersebut dibuat, maka semakin pasti bahwa norma-norma tersebut merupakan suatu sumber yang dapat dipercaya sebagai bahan pembanding.

Norma penilaian yang dipakai saat ini bersifat **temporer** dan harus selalu ditinjau ulang secara periodik agar tidak usang (*out of date*). Contoh: norma-norma pada tes kebugaran jasmani remaja putri dari AAHPER pertama kali dikembangkan pada tahun 1958 kemudian di revisi dua kali yaitu tahun 1963 dan tahun 1975. Tujuan revisi tes AAHPER ini mengindikasikan bahwa norma-norma tes tersebut bersifat temporer.

Norma penilaian hasil belajar siswa hendaknya disajikan kedalam suatu format yang dapat dipahami dengan mudah. Tipikal norma penilaian dalam pendidikan jasmani yang mudah dipergunakan di antaranya: skor standard, norma persentil, dan juga norma-norma nilai dan usia.

Seringkali guru pendidikan jasmani perlu membandingkan skor-skor dari ter-tes yang berbeda untuk mengevaluasi prestasi siswa. Norma-norma pada tes yang terpisah harus komparabel. Dalam rangkaian tes kebugaran jasmani remaja putri dari AAHPER dimana setiap item tesnya menggunakan sampel yang sama dalam mengembangkan norma-normanya. Tanpa prosedur demikian maka norma-norma tes kebugaran jasmani bagi remaja putri dari AAHPER tersebut tidak representatif dan dipermasalahkan keabsahannya.

Pendekatan acuan norma (PAN), pada dasarnya menggunakan kurva normal dan hasil-hasil perhitungan statistika sebagai dasar penilaian. Ukuran statistika yang digunakan adalah nilai rata-rata (*mean*) dan simpangan baku (SB) yang diperoleh dari hasil kelompok.

Penilaian acuan norma bersifat relatif, karena berpautan pada besarnya nilai rata-rata atau simpangan baku yang diperoleh dari kelompok itu. Jika hasil perhitungan nilai rata-ratanya naik atau lebih tinggi, maka patokan menjadi bergeser ke atas. Namun sebaliknya jika nilai rata-ratanya menurun maka patokan penilaian menjadi bergeser ke bawah.

Kurva normal dibangun di atas sumbu datar X, dengan grafik kurvanya berbentuk genta dan simetris. Nilai rata-rata itu terletak pada titik sentral dari kurva itu dibagi dua, belahan yang sama (simetris). Satuan nilai yang dipakai dalam kurva normal umumnya yaitu rata-rata dan simpangan baku.

Contoh: hasil tes pengetahuan umum penjaskes 20 siswa

Tabel 3: hasil tes pengetahuan umum penjaskes

Skor Tes (X)	$(X_i - \bar{X})$	$(X_i - \bar{X})^2$
25	0	-
29	4	16
22	-3	9
21	-4	16
20	-5	25
26	1	1
27	2	4
23	-2	4
28	3	9
29	4	16
31	6	36
28	3	9
24	-1	1
20	-5	25
22	-3	9
30	5	25
26	1	1
25	0	-

24	-1	1
20	-5	25
$\Sigma = 500$	25	227

Dari hasil perhitungan data tersebut diperoleh hasil sebagai berikut:

Diketahui;

$$\begin{aligned} \Sigma &= 500 \\ \Sigma (X_i - \bar{X})^2 &= 227 \\ \bar{X} &= 25 \\ \text{Simpangan baku} &= 3,45 \end{aligned}$$

Nilai hasil tes pengetahuan umum penjasokes dari 20 siswa tersebut kemudian dibuat ke dalam norma penilaian dengan menggunakan dasar kurva normal, rerata (*mean*), dan simpangan baku (SB). Dalam contoh ini dibuat skala 1 – 10, sehingga skala kurva normal, dirancang dengan standar nilai 1 – 10. Adapun norma penilaian berdasarkan penilaian acuan norma dengan menggunakan standar penilaian 1 – 10 dapat digambarkan pada tabel 2.

Tabel 4 : Penilaian Acuan Norma (standar 1 – 10)

Skala	Batas Skor	Rentang Skor	Nilai
$\bar{X} + 2,4$ (S)	$25 + 2,4$ (3,5) = 33	33 ke atas	10
$\bar{X} + 1,8$ (S)	$25 + 1,8$ (3,5) = 31	31 – 32	9
$\bar{X} + 1,2$ (S)	$25 + 1,2$ (3,5) = 29	29 – 30	8
$\bar{X} + 0,6$ (S)	$25 + 0,6$ (3,5) = 27	27 – 28	7
$\bar{X} + 0,0$ (S)	$25 + 0,0$ (3,5) = 25	25 – 26	6
$\bar{X} - 0,6$ (S)	$25 - 0,6$ (3,5) = 23	23 – 24	5
$\bar{X} - 1,2$ (S)	$25 - 1,2$ (3,5) = 21	21 – 22	4
$\bar{X} - 1,8$ (S)	$25 - 1,8$ (3,5) = 19	19 – 20	3
$\bar{X} - 2,4$ (S)	$25 - 2,4$ (3,5) = 17	17 – 18	2
		16 ke bawah	1

Mengacu pada tabel 4 di atas, apabila Rifki mendapatkan skor 29, maka nilai Rifki berdasarkan penilaian acuan norma dengan standar 1 – 10 adalah = 8

Selain Penilaian Acuan Norma dengan menggunakan standar penilaian 1 – 10, Pendekatan Acuan Norma dapat dikembangkan pula dengan menggunakan 5 kategori, yaitu:

Tabel 3: Penilaian Acuan Norma dengan 1 - 5 kategori

Rentangan Norma	Nilai	% (teoritis)
$\bar{X} + 1,5 \text{ SD}$ Ke atas	A atau 4	6,68 %
$\bar{X} + 0,5 \text{ SD s/d } < \bar{X} + 1,5 \text{ SD}$	B atau 3	24,47 %
$\bar{X} - 0,5 \text{ SD s/d } < \bar{X} + 0,5 \text{ SD}$	C atau 2	38,30 %
$\bar{X} - 1,5 \text{ SD s/d } < \bar{X} - 0,5 \text{ SD}$	D atau 1	24,47 %
Kurang dari $\bar{X} - 1,5 \text{ SD}$	E atau 0	6,68 %

Pendekatan evaluasi dengan menggunakan Pendekatan Acuan Norma memiliki kelemahan-kelemahan, yaitu: tidak ada standar kelulusan yang pasti untuk siswa karena tingkat kelulusan siswa sangat dipengaruhi oleh berapa besarnya rerata kelompoknya, gambaran tingkat penguasaan materi siswa belum kita dapatkan (*mastery learning of student*), terlebih jika besarnya nilai rerata kelompoknya rendah, siswa cenderung puas jika memperoleh status lulus walaupun secara penguasaan materi belum tuntas. Sedangkan kelebihan yang dimiliki oleh PAN di antaranya, yaitu: kebanyakan tingkat kelulusan/keberhasilan siswa tinggi sehingga secara psikologis sangat menguntungkan siswa, guru tidak perlu menyediakan program remedial bagi siswa.

(2) Penilaian Acuan Patokan (*Criterion- Referenced Measured*)

Penilaian acuan patokan adalah penilaian yang membandingkan hasil belajar siswa kepada patokan yang telah ditetapkan sebelumnya. Sebelum penilaian dilaksanakan, terlebih dahulu ditetapkan patokan yang harus dipakai untuk membandingkan skor hasil pengukuran sehingga skor itu bermakna. Asumsi penilaian acuan patokan adalah: bahwa hampir semua orang bisa belajar apa saja namun waktunya yang berbeda. Konsekuensi dari penilaian acuan patokan (PAP) adalah adanya program

remedial. Penilaian acuan patokan memiliki kelebihan yaitu bermanfaat untuk menjajaki tingkat penguasaan materi pelajaran secara tuntas (*mastery of learning*). Dengan demikian kualitas lulusannya dapat terkendali dengan baik.

Patokan ini ditetapkan atas dasar pertimbangan logis mengenai tingkat penguasaan minimum. Para siswa yang mencapai atau melebihi patokan ini dinyatakan lulus, sedangkan para siswa yang belum berhasil mencapai “**batas lulus**” tersebut dinyatakan tidak lulus. Hal ini berarti siswa-siswa tersebut dianggap belum menguasai secara minimum kemampuan tersebut.

Dengan demikian patokan yang digunakan dalam penilaian acuan patokan ini bersifat tetap. Berbeda dengan penilaian acuan norma yang bersifat relatif. Penilaian acuan patokan diterapkan untuk mengetahui tingkat penguasaan minimum suatu mata pelajaran. Penentuan batas penguasaan minimum ini ditetapkan para ahli dalam bidangnya. Penetapan batas penguasaan minimum itu bermacam-macam, mungkin 55% 60% atau 65% dari jumlah skor minimum tes itu.

Menurut Safrit (1986: 466) metode yang seringkali dipergunakan untuk penilaian acuan patokan dalam pendidikan jasmani adalah dengan metode persentase. Adapun metode pemberian nilai dengan sistem persentase akan disajikan pada tabel di bawah ini.

Tabel 6: Contoh pemberian penilaian acuan patokan (PAP) dengan metode persentase

Tingkat Penguasaan Materi	Nilai
90% - 100%	A atau 4
80% - 89%	B atau 3
70 – 79%	C atau 2
60 – 69%	D atau 2
Kurang dari 60%	E atau 1

Misalnya: Ujian pengetahuan umum mata pelajaran penjaskes skor maksimumnya adalah 80, maka tingkat penguasaan materi keseluruhan adalah sebagai berikut: (a) 72 – 80 mendapat nilai A, (b) 64 – 71 mendapat nilai B, (c) 56 -63 mendapat nilai C, (d) 48 – 55 mendapat nilai D, dan (e) kurang dari 55 mendapatkan nilai E.

Berdasarkan tabel 3 di atas, seumpama Wibisono mendapatkan skor 74, maka Si Wibisono tersebut akan mendapatkan nilai A atau 4 karena nilai 74 berada pada interval kelas tingkat penguasaan materi antara 72 sampai 80.

BAB VI
SISTEM EVALUASI
DALAM PENDIDIKAN JASMANI OLAHRAGA DAN KESEHATAN (PJOK)

Sistem evaluasi adalah sistem pemberian nilai atau *grade* yang pada umumnya dilakukan oleh para guru penjasorkes pada tengah dan akhir semester. Nilai yang diberikan itu berdasarkan pada data yang diperoleh dari tes yang diberikan oleh guru penjasorkes kepada para siswa sesuai dengan tujuan-tujuan yang telah dirumuskan untuk dicapai dengan proses belajar yang direncanakan dan dilaksanakan oleh guru penjasorkes tersebut. Untuk dapat memperoleh data yang valid dan reliabel guru harus telah memiliki kemampuan untuk memilih dan menggunakan tes yang sesuai dengan tujuan yang ingin dicapai. Demikian pula jika tes yang diperlukan tidak ada atau tidak memenuhi syarat sebagai sebuah tes yang baik, maka guru harus memiliki kemampuan untuk membuat tes yang diperlukannya.

Nilai adalah simbol yang dipakai untuk menyatakan kemajuan dan pencapaian siswa dalam proses belajar-mengajar. Penilaian dalam pendidikan telah banyak menjadi perhatian dan juga kritikan dari hampir sebagian besar pendidik. Permasalahan yang berkaitan dengan perlunya penilaian tentang hasil belajar siswa tidak ada perbedaan pendapat di antara para pendidik, bahkan para pendidik sepakat bahwa peserta didik perlu mengetahui hasil belajar dan statusnya di dalam kelas. Namun yang menjadi akar permasalahannya adalah metode pemberian nilai atau penggunaannya dalam menilai hasil belajar siswa.

Menurut Safrit (1981: 327) ada dua kelemahan utama yang berkaitan dengan pemberian nilai (*grading*). (1) arti nilai belum didefinisikan secara jelas dan tepat. Nilai dapat bervariasi dalam kelas yang satu dengan kelas yang lain, dari satu pendidik ke pendidik yang lain. Contoh: nilai A untuk satu mata pelajaran tertentu barangkali sama dengan nilai C untuk mata pelajaran yang sama jenisnya. (2) nilai sering tidak berdasarkan pada bukti yang objektif karena adanya pengaruh "*halo effect*" yang sering berperan dalam pemberian nilai.

**A. STATUS DAN ESENSI PRAKTIK PEMBERIAN NILAI DALAM MATA
PELAJARAN PENJASORKES (PJOK)**

Safrit dan Wood (1989:309-310) menunjukkan suatu fakta bahwa selama 25-30 tahun belakangan ini hanya sedikit evaluasi yang sistematis untuk menilai pembelajaran pendidikan

jasmani. Bukti tersebut mereka peroleh berdasarkan kajian literatur yang telah mereka lakukan mengenai masalah-masalah yang berkaitan dengan penilaian dalam pembelajaran penjasorkes. Pernyataan senada juga dikemukakan oleh Fox (1959:) yang menyatakan bahwa hanya sedikit pengetesan yang digunakan untuk menilai kinerja siswa dalam pendidikan jasmani di Oregon. Coker (1972) menemukan fakta yang sama di kelas-kelas pendidikan jasmani sekolah negeri di Louisiana dimana jarang dijumpai penilaian penjasorkes terhadap hasil belajar siswa. Kedua peneliti tersebut melaporkan bahwa faktor kehadiran yang paling sering digunakan sebagai dasar untuk pemberian nilai dalam pelajaran pendidikan jasmani.

Di sekolah negeri di Colorado, Morrow (1978) menemukan hasil surveinya bahwa pemberian nilai dalam mata pelajaran penjasorkes terutama didasarkan pada partisipasi dan berpakaian seragam siswa. Lebih dari dua pertiga dari responden menyatakan bahwa mereka sedikit atau tidak mempunyai kesempatan untuk menggunakan teknik-teknik pengukuran yang mereka pelajari dalam mata kuliah pengukuran dan evaluasi pada saat dahulu mereka menempuh studi. Imwold, Rider dan Johnson (1982) telah mensurvei guru-guru pendidikan jasmani di sekolah-sekolah negeri di Florida dan mereka menemukan bahwa kurang lebih 60% responden menggunakan tes ketrampilan untuk menilai kinerja para peserta didik dan kurang lebih 40% menggunakan tes pengetahuan.

Hasil penelitian Hensley dkk (1978:) melaporkan bahwa ada satu peningkatan dalam penggunaan Penilaian Acuan Patokan (PAP) atau *criterion referenced evaluation*. 58% dari responden melaporkan penggunaan PAP dalam penilaian sedangkan 37% melaporkan menggunakan penilaian acuan normal (PAN) atau *norm referenced evaluation*. PAP lebih sering digunakan oleh para guru yang kurang pengalaman, sedangkan pendidik yang lebih berpengalaman menggunakan penilaian PAN. Responden juga diberi pertanyaan tentang faktor yang paling sering dipakai sebagai dasar pemberian nilai kepada siswa adalah (1) partisipasi siswa, (2) sikap siswa. Hensley juga menemukan bahwa pendidik perempuan lebih sering menggunakan tes ketrampilan dan pengetahuan dari pada pendidik laki-laki dalam pemberian nilai penjasorkes kepada siswa.

Dari hasil penelitian Safrit dan Wood dapat menyimpulkan bahwa pemberian nilai terhadap hasil belajar siswa bukan merupakan prioritas yang tinggi dari guru pendidikan jasmani olahraga dan kesehatan. Bagaimanakah pula status praktik pemberian nilai penjasorkes oleh

para guru penjasorkes di sekolah-sekolah negeri maupun swasta di Indonesia? Oleh sebab itu perlu diadakan survei atau penelitian untuk menjawab pertanyaan tersebut. Barangkali jawaban sementara yang dapat dikemukakan oleh penulis hasilnya mungkin tidak jauh berbeda dengan hasil temuan penelitian tentang penilaian penjasorkes di Amerika.

B. KEGUNAAN NILAI DALAM MATA PELAJARAN PENJASORKES (PJOK)

Tujuan pemberian nilai terhadap hasil belajar penjasorkes bagi siswa harus diketahui atau dengan kata-kata lain harus ada alasan mengapa seorang guru/pendidik memberi nilai. Arma Abdoellah (2001: 2) mengemukakan bahwa kegunaan umum pemberian nilai dalam penjasorkes sebagai berikut.

1. Informasi bagi orang tua mengenai status atau kemajuan siswa.
2. Promosi dan wisuda.
3. Motivasi untuk belajar.
4. Bimbingan belajar.
5. Bimbingan untuk perencanaan pendidikan dan vokasional.
6. Bimbingan untuk perkembangan pribadi.
7. Laporan dan rekomendasi bagi yang akan menjadi majikan siswa.
8. Laporan kepada sekolah yang nanti akan dimasukinya.

Adapun kegunaan pemberian nilai dalam penjasorkes di antaranya :

a. Bagi peserta didik

Dalam mempelajari satu keterampilan olahraga dalam mata pelajaran penjasorkes siswa segera dapat mengetahui tingkat keberhasilan dalam menguasai ketrampilan yang dipelajarinya dan dapat pula membandingkan keterampilan dirinya dengan keberhasilan teman-temannya dalam kelas.

Pemberian nilai yang dilakukan oleh guru penjasorkes setelah akhir semester dinamakan evaluasi sumatif. Jadi nilai tersebut dapat mencerminkan hasil akhir hasil proses hasil belajar siswa. Para siswa dapat membandingkan nilainya dengan nilai yang didapat oleh teman-temannya sehingga ia dapat mengetahui kedudukannya di dalam kelas. Nilai dapat digunakan siswa sebagai alat untuk memotivasi diri. Apabila mendapatkan nilai kurang maka siswa akan termotivasi untuk memperbaiki nilainya, namun apabila ia mendapat nilai baik ia

akan berusaha untuk mempertahankan nilai baik itu atau berusaha agar menjadi lebih baik lagi. Menurut Barrow dan Mc. Gee (1971:340) ada satu kekhawatiran yang berkaitan dengan pemberian nilai yang dilakukan oleh guru penjasorkes kepada peserta didik. Pemberian nilai hanya berorientasi pada pencapaian tujuan-tujuan yang telah ditetapkan. Karena nilai itu sendiri merupakan muara akhir yang lebih nyata dari pada tujuan yang diwakilinya. Nilai sering menjadi tujuan bagi siswa dari pada fungsi nilai sebagai alat untuk mencapai tujuan. Nilai menjadi objek utama yang dicari oleh siswa padahal ia harus lebih memberikan perhatian pada tingkat keberhasilannya dalam mencapai tujuan yang telah ditentukan sebelumnya.

b. Bagi orang tua

Nilai diperlukan oleh orang tua siswa karena orangtua dapat mengikuti kemajuan dan keberhasilan anak mereka. Nilai adalah media penting dalam menjalin hubungan antara sekolah dengan orang tua. Orang tua dapat diberi informasi mengenai tujuan pendidikan jasmani olahraga dan kesehatan dan fakta tentang kemajuan anak mereka mencapai tujuan tersebut. Nilai dapat pula berfungsi untuk memberi informasi tentang status anak mereka dalam pelajaran pendidikan jasmani olahraga dan kesehatan.

c. Bagi guru atau pendidik

Nilai dapat dimanfaatkan oleh guru/pendidik untuk berbagai tujuan. Pertama, nilai mendorong guru membuat evaluasi dari setiap siswa dengan baik, jadi memberikan satu pengertian yang lebih komprehensif tentang siswa. Kedua, nilai memberikan informasi ke guru untuk digunakan dalam memberikan bimbingan. Siswa sangat tertarik pada kemampuannya (*performance*). Tugas guru adalah membimbing siswa mencari/menggali informasi tentang cara untuk meningkatkan kemampuannya. Ketiga, guru mendapat data yang dapat digunakan untuk menilai efisiensi program dan kualitas cara ia mengajar. Selanjutnya, nilai dapat pula dipergunakan guru untuk tujuan menggolongkan siswa dalam kelompok yang homogen.

d. Bagi administrator

Nilai mutlak diperlukan bagi administrator sekolah. Karena kepercayaan masyarakat pada kesahihan nilai, sering nilai dijadikan dasar dalam pengambilan keputusan administratif. Nilai digunakan sebagai dasar untuk wisuda dan promosi, penghargaan akademik, dan masuk perguruan tinggi. Nilai menunjukkan kemajuan dan pencapaian siswa dalam proses belajar-mengajar.

C. DASAR-DASAR UNTUK MENENTUKAN NILAI PJOK

Para pakar pengukuran dalam pendidikan jasmani dan olahraga sependapat bahwa salah satu dasar penting untuk menentukan nilai adalah perbaikan pada ranah psikomotor. Ranah ini meliputi pencapaian dalam ketrampilan, kemampuan bermain dan kesegaran jasmani siswa. Pada umumnya pakar tersebut juga ada kesesuaian pendapat yang menyatakan bahwa yang ingin dicapai dari ranah kognitif dalam penjasorkes berkaitan dengan; pengetahuan penjasorkes, pemahaman asas-asas dan mekanika gerak, peraturan pertandingan, keselamatan dalam olahraga, kondisioning, sejarah olahraga dan lain-lainnya. Berkaitan dengan ranah afektif terdapat “ketidaksesuaian” pendapat yang sangat menonjol di antara para pakar pengukuran pendidikan jasmani dan olahraga sebagaimana telah diuraikan sebelumnya. Ranah ini meliputi perilaku sosial, seperti: usaha, sikap dan sportivitas. Seberapa besar bobot diberikan pada setiap ranah dalam memberi nilai tergantung pada filsafat dari masing-masing guru penjasorkes.

Dasar-dasar untuk memberi nilai penjasorkes menurut Safrit (1981:330-333) adalah sebagai berikut:

a. **Pencapaian dalam ketrampilan gerak sebagai dasar untuk menilai.**

Pentingnya pencapaian dalam ketrampilan gerak sebagai dasar menilai tergantung pada tujuan-tujuan dari mata pelajaran penjasorkes. Barang kali faktor ini mendapat bobot yang paling besar dalam pemberian nilai dalam pelajaran pendidikan jasmani. Untuk siswa tingkat pemula (satuan pendidikan paling rendah) akan memiliki beban yang lebih berat untuk meningkatkan keterampilan yang ia miliki. Pada tingkatan satuan pendidikan berikutnya tekanan lebih mengarah pada pengembangan kemampuan bermain. Salah satu tujuan lain dari program pendidikan jasmani olahraga dan kesehatan di sekolah adalah pengembangan kebugaran jasmani siswa. Jadi faktor ini harus menjadi bagian dari nilai akhir hasil belajar siswa.

b. **Perbaikan dalam ketrampilan gerak sebagai dasar untuk menilai**

Apabila kinerja siswa menjadi lebih baik dalam satu periode waktu tertentu dapat diasumsikan bahwa pembelajaran siswa telah terjadi. Guru penjasorkes dapat mempelajari tingkat pencapaian akhir siswa atau guru dapat melihat seberapa besar perbaikan atau kemajuan yang telah dicapainya. Perbaikan sangat diinginkan oleh siswa dan guru namun

barang kali banyak masalah yang timbul apabila menggunakan perbaikan sebagai dasar pemberian nilai daripada faktor-faktor lainnya. **Pertama**, skor perbaikan tidak terandal (*unreliable*). Untuk memperoleh skor perbaikan, skor tes akhir (*post test*) harus dibandingkan dengan skor tes awal (*pre test*). Perbedaan antara kedua skor itu adalah skor perbaikan. Akan dijumpai salah satu factor yang berkaitan dengan tingkat keterandalan terkait dengan jika dasar pemberian nilai siswa pada membandingkan skor awal dan skor akhir. Kesalahan pengukuran (*error of measurement*) yang terkait dengan tiap tes itu akan terkumpul dalam skor perbaikan. Terlebih lagi, jika korelasi antara tes awal dan tes akhir meningkat maka keterandalan dari skor perbaikan akan menurun. **Kedua**, jumlah perbaikan yang dapat dicapai oleh siswa akan tergantung pada tingkat keterampilan awalnya. Siswa yang tingkat keterampilan awalnya sudah tinggi diharapkan perbaikan yang akan dicapainya sebanyak siswa yang tingkat keterampilannya rendah. **Ketiga**, untuk dapat memberikan nilai berdasarkan perbaikan perlu diperoleh skor tes awal. Motivasi yang cukup harus diberikan kepada siswa agar melakukan awal dengan kemampuan maksimalnya. Karena dasar yang dipakai untuk menilai adalah perbaikan. Siswa yang cerdas tidak akan berusaha mencapai skor yang tinggi pada saat pelaksanaan tes awal. Pemberian nilai berdasarkan perbaikan walaupun diinginkan barangkali penggunaannya perlu diminimalkan sampai ada acara yang lebih baik untuk menangani skor perbaikan.

c. Pengembangan keterampilan kognitif sebagai dasar untuk menilai

Telah ada kesepakatan dari para pakar penjasorkes bahwa pengetahuan dan pemahaman tentang berbagai aspek pendidikan jasmani perlu menjadi bagian dari nilai siswa. Pengertian tentang asas-asas dan mekanika gerak serta pengaplikasian asas-asas dalam aktivitas tertentu merupakan keterampilan kognitif yang sangat penting dalam mata pelajaran penjasorkes. Keterampilan kognitif yang bertingkat rendah, seperti pengetahuan tentang peraturan, strategi permainan, dapat digunakan dalam menentukan nilai. Keterampilan kognitif lain yang dapat diukur di antaranya adalah faktor keselamatan, sejarah cabang olahraga dan asas-asas kondisioning.

d. Pengembangan keterampilan afektif sebagai dasar untuk menilai

Nilai yang didasarkan pada domain/ranah afektif ditentukan oleh faktor-faktor sikap, kehadiran, sportivitas, usaha dan kerja keras. Untuk tipe pemberian nilai ini faktor-faktor tertentu yang biasanya dianggap sebagai pencerminan dari sikap adalah apabila sikap

seorang siswa baik, ia akan masuk kelas, selalu antusias dalam mengikuti pelajaran penjasorkes, dan berusaha keras. Namun faktor-faktor tersebut sangat sulit untuk diukur dan pengaruh “*hallo effect*” akan banyak berperan dalam menilai ranah afektif. Keterampilan siswa yang baik dalam ranah afektif bukan menjamin pencapaian dalam keterampilan gerak atau pengertian tentang gerak.

e. Usaha

Usaha sangat sulit untuk dievaluasi. Apakah siswa yang kinerjanya baik dengan usaha yang sedikit dapat dianggap kurang daripada siswa yang ketrampilannya kurang tetapi berusaha keras? Apakah guru penjasorkes dapat menentukan jumlah usaha yang dilakukan oleh siswa berdasarkan ekspresi wajahnya? Apabila “**usaha**” menjadi dasar dalam pemberian nilai siswa, maka usaha harus dievaluasi dengan tingkat objektivitas yang dapat diterima. Kita dapat bersumsi bahwa usaha dapat tercermin dalam pencapaian namun apabila pencapaian tidak terjadi, apakah usaha tidak mempunyai arti?

f. Sportivitas

Setiap guru penjasorkes berharap dapat menanamkan nilai-nilai yang berkaitan dengan sportivitas pada siswa. Pengembangan kebiasaan yang mencerminkan sportivitas menjadi salah satu tujuan dari banyak program pendidikan jasmani olahraga dan kesehatan. Namun proses pemberian nilai untuk atribut ini (sportivitas) sering dilakukan secara sembrono dan tidak dapat diandalkan. Apabila satu pendekatan sistematis tidak dapat dilakukan unruk menilai sportivitas dan keterampilan afektif lainnya maka pemberian nilai yang mendasarkan pada tingkat sportivitas dan keterampilan afektif siswa tidak perlu dipertimbangkan.

g. Kehadiran

Penentuan faktor-faktor sebagai dasar memberikan nilai kepada siswa harus tergantung pada tujuan-tujuan penjasorkes. Apakah kehadiran siswa merupakan tujuan pelajaran atau apakah kehadiran itu bukan merupakan kebijakan (*policy*) dari sekolah? Kehadiran tidak tepat untuk dijadikan faktor sebagai dasar pemberian nilai penjasorkes.

