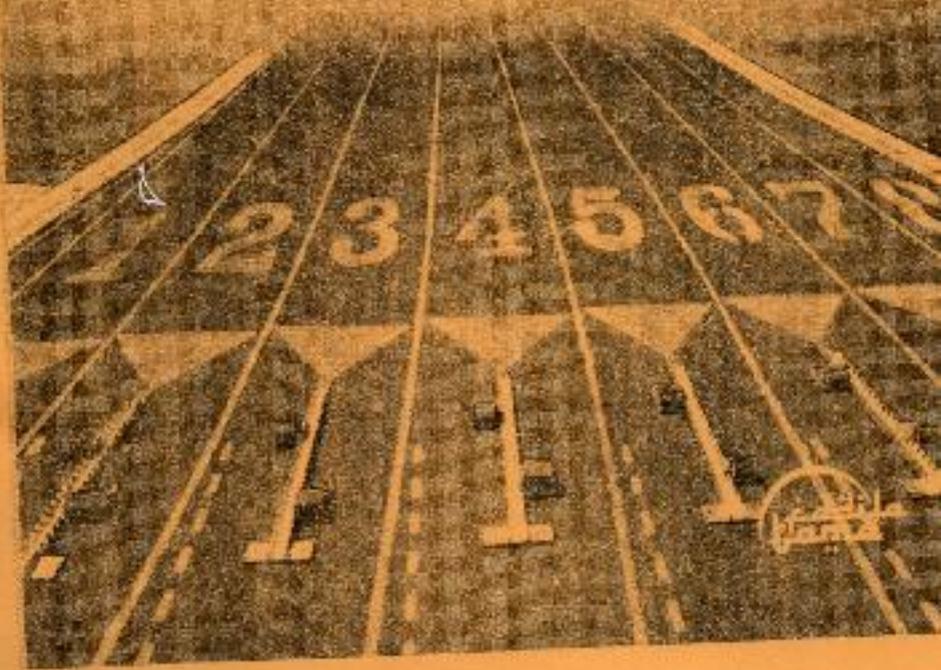


TES DAN PENGUKURAN
untuk Evaluasi dalam
PENDIDIKAN JASMANI
DAN OLIMPIADA





**TES DAN PENGUKURAN
UNTUK EVALUASI DALAM
PENDIDIKAN JASMANI
DAN OLAHRAGA**

Drs. Ngatman, M.Pd
Fitria Dwi Andriyani, M.Or.



Judul
**Tes dan Pengukuran untuk Evaluasi
dalam Pendidikan Jasmani dan Olahraga**
Cetakan Pertama, Maret 2017

Perpustakaan Nasional: Katalog Dalam Terbitan (KDT)
Drs. Ngatman, M.Pd dan Fitria Dwi Andriyani, M.Or. Tes
dan Pengukuran untuk Evaluasi dalam Pendidikan Jasmani
dan Olahraga. - Cet. 1 - Yogyakarta: Fadilatama, 2017
vi + 198 hlm; 140 x 200 mm
ISBN: 978-602-70296-7-5

Penulis : Drs. Ngatman, M.Pd
Fitria Dwi Andriyani, M.Or.
Editor : M. Fadillah
Tata Letak : Maryono Ahmad
Desain Cover : Tim Fadilatama

Penerbit:
FADILATAMA
Babadsari Rt. 05 Rw 36 Sendangtirto,
Berbah, Sleman, DI. Yogyakarta
Email: fadilatama@gmail.com
www.fadilatama.tk



KATA PENGANTAR

Kegiatan tes dan pengukuran merupakan hal yang tidak dapat terlepas dalam pembelajaran pendidikan jasmani di sekolah. Kegiatan tersebut memiliki makna penting baik bagi guru maupun siswa. Bagi guru misalnya, hasil dari tes dan pengukuran menjadi sumber informasi guru mengenai kemampuan awal dan kemampuan akhir siswa, serta peningkatan kemampuan siswa; yang pada akhirnya menjadi sumber bahan bagi guru dalam mengevaluasi program pembelajaran yang telah dilaksanakannya. Bagi siswa sendiri, hasil dari kegiatan tes dan pengukuran menyediakan informasi bagi siswa mengenai keterampilannya, peningkatan keterampilannya, dan posisi keterampilannya dibandingkan dengan teman sebayanya sehingga dapat menjadi bahan motivasi dan evaluasi diri.