D. KRITERIA-KRITERIA NILAI DALAM PJOK

Pemberian nilai hasil belajar siswa merupakan salah satu tugas guru penjasorkes. Oleh sebab itu pemberian nilai harus dilakukan dengan tepat dan cermat. Ada beberapa kriteria

yang harus diperhatikan oleh guru penjasorkes dalam pemberian nilai kepada siswa, di antaranya:

1. **Terkait dengan tujuan-tujuan**

Nilai penjasorkes harus ditentukan berdasarkan keterkaitan dengan tujuan-tujuan yang telah dirumuskan untuk dicapai dalam pelajaran penjasorkes. Jadi nilai siswa harus menyatakan tingkat profisiensi siswa dalam mencapai tujuan-tujuan program yang telah ditetapkan. Contoh: sekor tes kesegaran jasmani yang diperoleh siswa akan menggambarkan profisiensi siswa dalam pencapaian tujuan efisiensi perkembangan organik dan nilai harus mencerminkan pencapaian tujuan tersebut. Nilai bukan saja berhubungan dengan tujuan-tujuan tetapi juga berhubungan dengan pentingnya tujuan-tujuan tersebut.

Di antara pakar penjasorkes belum ada kesepakatan pendapat tentang cara menentukan bobot untuk masing-masing tujuan. Penentuan bobot masing-masing tujuan pelajaran hanya didasarkan pada penalaran atau logika yang harus diikuti. Nilai akhir dari pendidikan jasmani olahraga dan kesehatan adalah gabungan dari nilai-nilai untuk setiap tujuan yang menjadi indikasi pencapaian siswa dalam mata pelajaran penjasorkes.

2. **Kesahihan.**

Nilai harus mempunyai tingkat kesahihan atau ketepatan. Ini berarti bahwa nilai dengan sebenarnya telah mengukur kualitas atau faktor-faktor yang harus diukur. Ketepatan pemberian nilai guru penjasorkes menjadi rendah karena pada kenyataannya guru penjasorkes harus mengukur hal-hal yang sulit untuk diukur jika harus konsisten dan berpegang teguh pada ide dasar yang mengatakan bahwa nilai harus berkaitan dengan tujuan-tujuan pembelajaran penjasorkes. Sering hal-hal yang sulit diukur itu lebih bernilai bagi kehidupan daripada kualitas yang dapat diukur secara objektif. Misalnya; Semangat regu (*team spirit*) dan sikap mau berkorban untuk kepentingan regu adalah dua kualitas yang sangat penting dalam penjasorkes, akan tetapi hal ini sulit untuk diberi nilai tentang tingkat pencapaian dari kualitas tersebut. Berdasarkan kajian literasi, kesahihan nilai tidak akan pernah dapat dicapai seperti yang diharapkan.

3. **Keterandalan**

Nilai harus memiliki tingkat keterandalan yang dapat dipercaya. Sistem pemberian nilai harus melaporkan secara konsisten apa yang dilaporkannya. Keterandalan dari

sistem pemberian nilai dapat ditentukan dengan mengajukan pertanyaan, “Apakah guru penjasorkes menggunakan sistem pemberian nilai atau cenderung untuk memberikan nilai sama kepada siswa atas dasar kinerja yang sama?”. Bila jawabanya “**ya**” berarti sistem nilai guru tersebut mempunyai tingkat keterandalan yang dapat dipercaya.

4. Objektivitas

Nilai harus mempunyai tingkat objektivitas yang baik. Untuk mengetahui tingkat penguasaan materi pelajaran teori, guru penjasorkes sedapat mungkin menggunakan tes bentuk objektif apabila tes tersebut tersedia. Tes bentuk objektif ini berguna untuk mengeliminasi bias dan opini subjektif dari penilai. Apabila tes bentuk objektif tidak tersedia untuk mengukur tingkat pengetahuan siswa guru penjasorkes dapat menggunakan metode subjektif dalam memberikan nilai tetapi metode subyektif yang dipakai sedapat mungkin diobjektifkan. Ada beberapa teknik tertentu yang dapat digunakan, seperti sistem penilaian dalam bentuk rubrik.

5. Dapat dipahami

Nilai harus dapat dipahami oleh siswa dan orangtuanya serta dengan mudah dapat diinterpretasikan oleh guru penjasorkes. Siswa harus mengetahui dasar yang dipakai dalam memberi nilai dan sistem penilaian yang digunakan. Siswa sebelumnya harus dapat memprediksi nilai yang akan diperoleh sebelum siswa tersebut mendapat nilai dari gurunya.

6. Pembobotan

Karena faktor-faktor yang dinilai tidak sama pentingnya maka disarankan agar faktor-faktor tersebut diberi bobot sesuai dengan pentingnya secara relatif. Pemberian bobot untuk tiap-tiap faktor itu didasarkan pada logika yang disesuaikan dengan kebutuhan siswa dan penekanan materi yang diberikan pada pelajaran penjasorkes. Faktor-faktor yang telah diberi bobot dengan baik akan menghasilkan satu indeks nilai untuk tiap siswa.

Nilai ini akan mencerminkan pencapaian yang sesungguhnya dari kemampuan siswa. Misalnya; apabila guru penjasorkes akan memberikan nilai untuk kesegaran jasmani, keterampilan berolahraga, pengetahuan dan sikap, ia dapat memberi bobot sebagai berikut: 2, 3, 2, 1. Berdasarkan logika nilai untuk ranah psikomotor pada mata pelajaran penjasorkes pada umumnya diberi bobot paling banyak dan nilai untuk ranah

afektif paling sedikit. Pembobotan nilai akan mencerminkan penekanan materi pelajaran yang diberikan pada setiap tujuan pembelajaran. Pembobotan nilai juga akan mencerminkan filsafat dari pemberi nilai tersebut.

7. Diskriminasi

Nilai harus mendiskriminasi atau membedakan antara tingkatan-tingkatan kemampuan siswa. Kriteria ini akan mengeliminasi sistem dua kategori, seperti: lulus-gagal dan memuaskan-tidak memuaskan. Tingkatan-tingkatan kemampuan merupakan hal yang relatif dan harus ditangani secara empiris. Tingkatan-tingkatan kemampuan siswa harus ditetapkan cukup tinggi sehingga siswa yang berbakat akan tertantang untuk mencapainya sedangkan siswa yang memiliki kecakapan rendah (tidak berbakat) tidak putus asa. Jangan ada pengelompokan nilai baik pada tingkat kemampuan yang tinggi maupun kemampuan yang rendah. Jika di antara siswa ada yang berkelainan maka siswa tersebut harus diberi kesempatan untuk mendapat nilai yang adil (*fair*) sesuai dengan keadaan.

8. Ekonomis dan administrasi

Sistem pemberian nilai harus ekonomis dalam waktu, biaya dan personalia. Pertimbangan yang paling penting adalah tentang waktu. Sistem penilaian jangan terlalu banyak menyita waktu. Tugas guru penjasorkes yang pertama adalah mengajar. Jangan terlalu banyak waktu digunakan untuk pelaksanaan pemberian nilai yang rumit. Satu sistem pemberian nilai harus berfungsi sebagai alat untuk memperbaiki proses belajar-mengajar penjasorkes dan bukan justru menjadi penghambat.

E. SISTEM METODE PEMBERIAN NILAI PJOK

Ada beberapa metode pemberian nilai penjasorkes yang sering digunakan di sekolah-sekolah dan di perguruan tinggi keolahragaan, di antaranya:

1. Metode persentase

Apabila pemberian nilai menggunakan PAP guru penjasorkes dapat menentukan siswa mendapat A jika ia menjawab betul 90% atau lebih dari tes yang dikerjakannya. Menjawab betul 80% sampai 89% mendapat B. Menjawab 70% sampai 79% dengan betul mendapat C. Mendapat nilai D apabila menjawab betul 60% sampai 69%, dan memperoleh nilai F jika menjawab betul kurang dari 60%. Atas dasar pertimbangan-pertimbangan tertentu barangkali guru penjasorkes dapat menggunakan persentase lain.

Dalam mata pelajaran penjasorkes persentase yang ditentukan guru hendaknya mengacu pada persentase tugas yang harus dikuasai oleh siswa. Namun demikian penentuan persentase perlu dipertimbangkan dengan hati-hati, seperti; tingkat kesulitan tes atau variasi macam-macam tugas yang harus dikerjakan siswa. Misalnya; skor 60% pada tes yang satu sesungguhnya barangkali lebih baik dari pada 80% pada tes yang lain karena tes yang pertama lebih sulit daripada tes yang kedua. Kelemahan metode persentase ini adalah guru penjasorkes agak sulit menentukan persentase yang akan dipakai karena faktor tingkat kesulitan tes yang akan digunakan.

2. Metode angka.

Sistem pemberian nilai penjasorkes dengan metode angka sama dengan sistem presentase. Misalnya, apabila skor maksimal dari tes 100, maka siswa akan mendapat A jika siswa mendapat skor antara 90 sampai 100. Skor antara 80 dan 89 nilai B. Apabila skor antara 70 sampai 79 nilai C. Skor 60 sampai 69 mendapat nilai D, dan nilai F apabila skor yang diperoleh siswa kurang dari 59. Kelemahannya dari sistem pemberian nilai dengan metode angka ini adalah tingkat kesulitan tes yang terlalu tinggi. Konsekuensinya akan terjadi tidak seorang peserta didik yang mendapat nilai A, karena tes yang digunakan terlalu sulit. Sebaliknya apabila tes terlalu mudah akan banyak siswa yang memperoleh nilai A.

3. Metode kurva normal

Pemberian nilai dengan sistem Pendekatan Acuan Norma (PAN) menggunakan metode kurva normal. Dua contoh pendistribusian nilai berdasarkan distribusi kurve normal adalah sebagai berikut:

Contoh 1.

7%	A
18%	B
50%	C
18%	D
7%	E

Contoh 2.

10%	A
20%	B
40%	C
20%	D
10%	E

Penggunaan metode ini didasarkan pada asumsi bahwa kemampuan siswa dalam setiap kelas terdistribusi secara normal. Namun kenyataannya tidak akan selalu menggambarkan demikian dan hal inilah yang merupakan kelemahan dari sistem pemberian nilai dengan menggunakan metode kurve normal ini.

4. Beberapa contoh sistem memberi nilai dalam mata pelajaran penjasorkes

Contoh 1.

Faktor	Bobot	Nilai	Points
Ketrampilan olahraga	3	A (5)	15
Kebugaran jasmani	2	B (4)	8
Pengetahuan	2	C (3)	6
Sikap	1	B (4)	4
Bobot total	8	Total	33
Nilai akhir	$\frac{33}{8} = 4.125$ atau nilai B		

Contoh 2.

A = 5 B = 4 C = 3 D = 2 F = 1

Tujuan-tujuan	Persen bobot	Nilai	Points
Perkembangan organik	30%	B (4)	$0.30 \times 4 = 1.2$
Perkembangan <i>neuromuscular</i>	30%	A (5)	$0.30 \times 5 = 1.5$
Perkembangan intelektual	20%	D (2)	$0.20 \times 2 = 0.4$
Perkembangan sosial-emosional	20%	B (4)	$0.20 \times 4 = 0.8$
Nilai akhir 3.9 = B			

Contoh 3. A = Skor T 63 atau lebih B = 52 – 62 C = 47 – 51 D = 41 – 46

E = 33 – 40 F = kurang dari 33

Faktor	Bobot	Skor T siswa	Points
Pencapaian ketrampilan	3	58	$3 \times 58 = 174$
Perbaikan kebugaran jasmani	2	65	$2 \times 65 = 130$
Pengetahuan	1	50	$1 \times 50 = 50$
Sikap	1	48	$1 \times 48 = 48$

Nilai akhir $\frac{402}{7} = 57.4 = B$
--

Contoh 4. A = 90% - 100% B = 80% - 89% C = 60% - 79%
 D = 50% - 59% E = kurang dari 50%

Ranah	Bobot dari ranah	Faktor	Bobot dari faktor	Nilai dari faktor	Point
Psikomotor	5	Kesegaran jasmani	1	A	12
		Keterampilan olahraga	2	C	12
		Rating keterampilan olahraga	1	B	9
		Rating keterampilan oleh guru penjasorkes	1	B	9
Kognitif	3	Tes pengetahuan	3	A	36
Afektif	2	Rating oleh teman	1	A	12
		Rating oleh guru penjasorkes	1	B	9
Jumlah	10		10		99

A+ - 13 B+ - 10 C+ - 7 D+ - 4
 A - 12 B - 9 C - 6 D - 3 F - 1
 A- - 11 B- - 8 C- - 5 D- - 2

Apabila seorang siswa mendapatkan A untuk semua faktor maka siswa tersebut akan memperoleh poin atau sekor yang telah dibobot 120. Nilai akhir dinyatakan dalam persen. Jadi sekor yang telah dibobot sebesar 99 sama dengan $\frac{99}{120} = 82.5\%$ atau sama dengan nilai B.

.....

Kecenderungan yang lazim dipakai dalam penilaian pendidikan jasmani adalah dengan mempergunakan huruf untuk perguruan tinggi atau skala 1 – 10 untuk jenjang pendidikan

yang lebih rendah, seperti di SD, SLTP, atau SMA. Menurut Rusli Lutan (2000: 192 – 197) sistem penilaian yang dipergunakan dalam pendidikan jasmani di antaranya adalah:

(1). Metode Kesenjangan Dalam Distribusi

Sebuah distribusi skor tes biasanya memiliki kesenjangan skor, maksudnya kadangkala ada rentangan skor tertentu dimana tidak seorangpun siswa mendapatkan skor pada rentangan tersebut. Beberapa guru pendidikan jasmani kadangkala memanfaatkan metode kesenjangan dalam distribusi ini untuk menentukan nilai siswa mereka. Sebagai contoh, perhatikan table 4 berikut ini.

Tabel 7: Distribusi skor bagi suatu kelas

Skor	Nilai	Skor	Nilai	Skor	Nilai
95		77		52	
95		77		51	E
94	A	76	C	48	
93		76			
92		76			
		74			
85		74			
84	B				
83		68	D		
82		67			
		67			

Penentuan nilai berdasarkan kesenjangan skor dapat berbeda-beda bagi setiap kelas yang berbeda penyebaran skornya. Oleh karena itu cara tersebut kurang disukai karena nilai tergantung pada kesenjangan skor yang terjadi. Dengan demikian, system pemberian penilaian dengan metode kesenjangan dalam distribusi ini sebaiknya jangan dipakai.

(2). Metode Persentase

Metode persentase dapat juga diterapkan di lingkungan perguruan tinggi atau SMA dalam kaitannya dengan penguasaan materi secara tuntas. Sebagai contoh: seorang mahasiswa memperoleh nilai A jika mampu menjawab butir tes sebanyak 90% yang benar, nilai B jika menjawab 80%, nilai C jika benar 70%-79%, nilai D jika benar 60%-69%, dan

nilai G jika kurang dari 60%. Penggunaan metode persentase tersebut merupakan acuan patokan. Dengan kata lain, berapa persen materi yang telah dikuasai oleh para mahasiswa.

Guru atau dosen yang tertarik untuk menerapkan metode persentase itu, tentu dihadapkan dengan tugas terutama menentukan batas lulus atau batas penguasaan materi. Dalam kinerja aktivitas jasmani persentase yang ditentukan mengacu pada tugas yang dikuasai oleh peserta didik. Namun demikian, penentuan persentase perlu dipertimbangkan dengan hati-hati kesulitan dari tes atau tugas bervariasi cukup besar. Penetapannya dapat dilakukan berdasarkan pengalaman yang sudah-sudah. Karena tidak ada patokan yang tegas maka batas persentase penguasaan bahan harus dipertimbangkan dengan cermat. Misalnya: skor 60% pada suatu tes mungkin sesungguhnya lebih baik daripada skor 80% pada tes lain, karena tes yang pertama lebih sulit daripada tes kedua. Kelemahan dari metode ini adalah agak sulit menentukan persentase yang akan dipakai karena faktor tingkat kesulitan dari tes yang digunakan.

(3). Metode Himpunan Angka atau Nilai

Prosedur metode himpunan angka ditempuh dengan menjumlahkan beberapa angka atau nilai yang diperoleh siswa dari setiap mata pelajaran. Beberapa angka atau nilai itu diperoleh berdasarkan komponen penilaian yang telah direncanakan dan bahkan telah dikomunikasikan kepada siswa. Komponen itu misalnya: penguasaan pengetahuan, penguasaan keterampilan, kerajinan mengikuti program dan lain-lain, sesuai dengan pertimbangan guru yang bersangkutan mengenai unsur penilaian yang dianggap amat penting untuk menggambarkan kemajuan belajar siswanya.

Untuk lebih jelas pemahaman penggunaan metode himpunan angka, perhatikan ilustrasinya pada tabel 5 berikut ini.

Tabel 8: Jumlah angka bagi setiap tugas dan penentuan nilai

Tugas	Nilai	Skala Penilaian
Tes Keterampilan	90	A = 482 – 530
Metodik I	50	B = 428 – 481
Metodeik II	50	C = 375 – 427
Ujian Akhir	100	D = 321 – 374

Teori I	20	G = < 321
Teori II	30	
Makalah I	40	
Makalah II	60	
Keterampilan bermain	20	
Keikutsertaan dan kehadiran	70	
Jumlah maksimal	530	

(4) Metode Kurva Normal

Pendekatan yang paling lazim dalam penentuan nilai adalah metode kurva normal. Dalam metode ini, sebuah distribusi normal dijadikan landasan penentuan nilai dengan asumsi bahwa kemampuan peserta didik dalam tiap kelas terdistribusi secara normal. Namun kenyataannya tidak akan selalu demikian, dan hal inilah yang merupakan salah satu kelemahan dari metode ini.

Tabel 9: Penentuan nilai berdasarkan kurva normal

Nilai	Skor - Z	Skor - T	Persentil
7 % = A	> 1,48	> 64,8	93 ke atas
18 % = B	0,67 s/d 1,48	56,7 s/d 64,8	75 s/d 92
50 % = C	-0,67 s/d 0,67	43,3 s/d 56,7	26 s/d 74
18 % = D	- 0,48 s/d -0,67	35,2 s/d 43,3	8 s/d 25
7 % = G	< -1,48	< 35,2	7 ke bawah

Apabila sudah diperoleh nilai rata-rata (*mean*) dan simpangan baku (*SB*), maka tidak begitu sukar bagi guru pendidikan jasmani untuk mengetahui batas skor bagi masing-masing kategori nilai sesuai dengan luas kurva normal seperti tertera pada tabel 6 di atas.

(5). Penilaian Berdasarkan Kontrak

Maksud penilaian dengan system berdasarkan kontrak adalah guru dan siswa yang bersangkutan mengikat kesepakatan tentang apa yang harus dilakukan oleh siswa untuk memperoleh nilai tertentu. Sebagai contoh, dalam pelajaran atletik di SMA misalnya; seorang siswa akan memperoleh nilai A jika dia mampu menempuh lari *sprint* 100 meter selama 12 detik (untuk putra), lompat jauh = 5 meter, tolak peluru berat 5 kg = 7 meter,

membaca 3 artikel tentang atletik, dan menyusun makalah singkat (3-4 halaman) tentang atletik. Untuk memperoleh nilai B, tentu beban tugasnya lebih rendah dari beban tugas untuk memperoleh nilai A.

(6). Pendekatan Portofolio

Pendekatan portofolio akhir-akhir ini sering dikembangkan oleh sebagian guru pendidikan jasmani untuk memberikan penilaian kepada siswanya. Perkembangannya didorong oleh kenyataan bahwa pendekatan tes objektif telah kehilangan konteks. Pengalaman dan kegiatan siswa di luar situasi sekolah tidak terekam dan tidak memperoleh penghargaan. Yang diandalkan hanya himpunan prestasi belajar yang terukur pada saat tes dan pengukuran di sekolah.

Pendekatan portofolio pada dasarnya menekankan penghargaan kepada seluruh pengalaman dan kemajuan siswa baik yang diperagakan di sekolah maupun di luar sekolah. Pendekatan ini tampaknya cocok dengan ide pendidikan jasmani yang bertujuan untuk membentuk kebiasaan melaksanakan budaya atau gaya hidup aktif. Dengan demikian seluruh aktivitas siswa memperoleh penghargaan, seperti: kegiatan di klub/perkumpulan olahraga, latihan, latihan mandiri secara teratur untuk membina kebugaran jasmani, kegiatan ekstrakurikuler di sekolah, mengikuti pertandingan resmi, dan pengalaman lainnya.

Pengalaman itu dapat diklasifikasi sesuai dengan ruang lingkup dan tujuannya dan diselaraskan dengan komponen tujuan pendidikan jasmani, meliputi: (a) kegiatan untuk pengembangan pengetahuan, seperti: mengikuti seminar, diskusi, studi klub, (b) kegiatan untuk pengembangan keterampilan motorik, seperti: latihan mandiri, mengikuti kompetisi resmi, latihan kebugaran jasmani, *camping*, kepramukaan, dan lain-lain, (c) rekaman kegiatan yang mengandung nilai bagi pembinaan aspek afektif/sikap. Informasi itu semua dilaporkan oleh siswa itu sendiri untuk kemudian dinilai oleh gurunya.

Ada dua pendekatan yang dapat diterapkan jika guru pendidikan jasmani ini menerapkan pendekatan portofolio dalam memberikan penilaian. **Pertama**, berbentuk laporan uraian tertulis untuk setiap kegiatan yang kemudian dihimpun dalam sebuah *file* untuk setiap siswa. Pekerjaan ini cukup banyak dan guru juga akan memperoleh pekerjaan tambahan. Kedua, laporan dalam bentuk pengisian formulir yang disediakan. Bentuk ini memang tidak lazim untuk laporan portofolio karena sudah dibatasi ruang lingkup dan

kepanjangan isinya. Namun dapat digunakan sebagai alternative meskipun bukan laporan portofolio yang sebenarnya.

Rangkuman masukan informasi pengalaman siswa itu selanjutnya dinilai oleh gurunya. Untuk itu perlu disusun kerangka penilaian. Hasilnya digunakan untuk melengkapi prestasi belajar yang direkam secara formal pada waktu pelaksanaan tes dan pengukuran.

Selain ke tujuh sistem sistem pemberian nilai dalam pendidikan jasmani sebagaimana yang diungkapkan oleh Rusli Lutan tersebut, menurut Arma Abdoellah (2002: 8) ada contoh sistem memberikan nilai lain yang tidak kalah pentingnya untuk pemberian penilaian dalam pendidikan jasmani yang dapat dipakai oleh guru pendidikan jasmani. Adapun contoh sistem memberikan nilai tersebut adalah:

Contoh: 1

Faktor	Bobot	Nilai	Poin
Keterampilan Olahraga	3	A (5)	15
Kebugaran Jasmani	2	B (4)	8
Pengetahuan	2	C (3)	6
Sikap	1	B (4)	4
Bobot Total	8		33
Nilai Akhir	$33 : 8 = 4,125$ atau nilai B		

Contoh: 2

Tujuan-tujuan	Bobot (%)	Nilai	Poin
Perkembangan Organik	30%	B (4)	$0,30 \times 4 = 1,2$
Perkembangan Neuromuskuler	30%	A (5)	$0,30 \times 5 = 1,5$
Perkembangan Intelektual	20%	D (2)	$0,20 \times 2 = 0,4$
Perkem. Sosio- Emosional	20%	B (4)	$0,20 \times 4 = 0,8$

Nilai Akhir	$1,2 + 1,5 + 0,4 + 0,8 = 3,9$ (B)
--------------------	---

BAB VII
PENYUSUNAN TES BENTUK OBJEKTIF
UNTUK EVALUASI RANAH KOGNITIF DALAM PJOK

Memiliki pengetahuan tentang pendidikan jasmani olahraga dan kesehatan (PJOK) merupakan salah satu tujuan PJOK mulai dari pendidikan dasar sampai perguruan tinggi. Oleh sebab itu harus dilakukan penilaian/evaluasi untuk mengetahui seberapa jauh tujuan pengajaran PJOK itu telah dicapai oleh peserta didik. Perkembangan kemampuan interpretif siswa yang dijajagi dengan kemampuan tes pengetahuan (ranah kognitif) juga merupakan fokus perhatian dalam penyelenggaraan PJOK. Dalam proses penilaian ranah kognitif ini alat evaluasi dapat menggunakan tes tertulis bentuk objektif (*objective test*) dengan berbagai jenisnya, tes tertulis bentuk uraian (*essay test*) maupun tes lisan. Ketiga jenis tes tersebut tidak hanya mengukur pengetahuan tetapi dapat mengukur pengertian atau bahkan pemahaman pembelajaran PJOK.

Penggunaan tes pengetahuan (tes ranah kognitif) dalam PJOK jangan hanya dipergunakan untuk menilai hasil belajar peserta didik akhir semester saja, namun guru PJOK dapat menggunakan tes tersebut di awal proses pembelajaran untuk mengetahui seberapa banyak cakupan materi pembelajaran yang akan diajarkan telah diketahui peserta didik sehingga ia dapat menentukan apa yang telah dikuasai maupun yang belum dikuasai oleh sebagian besar peserta didiknya. Tes pengetahuan yang singkat dapat diberikan dalam proses menyampaikan materi pelajaran untuk memberikan penekanan substansi yang lebih banyak pada hal-hal yang penting atau untuk membantu peserta didik dalam membuat kesimpulan atau melihat kaitan materi pembelajaran, serta untuk memberikan latihan dalam menggunakan pengetahuannya dalam situasi baru.

Tes pengetahuan dapat juga dipergunakan untuk penelitian, di antaranya: untuk mengukur pengaruh ujian tertulis terhadap prestasi belajar peserta didik, pengaruh pemberian kuis-kuis yang sering terhadap prestasi belajar. Hasil-hasil penelitian ini ternyata memberi rekomendasi kepada guru PJOK agar lebih memperbanyak pemberian tes tertulis untuk meningkatkan prestasi belajar pengetahuan PJOK.

A. Langkah-Langkah Pembuatan Tes Bentuk Objektif PJOK

Berdasarkan studi literasi terkini di Amerika Serikat telah banyak disusun tes pengetahuan yang dipergunakan untuk menilai ranah kognitif untuk mata pelajaran sejarah, matematika, bahasa, ilmu pengetahuan alam, dan mata pelajaran-mata pelajaran lainnya, namun baru sedikit dijumpai tes pengetahuan untuk mata pelajaran PJOK. Penyusun tes pengetahuan itu dibuat oleh para ahli penyusun tes (*expert of test construction*) dan ahli dalam mata pelajaran (*expert of subject matter test*) sehingga menghasilkan sebuah tes yang populer dikenal dengan tes standar (*standardized test*) dari masing-masing mata pelajaran tersebut. Pada uraian selanjutnya akan dibahas langkah-langkah bagi guru PJOK dsism menyusun tes pengetahuan PJOK khususnya tes bentuk objektif. Adapun langkah-langkah penyusunan tes bentuk objektif dalam PJOK adalah sebagai berikut: (1) distribusi isi atau substansi materi pelajaran, (2) memilih tipe item atau pertanyaan dalam tes, (3) melaksanakan tes yang telah dibuat, dan (4) menilai atau mengevaluasi tes yang dibuat.