Oleh karena itu, tes dan pengukuran merupakan hal yang krusial dalam pendidikan jasmani. Para calon guru dan guru pendidikan jasmani perlu memahami dengan seksama mengenai tes dan pengukuran yang baik serta bagaimana cara mendesain, menyusun, dan melakukan tes dan pengukuran yang baik sehingga dapat menilai siswa secara akurat (valid), ajeg (reliabel), dan objektif. Selanjutnya, pengetahuan mengenai tes-tes baku dalam pendidikan jasmani juga penting bagi guru dan calon guru pendidikan jasmani guna kepentingan penelitian. Jika penelitian yang dilakukan memerlukan tes dan pengukuran sementara peneliti tidak memahami kriteria tes dan pengukuran yang baik, niscaya

data yang dihasilkan akan dipertanyakan.

Maka dari itu, buku tes dan pengukuran pendidikan jasmani ini disusun, dengan harapan memperkaya literatur dalam pendidikan jasmani, khususnya mengenai tes dan pengukuran. Sajian substansi yang disusun secara lengkap, detail, dan cermat ditujukan supaya buku ini dapat membawa manfaat optimal bagi penggunanya. Buku ini dapat digunakan oleh guru penjas, mahasiswa, pelatih, dan *stakeholder* olahraga.

Selanjutnya, 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 mengutangi makna, maka istilah PJOK dalam buku ini ditulis sebagai pendidikan jasmani atau penjas. Akhir kata, semoga penulisan buku ini membawa manfaat nyata dalam dunia pendidikan jasmani dan olahraga tanah air.

Yogyakarta, 30 Mei 2016

Penulis

DAFTAR ISI

Kata Pengantar.....	iii
Daftar Isi	v
Bab 1 Hakikat Tes, Pengukuran dan Evaluasi	1
Bab 2 Tujuan/Alasan Melakukan Tes, Pengukuran dan Evaluasi.....	17
Bab 3 Kriteria Tes yang Baik	29
Bab 4 Langkah-Langkah Memilih dan Mengkonstruksi Tes.....	45
Bab 5 Validitas Tes dan Pengujiannya.....	53
Bab 6 Reliabilitas Tes dan Metode Estimasinya	81
Bab 7 Objektivitas Tes	101
Bab 8 Pendekatan Acuan Penilaian dan Sistem Penilaian	107
Bab 9 Pengukuran Kesehatan Jasmani.....	123
Bab 10 Tes-Tes Standar dalam Pendidikan Jasmani dan Olahraga	141
Daftar Pustaka.....	193
Tentang Penulis.....	197



BAB 1 HAKIKAT TES, PENGUKURAN DAN EVALUASI

A. Definisi Tes, Pengukuran dan Evaluasi

Tes, pengukuran, dan evaluasi (penilaian) merupakan istilah yang umum dijumpai dalam kegiatan evaluasi pembelajaran termasuk dalam pembelajaran pendidikan jasmani olahraga dan kesehatan. Dalam kehidupan sehari-hari, baik di sekolah dan dalam dunia pendidikan umumnya, ketiga istilah tersebut sering dicampuradukkan. Berikut ini akan dibahas mengenai ketiga istilah tersebut.

Pertama, kita akan membahas mengenai Tes. Kirkendall (1982) menyatakan bahwa tes adalah suatu instrument yang digunakan untuk memperoleh informasi mengenai individu-individu atau objek-objek, yang mana instrumen tersebut dapat dalam bentuk form pertanyaan pada kertas atau wawancara; atau berupa observasi dari penampilan fisik dan perilaku berupa checklist atau catatan anekdot. Selanjutnya, tes juga dapat diartikan sebagai instrumen atau alat yang digunakan untuk melakukan pengukuran tertentu dan dapat dilakukan secara tertulis, oral, fisiologis, psikologis, atau dapat berupa peralatan mekanik seperti treadmill (Morrow et al, 2005: 4-5). Phillips (1979) mendefinisikan pengertian tes sebagai suatu alat atau instrument pengukuran yang digunakan untuk memperoleh data mengenai ciri atau karakteristik spesifik dari individu atau kelompok.

Dari beberapa pendapat tersebut dapat ditarik suatu kesimpulan bahwa tes merupakan suatu instrumen

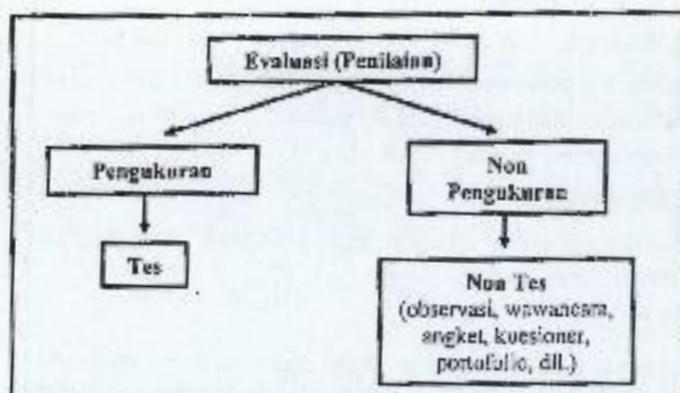
pengukuran yang dapat digunakan untuk memperoleh informasi mengenai karakteristik individu atau kelompok; yang diantaranya dapat dilakukan dengan cara: (1) tertulis dalam bentuk form pertanyaan, (2) oral (wawancara), (3) observasi menggunakan checklist atau catatan anekdot, dan (4) unjuk kerja dengan bantuan penggunaan peralatan mekanik seperti *treadmill*, lembing, bola, dan sebagainya.

Selanjutnya, kita akan membahas mengenai **Pengukuran**. Pengukuran adalah pengumpulan informasi sebagai dasar pengambilan keputusan (Baumgartner et al., 2007: 3). Tidak semua pengukuran harus dilakukan dalam bentuk tes. Ada juga pengukuran melalui non tes. Dalam pendidikan jasmani, pengukuran dalam bentuk tes dapat dilakukan melalui tes tertulis dan tes unjuk kerja; sedangkan pengukuran secara non tes dapat dilakukan melalui pengamatan, wawancara, angket, portofolio atau bentuk penilaian alternatif yang lain.

Berikutnya, kita akan membahas mengenai **Evaluasi**. Evaluasi adalah suatu pernyataan mengenai kualitas, kebaikan, kelebihan, nilai, atau kelayakan mengenai apa yang telah diukur dan berimplikasi pada pembuatan keputusan (Morrow, et al. 2005: 5). Evaluasi juga bermakna sebagai penggunaan informasi untuk membuat penilaian atau keputusan mengenai produk dan proses pada kegiatan pembelajaran (Rink, 2010:241).

Berdasarkan pengertian-pengertian yang telah dibahas sebelumnya, dengan demikian, dapat dikatakan bahwa tes merupakan salah satu bagian dari pengukuran, sedangkan pengukuran merupakan bagian dari evaluasi. Jadi evaluasi mempunyai pengertian yang lebih luas dibandingkan pengukuran maupun tes.

Hubungan antara penilaian, pengukuran, dan tes digambarkan dalam bentuk diagram sebagai berikut.



Gambar 1. Hubungan tes, pengukuran, dan evaluasi

Selain istilah tes, pengukuran, dan evaluasi, kita juga sering mendengar istilah *assessment*. Apa yang dimaksud dari *assessment*? *Assessment* adalah suatu proses untuk mengumpulkan informasi untuk membuat penilaian mengenai produk dan proses pembelajaran (Safrit & Wood dalam Rink, 2010: 241). Berdasarkan pengertian tersebut dapat disimpulkan bahwa *assessment* memiliki posisi dan makna yang sama dengan pengukuran. Saat dilakukan pengumpulan informasi melalui tes maka hal tersebut disebut sebagai *assessment*. Ketika guru menggunakan informasi yang dikumpulkan saat tes untuk membuat penilaian atau keputusan, maka guru tersebut melakukan kegiatan *mengevaluasi*.

Untuk membedakan antara tes, pengukuran, dan evaluasi secara lebih jelas, perhatikan ilustrasi berikut!

ILUSTRASI

- Anda hendak mengukur kapasitas VO_{2max} siswa anda.
- Anda meminta siswa melakukan lari 12 menit atau lari 2,4 km.
- Anda mengukur jarak tempuh atau waktu tempuh.
- Anda mengkategorikan siswa berdasarkan perolehan data tersebut, apakah baik, sedang, atau buruk sesuai dengan norma yang telah ada.

PERTANYAAN

Berdasarkan ilustrasi tersebut, yang manakah yang disebut tes, pengukuran, dan evaluasi?