1. Distribusi isi atau materi pelajaran

Langkah pertama dalam penyusunan tes pengetahuan dalam mata pelajaran PJOK adalah menentukan tujuan untuk apa tes itu akan digunakan. Apabila tujuan tes itu digunakan untuk ujian semesteran atau untuk kepentingan klasifikasi/penggolongan maka tes yang disusun harus komprehensif. Dalam hal ini harus dipelajari garis besar dari substansi materi pelajaran dan apabila daftar garis besar isi materi pelajaran terlalu ringkas penyusun tes perlu membuat sebuah daftar yang lebih lengkap sehingga semua cakupan materi yang penting dari pelajaran sudah tercakup di dalamnya. Pada umumnya sebagian besar pertanyaan/soal harus berkaitan dengan konsep-konsep yang dianggap penting. Berpedoman pada daftar isi pelajaran yang terdapat dalam buku materi pelajaran akan menghasilkan distribusi pertanyaan yang baik dan juga akan meniadakan penggunaan bahasa buku.

Berikut disajikan sebuah ilustrasi dalam usaha mendistribusikan jumlah pertanyaan sesuai dengan substansi isi materi pelajaran tentang pengetahuan permainan bola voli untuk siswa yang memilih olahraga sebagai bidang keahliannya dapat dilihat pada tabel 10.

Pengetahuan Permainan Bola Voli	Jumlah Item tes	(%) diestimasikan
Analisa teknik-teknik dan gaya (<i>style</i>) bermain bola voli	3	5
Analisis situasi permainan dan penggunaan keterampilan bola voli	8	13
Pengetahuan Umum (sejarah, pemeliharaan alat-alat dan fasilitas bola voli, tindakan pencegahan cedera bola voli)	3	5
Bagaimana menghindari pelanggaran dalam permainan bola voli	3	5
Cara-cara mengoper bola (<i>passing</i>) dan melakukan smash	2	4
Peraturan-peraturan penting untuk bermain bola voli	15	25
Taktik dan strategi bermain	22	37
Istilah-istilah dalam permainan bola voli	4	6
Jumlah	60	100

2. Memilih tipe item atau pertanyaan dalam tes

Setelan menentukan proporsi jumlah pertanyaan atau soal, langkah berikutnya adalah memilih tipe item atau pertanyaan yang paling tepat dan sesuai dengan materi pelajaran. Menurut *Scott* dan *French* ada kurang lebih 50 buah macam atau tipe item yang tepat digunakan oleh penyusun tes jawaban-pendek. Sungguhpun demikian item tes tersebut dapat digolongkan ke dalam dua golongan besar, yaitu: (1) tipe isian (*supply type*) dan (2) tipe pilihan (*selection type*). Kedua tes tipe isian dan tipe pilihan sering juga disebut dengan tes tipe ingatan (*recall*) dan tipe pengenalan (*recognition*). Kedua tes tipe isian dan

tes tipe pilihan dengan jelas menunjukkan bahwa tugas siswa apakah harus mengisi kata-kata, angka, huruf atau simbol lainnya sebagai jawaban (*supply type*) atau siswa harus memilih jawabannya dari sejumlah pilihan atau alternatif yang diberikan (*selection type*).

3. Melaksanakan tes yang telah dibuat

Dalam melaksanakan tes yang telah disusun, seorang guru PJOK harus memperhatikan beberapa kaidah-kaidah agar dalam pelaksanaannya tidak mengalami hambatan. Kaidah-kaidah yang perlu menjadi atensi, di antaranya:

- a. Apakah telah dipikirkan dari sudut waktu yang tersedia, siswa dalam menjawab tes dapat mempergunakannya seekonomis mungkin?
- b. Apakah petunjuk-petunjuk bagi siswa telah dibuat sesederhana mungkin dan mudah dapat dipahami?
- c. Apakah cara dan tempat siswa dalam memberikan jawaban telah dipikirkan agar dalam pemberian nilai nanti dapat dilakukan secara efisien, teliti, dan seekonomis mungkin?
- d. Apakah tipografi dari susunan pertanyaan masih dapat diperbaiki?
- e. Apakah pertanyaan dalam tes telah disusun secara hirarkis dimulai dari soal yang mudah meningkat kepada yang sulit?
- f. Apakah waktu yang disediakan untuk mengerjakan soal tes cukup memadai sesuai dengan jumlah soal yang harus dikerjakan siswa?

Parameter-parameter di atas sudah jelas tidak akan mungkin dapat dipenuhi semuanya, namun kriteria tersebut merupakan petunjuk yang sedapat mungkin dipenuhi oleh guru PJOK dalam menyusun tes untuk menilai ranah kognitif dalam pendidikan jasmani. Walaupun prosedur pelaksanaan tes itu mudah tetapi perlu direncanakan dengan seksama dan hati-hati. Hal-hal yang perlu diperhatikan oleh guru PJOK dalam melaksanakan tes pengetahuan untuk mengukur ranah kognitif meliputi;

- a. Ruangan harus tenang, cukup ventilasinya dan sistem penerangannya cukup memadai.
- b. Tempat duduk harus berjauhan satu dengan yang lain atau siswa disuruh duduk dengan antara satu kursi satu dengan yang lain. Hal ini tidak begitu penting apabila menggunakan kertas jawaban yang terpisah.
- c. Buku-buku dan catatan tidak boleh dibawa masuk ke dalam ruang ujian.
- d. Jangan diberikan petunjuk untuk membuat tes secara lisan tetapi harus dimuat secara eksplisit dalam tes.

- e. Usahakan sungguh-sungguh bahwa setiap siswa hanya mendapatkan satu berkas tes dan semua tes harus diserahkan kembali apabila ujian sudah selesai. Jika soal tes telah diberi nomor pengecekan jumlah sal tes yang harus kembali lebih mudah. Setiap siswa diminta untuk menyerahkan sendiri kertas ujian.
- f. Siswa diminta segera menyerahkan kertas ujian apabila ia telah selesai mengerjakannya.
- g. Bagikan kertas ujian secara terbalik dan mulai mengerjakan soal ujian secara bersama-sama.
- h. Tidak boleh membantu dalam memberikan penjelasan mengenai pertanyaan soal tes.

4. Menilai atau mengevaluasi tes yang dibuat.

Setelah item-item tes atau pertanyaan selesai dibuat perlu item-item tersebut dinilai terlebih dahulu tingkat kelayakan tes dengan mempergunakan daftar cek (check list). Beberapa hal yang perlu diperhatikan oleh guru PJOK dalam menilai tes yang dibuat di antaranya;

- a. Apakah sebenarnya yang ingin diukur dari item tes?
- b. Apakah tujuan yang ingin diukur dari item tes itu dapat diterima? Apakah penting item tes itu dimasukkan ke dalam soal tes.
- c. Apakah pernyataan/pertanyaan dalam item tes itu meragukan? Apakah item tes itu tidak dapat dirumuskan lebih jelas lagi?
- d. Apakah item tes berisikan petunjuk yang tidak disengaja sehingga jawabannya mudah diberikan?
- e. Dalam item pilihan berganda apakah ada di antara alternatif jawaban (*option*) yang salah, nyatanya lebih masuk akal dari jawaban yang benar ditinjau dari siswa-siswa yang pandai dalam kelompok testi? Apakah item tes itu terlalu sulit untuk siswa yang pandai dalam kelompok testi?
- f. Apakah item tes telah dirumuskan seekonomis mungkin? Apakah item tes secara langsung menanyakan maksudnya?
- g. Apakah tipe item tes telah disesuaikan dengan baik maksud yang akan dinilai? Apakah dengan penyajian diagram makna item tes dapat lebih jelas lagi?
- h. Apakah siswa yang suka menghafal memperoleh keuntungan yang mencolok dalam menjawab pertanyaan-pertanyaan? Apakah bahasa buku telah dihindari dalam penyusunan item tes?

Cara lain untuk menilai item-item tes sebelum dipergunakan adalah meminta bantuan kepada orang yang ahli (*expert*) dalam penyusunan tes untuk menilai item tes yang dibuat. Dengan minta masukan dari orang yang lebih ahli diharapkan tingkat ketepatan dan kualitas item tes akan meningkat. Mengulangi sendiri dengan membaca pertanyaan-pertanyaan setelah item tes itu dibuat dalam selang beberapa hari merupakan prosedur yang baik untuk meningkatkan kualitas item tes yang dibuat oleh guru PJOK.

B. Konstruksi Tes PJOK Bentuk Objektif

Disamping tes uraian (*essay test*), ada item tes jenis lain yang juga sering dipergunakan oleh guru PJOK dalam kegiatan belajar-mengajar. Item tes yang dimaksud adalah tes bentuk objektif. Tes ini dikatakan objektif karena para siswa tidak dituntut merangkai jawaban atas dasar informasi yang dimiliki seperti pada tes esai. Pada tes bentuk objektif, jawaban pada umumnya sudah disediakan atau sudah diarahkan dan lebih bersifat pasti. Secara garis besar bentuk tes objektif dapat dibedakan menjadi dua kelompok, yaitu: (1) tes objektif jenis isian (*supply type*), dan (2) tes objektif jenis pilihan (*selection type*) yang terdiri dari tiga tipe, yaitu: (1) tes pilihan berganda/*multiple choice*, (2) tes benar-salah/*true-false*, dan (3) tes menjodohkan/*matching*.

1. Tes Pilihan Berganda (PB)/*multiple choice test*

Item tes pilihan berganda (PB) merupakan jenis tes objektif yang paling banyak dipergunakan oleh para guru PJOK. Tes ini dapat mengukur pengetahuan yang luas dengan tingkat domain yang bervariasi mulai dari soal-soal pengetahuan, pemahaman, aplikasi, analisa, dan evaluasi. Tes pilihan berganda memiliki semua persyaratan sebagai tes PJOK yang baik, yakni dilihat dari segi objektivitas, reliabilitas, dan daya pembeda antara siswa yang berhasil dengan siswa yang gagal. Untuk mengetahui sejauh mana tingkat penguasaan materi pembelajaran siswa, tes ini cocok dipergunakan sebagai instrumen penilaian ranah kognitif PJOK mulai dari tingkat sekolah dasar sampai perguruan tinggi.

Berikut akan disajikan rambu-rambu pembuatan pertanyaan item tes PB

- a. Untuk pokok kalimat gunakan pertanyaan langsung, singkat, dan sederhana.
- b. Hindari alternatif-alternatif jawaban yang tidak masuk akal atau yang tidak terlalu jelas.
- c. Susunlah semua alternatif jawaban dengan bahasa yang ajeg, panjang, dan variasi dan tepat.

- d. Hindari adanya dua jawaban yang benar, apabila petunjuk tes hanya mengatakan memilih satu jawaban yang benar. Seharusnya dalam petunjuk dinyatakan memilih jawaban yang terbaik, dan kemungkinan ada beberapa jawaban yang juga benar tetapi hanya ada satu jawaban yang terbaik dari yang lainnya.
- e. Hindari petunjuk-petunjuk yang tidak disengaja, seperti: menempatkan jawaban yang benar terus-menerus pada nomor atau urutan yang sama selalu dari deretan alternatif, membuat jawaban yang benar selalu lebih panjang atau lebih pendek daripada alternatif-alternatif lainnya.
- f. Hindari penggunaan bahasa buku, apabila maksud pertanyaan itu untuk mengukur kemampuan menggunakan satu konsep atau untuk mengukur pengertian dan bukan hafalan.
- g. Alternatif jawaban sebaiknya seragam dalam bentuk dan struktur bahasa.
- h. Butir soal diberi nomor dengan angka, sedangkan alternatif jawaban dinyatakan dengan huruf.

Contoh: bentuk soal pilihan berganda

- 1. Fungsi seorang pengumpan (*set upper*) dalam permainan bola voli adalah:
 - a. Meniadakan putaran bola agar lebih mudah dan tepat dimainkan
 - b. Menempatkan bola sehingga mudah untuk dilakukan smash
 - c. Mengoper bola ke teman yang akan melakukan smash
 - d. Mendorong kerja sama yang baik dengan rekan satu tim

Kunci Jawaban: b

- 2. Kemampuan seseorang dalam melakukan aktivitas fisik dalam waktu yang cukup lama bergantung pada:
 - a. Efektivitas kerja paru, jantung dan pembuluh darah.
 - b. Kemampuan kerja paru, jantung dan pembuluh darah
 - c. Kapasitas paru, jantung, dan pembuluh darah.
 - d. Kekuatan kerja paru, jantung, dan pembuluh darah.

Kunci jawaban: a

Guru PJOK memilih dan menggunakan tes objektif pilihan berganda sebagai instrumen untuk menilai hasil belajar penjasorkes dengan argumentasi bahwa tes ini memiliki beberapa kelebihan, di antaranya;

- a. Memiliki *validity* atau tingkat keakuratan yang tinggi terutama karena *abilities* yang diukur jelas dan jawaban yang diharapkan juga pasti.
- b. Memiliki *reliability* atau tingkat kepercayaan/konsistensi yang tinggi, terutama karena menjawab benar ataupun salah dapat diberikan secara pasti oleh siapapun dan kapanpun.
- c. Cakupan materi ajar yang mungkin untuk diujikan dapat lebih luas dan lengkap khususnya keunggulan ini sulit dicapai oleh tipe tes esai.
- d. Item tes dapat dibuat singkat dan jelas. Keunggulan ini akan menghindarkan pemahaman yang ambiguous sehingga mudah dipahami serta mudah dikerjakan oleh peserta tes.
- e. Cara pemberian skor sangat mudah dilakukan oleh guru PJOK. Waktu yang dibutuhkan oleh guru untuk memeriksa item-item tes dan memberi skor akhir tidak perlu terlalu lama.
- f. Kualitas item tes dapat dianalisis dengan berbagai teknik statistik, seperti: *item difficulty analysis*, *item discriminant analysis*, *gessing* dan lain-lain baik berdasarkan *classical theory* maupun *item response theory*.
- g. Tingkat *item difficulty* dapat dikendalikan dengan mengubah homogenitas alternatif jawaban. Semakin homogen alternatif jawaban suatu item akan semakin tinggi tingkat kesulitan soalnya. Sebaliknya jika semakin heterogen alternatif jawaban suatu item akan menyebabkan tingkat kesukaran soal semakin rendah.

Bagi guru PJOK yang memilih bentuk tes objektif tipe pilihan berganda sebagai alat menilai hasil belajar siswanya perlu juga diingatkan bahwa disamping memiliki kelebihan-kelebihan tersebut tes tipe pilihan berganda juga memiliki berbagai kelemahan, di antaranya;

- a. Membuat item tes pilihan berganda tidak mudah, dibutuhkan ketekunan karena menyita tenaga dan pikiran serta membutuhkan waktu yang panjang.
- b. Tidak semua sasaran belajar dapat diukur dengan tes pilihan berganda. Tes tipe pilihan berganda lebih sering dan tepat dipergunakan untuk mengukur sasaran belajar ranah kognitif pada tataran C1 (*ingatan/recall*) saja
- c. Item tes pilihan berganda memerlukan jawaban yang absolut, padahal secara teoritis hasil belajar siswa kebanyakan tidak bersifat absolut tetapi kondisional.

- d. Jawaban yang diberikan siswa mengandung unsur *gambling* atau untung-untungan yang cukup tinggi, artinya jawaban yang diberikan dapat ditebak sehingga hasil tes belum tentu menunjukkan kemampuan siswa yang sebenarnya.
- e. Membutuhkan banyak biaya, terutama karena item tes pilihan berganda biasanya cukup banyak sehingga tidak ekonomis.
- f. Makin terbiasa siswa mengerjakan item tes pilihan berganda akan semakin besar kemungkinannya untuk mendapatkan skor yang besar (*testwise*)

2. Tes Betul-Salah (B-S)/*true-false*

Penyusunan item tes B-S dalam pembelajaran PJOK terdiri dari dua komponen pernyataan dan alternatif jawaban yang dapat berbentuk betul-salah, setuju-tidak setuju, baik-tidak baik atau pernyataan lain yang harus bersifat *mutual-exclusive*. Item tes betul-salah di kalangan ahli penyusun item tes juga mempunyai beberapa nama, misalnya jawaban berganti-ganti (*alternate response*). Item tes betul-salah secara garis besar dapat dibedakan menjadi dua macam bentuk, yaitu; (a) item tes bentuk regular atau tidak dimodifikasi, dan (b) item tes bentuk modifikasi. Dalam konteks pembelajaran PJOK, item tes B-S bentuk regular atau tidak dimodifikasi adalah bentuk yang paling banyak dikuasai dan digunakan oleh para guru PJOK. Namun sebaliknya, item tes B-S bentuk modifikasi jarang dipergunakan karena dibutuhkan pengalaman dan pemahaman dalam penyusunan soal serta kegunaannya secara umum terbatas pada siswa yang lebih senior (mahasiswa).

Untuk mencapai hal tersebut guru PJOK perlu menguasai dan mampu mengonstruksi item tes B-S dengan memperhatikan beberapa aturan yang perlu dipenuhi oleh para evaluator atau guru PJOK. Kaidah-kaidah penyusunan item tes B-S ini pada prinsipnya dapat digunakan untuk mengonstruksi item tes B-S baik yang bentuk regular maupun bentuk modifikasi. Adapun kaidah-kaidah penyusunan item tes B-S yang perlu diperhatikan oleh guru PJOK adalah:

- a. Pernyataan item tes B-S sebaiknya dibangun secara cermat, sehingga para siswa tidak secara jelas dapat menerka item tes tersebut
- b. Item tes B-S tidak menanyakan hal-hal yang *trivial* tetapi harus benar-benar mampu mengukur hasil belajar yang bermakna.

Contoh:

- B – S Kompetisi sepakbola Indonesia Super Liga (ISL) bukan hanya merupakan sebuah olahraga yang cukup memasyarakat tetapi sudah merupakan industri olahraga yang menjanjikan.
- c. Setiap item tes B-S harus menguji hasil belajar lebih tinggi dari sekedar pengetahuan yang hanya mengandalkan ingatan apalagi hanya mengingat kata atau *frase*.
- Contoh:
- B – S Seorang perenang ingin meningkatkan kemampuan *VO₂ Max* maka salah satu metode latihan yang tepat dan harus digunakan adalah metode latihan interval pendek.
- d. Kunci jawaban yang dibuat guru PJOK harus pada kategori yakin benar (berkaitan dengan *content validity*). Jangan sampai kunci jawaban benar yang dibuat guru berbeda dengan jawaban benar yang dipersepsikan oleh siswa. Hal ini dapat saja terjadi terutama jika subjektivitas guru masuk dan mewarnai kunci jawaban.
- Contoh:
- B – S Seorang guru PJOK tipe otoriter akan lebih berhasil dalam menghasilkan siswa berprestasi olahraga dibandingkan dengan guru PJOK tipe demokratis.
- e. Hindari penggunaan kalimat yang terlalu panjang dan kompleks dengan kata-kata yang mempunyai makna ganda.
- f. Item tes dengan kunci jawaban salah sebaiknya jumlahnya lebih banyak dari item dengan kunci jawaban yang betul. Hal ini penting dilakukan untuk membalik proses berfikir hafalan yang biasanya hanya menghafal hal-hal yang benar tanpa diimbangi oleh hal-hal yang salah. Jadi apabila siswa dapat menjawab item tes B-S dengan benar, dapat diduga siswa tersebut juga dapat memahami hal-hal yang salah.
- g. Dalam menyusun item tes B-S, guru PJOK hendaknya menuliskan huruf (B-S) pada setiap permulaan item atau persoalan.
- h. Pernyataan pada setiap item sebaiknya tidak diambil dari kata-kata yang terdapat pada buku (*tex book*) langsung.

Bagi guru PJOK yang menggunakan item tes B-S perlu mengetahui dan memahami kelebihan dan kelemahannya agar dihasilkan item tes yang baik dalam mengeksplorasi kemampuan siswa. Beberapa **kelebihan item tes B-S** ini di antaranya: (1) seperti bentuk objektivitas tes lainnya, item tes B-S hasil akhir penilaian dapat menjadi objektif, (2)

perangkat item tes dapat mewakili seluruh pokok bahasan, (3) mudah diskor dan merupakan instrument yang baik untuk mengukur fakta dan hasil belajar langsung khususnya sasaran belajar yang memerlukan kemampuan mengingat, (4) dibandingkan dengan bentuk tes objektif lainnya, pekerjaan penyusunan soal yang digunakan dalam tes ini relative lebih mudah, (5) item tes B-S memiliki karakteristik yang menguntungkan, yaitu mudah dan cepat dalam menilai, (6) item tes B-S yang dikonstruksi secara cermat, membawa implikasi kepada siswa, yaitu waktu mengerjakan soal lebih cepat diselesaikan. Sedangkan beberapa kelemahan item tes B-S adalah: (1) item tes B-S memiliki tingkat reliabilitas yang rendah, (2) hasil belajar yang diukur lebih banyak kognitif yang mengandalkan pada kemampuan mengingat, (3) mengkonstruksi item tes B-S pada umumnya diperlukan waktu yang lebih lama jika dibandingkan dengan tes esai, (4) penggunaan pertanyaan alternatif lebih memungkinkan peserta didik mengira-ngira jawabannya, (5) item tes B-S kurang dapat membedakan siswa yang pandai dan yang kurang pandai. Ini berarti item tes B-S memiliki daya diskriminasi yang rendah (6) jika konstruksi kalimat pada item tidak dibuat secara cermat, dapat membingungkan siswa yang dites, (7) probabilitas antara jawaban salah dengan benar sama-sama 50 % sehingga siswa sangat mungkin menggunakan pola menebak dalam menentukan jawabannya.

3. Tes Menjodohkan/*matching*

Item tes menjodohkan sering disebut dengan *matching test item*. Item tes menjodohkan ini termasuk dalam kelompok tes objektif. Secara fisik, bentuk item tes menjodohkan terdiri dari dua kolom yang sejajar. Pada kolom pertama berisi pernyataan (*stem*) yang sering disebut dengan istilah **daftar stimulus**, dan kolom kedua berisi kata atau *frasa* yang disebut juga **daftar respon** atau **jawaban**.

Item tes menjodohkan ini pada umumnya dipergunakan terbatas pada pengukuran pengetahuan yang mencakup istilah/terminologi olahraga, definisi atau batasan, fakta, dan asosiasi konsep yang memiliki kaitan sederhana. Hal ini sesuai dengan pernyataan Cross (1982) bahwa "*matching test items are appropriate for identifying the relationship things*", artinya item tes menjodohkan adalah tepat untuk mengidentifikasi hubungan antar sesuatu.

Item menjodohkan jika disusun dengan cermat mampu mengukur pengetahuan batasan atau terminologi. Batasan definisi maupun asas pengetahuan merupakan hal yang sangat penting untuk dipahami oleh siswa. Kemampuan mengukur batasan dan asas tersebut

merupakan karakteristik penting yang mampu membuat guru PJOK dan evaluator banyak mempertimbangkan penggunaannya dalam mengukur pengetahuan para siswa.

Item menjodohkan pada prinsipnya dapat mengevaluasi pengetahuan tentang fakta yang memiliki makna spesifik. Agar dapat digunakan sebagai materi premis atau kolom respon, fakta harus ditulis sederhana dan jelas. Apabila kedua kriteria tersebut tidak dapat dipenuhi maka tipe tes lain perlu dipertimbangkan penggunaannya.

Tugas siswa dalam menjawab item tes tipe menjodohkan adalah mencarikan pasangan yang tepat bagi pokok soal dengan jawaban-jawaban yang tersedia. Menyusun item tes tipe menjodohkan tidaklah mudah khususnya untuk mengukur kemampuan belajar afektif dan psikomotor. Item tes dapat dikonstruksi dengan baik apabila penyusun tes memperhatikan hal-hal berikut:

- a. Pernyataan pokok soal dan masing-masing pernyataan jawaban harus terdiri dari permasalahan yang homogen, dengan ruang lingkup luas. Jika tidak memenuhi persyaratan tersebut maka tingkat kemudahan item tes menjadi tinggi. Dengan demikian tentu tidak akan memberikan manfaat dalam mengungkap kemampuan siswa yang sebenarnya.
- b. Pernyataan jawaban harus berjumlah lebih banyak daripada pernyataan pokok soal. Perhatikan contoh item tes tipe menjodohkan berikut:

Petunjuk: tuliskan huruf yang cocok dengan pernyataan pada kolom sebelah kanan pada titik-titik yang disediakan di sebelah kiri.

Jawaban	Pokok Soal	Alternatif Jawaban
.....	Latihan meningkatkan daya tahan	A. Asam laktat
.....	Penyebab terjadinya kelelahan	B. <i>VO2 Max</i>
.....	Beban latihan meningkat terus	C. <i>Weight training</i>
.....	Kemampuan tubuh mengkonsumsi O ₂	D. <i>Over training</i>
		E. <i>Interval training</i>
		F. Superkompensasi

- c. Lajur sebelah kanan harus berisikan jawaban-jawaban dan jumlahnya harus lebih banyak dua buah daripada yang ada pada lajur sebelah kiri.

- d. Ruang untuk menuliskan jawaban yang berupa huruf atau angka dibuat di sebelah kiri dari lajur sebelah kiri.
- e. Isi pernyataan/soal dalam tiap-tiap lajur (pokok soal dan alternatif jawaban) harus sejenis (homogen).
- f. Dalam petunjuk mengerjakan tes harus dijelaskan apakah jawaban di lajur sebelah kanan dapat dipergunakan lebih dari satu kali atau tidak.
- g. Semua item untuk satu set tes menjodohkan sebaiknya ditempatkan pada satu halaman. Penempatan kedua kolom pada halaman lain atau terpisah akan mengakibatkan siswa membaca sambil membolak-balik halaman.
- h. Isi dalam kedua lajur harus disusun menurut abjad atau apabila isinya angka-angka diurutkan dari angka yang terkecil ke angka yang terbesar.

Contoh:

Jawaban	Cabang olahraga	Negara asalnya
.....	Anggar	A. Amerika Serikat
.....	Atletik	B. India
.....	Bola Basket	C. <u>Inggris</u>
.....	Bulu Tangkis	D. <u>Italia</u>
.....	Golf	E. <u>Jerman</u>
.....	Polo Air	F. Perancis
.....	Tenis Lapangan	G. Skotlandia
		H. Spanyol
		I. Yunani

Sebagai salah satu alat evaluasi tes bentuk objektif tipe tes menjodohkan ini memiliki beberapa keunggulan dan kelemahan. Keunggulan item tes tipe menjodohkan, di antaranya; (1) tepat untuk mengukur hasil belajar berupa pengetahuan tentang definisi, istilah, fakta, peristiwa, atau penanggalan, (2) menguji dengan tepat kemampuan menghubungkan dua hal yang berhubungan langsung maupun tidak langsung, (3) mudah dikonstruksi sehingga guru PJOK dapat menghasilkan item tes untuk pokok bahasan tertentu dalam waktu yang cukup singkat, (4) dapat dipergunakan untuk menguji seluruh

isi cakupan materi PJOK, (5) mudah dalam melakukan pensekoran. Sedangkan kelamahan utama dari tipe tes menjodohkan adalah: tipe tes ini hanya mampu dipergunakan untuk mengukur sasaran belajar aspek kognitif pada strata yang paling rendah, yaitu pengetahuan (C-1) yang sangat mengandalkan kemampuan mengingat (*recall*). Dengan demikian sasaran belajar yang lebih tinggi (pemahaman, aplikasi, analisa, sintesa, dan evaluasi) dalam ranah kognitif tidak dapat diukur dengan item tes menjodohkan.

4. Tes Isian (*Supply test*)

Tes objektif jenis isian lebih mirip dengan tes esai sederhana. Pada item tes jenis isian memerlukan siswa untuk mengingat materi pembelajaran, dan menyimpan informasi materi dalam pikiran, kemudian menuangkannya dalam bentuk memberikan jawaban atas suatu pertanyaan. Tes objektif jenis isian pada prinsipnya mencakup tiga macam tes, yaitu; (1) tes jawaban bebas atau jawaban terbatas, (2) tes melengkapi, dan (3) tes analogi/asosiasi. Tes jawaban bebas mengungkap kemampuan siswa dengan cara bertanya, tes melengkapi mengungkap kemampuan siswa dengan memberikan ruang kosong atau spasi untuk diisi dengan jawaban (kata atau suku kata) yang tepat, sedangkan tes analogi atau asosiasi mengungkap kemampuan siswa dengan menyediakan ruang kosong atau spasi yang diisi dengan satu jawaban atau lebih.

Jawaban dalam tes analogi tersebut masih memiliki keterkaitan dan bersifat homogen antara satu dengan yang lainnya. Ketiga tes objektif jenis isian mempunyai kemiripan, khususnya dalam tiga hal, yaitu; (1) masing-masing tes memerlukan kemampuan siswa dalam menghafal, (2) masing-masing tipe menuntut jawaban singkat dari para siswa, (3) masing-masing tes pada umumnya direncanakan untuk mengungkap pemikiran siswa tentang materi pembelajaran yang dikategorikan sebagai definisi atau batasan, pengetahuan tentang fakta dan prinsip-prinsip pengetahuan.

Tes objektif jenis isian ini walaupun sudah dikelompokkan sebagai tes objektif namun sebenarnya masih berkaitan dengan tes esai karena tes ini masih menuntut jawaban bebas dan singkat dari para siswa. Namun karena tes isian hanya memberikan kesempatan kepada siswa menjawab dengan satu kata atau satu suku kata dan biasanya telah terikat dalam definisi, fakta, dan atau prinsip-prinsip pengetahuan maka tes tersebut dikatakan sebagai tes objektif jenis isian (*supply objective test*).

Mengkonstruksi item tes merupakan langkah penting yang harus dikuasai dengan baik oleh guru PJOK. Mengenai kemampuan apa yang diperlukan untuk mengkonstruksi item tes objektif bentuk isian (*supply*) dan pilihan (*selection*) pada umumnya dinyatakan oleh Grondlund (1995: 143), “*the construction of good test items is an art. The skill is requires, however are the same as those found in effective teaching*”, maknanya penyusunan item tes yang baik pada prinsipnya adalah seni. Keterampilan yang diperlukan pada prinsipnya sama dengan keterampilan yang diperlukan dalam pengajaran yang efektif.

Agar mendapatkan tes jenis isian yang memiliki kualitas susunan dan penampilan yang baik, maka guru PJOK dapat mempertimbangkan beberapa petunjuk sebagai berikut:

- a. Nyatakan petunjuk tes yang singkat dan jelas dengan cara memberikan garis bawah pada kata-kata kunci
- b. Harus diusahakan sungguh-sungguh bahwa hanya ada satu jawaban yang benar.
- c. Tempat yang disediakan untuk menulis jawaban harus sama panjang dan cukup panjang untuk menulis jawaban yang terpanjang.
- d. Untuk memudahkan mensekor tempat untuk menulis jawaban ditempatkan di sebelah kiri atau kanan kertas ujian.
- e. Pertanyaan diusahakan sedemikian rupa sehingga jawabannya sesingkat mungkin.
- f. Tempatkan spasi atau ruang kosong pada akhir kalimat dari item tes melengkapi.
- g. Pilih batasan atau terminologi dari suatu pengetahuan dengan menghilangkan kata kuncinya. Kata kunci tersebut menjadi jawaban yang harus diisi oleh para siswa.

Berikut akan disajikan contoh-contoh item bentuk tes objektif tipe isian (tes jawaban bebas atau terbatas/melengkapi, enumerasi, dan asosiasi/analogi)

Tipe Tes Isian	Contoh soal
1. Jawaban bebas atau terbatas (melengkapi)	Alat pengukur lebar panggul: Jawaban: <i>caliper</i> Pengukuran bagain-bagian tubuh: Jawaban: <i>antropometri</i>
2. Enumerasi	Berikan empat keuntungan dari perpindahan pertahanan daerah (<i>zone defense</i>) ke pertahanan satu lawan satu (<i>man to man defence</i>) dalam permainan bola basket. 1.

	2. 3. 4.
3. Analogi/Asosiasi	Dalam kejuaraan beregu suatu cabang olahraga setiap tim memperebutkan piala kejuaraan. sepak bola: <i>world cup</i> , bulu tangkis beregu putra: <i>thomas cup</i> , bulu tangkis beregu putri: <i>uber cup</i> , tenis lapangan beregu putra:, tenis lapangan beregu putri:, tenis meja beregu putra:, tenis meja beregu putri:

Bagi penyusun tes, khususnya guru PJOK yang nantinya akan menggunakan item tes objektif bentuk isian ini sebagai instrumen penilaian maka perlu mengetahui dan memahami kelebihan dan kelemahannya agar dihasilkan item tes yang baik dalam menggali kemampuan siswa. Kelebihan item tes bentuk isian ini adalah: (1) sangat baik untuk tujuan mengungkap kemampuan kognitif yang rendah dan berguna bagi siswa yang sejak awal ingin dididik dan dikembangkan melalui latihan-latihan yang secara periodik dan bertahap meningkat ke arah jawaban yang lebih kompleks, (2) dapat dipergunakan untuk mengetes batasan atau definisi pengetahuan yang sering dipergunakan oleh para guru mata pelajaran umum maupun para guru PJOK. Adapun kelemahan tes objektif bentuk isian ini, di antaranya: (1) kurang cocok untuk tes dengan sasaran pengukuran domain kognitif yang lebih tinggi, seperti: aplikasi (C3), analisa (C4), sintesa (C5), dan evaluasi (C6), (2) tipe tes ini hanya tepat dipergunakan untuk tujuan mengajar, jika penggunaan tes tipe ini yang terlalu banyak akan berpengaruh pada pengembangan kemampuan siswa pada sasaran pengukuran domain kognitif yang paling rendah (C1).

BAB VIII
PENYUSUNAN TES BENTUK ESAI (URAIAN)
UNTUK EVALUASI RANAH KOGNITIF
DALAM PJOK

Secara ontologis tes esai adalah salah satu tes tertulis yang susunannya terdiri atas item-item pertanyaan yang masing-masing mengandung permasalahan dan menuntut jawaban siswa melalui uraian-uraian kata yang merefleksikan kemampuan berfikir siswa. Tes esai dapat juga disebut sebagai tes dengan menggunakan pertanyaan terbuka, dimana dalam tes tersebut siswa diharuskan menjawab sesuai dengan tingkat pengetahuan yang dimiliki. Tes esai dikatakan tes dengan jawaban panjang, apabila dalam aplikasi tes memerlukan jawaban siswa secara luas. Evaluator memberikan kesempatan kepada siswa (*evaluand*) untuk memberikan jawaban secara tuntas dan jelas.

Bentuk pertanyaan atau item tes esai dapat dikonstruksi dengan menggunakan kata bantu tertentu yang mengandung unsur singkatan $4 W + 1 H$, yang maknanya; *who* (siapa), *where* (dimana), *what* (apa), *why* (mengapa), dan *how* (bagaimana). Di samping itu pertanyaan tes esai direncanakan secara sistematis untuk mendorong para siswa agar memiliki kemampuan mengekspresikan ide-ide mereka dengan menggunakan bahasa atau kata-kata mereka sendiri, menggunakan informasi dari pengetahuan mereka sendiri, kemudian menuangkannya secara bebas dalam lembar jawaban yang tersedia. Jadi dapat dibayangkan betapa beragamnya jawaban yang diberikan oleh siswa dalam menjawab soal, akibatnya dalam penilaianpun terkesan cenderung subjektif. Jawaban siswa pada tes esai dapat diberikan dengan cara: menguraikan, menjelaskan, mendiskusikan, membandingkan, memberi alasan atau bentuk lain yang sejenis. Ciri khas tes esai adalah jawaban item-item tes tidak disediakan pembuat tes namun sepenuhnya harus dipikirkan oleh peserta tes untuk selanjutnya di ekspresikan dalam bentuk tulisan.

Dewasa ini kecenderungan untuk kembali menggunakan tes esai untuk mengukur tingkat keberhasilan pembelajaran khususnya di perguruan tinggi dilandasi oleh beberapa pertimbangan, di antaranya: (1) meningkatkan kualitas hasil belajar yang cenderung menurun karena penggunaan tes objektif, (2) meningkatkan kemampuan siswa dalam menggunakan bahasa tulisan yang akhir-akhir ini cenderung menurun sebagai dampak dari penggunaan tes bentuk objektif yang berlebihan, (3) meningkatkan kemampuan analisis (C4), sintesis (C5), dan evaluasi

(C6) siswa yang rendah sebagai akibat membudayanya penggunaan tes bentuk objektif di kalangan siswa.

A. KARAKTERISTIK TES ESAI

Menurut Gronlund (1990: 216) tes bentuk esai dapat dipergunakan untuk mengidentifikasi kecakapan-kecakapan siswa, di antaranya:

1. Kecakapan mengaplikasikan prinsip-prinsip,
2. Kecakapan menginterpretasikan hubungan-hubungan,
3. Kecakapan mengenal dan menyatakan inferensi,
4. Kecakapan mengenal relevansi berbagai informasi,
5. Kecakapan mengenal dan merumuskan hipotesis,
6. Kecakapan mengenal dan merumuskan kesimpulan-kesimpulan secara tepat,
7. Kecakapan mengidentifikasi asumsi berdasarkan suatu kesimpulan,
8. Kecakapan mengenal keterbatasan data,
9. Kecakapan mengenal dan menyatakan masalah, dan
10. Kecakapan mendesain prosedur eksperimen.

Penggunaan tes esai dengan jelas menampakkan kemampuan dan kecakapan siswa dalam mengintegrasikan berbagai ide dan berbagai sumber informasi ke dalam suatu pola pikir tertentu yang disertai dengan pemecahan masalahnya. Integrasi ide-ide tersebut mensyaratkan kemampuan setiap siswa untuk mengekspresikannya secara teratur dan taat asas. Tanpa dukungan kecakapan prasyarat tersebut maka kemampuan sesungguhnya tidak dapat dimunculkan dengan baik, yang dapat dilihat dari susunan kalimat dan penyusunan paragraf jawaban.

Tes esai sebenarnya merupakan bentuk tes yang mampu untuk meningkatkan motivasi belajar siswa, karena dalam menjawab soal siswa akan menggunakan kalimat yang disusunnya sendiri sehingga penguasaan materi pelajaran secara komprehensif akan sangat menentukan. Siswa yang kurang menguasai materi pelajaran, akan nampak dengan jelas pada jawaban yang diberikan. Kemampuan siswa dalam mengekspresikan ide-ide dalam bentuk tulisan menuntut keterampilan menulis, oleh sebab itu tes esai akan mendorong siswa untuk belajar lebih keras menuangkan ide-idenya dalam bentuk tulisan. Siswa yang pada awalnya tidak memiliki kemampuan yang memadai untuk menuangkan ide dalam bentuk bahasa tulis,

maka siswa tersebut akan sungguh-sungguh belajar menuangkan ide-idenya dalam bentuk tulisan.

Namun demikian, bagi guru PJOK yang akan menggunakan tes esai sebagai alat evaluasi hasil belajar siswa perlu mempertimbangkan hal-hal sebagai berikut: (1) tes esai tingkat reliabilitasnya rendah, hal ini disebabkan karena soal-soal tes esai merupakan sampel dari materi ajar yang terbatas dan subjektifitas penskoran, (2) waktu yang dibutuhkan untuk menyelesaikan soal-soal tes esai relatif lama termasuk waktu yang dibutuhkan guru untuk memeriksa hasil tes, (3) jawaban yang diberikan oleh siswa kadang-kadang penuh karangan (membual) terutama bagi siswa yang tidak menguasai materi pelajaran dengan baik.

Bentuk pertanyaan tes esai menyebabkan jawaban yang dapat diberikan oleh siswa tidak memiliki batas, tergantung pada kemampuan analisis dan sintesis serta pandangan siswa terhadap masalah yang ditanyakan. Berikut disajikan contoh-contoh soal tes esai pada mata pelajaran PJOK.

1. Apa yang anda ketahui tentang kesegaran jasmani?. Jelaskan !
2. Jelaskan mengapa terjadi kerusuhan dalam pertandingan sepak bola?
3. Bagaimana cara mengatasi anak yang mengalami ketergantungan pada obat-obat psikotropika?
4. Bagaiman usaha untuk meningkatkan derajat kesehatan badan?
5. Mengapa klub-klub bola voli yang mengikuti ajang proliga bola voli menggunakan pemain asing?

Jika dicermati kelima pertanyaan di atas nampak bahwa kelimanya memberi peluang kepada siswa untuk memberikan jawaban yang luas, bergantung pada opini dan kemampuan masing-masing siswa. Jawaban yang diberikan siswa sangat bergantung pada pandangan dan kemampuannya menguasai permasalahan yang ditanyakan, serta kemampuan dalam mengekspresikan ide-ide ke dalam bentuk tulisan.

Karena jumlah soal tes esai pada umumnya sangat terbatas dan disesuaikan dengan ketersediaan waktu maka soal yang dibuat harus direncanakan sedemikian rupa dengan memperhatikan dua hal berikut: (1) kesesuaian soal dengan materi pelajaran dan tujuan pembelajaran yang ingin dicapai, (2) keseimbangan alokasi materi yang diujikan. Dengan mempertimbangkan dua hal tersebut diharapkan jangan sampai justru materi yang kurang

penting banyak diumunculkan dalam soal. Pegangan utama adalah munculkan soal untuk mengukur materi yang harus dikuasai oleh siswa.

B. KELEBIHAN DAN KELEMAHAN TES ESAI

Dalam proses belajar-mengajar teori PJOK di kelas, tes esai masih banyak dipergunakan para guru PJOK, karena tes esai memiliki beberapa **kelebihan**, di antaranya:

1. Pembuatan soal tes esai relatif lebih mudah jika dibandingkan dengan tes objektif.
2. Soal-soal tes dapat ditulis di papan tulis atau ditayangkan lewat LCD.
3. Cocok untuk mengukur kemampuan berfikir divergen.
4. Mengukur kemampuan siswa dalam menjawab melalui kata dan bahasa mereka sendiri.
5. Mendorong siswa untuk mempelajari, menyusun, merangkai, dan menyatakan gagasan siswa secara aktif.
6. Mendorong siswa untuk berani mengemukakan pendapat serta menyusun dalam bentuk kalimat mereka sendiri.
7. Tes esai mampu mengukur kemampuan yang lebih kompleks dan tingkat pemahaman yang lebih tinggi (C3, C4, C5, dan C6).

Di samping beberapa kelebihan seperti yang telah di uraikan di atas, ternyata tes esai juga memiliki beberapa **kelemahan** yang perlu diperhatikan oleh guru PJOK, di antaranya:

1. Pemberian nilai tidak ajeg antara penilai yang satu dengan penilai yang lainnya.
2. Tingkat kejelasan atau kerapian tulisan siswa dalam menjawab soal akan berpengaruh terhadap pemberian nilai.
3. Penilai dipengaruhi oleh unsur "*hallo effect*", ada kecenderungan pengaruh subjektif yang selalu muncul dalam pribadi seorang guru.
4. Validitas dan reliabilitas penilaian tes esai sulit didapat.
5. Cakupan materi pelajaran yang diujikan terbatas karena keterbatasan waktu yang dipergunakan untuk ujian.
6. Waktu yang dipergunakan untuk menilai pekerjaan siswa banyak.
7. Bentuk pertanyaan yang memiliki arti ganda, sering membuat kesulitan pada siswa sehingga memunculkan unsur-unsur meneka dan menjawab dengan ragu-ragu.

Berdasarkan uraian kelebihan dan kekurangan tes esai sebagaimana telah di uraikan di atas, guru PJOK dapat mempertimbangkan kelebihan dan kelemahan masing-masing tipe tes.

Yang lebih hakiki adalah seorang guru PJOK tersebut dapat memutuskan kapan ia menggunakan tes esai dan tes objektif.

C. MENGOREKSI TES ESAI

Memberikan sekor tes esai dipandang gampang-gampang susah oleh sebagian guru PJOK karena ada beberapa faktor yang terlibat di dalamnya. Menurut Sukardi (2010: 102) faktor-faktor tersebut di antaranya: (1) subjektivitas, 2) pertimbangan, dan (3) pengaruh interaksi antara guru dengan para siswa selama dalam proses belajar-mengajar berlangsung. Untuk mengatasi adanya pengaruh tiga faktor di atas, berikut ini ada beberapa petunjuk yang dapat digunakan sebagai acuan oleh guru PJOK dalam memberikan nilai tes esai, yaitu:

1. Ambillah beberapa pekerjaan siswa sebagai sampel. Bacalah jawaban siswa untuk memperoleh gambaran umum tentang kualitas jawaban yang diharapkan.
2. Periksa kembali kerangka jawaban atau jika perlu sumber yang telah ditetapkan seperti buku ajar. Cek kembali bobot angka bagi setiap soal.
3. Setelah mengecek kerangka jawaban dan sumber yang telah ditetapkan selesai, barulah mulai menilai pekerjaan siswa. Jika dalam memeriksa jawaban siswa sudah lelah atau jenuh sebaiknya jangan dilanjutkan karena penilai cenderung kurang teliti lagi.
4. Menilai satu pertanyaan pada semua lembar jawaban sebelum pindah ke pertanyaan lain.
5. Usahakan dalam proses penilaian jawaban soal tidak melihat nama siswa penjawabnya.
6. Menentukan nilai dari setiap pertanyaan berdasarkan bobot permasalahan, kompleksitas jawaban, dan waktu yang diperlukan untuk menyelesaikan jawaban.

BAB IX
PENGUKURAN DAN PENILAIAN RANAH AFEKTIF
DALAM PJOK

Dalam pengajaran pendidikan jasmani terdapat sejumlah rumusan tujuan pendidikan jasmani yang ingin dicapai. Rumusan-rumusan tujuan pendidikan jasmani tersebut terdiri dari: (1) untuk mengembangkan alat-alat atau organ-organ tubuh, (2) untuk mengembangkan neuro-muskuler (keterampilan dalam aktivitas jasmani), (3) untuk mengembangkan interpretif (pengetahuan, pertimbangan, pengertian dan lain-lain yang sejenis), (4) untuk mengembangkan personal-sosial (sikap, sportivitas, kerjasama, percaya diri, dan lain-lain).

Apabila rumusan-rumusan tujuan pendidikan jasmani tersebut diklasifikasikan ke dalam ranah kognitif, ranah psikomotor dan ranah afektif, maka tujuan nomor 1 digolongkan ke dalam ranah fisik, tujuan nomor 2 digolongkan ke dalam ranah psikomotor, tujuan nomor 3 digolongkan ke dalam ranah kognitif, dan tujuan nomor 4 digolongkan ke dalam ranah afektif (Annarino dkk., 1980: 65).

Pengajaran pendidikan jasmani yang dirancang dan dilaksanakan mulai dari TK sampai tingkat Perguruan Tinggi pada prinsipnya berupaya untuk mencapai tujuan-tujuan ke empat ranah tersebut. Untuk mengetahui seberapa jauh tujuan pengajaran pendidikan jasmani tersebut sudah mencapai ranah-ranah tersebut, perlu diadakan evaluasi atau penilaian. Secara ringkas langkah-langkah dalam penilaian menurut Arma Abdoellah (1988: 2) terdiri dari: (1) merumuskan tujuan, (2) membuat atau memilih alat penilai sesuai dengan tujuan, (3) menggunakan alat penilai untuk memperoleh data, (4) mengolah data untuk memperoleh satu penilaian.

Dalam pengajaran pendidikan jasmani, pengembangan ranah afektif ini perlu di angkat sebagai salah satu tujuan yang sangat penting untuk diperhatikan dan ditangani secara sungguh-sungguh sebagai langkah awal pembentukan *character building*, meskipun dalam realitanya masih ada beberapa guru pendidikan jasmani yang memahaminya hanya sebagai tujuan pengiring. Hal ini dipertegas oleh pendapat Arma Abdoellah (1989: 7) yang mengatakan bahwa: ranah afektif harus dinilai karena telah disepakati sebagai salah satu tujuan pendidikan jasmani dan program pendidikan jasmani telah dirancang untuk mencapai tujuan ranah afektif.

Tingkat keberhasilan pencapaian tujuan ranah afektif memang harus diakui “**sukar untuk diajarkan**” dan “**sukar diukur**”, namun dalam kenyataannya rangsangan begitu melimpah untuk mengembangkannya apabila pelaksanaan pengajaran pendidikan jasmani ini dikelola dengan baik. Tanpa disadari, melalui pengelolaan perilaku siswa banyak adegan pengembangan ranah afektif itu benar-benar tersaji dalam lingkup pelaksanaan pendidikan jasmani, seperti: dalam cetusan aspirasi tentang pentingnya “pembentukan watak” (*character building*) atau perlunya pengembangan “*fairplay*” dalam pertandingan.

Permasalahan yang timbul sampai saat ini adalah: pengembangan ranah afektif itu menjadi terlantar dan seakan-akan terabaikan. Hal ini disebabkan oleh falsafah yang keliru, yakni: “**pengembangan dan penilaian dalam ranah afektif sulit dilaksanakan dan sukar untuk diukur kebenarannya**”. Terlebih karena kebiasaan sebagian besar guru pendidikan jasmani yang **hanya** memfokuskan penilaian pada kemampuan kognisi (ranah kognitif) dan penguasaan keterampilan gerak (ranah psikomotor) semata-mata sebagai tujuan instruksionalnya.

Oleh karena itu, melalui artikel ini akan dibahas lebih mendalam mengenai aspek pengukuran dan penilaian ranah afektif dalam pengajaran pendidikan jasmani.

A. HIRARKISME PENGUKURAN RANAH AFEKTIF

Salah satu prinsip dasar yang harus senantiasa diperhatikan dan dipegang teguh dalam memberikan penilaian adalah prinsip kebulatan (komprehensif). Dengan menggunakan dasar ini guru pendidikan jasmani dalam memberikan penilaian terhadap hasil belajar siswa dituntut untuk mengevaluasi secara menyeluruh, baik dari ranah kognitif, ranah psikomotor, ranah fisik maupun ranah afektif.

Pengembangan ranah afektif merupakan sumbangan paling unik dari pendidikan jasmani. Menurut Rusli Lutan dkk., (2000: 124) melalui pengelolaan perilaku siswa beberapa adegan pendidikan ranah afektif dalam pendidikan jasmani dapat berlangsung, seperti: (1) belajar dan mematuhi peraturan yang dibutuhkan dalam pelaksanaan beberapa permainan atau pertandingan olahraga, (2) aktivitas dalam kelompok dan pengalaman kreatif dalam pendidikan jasmani merupakan sebuah laboratorium yang menyenangkan untuk mengembangkan sikap kerja sama, kepemimpinan, dan kesiapan untuk mematuhi kepemimpinan, (3) pelaksanaan tugas gerak menyadarkan siswa terhadap kelebihan dan keterbatasannya, hal ini merupakan landasan bagi pembentukan konsep dirinya, (4) aktivitas jasmani dapat mengajarkan siswa tentang pentingnya makna ketekunan yang melandasi berbagai kegiatan.

Pada ranah afektif ini, seorang individu memiliki sistem nilai yang telah mengontrol perilakunya cukup lama untuk mengembangkan suatu “**gaya hidup**”. Hasil belajar pada tingkat ini mencakup kegiatan yang sangat luas, tetapi penekanan yang utama adalah pada kenyataan bahwa tingkah laku tersebut adalah khas bagi siswa-siswa tersebut. Tujuan instruksional yang berkenaan dengan pola umum penyesuaian siswa yang meliputi: personal, sosial, dan emosional merupakan cakupan utama dalam ranah afektif ini.

Menurut Kirkendall dkk., (1982: 319) ada 5 jenjang pengukuran ranah afektif dalam pendidikan jasmani dalam usaha guru pendidikan jasmani mengembangkan jiwa sportivitas atau “*sense of fairplay*”. Adapun hirarkisme ranah afektif tersebut terdiri dari: (1) penerimaan atau *receiving*, (2) jawaban atau *responding*, (3) penilaian atau *valuing*, (4) organisasi atau *arganization*, dan (5) karakterisasi sebuah nilai atau seperangkat nilai atau *characterization by a value or value complex*.

1. Penerimaan atau *receiving*

Berarti kesediaan siswa-siswa mengikuti gejala atau rangsangan tertentu. Dari sudut pandang kegiatan belajar-mengajar hal ini berkenaan dengan siswa mendapat contoh/peragaan langsung dari guru pendidikan jasmani, kemudian siswa memfokuskan perhatiannya dalam mengikutipelajaran pendidikan jasmani sehingga siswa memperoleh pengetahuan dan pengalaman. Hasil belajar dalam bidang ini bergerak mulai dari kewaspadaan yang bersifat sederhana sampai perhatian siswa bersifat selektif. Penerimaan merupakan hasil belajar dari tingkat yang paling rendah dalam ranah afektif ini.

2. Jawaban atau *responding*

Berarti siswa berpartisipasi aktif dalam kegiatan belajar-mengajar pendidikan jasmani. Dalam jenjang ranah efektif yang kedua ini siswa tidak hanya berpartisipasi saja tetapi juga bereaksi dalam mengikuti pelajaran pendidikan jasmani mulai dari mengamati semua peraturan yang harus dipatuhi termasuk pelanggarannya, sampai siswa mulai terlibat secara emosional dan mulai merasa senang dengan berbuat baik kepada orang lain.

Hasil belajar dalam bidang ini mungkin menekankan kemauan siswa untuk menjawab pertanyaan dan kepuasan dalam menjawab siswa. Tingkat lebih tinggi dalam kategori ini mencakup tujuan instruksional yang biasanya diklasifikasikan dibawah “**minat**”, yaitu: jawaban yang menekankan pencarian dan penikmatan kegiatan tertentu dalam *mengikuti pelajaran pendidikan jasmani*.

3. Penilaian atau *valuing*

Penilaian berkenaan dengan usaha siswa memberikan nilai terhadap suatu objek, gejala atau tingkah laku yang satu dengan yang lain mulai dari penerimaan yang bersifat sederhana sampai kepada komitmen yang kompleks. Menilai didasarkan atas internalisasi dari sejumlah nilai-nilai tetapi petunjuk atas nilai-nilai ini dinyatakan dalam tingkah laku siswa yang nyata.