JAWABAN

Dari pertanyaan di atas, saat anda meminta siswa melakukan lari 12 menit atau lari 2,4 km, maka kegiatan lari 12 menit atau lari 2,4 km tersebut disebut sebagai tes. Saat anda mengukur jarak tempuh atau waktu tempuh, maka anda melakukan pengukuran. Sedangkan saat anda mengkategorikan siswa berdasarkan perolehan data pengukuran, apakah termasuk kategori baik, buruk, sedang sesuai dengan norma yang telah ada, maka hal tersebut dinamakan sebagai evaluasi.

B. Permasalahan Pada Evaluasi Pembelajaran Pendidikan Jasmani

Penilaian merupakan salah satu komponen yang sangat penting dalam penyelenggaraan pendidikan. Upaya peningkatan kualitas pendidikan termasuk pendidikan jasmani salah satu di antaranya dapat dicapai melalui peningkatan kualitas pembelajaran dan sistem penilaiannya. Kualitas sistem pembelajaran dan kualitas sistem penilaian

merupakan mata rantai yang saling berkaitan. Sistem pembelajaran yang baik akan menghasilkan kualitas belajar yang baik. Kualitas pembelajaran ini dapat dilihat dari hasil penilaiannya. Selanjutnya sistem penilaian yang baik akan mendorong pendidik untuk menentukan strategi mengajar yang baik dan memotivasi peserta didik untuk belajar yang lebih baik. Oleh karena itu, dalam upaya peningkatan kualitas pendidikan jasmani olahraga dan kesehatan diperlukan perbaikan sistem penilaian yang diterapkan.

Dalam skala mikro, penilaian terhadap pencapaian hasil belajar siswa dalam pendidikan jasmani olahraga dan kesehatan merupakan langkah untuk mengetahui seberapa jauh tujuan kegiatan belajar-mengajar (KBM) telah dicapai. Salah satu tugas yang harus dilaksanakan oleh guru pendidikan jasmani olahraga dan kesehatan adalah menilai hasil belajar peserta didik. Pemberian nilai merupakan bagian yang tidak dapat dipisahkan dari program instruksional keseluruhan.

Pada proses evaluasi pembelajaran penjas, guru perlu melakukan evaluasi tidak hanya pada ranah psikomotor, tetapi juga pada ranah afektif dan kognitif, yang besaran proporsi bobotnya sesuai dengan kebijakan guru. Morrow (2000: 130) menyatakan bahwa untuk menilai hasil belajar siswa, penilaian yang diberikan oleh guru pendidikan jasmani harus bersifat menyeluruh/*holistic* yang terdiri atas tiga ranah, yaitu: *psychomotor domain*, *cognitive domain*, dan *affective domain*.

Proses evaluasi pembelajaran bermuara pada pemberian nilai pada siswa. Terkait hal tersebut, masih terdapat masalah-masalah terkait pemberian nilai dalam mata pelajaran PJOK. Hal ini mungkin terjadi karena mata pelajaran ini melibatkan aspek unjuk kerja keterampilan yang seringkali melibatkan subjektivitas penilaian guru. Misalnya, merupakan hal wajar mendapat nilai 10 dalam pelajaran Matematika, namun

merupakan hal yang sangat sulit, jika tidak ingin mengatakan "tidak mungkin", seorang siswa dapat memperoleh nilai 10 dalam mata pelajaran PJOK. Bagaimanakah kriteria nilai 10 pada saat bermain bolabasket, lompat jauh, lembar lembing, ataupun renang?

Hal tersebut terkait dengan subjektivitas ataupun objektivitas guru. Bahkan ketika guru sudah berusaha seobjektif mungkin dengan kriteria penilaian mendetail, acapkali guru dapat dibingungkan dalam memberi keputusan saat menilai siswa. Misalnya ketika menilai keterampilan renang. Meskipun sudah ada panduan menilai dari sikap awal, pelaksanaan, dan sikap akhir; bukan merupakan hal yang mudah bagi guru untuk memutuskan penampilan siswa. Apalagi jika guru tersebut masih guru baru yang minim pengalaman. Daripada bingung, bisa jadi seorang guru menetapkan nilai secara normatif secara wajar berdasar waktu tempuh tanpa mempertimbangkan teknik sehingga kurang merepresentasikan penampilan siswa sesungguhnya.