Hasil belajar dalam bidang ini berkenaan dengan tingkah laku yang tetap sehingga nilai tersebut dapat diidentifikasi. Contoh: siswa menerima nilai bahwa berbuat sportif itu adalah baik dan diperlukan dalam pengajaran pendidikan jasmani. Tanpa timbul keraguan siswa tersebut kemudian berbuat untuk mematuhi peraturan dan berbuat sportif itu merupakan sesuatu yang benar. Selanjutnya siswa menghayati dengan jelas dan memperlihatkan bahwa pengamalan sportivitas itu bukan hanya untuk diri sendiri tetapi juga meyakinkan orang lain. Tujuan instruksional yang biasanya ditekankan pada kategori tahap penilaian atau *valuing* ini adalah “**penanaman sikap**” dan “**apresiasi**”.

4. Organisasi atau *organization*

Organisasi berkenaan dengan mempersatukan nilai-nilai yang berbeda, memecahkan pertentangan sesama mereka dalam mengikuti pembelajaran pendidikan jasmani dan mulai membangun suatu nilai batiniah yang tetap. Jadi penekanannya adalah menghubungkan dan mensintesakan nilai-nilai.

Pada tahap organisasi ini siswa melihat bahwa konsep keadilan dan sportivitas atau *fair play* itu dapat dipakai dalam aspek kehidupan lainnya serta menempatkan pengamalan sportivitas itu ke dalam sistem nilai. Tujuan instruksional yang berhubungan dengan “**pengembangan filsafat hidup**” termasuk dalam kategori ini.

5. Karakterisasi sebuah nilai atau seperangkat nilai atau *characterization by a value or value complex*

Karakterisasi sebuah nilai atau seperangkat nilai adalah keterpaduan semua sistem nilai yang telah dimiliki oleh siswa yang mempengaruhi pola kepribadian dan tingkah lakunya. Pada tahap ini proses internalisasi nilai telah menempati nilai tertinggi dalam suatu hirarki nilai. Nilai itu telah tertanam secara konsisten pada sistemnya dan telah mempengaruhi emosinya. Ini merupakan tingkatan ranah afektif tertinggi, karena sikap batin siswa telah benar-benar bijaksana. Siswa telah memiliki *philosophy of life* yang

mapan. Jadi pada jenjang ini siswa telah memiliki sistem nilai yang mengontrol tingkah lakunya dalam waktu yang cukup lama sehingga membentuk karakteristik “**pola hidup**”, tingkah lakunya yang menetap, konsisten, dan dapat diramalkan.

Dalam manifestasi pengajaran pendidikan jasmani, siswa selalu berperilaku konsisten dalam semua aspek kehidupan dengan memperlihatkan perilaku sportif dan selalu memiliki kesediaan untuk mengambil jenis aksi tertentu. Aksi tersebut terdiri dari: (1) *mental set* atau kesediaan mental beraksi, (2) *physical set* atau kesediaan fisik beraksi, serta (3) *emotional set* atau kemauan untuk beraksi.

B. PERSOALAN-PERSOALAN DALAM PENGUKURAN DAN PENILAIAN RANAH AFEKTIF

Walaupun alat penilaian /evaluasi telah banyak dibuat untuk mengukur perilaku afektif yang menjadi perhatian bagi guru pendidikan jasmani, namun hanya sedikit yang telah dibuat menurut standar untuk pembuatan tes pendidikan jasmani dan psikologi. Dalam banyak kasus, walaupun keterandalan tes sikap dapat diterima, namun dari aspek kesahihan isi (*content validity*) sangat minim dan kesahihan konsep (*construct validity*) tidak memadai dan sering tidak mendapat perhatian (Safrit, 1986: 372).

Sedangkan menurut Arma Abdoellah (1989: 173-176) ada beberapa permasalahan yang perlu dicermati oleh para guru pendidikan jasmani dalam melaksanakan pengukuran dan penilaian ranah afektif. Permasalahan-permasalahan tersebut terdiri dari: (1). kebanyakan instrumen untuk menilai perilaku afektif mengandung kemungkinan distorsi jawaban. Siswa mungkin tidak selalu menjawab dengan penuh kejujuran sesuai dengan apa yang ada dalam perasaan atau intuisinya, (2). kegagalan untuk menetapkan suatu kerangka kerja teoritis yang memadai untuk konsep minat, (3). alat penilaian untuk mengukur ranah afektif masih jarang. Jadi penilaian sikap sejauh ini agak diabaikan. Hal ini dimungkinkan oleh asumsi guru pendidikan jasmani yang menganggap penilaian sikap kurang begitu penting.

Persoalan-persoalan senada juga dikemukakan oleh Rusli Lutan, Dkk., (2000: 129-131) yang mengatakan bahwa: (1) ada perbedaan yang amat mendasar antara skor pengukuran dalam pendidikan jasmani, keterampilan gerak, skor tes kemampuan kognitif, dan skor tes kemampuan ranah afektif. Skor pengukuran terhadap aspek fisik seseorang dan kognitif mengungkap kemampuan maksimal seseorang sebagai implementasi tingkat penguasaan bahan, sedangkan skor afektif lebih menunjukkan tentang apa yang dirasakan seseorang dalam situasi tertentu.

Oleh karena itu, variasinya akan banyak dan bersifat situasional, (2). hasil pengukuran ranah afektif juga dipengaruhi oleh faktor situasional, terutama motif pemanfaatan hasilnya. Apabila hasil pengukuran dimaksudkan untuk membuat keputusan penting yang mempengaruhi keadaan diri responden atau apabila responden tidak mengerti tidak mengerti untuk apa data itu digunakan, maka hal itu dapat berakibat data yang diperoleh tidak valid, (3) pengukuran ranah afektif terkait dengan latar belakang budaya yang membentuk jawaban responden, seperti halnya ukuran-ukuran kepatutan perilaku di suatu daerah, (4) siswa cenderung mengemukakan jawaban yang tidak sesungguhnya sesuai dengan kenyataan yang ada apabila mengetahui bahwa dirinya sedang dinilai, (5) hasil pengukuran ranah afektif tidak stabil, hal ini terkait dengan variasi dari hari ke hari yang dialami oleh siswa, (6) konstruksi butir tes dalam menilai ranah afektif cenderung mengeceh responden, (7) konstruksi butir tes dalam penilaian ranah afektif cenderung mendorong responden menjawab bagaimana patutnya yang ideal dan terbaik, (8) pengukuran ranah afektif yang terdiri dari respon berupa kategori yang menunjukkan derajat, seperti: sering kadang-kadang, jarang sekali. Makna yang dikenakan oleh seseorang terhadap kata “**sering** “ atau “**jarang sekali**” sungguh dapat berbeda-beda, (9). kriteria pengukuran ranah afektif tidak memadai. Apa kriteria dan bagaimana memperoleh skor kriteria bagi sebuah instrumen pengukuran ranah afektif sungguh amat sulit, hal ini terkait dengan ketidakjujuran responden untuk mengemukakan jawabannya.

C. INSTRUMEN PENILAIAN RANAH AFEKTIF

Beberapa instrumen yang efektif telah disusun untuk mengukur perilaku sosial, sikap, dan penyesuaian diri terhadap pengajaran pendidikan jasmani. Menurut Miller (2002: 285-286) ada 7 jenis inventori yang dapat dipergunakan sebagai instrumen untuk menilai ranah afektif, di antaranya: (1). inventori untuk mengukur sikap/ *attitude inventories*, (2). inventori minat/*interest inventories*, (3). inventori sportivitas/*sportmanship*, (4). inventori kepemimpinan/*leadership*, (5). inventori perkembangan sosial/*social behaviour*, (6). Inventori perilaku/*behavior rating*, dan (7). inventori kepribadian/*personality inventories*.

1. Inventori sikap atau *attitude inventories*

Banyak penelitian-penelitian penelitian terdahulu dalam pendidikan jasmani tentang ranah afektif berfokus pada sikap dan pengukurannya. Sikap adalah kesiapan untuk berbuat yang masih pada taraf mental sehingga disebut juga dalam istilah predisposisi mental untuk berbuat. Namun demikian, perilaku nyata dapat diartikan sebagai cerminan dari sikap. Sikap

itu biasanya terentang antara dua kutub, mulai dari kutub negatif sampai kutub positif. Kutub negatif merupakan cerminan dari sikap seseorang yang tidak menyukai suatu objek, namun sebaliknya kutub positif merupakan cerminan dari sikap seseorang yang menyukai suatu objek.

Inventori sikap merupakan salah satu jenis inventori yang bertujuan mengungkap “**perasaan suka atau tidak suka** ” dan dijawab sendiri oleh responden. Objek sikap biasanya berupa perasaan tentang: objek-objek yang berkaitan dengan jasmani/fisik, tipe-tipe orang, orang/individu tertentu, institusi sosial, dan kebijakan pemerintah (Morrow, 2005: 357).

Inventori sikap biasanya tertuju pada pengungkapan kedua kecenderungan itu yang dipahami bergerak dalam sebuah garis kontinum. Kecenderungan sikap bergerak di antara kedua kutub dan kemudian berubah ke arah makin positif atau makin negative. Perubahan seperti itu disebut perubahan yang selaras (kongruen). Sebaliknya perubahan sikap dapat terjadi dari positif ke negatif atau sebaliknya dari negative ke positif (inkongruen).

Berkenaan dengan konteks pendidikan jasmani, inventori sikap sebenarnya dikembangkan untuk mengukur 6 (enam) dimensi keterlibatan aktif dan pasif dalam aktivitas jasmani, yaitu: (a). aktivitas jasmani sebagai pengalaman sosial, (b). aktivitas jasmani untuk kesehatan dan kebugaran, (c). aktivitas jasmani sebagai pengejaran kepentingan (sensasi atau kebahagiaan), (d). aktivitas jasmani sebagai pengalaman estetis, (e). aktivitas jasmani sebagai katarsis (perasaan terharu), dan (f). aktivitas jasmani sebagai pengalaman asetis atau pertapa (Morrow, 2005: 358).

(2) Inventori minat atau *interest inventories*

Bagaimana kecenderungan minat siswa terhadap proses belajar-mengajar pendidikan jasmani dapat diungkap melalui inventori minat yang bisa dijawab sendiri oleh responden. Instrumen ini bermanfaat untuk memotret kecenderungan minat siswa sehingga didapatkan umpan balik yang dapat dipakai sebagai bahan untuk merencanakan program pengajaran pendidikan jasmani yang sesuai dengan minat para siswa. Program tersebut dapat berkenaan dengan kegiatan intra kurikuler maupun ekstra kurikuler. Berdasarkan hasil evaluasi tersebut juga dapat diungkapkan faktor-faktor penyebab mengapa siswa menyukai atau tidak menyukai terhadap suatu program, seperti: karena pengaruh pengalaman masa lalu, tekanan

teman sepermainan,, rangsangan guru atau orang tua, dan faktor social budaya yang lain. Jawaban terhadap inventori minat dapat ditafsirkan sebagai gambaran sikap seseorang.

(3). Inventori sportivitas (*sportmanship*)

Instrumen ini tertuju pada beberapa aspek perilaku sportif, seperti: ketaatan seseorang terhadap peraturan, penerimaan kekelahan dan kemenangan secara terhormat, dan puas apabila berbuat sesuatu hal yang baik bagi kelompok maupun terhadap pihak lawan (Miller,2002: 267). Sportivitas seringkali sebagai bagian dari cerminan watak seseorang yang ditampilkannya ketika terlibat dalam situasi pertandingan.atau perlombaan.

Guru pendidikan jasmani dapat menerapkan penilaian berskala untuk mengamati perkembangan sikap sportivitas para siswa. Inventori sportivitas dapat dipergunakan oleh guru pendidikan jasmani untuk mengamati perilaku siswanya manakala mengikuti mengikuti berbagai perlombaan dan pertandingan antar kelas (*class meeting*) maupun pada waktu mengikuti festival olahraga seperti: porseni, popda, maupun popnas.

4. Inventori kepemimpinan (*leadership*)

Inventori kepemimpinan dimaksudkan untuk mengidentifikasi siapa di antara siswa yang dapat ditunjuk sebagai pemimpin seperti: ketua kelas, ketua regu, kapten tim, dan kepentingan lain yang berkaitan dengan kepemimpinan di kelas maupun di sekolah. Yang ingin diukur dari inventori ini adalah tingkat kecakapan atau keterampilan memimpin.

Kemajuan keterampilan memimpin dapat dinilai dengan penilaian berskala, hasilnya dapat dibahas dan didiskusikan bersama-sama dengan para siswa. Pengisian inventori kepemimpinan dilakukan sendiri oleh siswa agar bisa objektif dan tidak ada unsur campur tangan dari pihak luar (guru pendidikan jasmani).

5. Inventori perkembangan sosial (*social behaviour*)

Inventori perkembangan sosial bertujuan untuk mengukur sejauh mana seorang siswa itu dapat diterima oleh siswa lainnya, atau dapat pula digunakan untuk mengukur seberapa jauh siswa mampu menyesuaikan diri (adaptasi) dengan lingkungan di kelas atau di sekolah.

Siswa yang baru masuk ke lingkungan yang baru, seperti: awal masuk di kelas, sekolah, ataupun pindah sekolah, biasanya ada yang mengalami kesulitan untuk beradaptasi dengan lingkungan yang baru. Kecemasan siswa seringkali timbul karena siswa tersebut bukan hanya menyesuaikan dengan lingkungan baru namun mungkin juga teman-temannya

mengalami kesulitan untuk menerimanya secara tulus sebagai kawan baru. Pengisian inventori perkembangan sosial dapat dilakukan sendiri oleh siswa, teman sejawat, maupun *rating* yang dilakukan oleh guru pendidikan jasmani.

6. Rating perilaku (*behavior rating*)

Setelah melakukan pengamatan/observasi terhadap siswa dalam waktu yang cukup lama guru pendidikan jasmani mengamati perilaku siswa. *Rating* yang dilakukan oleh guru pendidikan jasmani meliputi beberapa aspek perilaku yang berbeda, di antaranya: sportivitas, kepemimpinan, sikap, pengendalian diri (*self control*), motivasi intrinsik, kerjasama, dan hubungan antara teman sejawat (Miller, 2002: 267).

Validitas, reliabilitas, dan objektivitas hasil penilaian memang dapat dipertanyakan. Seringkali hasil penilaian itu dipengaruhi oleh "*hallo effect*", yaitu: penilaian dipengaruhi oleh sifat-sifat lain yang telah diketahui berkenaan dengan siswa yang bersangkutan. Contoh: "Si A" diketahui sejak dahulu identik dengan sikap dan perilakunya yang sering bikin gaduh/ribut di kelas, meskipun ia berbuat baik dan menunjukkan kerjasama yang bagus dalam permainan, bisa saja penilaian guru pendidikan jasmani dipengaruhi oleh perilakunya terdahulu yang seolah-olah sudah menjadi *trade mark* pada diri "Si A" tersebut.

7. Inventori kepribadian (*personality inventory*)

Inventori kepribadian diarahkan untuk mengukur sifat-sifat atau ciri-ciri, seperti: kontrol emosi, kepercayaan diri, motivasi, sifat agresif, determinasi, ketenangan (*poise*), dan ketegaran mental (*mental toughness*) (Miller, 2002: 267). Inventori ini terdiri dari sejumlah butir tes yang diisi atau diselesaikan sendiri oleh siswa.

Instrumen ini tergolong sangat sulit dalam pelaksanaannya, oleh karena itu seyogyanya guru pendidikan jasmani sebaiknya dibantu oleh tenaga ahli/professional dalam bidangnya dan sudah memiliki bekal pengalaman yang cukup memadai dalam pelaksanaannya.

D. TIPE-TIPE SKALA PENGUKURAN DAN PENILAIAN RANAH AFEKTIF

Untuk mengungkap data yang berkenaan dengan ranah afektif lazimnya dapat menggunakan skala. Skala adalah seperangkat set angka-angka yang menyatakan nilai-nilai tentang suatu objek atau perilaku, dan atas dasar itu nilai kualitatif dinyatakan secara kuantitatif (Nurhasan, 2001: 113). Skala sering dipergunakan untuk mengukur sikap dan

karakteristik lainnya, misalnya: skala dipergunakan untuk mengukur sikap para siswa terhadap pendidikan jasmani atau topik-topik lainnya.

Menurut Arma Abdoellah (1988: 193) ukuran perilaku afektif/sikap dalam pendidikan jasmani telah digunakan paling kurang dalam tiga cara, yaitu: (1). untuk menilai tingkat perilaku afektif, (2). untuk menilai tingkat perilaku afektif dalam satu kelas, (3). untuk mengontrol perasaan dalam melakukan satu proyek penelitian.

Pengukuran sikap bertujuan ingin mengungkapkan kecenderungan seseorang dalam menetapkan pilihannya dalam sebuah rentang penilaian, “**sangat suka**” sampai dengan “**sangat tidak suka**” terhadap suatu objek. Terpusatnya perhatian pada pengukuran sikap terhadap pendidikan jasmani tidaklah mengherankan, karena banyak guru pendidikan jasmani berasumsi bahwa seorang siswa dengan sikap yang positif lebih cenderung untuk berpartisipasi dalam aktivitas jasmani.

Untuk mengumpulkan data yang mengungkapkan aspek dari ranah afektif, dapat mempergunakan instrumen yang bervariasi. Dua metode yang paling populer dan semakin sering dipergunakan sebagai instrumen untuk pengukuran dan penilaian ranah afektif dalam pendidikan jasmani adalah: (1) metode kuantitatif (jenis kuesioner) dengan berbagai tipe skalanya, dan (2). metode kualitatif. Metode kuantitatif jenis kuesioner yang paling sering dipergunakan adalah: (a). skala likert/*summated rating scales*, (b). skala diferensial semantik, dan (3). skala urutan peringkat (*rank order*), dan (4). skala butir pilihan-paksa (*forced-choice*). Sedangkan metode kualitatif terdiri dari: (a) wawancara, dan (b). observasi (Morrow, 2005: 345-349).

(1). Metode Kuantitatif

(a). Skala Likert

Salah satu di antara bentuk skala yang paling banyak dipergunakan untuk mengukur sikap adalah Skala Likert. Skala Likert merupakan skala nilai 5 atau 9 (nilai 1-5 atau nilai 1-9) yang mengasumsikan sama di antara tanggapan. Responden menentukan jawabannya sendiri sesuai dengan pendapat atau perasaannya. Pada contoh ini, perbedaan antara **sangat setuju** dengan **setuju** dianggap *equivalent*/sama dengan perbedaan antara **tidak setuju dengan sangat tidak setuju**. Tipe skala ini dipergunakan untuk menilai **tingkat setuju** dan **tidak setuju** dengan pernyataan dan dipergunakan secara luas pada inventori sikap. Contoh *item* yang digunakan pada Skala Likert disajikan pada tabel 11.

Tabel 11: “Semua siswa SMA diharuskan untuk mengikuti pendidikan jasmani selama 2 tahun”.

Sangat Setuju	Setuju	Ragu-Ragu	Tidak Setuju	Sangat Tidak Setuju
5	4	3	2	1

Keunggulan utama tanggapan/jawaban yang diskala adalah bahwa tanggapan tersebut memungkinkan pilihan pernyataan yang lebih luas dibandingkan dengan tanggapan dalam bentuk kategori yang pada umumnya bersifat dikotomi, yaitu: yang memberikan pilihan jawaban seperti “*ya*” dan “*tidak*” atau jawaban “*benar*” dan “*salah*”. Dengan menggunakan lima, tujuh, atau lebih interval jawaban, juga akan semakin meningkatkan reliabilitas instrument tersebut. Disamping itu, penggunaan kata-kata tanggapan yang berbeda dapat digunakan dalam tanggapan berskala, seperti: *sangat bagus, bagus, cukup, jelek, dan sangat jelek*, atau *sangat penting, penting, agak penting, tidak penting, dan sangat tidak penting*.

Berikut akan diberikan contoh tanggapan berskala dengan menggunakan interval 5, interval 7, dan interval 11.

Tabel 12: Contoh penggunaan kata-kata tanggapan yang berbeda dengan 5 interval.

1	2	3	4	5
Tidak Pernah	Kadang-Kadang	Agak Sering	Sering	Selalu

Tabel 13: Contoh penggunaan kata-kata tanggapan yang berbeda dengan 7 interval

1	2	3	4	5	6	7
Selalu						Tidak Pernah

Tabel 14: Contoh penggunaan kata-kata tanggapan yang berbeda dengan 11 interval

1	2	3	4	5	6	7	8	9	10	11
Setuju										Tidak Setuju

(Morrow, 2005: 346).

(b). Skala Diferensial Semantik (*Semantic Differential Scales*)

Tipe pengukuran skala terkenal lainnya adalah penggunaan Skala Diferensial Semantik. Pada skala ini siswa/responden diminta untuk menanggapi kata sifat bipolar,

pasangan kata sifat dengan makna yang berbeda, seperti: lemah-kuat, rilek-tegang,cepat-lambat, berat-ringan, aktif-pasif dan lain-lain. Dengan menggunakan kata sifat bipolar, responden diminta untuk memilih nilai pada rangkaian tersebut yang paling menggambarkan perasaan mereka tentang konsep tertentu. Berdasarkan hasil penelitian, Skala Diferensial Semantik mengukur tiga faktor utama, yaitu: (a).komponen evaluasi baik-buruk), (b). komponen potensi (kuat-lemah), dan (c). komponen aktivitas (aktif-pasif) (Morrow, 2005: 346).

Daftar berikut ini menunjukkan beberapa contoh kata sifat bipolar yang mengukur ketiga komponen tersebut.

(a). komponen evaluasi

Menyenangkan	Tidak Menyenangkan
Adil	Tidak Adil
Jujur	Tidak Jujur
Baik	Buruk
Berhasil	Gagal
Berguna	Tidak Berguna

(b). komponen potensi

Kuat	Lemah
Keras	Lunak
Berat	Ringan
Dominan/Otoriter	Pengikut/ <i>Submissive</i>
Lembut	Kasar
Kotor	Bersih

(c). komponen aktivitas

Siap	Gemetar/ <i>Nervous</i>
Bahagia	Sedih
Aktif	Pasif
Dinamis	Statis

Tetap	Bergerak
Cepat	Lambat

Pada tabel 15 berikut akan disajikan contoh Skala Diferensial Semantik yang dipergunakan untuk mengukur sikap terhadap aktivitas jasmani.

Tabel 15: Skala Diferensial Semantik untuk mengukur sikap terhadap aktivitas jasmani.

Aktivitas Jasmani								
	1	2	3	4	5	6	7	
Jelek								Bagus
Tidak Menyenangkan								Menyenangkan
Tegang								Santai
Dingin								Panas
Tidak Sehat								Sehat
Mengerikan								Baik
Kasar								Lembut
Pasif								Aktiv

Skala tersebut di atas terdiri dari 8 pasangan sifat dengan rentang skor 1 sampai dengan skor 7, apabila jawaban/respon negatif berada di sebelah kiri dan jawaban/respon positif di sebelah kanan (misalnya: jelek – bagus). Dengan cara ini maka akan diperoleh skor keseluruhan. Dengan demikian, maka akan diperoleh skor yang berbeda-beda dari setiap siswa.

(c). Skala Urutan-Peringkat (*Rank Order*)

Skala urutan-peringkat mensyaratkan bahwa butir-butir harus diberi peringkat oleh penjawab sesuai dengan kriteria tertentu. Sebuah contoh dari skala urutan-peringkat adalah *Rokeach Value Survey*, yang telah dipakai untuk mempelajari nilai dari pelatih dan administrator. Dalam instrument ini, satu seri konsep satu-nilai (*single value*), seperti: kesamaan, kebebasan, pengakuan social diberikan individu untuk diberi peringkat menurut pentingnya bagi individu (Arma Abdoellah, 1988: 175-176).

(d). Skala Butir Pilihan-Paksa (*Forced-Choice*)

Menurut Arma Abdoellah (1988: 176) dengan skala pilihan-paksa individu diminta untuk memilih di antara dua atau lebih alternatif yang kelihatannya sama baik atau sama tidak baik. Diskriminasi dan nilai butir yang disenangi ditentukan dan alternatif digabungkan sedemikian rupa untuk menyeimbangkan nilai-nilai.

(2). Metode Kualitatif

Metode yang semakin populer akhir-akhir ini untuk mengukur dan menilai ranah afektif adalah metode kualitatif karena metode tersebut memberikan kekayaan informasi yang biasanya tidak dimanfaatkan ketika digunakan kuesioner tradisional. Dua cara yang paling sering dipergunakan oleh metode kualitatif untuk pengukuran dan penilaian ranah afektif dalam pendidikan jasmani adalah melalui: (a). wawancara (*interview*), dan (b). observasi (Morrow, 2005: 348).

(a). wawancara (*interview*)

Tidak diragukan lagi bahwa wawancara merupakan sumber data yang paling umum pada metode kualitatif. Wawancara berkisar dari jenis wawancara yang terstruktur, dimana butir-butir pertanyaan ditentukan sebelum wawancara benar-benar dimulai, sampai wawancara terbuka yang memungkinkan responden memberikan tanggapan secara bebas.

Interviewer atau pewawancara yang bagus pertama kali harus menciptakan hubungan dengan responden untuk memungkinkan mereka saling terbuka dan menggambarkan pemikiran dan perasaan mereka yang sesungguhnya. Pewawancara tidak diperkenankan menilai terkait dengan isi tanggapan responden, serta harus menjadi seorang pendengar yang baik.

Menggunakan alat bantu *tape-recorder* mungkin metode perekaman wawancara yang paling umum, karena metode tersebut mempertahankan keseluruhan wawancara untuk analisis data selanjutnya. Walaupun persentase subjek yang kecil awalnya tidak nyaman untuk direkam, kesulitan ini biasanya hilang dengan cepat.

Mencatat selama melakukan wawancara merupakan metode yang sering digunakan lainnya untuk merekam data, kadang-kadang digunakan untuk bantuan

perekaman ketika pewawancara ingin menekankan hal penting tertentu. Salah satu kekurangan mencatat tanpa perekaman akan mengakibatkan pewawancara sangat sibuk sehingga mengganggu pemikiran dan pengamatan terhadap responden.

(b). observasi

Observasi juga merupakan salah satu teknik pengukuran dan penilaian ranah afektif dalam pendidikan jasmani dari metode kualitatif. Walaupun sebagian besar penelitian sebelumnya sangat bergantung pada observasi langsung dengan pencatatan dan pengkodean kategori tingkah laku tertentu, kecenderungan yang sekarang adalah melakukan observasi dengan menggunakan *video-tape*.

Seperti halnya penggunaan *tape-recorder*, untuk melakukan wawancara yang memungkinkan observer meninjau ulang setiap apa yang dikatakan oleh subjek, peranan pemakaian *video-tape* berfungsi mengambil tingkah laku subjek untuk dianalisis lebih lanjut. Dalam observasi lapangan sangatlah penting bahwa observer berusaha bersikap rendah hati terhadap subjek yang sedang diteliti. Subjek yang mengetahui bahwa dia sedang diamati dan direkam bisa mengubah perilaku mereka. Oleh karena itu, hal terpenting dari observasi ini adalah membuat suasana agar perilaku subjek yang diamati dan direkam itu benar-benar merupakan cerminan perilaku alamiah subjek.

BAB X
PENILAIAN ALTERNATIF (*ALTERNATIVE ASSESSMENT*)
DALAM PJOK

Dalam proses pembelajaran pendidikan jasmani, kegiatan evaluasi merupakan salah satu aspek yang sangat penting, disamping aspek-aspek yang lain, seperti: tujuan, materi, sarana-prasarana, dan metode pembelajaran. Data yang dikumpulkan melalui pengukuran sangat berguna untuk keperluan evaluasi dan pengambilan keputusan dalam proses pendidikan. Upaya peningkatan kualitas pembelajaran pendidikan jasmani dapat ditempuh melalui peningkatan kualitas alat evaluasinya. Keduanya saling berkaitan erat, sistem pembelajaran yang baik akan menghasilkan kualitas belajar yang baik. Salah satu indikator kualitas pembelajaran yang baik dapat dilihat dari hasil penilaiannya. Selanjutnya sistem penilaian yang baik akan mendorong pendidik untuk menentukan strategi mengajar yang baik dan memotivasi anak didik untuk belajar yang lebih baik.

Penilaian berfokus pada siswa, yaitu prestasi belajar yang dicapai oleh anak didik. Proses penilaian meliputi pengumpulan bukti-bukti tentang pencapaian belajar anak didik. Bukti ini tidak selalu diperoleh melalui “**alat evaluasi tes**” saja, tetapi juga dapat dikumpulkan melalui “**alat evaluasi alternatif yang pada umumnya berbentuk alat evaluasi nontes**”, seperti: pengamatan (*observation*), wawancara (*interview*), angket (*questionnaire*), pemeriksaan dokumen (*documentary analysis*), serta penilaian terhadap kumpulan hasil karya siswa (portofolio). Hal ini didukung oleh pendapat James Morrow (2005: 128) bahwa pencapaian hasil belajar siswa dapat diukur bukan saja melalui *assesment* yang berupa tes, namun dapat juga ditempuh melalui *assesment* yang berupa penilaian alternatif yang berbentuk nontes, seperti: observasi, proyek individu atau kelompok, portofolio, pertunjukan/pameran, serta melalui catatan siswa atau membuat jurnal. Dengan penilaian alternatif yang pada umumnya dengan menggunakan teknik nontes maka penilaian hasil belajar siswa dilakukan dengan tanpa “**menguji**” siswa, melainkan dilakukan dengan dengan melakukan pengamatan secara sistematis, melakukan wawancara, menyebarkan angket, dan memeriksa atau meneliti dokumen-dokumen dan pengumpulan hasil karya siswa melalui portofolio, pertunjukan/pameran, serta melalui catatan siswa atau membuat jurnal.

Alat evaluasi alternatif berupa instrumen nontes ini pada umumnya memegang peranan yang sangat penting dalam rangka mengevaluasi hasil belajar siswa dari segi ranah sikap hidup (*affective domain*) dan ranah keterampilan (*psikomotoric domain*), sedangkan alat evaluasi dengan menggunakan tes lebih banyak digunakan untuk mengevaluasi hasil belajar siswa dari segi ranah proses berfikir (*cognitive domain*). Dalam memberikan penilaian hasil belajar siswa, guru pendidikan jasmani memerlukan data yang baik mutunya sehingga perlu didukung oleh proses pengukuran yang baik pula.

Menurut Chittenden yang dikutip oleh Djemari Mardapi (2004: 17) kegiatan evaluasi dalam proses pembelajaran baik dengan tes maupun nontes sebelumnya mesti didahului oleh kegiatan pengukuran. Dalam kegiatan evaluasi seyogyanya di arahkan pada empat hal, yaitu: (1). Penelusuran, (2). Pengecekan, (3). Pencarian, dan (4). Penyimpulan. **Penelusuran**, yaitu: kegiatan yang dilakukan untuk menelusuri apakah proses pembelajaran telah berlangsung sesuai dengan yang direncanakan atau tidak. Untuk kepentingan ini, guru mengumpulkan berbagai informasi sepanjang semester atau tahun pelajaran melalui berbagai bentuk pengukuran untuk memperoleh gambaran tentang pencapaian kemajuan belajar anak.

Pengecekan, yaitu untuk mencari informasi apakah terdapat kekurangan-kekurangan pada anak didik selama berlangsungnya proses pembelajaran. Dengan melakukan berbagai bentuk pengukuran, guru berusaha untuk memperoleh gambaran yang berkaitan dengan kemampuan peserta didiknya mengenai apa yang telah berhasil dikuasai dan apa yang belum dikuasai.

Pencarian, yaitu untuk mencari dan menemukan penyebab kekurangan yang muncul selama proses pembelajaran berlangsung. Dengan jalan ini guru dapat segera mencari solusi untuk mengatasi kendala-kendala yang timbul selama proses belajar berlangsung. **Penyimpulan**, yaitu untuk menyimpulkan tentang tingkat pencapaian belajar yang telah dimiliki anak didik. Hal ini sangat penting bagi guru untuk mengetahui tingkat pencapaian yang diperoleh anak. Selain itu, hasil penyimpulan ini dapat digunakan sebagai laporan hasil tentang kemajuan belajar anak, baik untuk anak itu sendiri sekolah, orang tua, maupun pihak-pihak lain yang membutuhkan.

Menurut Stark dan Thomas (1994: 19) dalam melakukan evaluasi terhadap anak didik sebaiknya menggunakan pendekatan *goal free*, artinya walaupun tujuan utama dari evaluasi adalah untuk meningkatkan prestasi belajar siswa, namun dapat diperoleh hasil lain dalam diri siswa yaitu berupa: pembentukan rasa percaya diri, mengembangkan kecakapan/keterampilan, kreativitas, kemandirian, kepekaan sosial, dan sebagainya. Hasil-hasil lain yang berupa

pembentukan rasa percaya diri, pengembangan kecakapan atau keterampilan, kreativitas, kemandirian, dan kepekaan sosial siswa tersebut dapat dievaluasi dengan menggunakan alat penilaian yang berupa non tes.

Dari pendapat tersebut dapat ditarik suatu makna bahwa dalam mengevaluasi prestasi belajar siswa seorang guru baik untuk keperluan penelusuran, pengecekan, pencarian, penyimpulan hendaknya tidak hanya terpaku pada alat evaluasi dengan menggunakan tes semata-mata, namun alternatif penilaian lain dapat digunakan, yaitu dengan alat evaluasi nontes. Melalui artikel ini akan diuraikan evaluasi hasil belajar siswa melalui alat evaluasi non tes dalam pendidikan jasmani.

A. PENGEMBANGAN INSTRUMEN ALAT EVALUASI

Guru pendidikan jasmani apabila ingin mengevaluasi kemampuan hasil belajar siswa-siswinya kadangkala dihadapkan pada permasalahan bahwa alat evaluasi (tes) yang akan dipergunakan untuk mengadakan penilaian belum tersedia sehingga guru tersebut perlu menyusun sendiri instrumen yang dimaksud. Tujuan keseluruhan dari evaluasi yang dilakukan oleh guru adalah untuk menyediakan informasi yang valid guna dipakai untuk pengambilan keputusan. Informasi penilaian yang baik menyediakan indikator yang akurat dari penampilan peserta didik dan memudahkan guru pendidikan jasmani, pelatih, dan pendidik (instruktur, widyaiswara, dll.) untuk membuat keputusan yang tepat. Beberapa rambu-rambu yang perlu diperhatikan dalam pengembangan alat evaluasi di antaranya adalah: (1). Menentukan tujuan tes dan pengukuran dan memilih kriteria penilaian yang akan dipergunakan, apakah menggunakan PAN atau PAP, (2). Memilih atau mengembangkan butir keterampilan yang dapat mengetes kemampuan yang diinginkan sesuai dengan kriteria, (3). Memberikan jaminan bahwa alat-alat yang dipergunakan aman, bebas dari potensi mencelakakan siswa, (4). Buatlah percobaan dalam skala kecil untuk mengkaji keterpakaian tes, (5). Membuat revisi dan petunjuk penggunaan tes, (6). Menentukan validitas internal tes, (7). Mengembangkan norma penilaian, (8). Menyusun pedoman pelaksanaan tes, serta (9). Secara periodik tes itu harus disempurnakan agar tidak ketinggalan zaman (Rusli Lutan, dkk, 2000: 177-178). Rambu-rambu yang harus diperhatikan dalam penyusunan alat evaluasi nontes memiliki kesamaan dengan rambu-rambu penyusunan alat evaluasi dalam menyusun tes, di antaranya adalah: (1). Menentukan tujuan dari penilaian, (2). Menentukan sasaran atau hasil yang diharapkan dari siswa/outcome, (3). Memilih atau membuat tugas penilaian

yang tepat, (4). Menetapkan kriteria penilaian keterampilan yang ditampilkan siswa, serta (5). Menentukan kualitas dari penilaian (James R. Morrow, 2005: 148). Dengan adanya rambu-rambu tersebut diharapkan dapat dipakai sebagai pedoman guru dalam menyusun alat evaluasi yang dipergunakan untuk menilai prestasi belajar siswa.

Setiap mata pelajaran tidak terkecuali mata pelajaran pendidikan jasmani memerlukan kemampuan berfikir. **Kemampuan berfikir** termasuk dalam cakupan ranah kognitif dengan tataran kemampuan proses berfikirnya dimulai dari: kemampuan menghafal, kemampuan memahami, kemampuan menerapkan, kemampuan menganalisis, kemampuan mensintesis, dan kemampuan mengevaluasi. Kemampuan yang penting pada ranah kognitif adalah kemampuan menerapkan konsep-konsep untuk memecahkan masalah yang terjadi di lapangan. Kemampuan ini sering disebut dengan kemampuan **mentransfer** pengetahuan ke berbagai situasi sesuai dengan permasalahan aktual yang terjadi. Dalam konteks pendidikan jasmani hal ini sering dialami oleh siswa manakala terlibat pertandingan baik beregu maupun perseorangan. Pada pelaksanaan pelajaran pendidikan jasmani khususnya cabang-cabang permainan, tuntutan untuk menganalisis kelebihan dan kelemahan serta penerapan taktik maupun strategi untuk memenangkan pertandingan merupakan salah satu wujud nyata implementasi dari hirarkisme kemampuan proses berfikir dalam ranah kognitif.

Kemampuan yang kedua yang perlu dimiliki oleh peserta didik adalah **keterampilan psikomotor**, yaitu kemampuan yang berkaitan dengan gerak yang menggunakan otot, seperti: jalan, lari, lempar, maupun lompat. Hirarkisme kemampuan psikomotorik menurut Sax (1980: 76) ada lima, yaitu: (a). gerakan refleks, (b). gerakan dasar, (c). kemampuan persepsual, (d). kemampuan fisik (gerakan terampil), dan (e). komunikasi non diskursip. Gerakan refleks adalah respon gerak tanpa sadar yang muncul ketika bayi lahir. Gerakan dasar adalah gerakan yang mengarah pada keterampilan kompleks yang khusus. Siswa yang telah mencapai kompetensi dasar pada ranah psikomotorik ini mampu melakukan tugas dalam bentuk keterampilan sesuai dengan standar atau kriteria yang semestinya, seperti: siswa mampu melakukan lompat jauh gaya jongkok dengan teknik yang benar, siswa mampu melakukan pas bawah dalam bola voli, siswa mampu mendribel bola melewati beberapa pancang, dan sebagainya.

Kemampuan persepsual adalah kombinasi antara kemampuan kognitif dan kemampuan gerak. Kemampuan fisik adalah kemampuan untuk mengembangkan gerakan

yang paling terampil. Karakteristik gerakan yang paling terampil adalah gerakan yang mampu dilakukan oleh siswa sehingga menghasilkan produk yang optimal, seperti: keterampilan melakukan salto dalam senam lantai, kemampuan melakukan *lay up and shoot* dalam permainan bola basket, kemampuan melakukan pukulan *forehand groundstroke* ke arah sasaran tertentu dalam tenis, dan gerakan-gerakan terampil lainnya. Untuk mencapai gerakan yang terampil tersebut siswa harus belajar secara sistematis melalui tahapan-tahapan tertentu.

Komunikasi nondiskursip adalah kemampuan berkomunikasi dengan menggunakan gerakan. Hal ini berkaitan dengan kemampuan mengucapkan kata-kata dalam mempelajari bahasa asing, seperti ketika mahasiswa belajar mengucapkan/menghafal istilah-istilah dalam anatomi, seperti: gerakan *pronasi, supinasi, flexi, ekstensi, abduksi, adduksi* dan lain-lain. Gerakan ini mencakup gerakan lidah, penempatan lidah dan intonasi suara sehingga mahasiswa dapat mengucapkan berbagai istilah anatomi dengan benar.

Mata pelajaran yang banyak berhubungan dengan dengan ranah psikomotor salah satu di antaranya adalah pendidikan jasmani, karena dalam pembelajaran pendidikan jasmani selalu berhubungan dengan gerak anggota badan atau panca indera. Gerakan anggota badan peserta didik melalui tahapan tertentu, dimana setiap tahapan gerakan memiliki kunci gerakan, seperti: gerakan menendang bola dalam sepak bola, gerakan memukul bola tenis dalam *forehand groundstroke*, gerakan servis dalam permainan bola voli, gerakan guling depan dalam senam dan sebagainya.

Kemampuan ketiga yang harus dimiliki oleh peserta didik adalah ranah afektif atau kemampuan sikap peserta didik. Ranah afektif mencakup watak perilaku, seperti: perasaan, minat, sikap, emosi, atau nilai. Menurut Djemari Mardapi (2008: 101), ranah afektif sangat menentukan tingkat keberhasilan belajar peserta didik. Siswa yang tidak memiliki minat pada pelajaran tertentu maka akan sulit untuk mencapai keberhasilan studi secara optimal. Siswa yang berminat dalam suatu mata pelajaran diharapkan akan mencapai hasil pembelajaran yang optimal. Dengan mengingat begitu pentingnya peran ranah afektif bagi peserta didik, maka seyogyanya guru pendidikan jasmani harus mampu membangkitkan minat semua peserta didik belajar terhadap mata pelajaran yang diampu guru. Di samping itu ikatan emosional sering diperlukan untuk membangun semangat kebersamaan, semangat persatuan, semangat nasionalisme, rasa sosial dalam pelaksanaan pembelajaran jasmani. Untuk itu

semua lembaga pendidikan dalam merancang program pembelajaran harus memperhatikan ranah afektif.

B. KRITERIA-KRITERIA PENILAIAN ALTERNATIF

Proses pengembangan penilaian alternatif agar menjadi penilaian yang lebih bermakna pada prinsipnya langkah-langkah yang dilakukan sama dengan proses pengembangan penilaian konvensional pada umumnya, yaitu: (a) menentukan tujuan penilaian, (b) menentukan target atau hasil yang diharapkan dari peserta didik, (c) memilih atau menentukan tugas yang sesuai, (d) menentukan kriteria untuk menilai penampilan peserta didik, dan (e) menentukan kualitas penilaian.

Ada 3 parameter/persyaratan utama yang harus dipertimbangkan jika seorang guru pendidikan jasmani ingin menyusun penilaian alternatif. Menurut Griffin dkk. (1991: 133), parameter-parameter tersebut terdiri dari: (a). apakah penilaian alternatif menyediakan informasi yang akurat untuk pengambilan keputusan, (b). apakah hasil penilaian alternatif memberikan kesimpulan yang akurat dan adil tentang penampilan peserta didik, (c). apakah penggunaan hasil penilaian alternatif memberikan kontribusi pada pengambilan keputusan yang baik. Untuk dapat menjawab pertanyaan yang ke tiga (butir c) maka ada 3 kriteria penilaian alternatif yang harus dipenuhi, yaitu: reliabilitas, validitas, dan objektivitas.

Reliabilitas berkaitan dengan konsistensi skor atau observasi. Sebuah tes yang tidak reliabel (andal) secara esensial tidak berguna karena tidak menyediakan informasi yang berguna bagi pengguna (*user*). Karena penilaian alternatif sangat tergantung pada keputusan subjektif guru (penilai) untuk menilai atau menginterpretasikan penampilan peserta didik, maka guru harus memperhatikan reliabilitas dan objektivitas pengukuran yang dilakukan.

Untuk memastikan agar penilaian alternatif yang dilakukan oleh guru pendidikan jasmani memenuhi persyaratan sebagai instrumen yang reliabel/handal, maka harus memenuhi 3 persyaratan berikut:

- 1) Definisikan dengan baik penampilan apa yang akan dinilai, kriteria penilaian harus dituangkan secara jelas dan rinci tentang kualitas penampilan peserta didik yang diharapkan.
- 2) Pembuat keputusan (teman sejawat, guru, pelatih, *expert judgement*) harus memahami secara menyeluruh kriteria penilaian yang dipergunakan untuk menilai.

- 3) Pembuat keputusan sudah belajar bagaimana mengaplikasikan kriteria pemberian skor dengan cara yang konsisten. Keseragaman aplikasi kriteria pemberian skor dapat dicapai dengan baik melalui latihan dan praktik menilai.

Validitas adalah indikasi seberapa baik penilaian benar-benar mengukur apa yang sedang diukur. Keterandalan sangat diperlukan dalam penilaian alternatif, namun tidak menjamin bahwa penilaian yang handal juga merupakan indikasi bahwa penilaian tersebut juga penilaian yang valid. Sebuah penilaian dapat dengan sempurna dikatakan handal, namun tidak relevan dengan keputusan yang dimaksudkan. Apabila yang dihasilkan tidak berkaitan dengan karakteristik yang sedang diukur, maka hal ini dapat membahayakan kesimpulan yang akurat tentang penampilan peserta didik dan keputusan yang mengikutinya. Sayangnya sampai saat ini hanya mempunyai informasi empiris yang sedikit tentang validitas penilaian alternative. Bentuk utama validitas penilaian alternatif baru sebatas validitas tampak (*face validity*) dan berhubungan dengan validitas isi/*content validity*. Hal ini didasarkan pada hubungan yang diasumsikan bahwa jika pengajaran yang dilakukan baik, maka penilaian yang dilakukanpun juga akan baik. Walaupun kaitan antara pengajaran dan penilaian itu penting, validitas isi (*content validity*) sendiri tidak harus diterima sebagai bukti yang cukup untuk dipergunakan sebagai metode penilaian. Guru pendidikan jasmani perlu mencari bukti-bukti yang lebih memperkuat bahwa penilaian alternatif dapat memberikan kesimpulan yang valid (akurat). Linn dkk (1995: 134) menyarankan kriteria-kriteria untuk menentukan validitas penilaian alternatif, di antaranya adalah:

- 1) Konsekuensi (*consequences*): apakah penilaian mendorong pada dampak yang diharapkan ?. Demikian juga sebaliknya apakah penilaian mendorong dampak yang tidak diharapkan.
- 2) Keadilan (*fairness*): apakah penilaian memudahkan peserta didik untuk mendemonstrasikan kecakapan mereka yang sebenarnya, atau apakah merugikan peserta didik dengan tidak adil ?
- 3) Transfer dan generalisasi (*transfer and generalizability*): apakah hasil penilaian dipergunakan untuk menggeneralisasikan permasalahan lain atau situasi ranah yang lain?
- 4) Pengetahuan yang kompleks (*cognitive complexity*): apakah penilaian cukup memadai sebagai pertimbangan tingkat pemahaman dan pemikiran ranah kognitif yang lebih tinggi/lebih kompleks ?

- 5) Kualitas isi (*content quality*); apakah penilaian benar-benar dipergunakan untuk mengukur isi/materi ranah yang menjadi fokus untuk dinilai.
- 6) Ketercakupan isi (*content coverage*): apakah penilaian cukup mewakili keluasan isi dari materi yang diinginkan ?
- 7) Keberartian (*meaningfulness*): apakah penilaian berarti bagi peserta didik dan apakah hal itu dapat memotivasi mereka untuk menunjukkan kemampuan terbaik peserta didik ?
- 8) Biaya dan efisiensi (*cost and efficiency*): apakah biaya pengumpulan data dan pemberian skor penilaian dapat diterima nalar dan efisien ?

Walaupun kriteria-kriteria ini sebenarnya lebih cocok diaplikasikan pada penilaian di kelas, pertimbangan cermat dari setiap desain dan penggunaan penilaian alternatif pada pendidikan jasmani dan kemampuan penilai tidak diragukan lagi akan meningkatkan kualitas maupun validitas penilaian alternatif.

Objektivitas penilaian adalah kriteria yang harus dipertimbangkan dalam penilaian alternatif selain validitas dan reliabilitas. Penilaian dikatakan objektif/adil apabila penilaian tersebut memperlakukan peserta didik tanpa memandang jenis kelamin, suku, atau latar belakang dengan kesempatan yang sama untuk melakukan yang terbaik. Walaupun banyak keberagaman di dalam masyarakat kita dan peserta didik tidak datang ke sekolah dengan latar belakang, pendapatan, motivasi, atau nilai yang sama, tetapi semua peserta didik mempunyai kesempatan yang sama untuk menampilkan kemampuan dan pengetahuan yang akan dinilai.

Apakah penilaian itu objektif dan bebas dari bias ? Apakah penilaian memperlakukan lebih baik pada peserta didik laki-laki atau perempuan, peserta didik dari suku tertentu, peserta didik yang tinggal di lokasi tertentu atau yang keluarga kaya?. Untuk dapat melakukan penilaian dengan objektif, penilaian harus mencerminkan pengetahuan, nilai, dan pengalaman yang mencukupi untuk semua peserta didik dan harus mencari supaya mengukur pengetahuan dan kemampuan yang dimiliki peserta didik. Selain itu, ada hal yang substansial yang harus diperhatikan penilai bahwa prosedur pemberian skor dan kriteria untuk memutuskan suatu prestasi hasil belajar peserta didik terbebas dari bias. Hal ini dapat membantu memastikan bahwa penilaian kecakapan peserta didik mencerminkan kemampuan sebenarnya dari peserta didik dan bukan fungsi bias dan persepsi dari orang yang menilai kemampuan peserta didik.

C. ALAT EVALUASI ALTERNATIF (*ALTERNATIVE ASSESSMENT*) DALAM PENDIDIKAN JASMANI

Kegiatan melakukan pengukuran adalah merupakan kegiatan yang paling umum dilakukan dan merupakan tindakan yang mengawali kegiatan evaluasi dalam penilaian hasil belajar peserta didik. Kegiatan mengukur itu pada umumnya tertuang dalam bentuk tes dengan berbagai variasinya. Dalam praktiknya, teknik tes inilah yang lebih sering dipergunakan dalam rangka mengevaluasi hasil belajar peserta didik jika dibandingkan dengan penilaian alternatif (teknik nontes).

Pernyataan sebagaimana tertuang seperti di atas tidaklah diartikan bahwa teknik tes adalah satu-satunya teknik untuk melakukan evaluasi hasil belajar, sebab masih ada teknik lainnya yang dapat dipergunakan, yaitu penilaian alternatif melalui teknik nontes. Menurut Wood (2003: 131), alat evaluasi yang selama ini dipergunakan oleh guru pendidikan jasmani untuk menilai ranah kognitif (seperti: tes pilihan banyak, tes benar-salah, tes penjadohan) dan ranah psikomotorik (seperti tes keterampilan olahraga) dipandang sudah tidak mencukupi untuk dipakai sebagai dasar menilai peserta didik. Hal ini sebagai akibat dari perubahan/ akselerasi pendidikan yang begitu cepat disesuaikan dengan karakterisasi hasil belajar, kemampuan kognitif yang lebih tinggi, dan pembelajaran terintegrasi peserta didik. Di samping itu instrumen dan teknik penilaian tradisional (tes pengetahuan dan beberapa tes keterampilan) cenderung mengukur secara dangkal karakteristik yang didefinisikan, tidak memfasilitasi integrasi kemampuan atau proses, dan kadang-kadang bersifat buatan. Hal ini menghasilkan pergeseran baru untuk meningkatkan penggunaan bentuk penilaian alternatif sebagai alat untuk mengevaluasi hasil belajar peserta didik.

Menurut Morrow (2005: 128) dengan teknik penilaian alternatif yang berbentuk nontes, maka penilaian atau evaluasi hasil belajar siswa dilakukan dengan tanpa “menguji” peserta didik melainkan dilakukan dengan: (1). pengamatan secara sistematis (*observation*), (2). melakukan wawancara (*interview*), (3). menyebarkan angket (*questionnaire*), (4). memeriksa/meneliti dokumen-dokumen (*documentary analysis*), (5). Pertunjukan atau pameran (*exhibition*), serta (6). penilaian terhadap kumpulan semua hasil karya siswa (*portofolio*).

1. PENGAMATAN (*OBSERVATION*)

Secara umum pengertian observasi adalah cara menghimpun bahan-bahan yang dilakukan dengan mengadakan pengamatan dan pencatatan secara sistematis terhadap fenomena-fenomena yang dijadikan sebagai sasaran observasi. Observasi sebagai alat evaluasi nontes banyak dipergunakan untuk menilai tingkah laku individu atau proses terjadinya suatu kegiatan yang dapat diamati, baik dalam situasi yang sebenarnya (*real situation*) maupun dalam situasi buatan (*artificial situation*). Observasi dapat mengukur atau menilai hasil dan proses belajar, misalnya: perilaku siswa pada waktu guru pendidikan jasmani menyampaikan pelajaran praktik permainan di lapangan, perilaku siswa pada saat jam istirahat di antara jam pelajaran, saat jam teori kesehatan pendidikan jasmani di dalam kelas, perilaku siswa pada saat terjadinya kekosongan pelajaran, perilaku pada saat praktikum di laboratorium, dan sebagainya.

Observasi dapat dilakukan baik secara partisipatif (*participant observation*) maupun nonpartisipatif (*nonparticipant observation*). Observasi dapat pula berbentuk observasi eksperimental (*experimental observation*), yaitu: observasi yang dilakukan dalam situasi buatan atau berbentuk observasi yang dilakukan dalam situasi yang wajar (*nonexperimental observation*). Pada observasi partisipatif, *observer* (dalam hal ini guru pendidikan jasmani, dosen, dan sebagainya) melibatkan diri/membaur di tengah-tengah

siswa, mahasiswa, dan sebagainya), sedangkan pada observasi nonpartisipatif, *observer* berada “**di luar garis**” sehingga *observer* bertindak sebagai penonton belaka atau seorang *visitor*.

Pada observasi eksperimental di mana perilaku yang diharapkan muncul karena siswa dikenai perlakuan (*treatment*), maka observasi memerlukan perencanaan dan persiapan yang benar-benar matang, sedangkan pada observasi yang dilaksanakan pada situasi yang wajar pelaksanaannya jauh lebih sederhana karena observasi semacam ini dapat dilakukan secara sepintas lalu walaupun garis besar tentang hal-hal yang diobservasi juga harus dituliskan terlebih dahulu.

Anas Sudijono (1996: 77) mengatakan bahwa apabila observasi dipergunakan sebagai alat evaluasi maka harus diingat bahwa pencatatan hasil observasi itu pada umumnya jauh lebih sukar daripada mencatat jawaban-jawaban yang diberikan oleh siswa terhadap pertanyaan yang diberikan dalam suatu tes, ulangan maupun ujian sebab respon

yang diperoleh dalam observasi adalah berupa unjuk kerja/*performance* siswa. Mencatat *performance*/unjuk kerja siswa adalah merupakan pekerjaan yang sulit sebab observer selaku evaluator harus dapat dengan secara cepat mencatatnya. Pencatatan terhadap segala sesuatu yang dapat disaksikan dalam observasi itu penting sekali sebab hasilnya akan dijadikan landasan untuk menilai makna yang terkandung dibalik unjuk kerja siswa tersebut.