Ngatman (2006: 70) mengemukakan bahwa permasalahan yang sering timbul dan banyak dilontarkan berkenaan dengan pemberian nilai (*grading*) dalam pendidikan jasmani di antaranya adalah: (1) penentuan nilai yang dilakukan oleh sebagian guru pendidikan jasmani belum cermat dan cenderung bersifat subjektif, misalnya belum ada patokan pasti mengenai bagaimana penampilan siswa supaya mendapat nilai 6, 7, 8, bahkan 10, sehingga guru cenderung menilai dengan cara mengira-ngira, (2) para siswa yang memperoleh nilai rendah cenderung kurang termotivasi dalam kegiatan belajar, (3) penentuan nilai cenderung membeda-bedakan seseorang, yang jika tidak cermat pelaksanaannya dapat menimbulkan masalah sosial dan psikologis, dan (4) masih perlu ditingkatkannya pengetahuan

dan pemahaman guru mengenai pendekatan acuan penilaian dan sistemnya dalam pendidikan jasmani. Oleh karena itu, penilaian yang merupakan bagian dari pelaksanaan evaluasi perlu benar-benar dicermati guru supaya tidak merugikan siswa dan menjadi sarana pemacu siswa dalam meningkatkan kualitas belajarnya.

C. Prinsip-prinsip Evaluasi

Teknik dan prosedur evaluasi dalam pendidikan jasmani olahraga dan kesehatan menjadi bagian penting, namun untuk dapat melakukannya dengan baik dan benar memerlukan panduan berupa prinsip-prinsip evaluasi. Menurut I Made Suciandy Mahardika (2014) prinsip umum yang harus dipenuhi menyangkut keterkaitan antar komponen-komponen penting pengajaran, di antaranya:

1. Prinsip Keterpaduan

Perencanaan evaluasi harus dilakukan bersamaan dengan perencanaan tujuan, perencanaan materi ajar dan perencanaan proses belajar mengajar yang menyangkut strategi instruksional, kegiatan guru, kegiatan peserta didik dan sarana prasarana. Hindari kegiatan merencanakan evaluasi setelah PBM selesai dilaksanakan karena menyebabkan guru cenderung akan mengevaluasi hasil belajar, dan bukan mengevaluasi tujuan pembelajaran. Jika evaluasi hanya dilakukan terhadap hasil pengajaran maka prinsip keterpaduan sering tidak dapat tercapai. Akibatnya kecakapan yang akan dievaluasi cenderung kecakapan yang berhasil diajarkan guru atau kecakapan-kecakapan yang berhasil dikuasai peserta didik saja dan bukan kecakapan-kecakapan yang telah direncanakan untuk dikuasai oleh peserta didik melalui PBM.

2. Prinsip Partisipasi Aktif

Prinsip ini mewajibkan keterlibatan secara aktif pihak-pihak yang terlibat dalam evaluasi, seperti guru dan peserta didik. Prinsip ini menyebabkan evaluasi bukan hanya menjadi kepentingan lembaga atau guru saja, tetapi juga harus menjadi kebutuhan yang penting bagi peserta didik. Peserta didik harus merasa nyaman dengan adanya kegiatan evaluasi bukan justru sebaliknya menjadikan kegiatan evaluasi sebagai kegiatan yang menakutkan dan cenderung dihindari.

Guru harus berperan memfasilitasi secara aktif sedemikian rupa agar apa yang menjadi tujuan evaluasi dan apa yang akan dievaluasi dipahami peserta didik. Dengan demikian evaluasi merupakan kegiatan yang menjadi perhatian dan menyenangkan. Jika prinsip ini terpenuhi maka peserta didik akan merasakan kegiatan evaluasi sebagai kebutuhan untuk mengukur dan menguji kemampuannya, serta menganggap guru hanya membantu melaksanakannya. Bagaikan seorang petenis yang sudah latihan berbulan-bulan dan mempersiapkan fisik serta mental secara matang untuk menyongsong pertandingan. Sudah barang tentu petenis tersebut tidak sabar untuk bertanding guna menguji kemampuannya dari hasil latihan yang telah dilakukan.