Observasi yang dilaksanakan dengan didahului oleh pembuatan perencanaan yang matang dikenal dengan istilah observasi sistematis (*systematic observation*). Pada observasi sistematis pelaksanaan observasi dilaksanakan berlandaskan pada kerangka kerja yang memuat faktor-faktor yang telah diatur kategorisasinya. Isi dan luas materi observasinya pun telah ditetapkan dan dibatasi secara tegas, sehingga pengamatan dan pencatatan yang dilakukan oleh evaluator dalam rangka evaluasi hasil belajar siswa sifatnya selektif. Faktor-faktor apa saja yang tercantum dalam pedoman observasi itulah yang diamati dan dicatat. Di luar itu tidak perlu dilakukan pengamatan dan pencatatan.

Berikut ini diberikan contoh instrumen evaluasi berupa rubrik penilaian untuk mengobservasi kemampuan melakukan *forehand groundstroke* dalam permainan tenis.

Mata Pelajaran : Penjas (Tenis Lapangan)											
Topik : <i>forehand goundstroke</i>											
Kelas :											
Cawu / Semester :											
No	Nama Siswa	Skor / Nilai untuk tiap-tiap Gerakan / Aspek:								Jumlah	Rata-rata:
		1	2	3	4	5	6	7	8		

Hasil penilaian dengan menggunakan instrumen tersebut di atas sifatnya individual. Setelah observasi selesai dilakukan, nilai-nilai individual itu dimasukkan ke dalam daftar nilai yang bersifat kolektif.

Selain observasi yang bersifat sistematis kadang kala untuk menilai *performance*/unjuk kerja peserta didik, guru pendidikan jasmani dapat melakukan observasi nonsistematis, yaitu: observasi yang dilakukan di mana *observer* atau *evaluator* dalam melakukan pengamatan dan pencatatan terhadap hasil unjuk kerja siswa tidak dibatasi oleh kerangka kerja yang pasti. Dengan demikian observasi yang dilakukan tersebut semata-mata hanya dibatasi oleh tujuan dari observasi itu sendiri. Contoh: seorang guru pendidikan jasmani mengobservasi jalannya simulasi pertandingan bola voli antar siswa pada waktu pelajaran pendidikan jasmani guna mengetahui dan menilai partisipasi dan keaktifan siswa dalam pembelajaran pendidikan jasmani tersebut. Penilaian atau evaluasi hasil belajar yang dilaksanakan dengan melakukan observasi itu disamping memiliki kebaikan, juga tidak terlepas dari kekurangan-kekurangan.

Di antara segi kebaikan yang dimiliki oleh observasi itu ialah, bahwa: data observasi itu diperoleh secara langsung di lapangan, yakni dengan jalan melihat dan mengamati kegiatan atau ekspresi peserta didik di dalam melakukan sesuatu tugas gerak, seperti: lompat jauh, lompat tinggi, guling depan, guling belakang, dan sebagainya sehingga dengan demikian data tersebut dapat lebih bersifat obyektif dalam melukiskan aspek-aspek keterampilan teknik gerakan peserta didik menurut keadaan yang senyatanya.

Data hasil observasi dapat mencakup berbagai aspek kepribadian maupun keterampilan gerak masing-masing individu peserta didik, dengan demikian maka di dalam pengolahannya tidak berat sebelah atau hanya menekankan pada salah satu segi saja dari kecakapan atau prestasi belajar mereka.

Adapun segi-segi kelemahannya di antara lain adalah, bahwa:

- a. Observasi sebagai salah satu alat evaluasi hasil belajar tidak selalu dapat dilakukan dengan baik dan benar oleh para pengajar. Guru yang tidak atau kurang memiliki kecakapan atau keterampilan dalam melakukan observasi, maka hasil observasinya menjadi kurang dapat diyakini kebenarannya. Untuk menghasilkan data observasi yang baik, seorang guru harus mampu membedakan antara; apa yang tersurat, dengan apa yang tersirat.
- b. Kepribadian (*personality*) dari *observer* atau *evaluator* juga acapkali mewarnai atau menyelinap masuk dala penilaian yang dilakukan dengan cara observasi. Prasangka-prasangka yang memungkinkan melekat pada diri *observer* (*evaluator*) dapat

mengakibatkan sulit dipisahkan secara tegas mengenai perilaku peserta didik yang diamati.

- c. Data yang diperoleh dari kegiatan observasi pada umumnya baru dapat mengungkap kulit luarnya saja (kurang mendetail). Adapun apa-apa sesungguhnya terjadi di balik hasil pengamatan itu belum dapat diungkap secara tuntas hanya dengan melakukan observasi saja. Oleh karena itu observasi harus didukung dengan cara-cara lainnya, seperti: penggunaan media *audio visual/cinematografi*, menyiapkan rubrik-rubrik teknik gerakan yang akan diobservasi maupun dengan melakukan wawancara.

4) WAWANCARA (*INTERVIEW*)

Secara umum yang dimaksud dengan wawancara adalah: cara menghimpun bahan-bahan keterangan yang dilaksanakan dengan melakukan tanya jawab lisan secara sepihak, berhadapan muka, dan dengan arah serta tujuan yang telah ditentukan.

Ada dua jenis wawancara yang dapat dipergunakan sebagai alat evaluasi, yaitu:

- a. **Wawancara terpimpin** yang juga sering dikenal dengan istilah wawancara berstruktur atau wawancara sistematis.
- b. **Wawancara tidak terpimpin** yang sering dikenal dengan istilah wawancara sederhana atau wawancara tidak sistematis, atau wawancara bebas

Dalam wawancara terpimpin, evaluator melakukan tanya jawab lisan dengan pihak-pihak yang diperlukan; misalnya dengan peserta didik, wawancara dengan orang tua, dan lain-lain, dalam rangka menghimpun bahan-bahan keterangan untuk penilaian terhadap peserta didiknya. Wawancara ini sudah dipersiapkan secara matang, yaitu dengan berpegangan pada panduan wawancara yang butir-butir itemnya terdiri dari hal-hal yang dipandang perlu guna mengungkapkan kebiasaan/aktivitas sehari-hari dari peserta didik setelah kuliah, cabang-cabang olahraga yang disukai dan tidak disukai, keinginan atau cita-citanya setelah lulus kuliah, cara belajar keterampilan olahraga yang efektif, cara menggunakan waktu luangnya, bacaan atau artikel olahraga yang paling disukai, dan sebagainya. Dalam konteks pendidikan jasmani penggunaan teknik wawancara sebagai salah satu alat evaluasi nontes seringkali dipergunakan untuk memberikan ujian lisan mata kuliah anatomi.

Di antara kelebihan yang dimiliki oleh wawancara adalah, bahwa dengan melakukan wawancara, pewawancara sebagai evaluator dapat melakukan kontak

langsung dengan peserta didik yang akan dinilai, sehingga dapat diperoleh hasil penilaian yang lebih lengkap dan mendalam. Dengan melakukan wawancara, peserta didik dapat mengeluarkan isi hatinya secara lebih bebas. Melalui wawancara, data dapat diperoleh baik dalam bentuk kualitatif maupun kuantitatif; pertanyaan-pertanyaan yang kurang jelas dapat diulang dan dijelaskan lagi dan sebaliknya jawaban-jawaban yang belum jelas dapat diminta lagi dengan lebih terarah dan lebih bermakna, asalkan tidak mempengaruhi atau mengarahkan jawaban peserta didik.

Wawancara juga dapat dilengkapi dengan alat bantu berupa *tape recorder*, sehingga jawaban atas pertanyaan-pertanyaan yang diajukan dapat dicatat dengan secara lebih lengkap. Penggunaan pedoman wawancara dan alat bantu perekam suara itu akan sangat membantu kepada pewawancara dalam mengkategorikan dan menganalisis jawaban-jawaban yang diberikan peserta didik untuk ditarik kesimpulannya.

Dalam wawancara bebas, pewawancara selaku evaluator mengajukan pertanyaan-pertanyaan kepada peserta didik atau orang tuanya tanpa dikendalikan oleh pedoman tertentu. Hanya saja pada saat menganalisis dan menarik kesimpulan hasil wawancara bebas ini pewawancara akan dihadapkan pada kesulitan-kesulitan, terutama apabila jawaban mereka yang beraneka ragam. Dalam pada itu, mengingat bahwa daya ingat manusia itu dibatasi oleh ruang dan waktu, maka sebaiknya hasil-hasil wawancara itu dicatat seketika. Mencatat hasil wawancara terpimpin tidaklah terlalu sulit, sebab pewawancara sudah dilengkapi dengan alat bantu berupa pedoman wawancara, namun sebaliknya mencatat hasil wawancara bebas adalah jauh lebih sulit, karena pewawancara harus memerlukan keterampilan khusus dalam mencatat pokok-pokok jawaban yang diberikan oleh *interviewee*.

5) **ANGKET (*QUESTIONNAIRE*)**

Angket juga dapat digunakan sebagai alat bantu dalam rangka penilaian hasil belajar. Berbeda dengan wawancara di mana penilai berhadapan langsung dengan peserta didik atau pihak lainnya, maka dengan menggunakan angket pengumpulan data sebagai bahan penilaian hasil belajar jauh lebih praktis, menghemat waktu dan tenaga. Hanya saja acapkali jawaban-jawaban yang diberikan tidak sesuai dengan kenyataan yang sebenarnya dan cenderung bias. Apalagi jika pertanyaan-pertanyaan yang diajukan dalam angket tersebut kurang tajam, sehingga memungkinkan bagi siswa (responden)

untuk memberikan jawaban yang diperkirakan akan melegakan atau memberikan kepuasan kepada pihak penilai.

Pada umumnya tujuan penggunaan angket dalam proses pembelajaran adalah untuk memperoleh data mengenai latar belakang peserta didik sebagai salah satu bahan dalam menganalisis tingkah laku dan proses belajar mereka. Disamping itu juga dimaksudkan untuk memperoleh data sebagai bahan dalam menyusun kurikulum dan program pembelajaran.

Data yang dapat dihimpun melalui kuesioner misalnya adalah data yang berkenaan dengan sikap para siswa berkenaan dengan kesulitan-kesulitan yang dihadapi oleh peserta didik dalam proses belajar-mengajar, seperti: kesulitan penguasaan *groundstroke backhand* dalam permainan tenis lapangan,, cara belajar yang efektif dalam senam ketangkasan, sikap yang berkenaan dengan efektifitas pemanfaatan fasilitas pembelajaran pendidikan jasmani, dan lain-lain. Kuesioner sering digunakan untuk menilai hasil belajar ranah afektif. Kuesioner dapat berbentuk pilihan ganda dan dapat pula berbentuk skala sikap. Skala yang mengukur sikap, sangat terkenal dan sering digunakan untuk mengungkap sikap peserta didik dalam pendidikan jasmani adalah Skala *Likert*, walaupun skala-skala pengukuran sikap yang lain juga dapat dipergunakan, seperti: Skala *Thurnstone*, Skala *Guttman*, dan *Semantic Differential Scale*.(Baumgartner dan Jackson, 1982: 137)

6) PEMERIKSAAN DOKUMEN (*DOCUMENTARY ANALYSIS*)

Evaluasi mengenai kemajuan, perkembangan atau keberhasilan belajar peserta didik tanpa menguji juga dapat dilengkapi dengan cara melakukan pemeriksaan terhadap dokumen-dokumen; misalnya dokumen yang memuat informasi riwayat hidup, agama yang dianut, kapan dan dimana dilahirkan, kedudukan anak dalam keluarga, dan lain-lain. Selain itu juga dokumen yang memuat informasi tentang orang tua peserta didik, seperti: nama, tempat tinggal, agama yang dianut, pekerjaan, dan sebagainya. Juga dokumen yang memuat tentang lingkungan nonsosial seperti: kondisi bangunan rumah, ruang belajar, lampu penerangan, dan sebagainya.

Berbagai informasi, baik mengenai peserta didik, orang tua, dan lingkungannya itu bukan tidak mungkin pada saat-saat tertentu sangat diperlukan sebagai bahan pelengkap bagi pendidik dalam melakukan evaluasi hasil belajar terhadap peserta

didiknya. Informasi-informasi seperti telah dikemukakan contoh diatas, dapat direkam melalui sebuah dokumen berbentuk formulir atau blangko isian, yang harus diisi pada saat peserta didik untuk pertama kali diterima sebagai siswa di sekolah yang bersangkutan.

Dari uraian tersebut di atas dapatlah dipahami, bahwa dalam rangka evaluasi hasil belajar peserta didik, evaluasi itu tidak harus semata-mata dilakukan dengan menggunakan alat berupa tes-tes hasil belajar. Teknik-teknik nontes juga menempatkan kedudukan yang penting dalam rangka evaluasi hasil belajar, lebih-lebih evaluasi yang berhubungan dengan kondisi kejiwaan peserta didik seperti persepsinya terhadap mata pelajaran tertentu, minatnya, bakatnya, dan sebagainya, yang kesemuanya tidak mungkin dievaluasi dengan menggunakan tes sebagai alat pengukurnya.

7) PERTUNJUKAN/PAMERAN (*EXHIBITION*)

Sebuah pertunjukan/pameran kepada masyarakat umum juga dipergunakan sebagai salah satu alat evaluasi hasil belajar siswa. Dengan pertunjukan ini siswa dapat mempertontonkan keterampilan/keahlian gerak yang dimiliki. Keterampilan gerak yang diperagakan siswa akan dinilai dengan menggunakan rubrik penilaian khusus yang sudah disiapkan oleh guru pendidikan jasmani. Pertunjukan-pertunjukan gerak yang dapat dilakukan siswa di antaranya adalah:

- a. Menampilkan kegiatan aerobik secara rutin di lingkungan sekolah/lingkungan masyarakat sekitar sekolah.
- b. Mengorganisasi dan menampilkan lompat tali pada saat jeda/disela-sela waktu istirahat antar babak dalam permainan bola basket.
- c. Menampilkan sebuah seni pencak silat/pomsae (taekwondo) pada saat di sela-sela kejujda maupun kejurnas pelajar/tingkat nasional.
- d. Mendemonstrasikan olahraga wushu di pusat-pusat perbelanjaan lokal.
- e. Membuat poster tentang manfaat aktivitas jasmani yang ditampilkan dalam pameran *fitness* tahunan.
- f. Menggambar dan mewarnai gambar berbagai aktivitas olahraga untuk pertunjukan pada rumah makan lokal dan perkantoran (Morrow, 2005: 147).

8) PENILAIAN PORTOFOLIO (*PORTFOLIOS ASSESSMENT*)

Penilaian portofolio adalah sebuah penilaian alternatif teknik nontes yang terdiri dari kumpulan hasil karya siswa/peserta didik yang disusun secara sistematis yang menunjukkan dan membuktikan upaya belajar, hasil belajar, proses belajar, dan kemajuan yang dilakukan peserta didik/siswa/mahasiswa dalam jangka waktu tertentu. Portofolio bersifat sistematis, penuh tujuan, dan berguna dalam mengumpulkan pekerjaan peserta didik yang telah dirakit selama beberapa waktu. Dalam banyak cakupan, bahan-bahan yang termasuk dalam portofolio dapat mirip dengan produk proyek peserta didik yang telah dikumpulkan selama beberapa waktu.

Menurut Morrow (2005: 144) ada 2 tipe portofolio, yaitu: portofolio model (*showcase or model portfolios*), yaitu sebuah portofolio yang terdiri dari contoh kerja yang dipilih oleh peserta didik yang mendokumentasikan kerja terbaik peserta didik. Dengan portofolio model ini peserta didik dapat secara sadar mengevaluasi pekerjaannya dan dipilih hanya pada hasil yang memiliki kualitas terbaik. dan portofolio deskriptif atau representatif (*descriptive or representative portfolios*) yaitu: sebuah portofolio yang terdiri dari pekerjaan representatis dari peserta didik, namun peserta didik tidak dapat mengevaluasi hasil/produk mereka.

Tipe portofolio, format, dan kandungan umum dalam portofolio pada umumnya biasanya ditentukan sebelumnya oleh guru. Hal ini untuk membatasi portofolio, untuk memastikan jumlah halaman kerja peserta didik, untuk menghindari agar tidak menjadi koleksi besar namun tanpa makna bagi peserta didik, serta dapat mempresentasikan penilaian yang sangat penting bagi guru. Portofolio itu sendiri pada umumnya berupa file atau folder yang berisi kumpulan kerja peserta didik. kandungan dalam portofolio dapat berupa item-item, seperti: catatan latihan, jurnal atau buku harian peserta didik, laporan tertulis, foto atau sketsa, surat, diagram atau grafik, fotokopi sertifikat/piagam kejuaraan, hasil tes kebugaran, statistik pertandingan, perencanaan latihan, analisis laporan harian, dan bukti-bukti melalui rekaman video.

Penilaian portofolio cenderung lebih berbentuk *self-assessment* (penilaian hasil karya siswa yang bersifat individual). Kumpulan hasil karya siswa/peserta didik tersebut menuntut partisipasi penuh peserta didik/mahasiswa untuk turut menentukan kriteria dan pemilihan bahan yang akan dimasukkan dalam portofolio. Salah satu isu dalam penilaian

portofolio adalah keharusan untuk dapat membedakan antara koleksi hasil karya yang ditempatkan dalam satu *folder* yang biasanya disebut sebagai **portofolio** dengan suatu model penilaian untuk memantau dan meningkatkan kinerja peserta didik/mahasiswa dalam pendidikan persekolahan yang biasanya disebut sebagai **assessment portofolio**. Beberapa elemen penting harus ditambahkan pada portofolio untuk dapat dikatakan sebagai penilaian portofolio. Dengan kata lain, tidak semua portofolio dapat digunakan untuk penilaian portofolio.

Perbedaan pokok antara **portofolio sebagai koleksi karya** dan **assessment portofolio** dapat dilihat dari tujuan, sebagaimana yang tercantum pada tabel 17 berikut ini:

Tabel 17: Perbedaan Portofolio dan *Assesment Portofolios*

Portofolio Sebagai Koleksi Karya	<i>Assesmen Portofolio</i>
▪ Sebagai contoh keterampilan yang representatif	▪ Sebagai landasan untuk mencapai level penguasaan berikutnya
▪ Sebagai ranah yang telah dikembangkan	▪ Sebagai ranah yang harus dikembangkan
▪ Sebagai bukti kemampuan yang dimiliki	▪ Sebagai pencatatan kemampuan yang telah dicapai
▪ Sebagai bahan yang akan dibahas	▪ Sebagai bahan untuk penyempurnaan instrumen
▪ Sebagai bahan laporan	▪ Sebagai bahan untuk menyesuaikan kurikulum

Sumber: Shaklee, B.D., dkk. (1997: 26)

Menurut Azmawi Zainul (2005: 44) portofolio sebagai alat untuk penilaian alternatif hasil belajar (*assesmen portofolio*) haruslah memperhatikan hal-hal sebagai berikut: (1) portofolio hendaknya memiliki kriteria penilaian yang jelas; (2) informasi atau hasil karya yang didokumentasikan dapat berasal dari semua orang yang mengetahui mahasiswa secara baik seperti dosen, rekan sesama mahasiswa, dosen dalam mata kuliah lain, dan sebagainya; (3) portofolio dapat terdiri dari berbagai bentuk informasi atau hasil karya seperti karangan, hasil lukisan, skor tes, foto hasil karya, dan lain-lain; (4) kualitas portofolio harus senantiasa ditingkatkan dari waktu ke waktu berdasarkan hasil karya

yang memenuhi kriteria; (5) setiap mata kuliah mungkin mempunyai bentuk portofolio yang sangat berbeda dengan mata kuliah lainnya; dan (6) portofolio harus terbuka bagi orang-orang yang secara langsung berkepentingan dengan hasil karya mahasiswa seperti dosen, jurusan, orang tua mahasiswa, dan mahasiswa itu sendiri. Dari karakteristik portofolio yang dikemukakan, setiap portofolio yang digunakan sebagai alat asesmen hasil belajar, secara langsung dapat dijadikan landasan bagi pengembangan kegiatan pembelajaran berikutnya. Dengan demikian, portofolio dapat dijadikan sebagai dasar untuk menilai maupun memperbaiki interaksi belajar mengajar, dan dapat pula dijadikan sebagai dasar perencanaan, baik bagi dosen/guru maupun bagi siswa/mahasiswa.

BAB XI
PENILAIAN OTENTIK (*AUTHENTIC ASSESSMENT*)
DALAM PJOK

Peningkatan kualitas proses belajar mengajar merupakan permasalahan penting dalam pendidikan jasmani olahraga dan kesehatan (penjasorkes) di semua jenjang satuan pendidikan mulai dari tingkat usia dini sampai ke perguruan tinggi. Setiap proses belajar mengajar (penjasorkes) memerlukan proses evaluasi. Proses belajar tidak akan diketahui secara pasti hasilnya manakala tidak ada evaluasi. Evaluasi merupakan bagian integral dari suatu proses belajar mengajar penjasorkes. Evaluasi berfungsi sebagai salah satu cara untuk memantau perkembangan belajar peserta didik dan untuk mengetahui seberapa jauh tujuan pengajaran dapat dicapai. Oleh sebab itu, dalam upaya peningkatan kualitas hasil belajar peserta didik harus ada kesinergian antara: materi pembelajaran, strategi pembelajaran, dan instrumen penilaian yang dipergunakan.

Untuk mendapatkan model penilaian yang tepat terhadap hasil belajar peserta didik diperlukan suatu instrumen penilaian yang dapat mengakses hasil belajar peserta didik secara kontekstual dan nyata (otentik). Penilaian yang kontekstual dan otentik dinamakan penilaian berbasis kinerja atau penilaian alternatif (Kirk dan Lund, 2010: 19). Menurut Mueller (2009: 102) penilaian otentik atau penilaian berbasis kinerja adalah penilaian yang meminta peserta didik untuk mendemonstrasikan keterampilan dan kompetensi tertentu yang merupakan penerapan dari pengetahuan yang telah dikuasai. Lebih lanjut Mueller (2009: 106) mengatakan bahwa penilaian berbasis kinerja merupakan suatu bentuk tugas yang menghendaki peserta didik untuk menunjukkan kinerja di dunia nyata secara bermakna sebagai esensi pengetahuan dan keterampilan yang sudah dikuasai. Sedangkan Suzann (2000: 3) menyatakan bahwa penilaian berbasis kinerja merupakan model penilaian yang diterapkan dalam situasi kehidupan nyata, yang mengharuskan peserta didik untuk menggunakan keterampilan berfikir tingkat tinggi (*hight order thinking*), seperti pemecahan masalah dan pengambilan keputusan dalam menggunakan keterampilan.

Dari beberapa pendapat ahli tersebut dapat disimpulkan bahwa penilaian berbasis kinerja adalah suatu model penilaian hasil belajar peserta didik yang bersifat multidimensional atau menyeluruh dan merujuk pada dunia “nyata”. Dari berbagai macam pendekatan tersebut dapat

memberikan kemungkinan bahwa pemecahan masalah dapat dilakukan dalam beberapa alternatif. Dengan kata lain, penilaian berbasis kinerja dapat memonitor dan mengukur kemampuan peserta didik dalam bermacam-macam kemungkinan pemecahan masalah yang dihadapi dalam dunia nyata.

Atas dasar uraian tersebut, maka seorang guru penjasorkes seharusnya tidak hanya terpaku pada satu model penilaian psikomotorik tradisional (penilaian kertas yang distandarisasi/tipe tes pilihan ganda, tes keterampilan olahraga) yang selama ini sudah mengakar dikalangan guru penjasorkes. Manakala model penilaian tradisional ini masih dipertahankan ada kekhawatiran penilaian ini tidak cukup informasi untuk mengikuti perubahan pendidikan yang begitu cepat yang menekankan pada hasil belajar, kemampuan kognitif, psikomotor, dan afektif yang lebih tinggi dan pembelajaran yang terintegrasi. Guru penjasorkes harus memiliki keberanian untuk mengembangkan suatu model penilaian terhadap prestasi hasil belajar peserta didik, salah satu model penilaian tersebut adalah penilaian berbasis kinerja (*authentic assessment*). Namun realita di lapangan menunjukkan bahwa sistem evaluasi pendidikan yang dilaksanakan selama ini (termasuk sistem penilaian penjasorkes) belum memberikan sumbangan untuk peningkatan kualitas pendidikan. Hal ini disebabkan oleh sistem evaluasi yang digunakan belum tepat sesuai dengan bidang kajian keilmuan masing-masing (Mardapi, 2008: 12). Berkenaan dengan mata pelajaran penjasorkes, ada suatu pertanyaan: apakah guru penjasorkes menguasai penilaian berbasis kinerja?. Ternyata hasil survei Tomoliyus dkk. (2010) dengan mewawancarai 43 guru penjasorkes SD di DIY menunjukkan bahwa guru penjasorkes di SD 96 % belum menguasai penilaian berbasis kinerja. Selain itu, 93 % instrumen penilaian yang dipergunakan guru penjasorkes belum sesuai dengan kompetensi yang diajarkan baik itu yang mengacu pada Kurikulum Tingkat Satuan Pendidikan (KTSP) maupun Kurikulum 2013.

Atas dasar latar belakang masalah sebagaimana di uraikan di atas terdapat kesenjangan guru penjasorkes dalam melakukan penilaian terhadap hasil belajar peserta didik. Kesenjangan tersebut di antaranya: secara teoritik dan praktik guru penjasorkes belum memahami dan menguasai penilaian kinerja peserta didik, guru penjasorkes belum memiliki kemampuan menyusun sendiri instrumen penilaian berbasis kinerja dan belum tersedianya instrumen penilaian terhadap hasil kinerja. Bertitik tolak dari kesenjangan yang terjadi maka perlu dicari solusi untuk dapat menjawab persoalan tersebut. Salah satu solusi dengan cara mengembangkan model penilaian otentik/berbasis kinerja atau *authentics assessment*.

A. HAKIKAT PENILAIAN (*ASSESSMENT*)

Penilaian merupakan komponen penting dalam penyelenggaraan pendidikan. Upaya meningkatkan kualitas pendidikan dapat ditempuh melalui peningkatan kualitas pembelajaran dan kualitas sistem penilaian yang tepat. Keduanya saling terkait, sistem pembelajaran yang baik akan menghasilkan kualitas belajar yang baik. Kualitas pembelajaran ini dapat dilihat dari hasil penilaian yang dicapai peserta didik. Selanjutnya sistem penilaian yang baik akan mendorong guru untuk menentukan strategi mengajar yang baik dan memotivasi peserta didik untuk belajar yang lebih baik. Oleh sebab itu, dalam upaya peningkatan mutu pendidikan diperlukan perbaikan sistem penilaian yang tepat.

Penilaian adalah proses pengumpulan data secara sistematis untuk membuat keputusan tentang peserta didik. Data yang diperoleh dengan menggunakan tes maupun nontes kemudian diolah menjadi informasi tentang peserta didik. Jadi proses penilaian meliputi pengumpulan bukti tentang hasil belajar peserta didik. Penilaian hasil belajar peserta didik pada rumpun mata pelajaran penjasorkes dilakukan melalui pengamatan terhadap perubahan perilaku dan sikap untuk menilai perkembangan pengetahuan, psikomotor, dan afektif peserta didik. Penilaian yang dilakukan oleh guru meliputi semua hasil belajar peserta didik yang terdiri dari: kemampuan kognitif, psikomotor, dan afektif, namun penekanan masing-masing ranah tidak sama sehingga harus diperhatikan karakteristik spesifikasi mata pelajaran yang akan dinilai (Djemari Mardapi, 2012: 17).

Menurut Adang Suherman (2001: 6) Penilaian adalah proses pengumpulan informasi. Penilaian berfungsi untuk membantu peserta didik dalam belajar. Data yang dihimpun melalui penilaian, dapat secara langsung dipakai sebagai umpan balik bagi perbaikan peningkatan pembelajaran. Melalui penilaian maka akan tersedia informasi yang valid untuk mengambil keputusan secara tepat. Atas dasar pendapat ahli tersebut dapat ditarik suatu makna bahwa peran penilaian terhadap hasil kinerja peserta didik begitu strategis, oleh karena itu pelaksanaan evaluasi dalam proses pembelajaran sebaiknya dikerjakan setiap hari dengan skedul yang sistematis dan terencana dengan baik. Ini dapat dilakukan oleh seorang guru penjasorkes dengan menempatkan secara integral evaluasi dalam perencanaan dan implementasi satuan pelajaran materi pembelajaran. Bagian penting lainnya yang perlu diperhatikan bagi seorang guru penjasorkes adalah perlu melibatkan peserta didik dalam

evaluasi sehingga mereka secara sadar dapat mengenali perkembangan pencapaian hasil belajar mereka.

B. PRINSIP DASAR PELAKSANAAN PENILAIAN OTENTIK PJOK

Agar evaluasi dapat memenuhi fungsinya untuk meningkatkan mutu proses belajar mengajar penjasorkes, maka pelaksanaan evaluasi harus mempertimbangkan prinsip sebagai berikut:

a. Proses pengumpulan data dilakukan melalui kerjasama secara alami (bersifat objektif).

Untuk meningkatkan aktivitas proses pembelajaran dan memotivasi peserta didik, hindarkanlah penggunaan standar yang baku, atau perbandingan dengan teman. Bagi peserta didik tertentu hal ini justru dapat menurunkan motivasi belajar. Sebagai penggantinya, lakukanlah kerjasama antara guru dengan peserta didik secara individu untuk mendiskusikan tujuan belajar yang ingin dicapai. Bimbinglah dan doronglah peserta didik untuk menentukan tujuan-tujuan yang maksimal, akan tetapi harus realistis sesuai dengan tingkat kemampuannya. Para peserta didik akan menjadi lebih termotivasi untuk melakukan tes serta menjadi tekun belajar manakala mereka bekerja sama dengan guru.

b. Proses pengumpulan data dilakukan secara terus-menerus dan berkesinambungan (kontinyu).

Bukti kemajuan hasil belajar peserta didik ada yang secara langsung dapat dilihat hasilnya dalam tempo waktu relatif singkat. Namun ada juga kemajuan belajar peserta didik yang hanya dapat dilihat hasilnya setelah menempuh program yang panjang, misalnya satu tahun atau lebih. Oleh sebab itu, kemajuan belajar peserta didik harus dicatat dan didokumentasikan agar dapat diperlihatkan peningkatannya kepada peserta didik, orang tua, atau kepada pihak pimpinan sekolah yang membutuhkan. Proses pengumpulan data dilakukan secara periodik, kalau penilaian hanya satu kali menjadi kurang tepat untuk melihat kelebihan dan kelemahan masing-masing peserta didik.

c. Data yang dihimpun meliputi berbagai aspek (bersifat holistik).

Kemajuan belajar, motivasi, dan proses belajar merupakan tiga komponen utama yang satu sama lain saling berintegrasi dan saling mempengaruhi dengan kuat. Data yang dikumpulkan harus meliputi berbagai aspek. Proses pengumpulan data harus dapat mengumpulkan informasi secara mendalam tentang: (1) apa yang harus diketahui oleh

peserta didik (aspek kognitif), (2) apa yang dirasakan dan bagaimana perasaan peserta didik terhadap makna pelajaran penjasorkes (aspek afektif), (3) kemampuan/keterampilan yang diperoleh peserta didik (aspek psikomotor), dan usaha yang dilakukan oleh peserta didik (proses) untuk memperoleh hasil dari tujuan belajar. Proses belajar sama pentingnya dengan hasil belajar.

- d. Laporan kemajuan belajar harus dibuat dan disampaikan kepada orang tua/wali peserta didik.

Laporan kemajuan belajar peserta didik tidak hanya cukup diketahui oleh guru penjasorkes saja, namun pihak sekolah, peserta didik, dan orang tua juga harus mengetahui. Agar pihak yang berkepentingan tersebut memahami secara mendalam tentang kemajuan belajar peserta didik, seyogyanya laporan tidak hanya memuat nilai numerik saja, misalnya skor 7, 8, atau 9. Akan lebih baik nilai tersebut disertai dengan data dari berbagai aspek kemajuan belajar, misalnya: tentang perkembangan disiplin, keterampilan gerak, dan tingkat kebugaran jasmani. Laporan hasil belajar peserta didik yang memuat berbagai aspek perkembangan kemajuan belajar, akan jauh lebih berharga bagi orang tua peserta didik. Dengan cara ini, orang tua peserta didik akan mengetahui kelemahan dan kelebihan yang dimiliki anak, seperti: keterampilan apa yang mesti ditingkatkan, aspek apa yang harus dipertahankan, dan bagaimana cara mengatasinya. Dengan laporan hasil belajar dengan model seperti ini pada akhirnya akan menanamkan rasa tanggung jawab orang tua peserta didik terhadap keberhasilan belajar anak.

C. PENILAIAN OTENTIK ATAU PENILAIAN BERBASIS KINERJA (*AUTHENTIC ASSESSMENT*)

Penilaian otentik/penilaian berbasis kinerja merupakan model penilaian dalam penjasorkes yang akhir-akhir ini mulai dikembangkan untuk merespon implementasi pelaksanaan kurikulum penjasorkes baik kurikulum KTSP maupun Kurikulum 2013. Penilaian otentik merupakan model penilaian terhadap kinerja peserta didik yang dirancang untuk terjadi dalam kehidupan nyata dan dapat langsung diobservasi, bukan keadaan tiruan sebagaimana sistem penilaian bentuk konvensional (penilaian soal pilihan berganda). Penilaian otentik bersifat multidimensional atau menyeluruh yang menuntut peserta didik menyatukan pemikiran tingkat tinggi ke dalam perilaku peserta didik tersebut (Alan C. Lacy, 2011: 111). Sebagai ilustrasi dari penilaian otentik ini adalah: dalam bermain tenis lapangan,

peserta didik tidak hanya dapat melakukan pukulan servis, *forehand*, *backhand*, maupun voli saja, namun lebih dari itu, peserta didik juga dituntut harus mampu memahami peraturan permainan maupun menerapkan taktik dan strategi bermain tenis lapangan.

Penilaian otentik adalah suatu penilaian belajar yang merujuk pada konteks “dunia nyata” yang memerlukan berbagai macam pendekatan untuk memecahkan masalah yang memberikan kemungkinan bahwa suatu masalah dapat mempunyai lebih dari satu cara pemecahan. Dengan kata lain, penilaian otentik memonitor dan mengukur kemampuan peserta didik dalam bermacam-macam kemungkinan pemecahan masalah yang dihadapi dalam konteks dunia nyata (Morrow, 2005 :122). Sedangkan menurut Lund (2010: 19) penilaian otentik adalah penilaian yang dirancang agar peserta didik mampu melaksanakan tugas bermakna dan dapat mempresentasikan kinerja secara nyata, menekankan berfikir tingkat tinggi, belajar lebih kompleks, serta melibatkan pemeriksaan proses maupun produk pembelajaran.

Dari pendapat tersebut dapat disimpulkan bahwa penilaian otentik adalah penilaian yang mengukur, memonitor, dan menilai semua aspek hasil belajar peserta didik secara holistic (meliputi ranah kognitif, psikomotor, dan afektif) baik yang nampak sebagai hasil akhir dari suatu proses pembelajaran maupun berupa perubahan dan perkembangan aktivitas selama proses pembelajaran di kelas maupun di lapangan.

D. KARAKTERISTIK PENILAIAN OTENTIK (*CHARACTERISTIC OF AUTHENTIC ASSESSMENT*)

Dalam konteks penilaian umum dalam pendidikan jasmani, olahraga dan kesehatan , penilaian otentik memiliki beberapa ciri-ciri yang lebih spesifik jika dibandingkan dengan model penilaian yang lain. Menurut Zainul (2005: 9-13) ciri-ciri penilaian otentik terdiri atas:

a. Menuntut peserta didik berfikir tingkat tinggi baik dalam proses maupun produk pembelajaran (kemampuan analisis, sintesis, dan evaluasi). Tuntutan penerapan taktik dan strategi dalam situasi yang sebenarnya (*like game situation*) merupakan pendekatan yang ditonjolkan dalam penilaian otentik.

b. Penilaian menggunakan kriteria

Dalam penilaian otentik, kinerja peserta didik diberikan rubrik yang dipergunakan untuk mengevaluasi. Peserta didik ditugaskan untuk mengambil suatu bentuk tugas. Hasil dari

menerjakan tugas tugas yang berupa pekerjaan peserta didik yang disebut sebagai produk, dan mencetak rubrik yang dirancang untuk dikembangkan agar sesuai dengan isi tugas dan dipergunakan untuk menilai produk peserta didik.

c. Penilaian yang memiliki reliabilitas

Sifat penting dalam penilaian domain apapun, diperlukan adanya reliabilitas instrumen penilaian. Dalam penilaian otentik untuk menentukan reliabilitas penilaian menggunakan metode reliabilitas antar penilai (*inter-tester reliability*). Metode ini mengharuskan penilai/*rater* melakukan pengamatan untuk menilai beberapa aspek perilaku peserta didik dalam melaksanakan tugas. Untuk mencapai tingkat keterandalan antar *rater* perlu dirancang dengan baik untuk menerapkan rubrik penilaian.

d. Penilaian dengan melibatkan rubrik

Pengembangan rubrik penilaian sangat diperlukan dalam penilaian kinerja peserta didik yang dipergunakan sebagai dasar pengukuran. Rubrik penilaian merupakan panduan untuk memberi nilai/skor yang jelas dan disepakati antara guru penjasorkes dan peserta didik. Lund dkk., (2010: 43) mengatakan bahwa "*rubric indicate the criteria a person scoring performance based assessment should use when doing evaluation*"

Contoh: Rubrik penilaian pukulan *groundstrokes* (*forehand* dan *backhand*) dalam tenis

Nilai	Norma Kategori	Deskripsi Rubrik
5	Sangat Baik	Menggunakan <i>grip</i> /pegangan yang memadai, keseimbangan dan gerakan kaki (<i>footwork</i>) bagus, gerakan <i>groundstrokes</i> yang dilakukan hampir sempurna. Ditinjau dari mekanika gerak secara konsisten mampu memperagakan teknik pukulan dengan baik. Pukulan dilakukan menggunakan power dan konsisten mengarahkan bola pada tempat yang tepat.
4	Baik	Menggunakan <i>grip</i> /pegangan yang memadai, keseimbangan bagus, gerakan kaki (<i>footwork</i>) cukup dan memadai tetapi gerakan <i>groundstrokes</i> yang dilakukan tidak cukup bagus. Dalam

3	Cukup	<p>memperagakan teknik pukulan ditinjau dari mekanika gerak kurang konsisten. Pukulan dilakukan dengan <i>power moderate</i> dan konsisten menempatkan bola masuk ke lapangan.</p> <p>Menggunakan <i>grip</i>/pegangan yang memadai, tetapi gerakan kaki (<i>footwork</i>) kurang. Gerakan <i>groundstrokes</i> yang dilakukan tidak ritmis/terputus-putus dan tidak efisien, pukulan yang dilakukan hasilnya tidak konsisten. Gaya/<i>style</i> pukulan secara alamiah cenderung lebih bertahan (<i>defensive</i>), tetapi tidak cukup memadai untuk dipergunakan melakukan <i>rally</i> dalam waktu singkat.</p>
2	Kurang	<p>Menggunakan <i>grip</i>/pegangan yang kurang memadai, gerakan kaki (<i>footwork</i>) kurang, dan secara teknik dasar gerakan <i>groundstrokes</i> yang dilakukan salah. Dalam melakukan pukulan tidak konsisten menggunakan mekanika gerak. Gaya/<i>style</i> pukulan lebih bertahan (<i>defensive</i>), dan hanya berusaha agar bola melewati net. Tidak mampu melakukan rally.</p>
1	Kurang Sekali	<p>Menggunakan <i>grip</i>/pegangan yang salah, keseimbangan hilang, gerakan kaki (<i>footwork</i>) kurang sekali. Bentuk gerakan <i>groundstrokes</i> yang dilakukan sangat jelek dan tidak ritmis/terputus-putus. Secara nyata tidak dapat mengontrol penempatan bola. Berulangkali bola tidak dapat melewati net.</p>

(Lund, dkk. 2010: 38)

e. Penilaian yang berorientasi pada proses dan produk

“Proses pembelajaran” peserta didik merupakan komponen penting dalam penilaian otentik. Dalam penjasorkes keterkaitan antara proses dan produk merupakan dua sisi mata rantai yang tidak dapat dipisahkan. Bagaimana proses dan cara peserta didik menyelesaikan tugas adalah bagian penilaian yang utama.

Contoh: dalam mengajarkan gerak dasar lompat jauh, penilaian tidak hanya didasarkan pada jauh lompatan, namun proses pengambilan awalan, tumpuan, saat melayang di udara, dan pendaratan merupakan komponen penilaian yang harus diutamakan.

Menurut Morrow (2005: 123) karakteristik penilaian otentik adalah (1) penilaian yang berbasis kinerja melalui lembar tugas untuk menerapkan pengetahuan yang telah dikuasai secara teoritis, (2) penilaian yang lebih menuntut peserta didik mampu mendemonstrasikan pengetahuan, keterampilan, dan strategi dengan mengkreasikan jawaban atau produk, (3) penilaian mengharuskan peserta didik untuk menggunakan keterampilan berfikir tingkat tinggi, seperti pemecahan masalah dan pengambilan keputusan dalam menggunakan keterampilan, (4) suatu penilaian hasil belajar peserta didik yang merujuk pada situasi dunia nyata, dapat menggunakan berbagai macam alternatif pendekatan untuk memecahkan masalah yang memungkinkan satu masalah dapat memiliki lebih dari satu pemecahan masalah, (5) proses penilaian harus merupakan bagian yang tidak terpisahkan dari proses pembelajaran, (6) penilaian harus bersifat holistik yang mencakup semua aspek pembelajaran dari tujuan ranah kognitif, afektif, dan psikomotor.

E. TIPE-TIPE PENILAIAN OTENTIK (*TYPES OF AUTHENTIC ASSESSMENT*)

Menurut Morrow (2005: 128) penilaian otentik merupakan teknik penilaian terhadap hasil belajar peserta didik yang berbentuk nontes. Berbagai macam tipe penilaian otentik yang berbentuk nontes ini dapat dipilih dan disesuaikan dengan tujuan penilaian yang diharapkan oleh guru penjasorkes. Menurut Lund (2010: 26-33) tipe-tipe penilaian otentik terdiri atas: 1) pengamatan guru/*teacher observation*, 2) observasi teman sejawat/*peer observation*, 3) observasi diri/*self observation*, 4) bermain game dan memodifikasi permainan/*game play and modified game play*, 5) bermain peran/*role plays*, 6) tugas menyelenggarakan suatu even/*event task*, 7) wawancara/*interviews*, 8) pertanyaan-pertanyaan/*essays*, 9) pertanyaan-pertanyaan terbuka/*open response questions*, 10) jurnal-

jurnal/*journals*, 11) proyek-proyek peserta didik/*student project*, 12) kinerja peserta didik/*student performance*, 13) jurnal atau buku harian/*student logs*, serta 14) portofolio/*portfolios*.

Selanjutnya menurut Morrow (2005: 128) tipe-tipe penilaian otentik yang dipergunakan untuk menilai hasil belajar penjasorkes peserta didik di antaranya dapat dilakukan melalui: 1) pengamatan secara sistematis/*observation*, 2) wawancara/*interview*, 3) angket/*questionnaire*, 4) analisis dokumen/*documentary analysis*, 5) pertunjukan atau pameran/*exhibition*, dan 6) penilaian terhadap kumpulan semua hasil karya peserta didik/*portofolio*.

Dari pendapat tersebut dapat disimpulkan bahwa tipe-tipe penilaian otentik merupakan penilaian yang menggunakan instrumen nontes yang dirancang untuk terjadi dalam kehidupan nyata, bukan keadaan yang dibuat-buat/imitasi dan dilakukan tanpa “menguji” seperti tipe penilaian konvensional pada umumnya.

F. Kriteria-Kriteria *Authentic Assessment*

Untuk mengevaluasi apakah penilaian otentik merupakan sebuah instrumen penilaian penjasorkes yang berkualitas, maka harus diperhatikan beberapa kriteria berikut. Menurut Lund dkk. (2005: 134) kriteria-kriteria tersebut, di antaranya:

a. Konsekuensi (*consequences*)

Apakah penilaian otentik mendorong agar terjadi dampak yang diharapkan? Demikian juga sebaliknya apakah penilaian mendorong kepada dampak yang tidak diharapkan.

b. Keadilan (*fairness*)

Apakah penilaian otentik memudahkan peserta didik untuk mendemonstrasikan kemampuan mereka yang sebenarnya, atau apakah merugikan peserta didik dengan tidak adil.

c. Transfer dan generalisasi (*transfer and generalizability*)

Apakah hasil penilaian otentik dapat dipergunakan untuk menggeneralisasikan permasalahan lain atau situasi ranah yang lain?

d. Pengetahuan yang kompleks (*cognitive complexity*)

Apakah penilaian otentik cukup memadai sebagai pertimbangan tingkat pemahaman dan pemikiran ranah kognitif yang lebih tinggi/lebih kompleks?

e. Kualitas isi (*content quality*)

Apakah penilaian otentik benar-benar dipergunakan untuk mengukur isi/materi ranah yang menjadi fokus untuk dinilai?

f. Ketercakupan isi (*content coverage*)

Apakah penilaian otentik cukup mewakili keluasan isi dari materi yang diinginkan?

g. Keberartian (*meaningfulness*)

Apakah penilaian otentik berarti bagi peserta didik dan apakah hal itu dapat memotivasi mereka untuk menunjukkan kemampuan terbaik peserta didik?

h. Biaya dan efisiensi (*cost and efficiency*)

Apakah biaya pengumpulan data dan pemberian skor penilaian otentik dapat diterima nalar dan efisien?

Sedangkan Popham (1995: 147) mengatakan bahwa tujuh kriteria yang dapat dipergunakan untuk mengevaluasi kualitas penilaian otentik adalah:

- a. **Generalizability**, artinya kinerja/*performance* peserta didik dalam melakukan tugas yang diberikan sudah memadai untuk digeneralisasikan kepada tugas-tugas lain. Semakin dapat digeneralisasikan tugas-tugas yang diberikan dalam rangka penilaian kinerja peserta didik atau semakin dapat dibandingkan dengan tugas-tugas lainnya, maka semakin baik tugas tersebut.
- b. **Authenticity**, artinya tugas yang diberikan ke peserta didik sudah serupa/relevan dengan apa yang sering dihadapinya dalam praktik kehidupan sehari-hari baik di sekolah maupun di masyarakat.
- c. **Multiple**, artinya tugas yang diberikan kepada peserta didik mengukur lebih dari satu kemampuan yang diinginkan.
- d. **Teachability**, artinya tugas yang diberikan merupakan tugas yang hasilnya semakin baik karena adanya usaha belajar. Jadi tugas-tugas yang diberikan dalam penilaian kinerja peserta didik adalah tugas-tugas yang relevan dengan materi pembelajaran.
- e. **Fairness**, artinya tugas yang diberikan sudah adil untuk semua peserta didik. Dengan demikian tugas-tugas yang diberikan harus dirancang agar tidak bias untuk semua kelompok.

- f. **Feasibility**, artinya tugas yang diberikan dalam penilaian kinerja peserta didik memang relevan untuk dilaksanakan dengan mempertimbangkan faktor-faktor, seperti biaya, tempat, waktu, atau peralatan yang dipergunakan.
- g. **Scorability**, artinya tugas yang diberikan ke peserta didik dapat diskor dengan akurat dan reliabel. Karena keakuratan dan reliabilitas merupakan salah satu yang sensitif dari penskoran dalam penilaian otentik.

G. KELEBIHAN DAN KELEMAHAN PENILAIAN OTENTIK

Salah satu sisi kelemahan yang melekat pada penilaian otentik adalah minimnya informasi empiris tentang validitas penilaian. Bentuk utama validitas penilaian otentik baru sebatas validitas tampang (*face validity*) yang berhubungan dengan validitas isi (*content validity*). Hal ini didasarkan pada hubungan yang diasumsikan bahwa jika pembelajaran yang dilakukan baik, maka penilaian yang dilakukanpun juga akan baik. Walaupun kaitan antara pembelajaran dan penilaian itu penting, validitas isi (*content validity*) sendiri tidak harus diterima sebagai bukti yang cukup untuk dipergunakan sebagai metode penilaian. Guru penjasorkes perlu mencari bukti-bukti yang lebih memperkuat bahwa penilaian otentik dapat memberikan kesimpulan yang valid/akurat.

Adapun kelebihan-kelebihan penilaian otentik menurut Lund (2010: 133-135) adalah:

1. **Objektif**, artinya penilaian otentik dilaksanakan dengan mengacu pada hal yang sebenarnya, tidak mencari kesalahan peserta didik, serta tidak dilandasi oleh perasaan suka atau tidak suka, tetapi lebih fokus pada realita tentang kemampuan peserta didik.
2. **Realistis**, artinya penilaian otentik dapat dilaksanakan oleh semua guru yang hasilnya sesuai dengan kemampuan masing-masing peserta didik.
3. **Tepat waktu**, artinya penilaian peserta didik mengacu pada prinsip tepat waktu dan dilaksanakan secara berkala sesuai dengan tahapan-tahapan proses penilaian kinerja berlangsung.
4. **Dapat dipertanggung jawabkan**, artinya sebagai bentuk akuntabilitas, proses penilaian terhadap peserta didik didasarkan pada kinerja yang sebenarnya dan tidak terjadi manipulasi selama proses penilaian berlangsung.
5. **Terukur**, artinya selama proses penilaian berlangsung mengacu pada instrumen yang telah tersusun dan digunakan sebagai dasar dalam pelaksanaannya.

6. **Terbuka**, artinya hasil penilaian otentik bersifat terbuka, ada peluang untuk melakukan klarifikasi bagi peserta didik yang dinilai, untuk menghindari subjektivitas penilai selama proses penilaian berlangsung.
7. **Tidak Deskriminatif**, artinya selama proses penilaian berlangsung tidak diperkenankan ada diskriminasi dalam hal, suku, ras, agama, dan jenis kelamin.

DAFTAR PUSTAKA

- Allen, M. J. dan Yen, W.M. (1979). *Introduction to Measurement Theory*. Monterey, CA: Brooks Cole Publishing Company
- Arma Abdoellah. (1985). *Evaluasi Hasil Belajar Dalam Pendidikan Olahraga*. Yogyakarta: IKIP Yogyakarta.
- Azwar, Saifuddin, (1997). *Reliabilitas dan Validitas* (edisi ke-3). Yogyakarta: Pustaka Pelajar.
- _____, (2005). *Dasar-Dasar Psikometri* (edisi ke-5). Yogyakarta: Pustaka Pelajar.
- Barrow Harold M. dan Rosemary, Mc. Gee. (1979). *A Practical Approach to Measurement in Physical Education*. Philadelphia: Lia & Febiger.
- Baumgartner, T.A., Jackson, A.S., Mahar, M.T., Rowe, D.A. (2007). *Measurement for Evaluation in Physical Education & Exercise Science (Eight Edition)*. New York: McGraw-Hill.
- Departemen Pendidikan Nasional (2000) *Pedoman dan Modul Pelatihan Kesehatan Olahraga Bagi Pelatih Olahragawan Pelajar*. Jakarta: Departemen Pendidikan Nasional Pusat Pengembangan
- Djemari Mardapi, (1996). *Penilaian Unjuk Kerja sebagai Usaha Meningkatkan Kemampuan Sumber Daya Manusia*. Yogyakarta: IKIP Yogyakarta.
- Fernandez, H. J. X. (1984). *Testing and Measurement*. Jakarta: National Educational Planning, Evaluation and Curriculum Development.
- Gronlund, N.E. dan Linn, R.L.. (1990). *Measurement and Evaluation in Teaching*. 6th edition. New York: Macmillan Publishing Company.
- I Made Sriundy Mahardika. (2014). *Evaluasi Pengajaran*. Surabaya. Unesa University Press.
- Ismaryati. (2009). *Tes dan Pengukuran Olahraga*. Surakarta: LPP UNS dan UNS Press.
- Jensen Clayne R. dan Hirst, Cynthia C. (1980). *Measurement in Physical Education And Athletics*. New York: Mac Millan Publishing Co. Inc.
- Kirkendall, E. G. dan Johnson, R. (1987). *Measurement and Evaluation for Physical Education*. Champaign, Illinois: Human Kinetics.
- Kravitz, Len. (2001). *Panduan Lengkap: Bugar Total*. Jakarta: PT Raja Grafindo Persada.
- Linn, R. L. dan Grondlund, N. E. (1995). *Measurement and Assesment in Teaching* (edisi ke-7). Englewood Cliffs, NJ: Prentice Hall.
- Miller, David K. (2002). *Measurement by the Physical Educator (Why and How)*. New York: The Mc. Graw-Hill Companies, Inc.

- Morrow, James R. (2000). *Measurement and Evaluation in Human Performance (Second Edition)*. United States of America: Human Kinetics.
- Morrow, J.R., Jackson, A.W., Disch, J.G., & Mood, D.P. (2005). *Measurement and Evaluation in Human Performance (Third Edition)*. United States of America: Human Kinetics.
- Ngatman. (2002). "Validitas, Reliabilitas, dan Obyektivitas Instrumen Penilaian dalam Pendidikan Jasmani." (*Majalah Ilmiah Olahraga, nomor: 3, 2002*). Yogyakarta: FIK Universitas Negeri Yogyakarta.
- _____. (2006). Pendekatan Acuan Penilaian Dalam Pendidikan Jasmani. *Jurnal Pendidikan Jasmani Indonesia* Volume 3, Nomor 3, November 2006, hal 69-82.
- Phillips, Allen D. dan Hornak, James E. (1979). *Measurement and Evaluation in Physical Education*: New York, Cichester, Brisbane, Toronto: John Wiley and Sons.
- Rink, J.E. (2010). *Teaching Physical Education for Learning (Sixth Edition)*. New York: McGraw-Hill.
- Rusli Lutan dan Adang Suherman. (2000). *Pengukuran dan Evaluasi Penjaskes*. Jakarta: Departemen Pendidikan dan Kebudayaan.
- Safrit, Margareth J. (1986). *Introduction To Measurement in Physical Education and Exercise Science*. St. Louis, Missouri: CV Mosby Company.
- Strand, Bradford N. dan Wilson, R. (1993). *Assesing Sport Skills*. United Stated of America: Human Kinetics Publishers.
- Sudijono, Anas. (2005). *Pengantar Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada.
- Sukardi. (2010). *Evaluasi Pendidikan Prinsip dan Operasionalnya*. Jakarta. PT Bumi Aksara/