3. Prinsip Kontinuitas

Secara umum evaluasi hasil belajar hanya dipahami sebagai kegiatan akhir dari sebuah proses pengajaran yang tidak perlu direncanakan dengan rinci. Namun sebenarnya tidak demikian, kegiatan evaluasi adalah kegiatan penting berikutnya yang akan menentukan masa depan peserta didik. Oleh sebab itu, seharusnya kegiatan evaluasi harus direncanakan secara sistematis. Guru perlu merencanakan tahapan-tahapan kegiatan evaluasi secara kontinyu dan pasti,

kapan dilakukan pretes, kapan tes formatif, dan kapan tes sumatif. Apabila semua pentahapan evaluasi direncanakan dengan baik maka semua pemangku kepentingan akan memiliki pedoman yang pasti dalam mengatur kontribusinya terhadap upaya belajar peserta didik.

Pemenuhan prinsip kontinuitas ini secara langsung akan menghasilkan model evaluasi yang lebih objektif, terutama karena peserta didik akan dapat mempersiapkan dirinya untuk menghadapi kegiatan evaluasi. Dengan demikian tujuan evaluasi untuk mengeksplorasi kemampuan peserta didik dalam menguasai materi pembelajaran akan berhasil dengan lebih baik karena melalui perencanaan yang matang dan terprogram.

4. Prinsip Koherensi (Kesesuaian)

Pada prinsipnya materi ajar yang dipelajari dalam proses belajar-mengajar adalah untuk mencapai tujuan-tujuan pembelajaran yang telah ditetapkan. Maka prinsip koherensi secara logis antara materi ajar yang menjadi bahan evaluasi dengan tujuan sangat penting. Sebagaimana diketahui tujuan pengajaran dirumuskan dalam bentuk tujuan instruksional umum yang kemudian di *break down* menjadi tujuan-tujuan instruksional khusus. Dari tujuan-tujuan instruksional khusus inilah ditetapkan materi pembelajaran yang relevan bagi peserta didik.

Materi pembelajaran tentu memiliki ruang lingkup yang luas, dan tidak mungkin sebuah alat evaluasi akan mampu menguji keseluruhan kecakapan yang dikuasai oleh peserta didik. Oleh karena itu sangat dibutuhkan kemampuan guru dalam menyeleksi kecakapan-kecakapan penting dari materi ajar. Ketepatan pemilihan materi ajar beserta sistem evaluasinya ini secara langsung akan berdampak pada

ketepatan pengukuran tujuan-tujuan instruksional.

5. Prinsip Akuntabilitas

Pada akhirnya semua hasil tes harus dapat dipertanggungjawabkan kepada berbagai pihak yang terkait (kepala sekolah, orang tua, masyarakat, dan instansi pemerintah). Secara moral pertanggungjawaban hasil tes harus dapat diberikan kepada Tuhan, karena guru melalui kegiatan evaluasinya mewarnai atau bahkan menentukan nasib peserta dididndi kemudian hari. Kekeliruan dalam mengevaluasi akan berdampak pada masa depan peserta didik. Oleh sebab itu, seorang guru harus memiliki moral yang baik dan tingkat kepercayaan kepada Tuhan yang tinggi, jika tidak maka nasib peserta didik dapat dipermainkan sebagai akibat praktik evaluasi seandainya. Pertanggungjawaban juga harus dapat diberikan kepada peserta didik yang dievaluasi, kepada *stake holder*, serta pemerintah.

D. Model Evaluasi pada Pendidikan Jasmani, Olahraga, dan Kesehatan

Tujuan utama pembelajaran adalah untuk menghasilkan perubahan perilaku yang terukur. Baungartner et al. (2007: 10-11) menjelaskan model evaluasi yang cocok digunakan pada setting sekolah adalah model evaluasi yang mengintegrasikan pengukuran dan evaluasi dengan proses pembelajaran. Komponen-komponen tersebut adalah: (1) tujuan, (2) pretest, (3) pembelajaran, (4) pengukuran, dan (5) evaluasi, yang akan dijelaskan berikut ini.

1. Tujuan

Penyiapan tujuan adalah langkah awal dalam proses evaluasi karena tujuan menentukan kecakapan apa yang

akan dicari. Tujuan menentukan arah pada pembelajaran dan mendefinisikan perilaku-perilaku apa yang ingin diubah.

2. Pretest

Dengan beberapa bentuk pretest, kita dapat menjawab tiga pertanyaan berikut: (1) berapa banyak yang telah dipelajari? (2) apa status dan kapabilitas individu terkini? (3) apa tipe aktivitas yang harus disusun untuk membantu mencapai tujuan? Melakukan pretest tidak selalu melibatkan pelaksanaan tes yang sama yang akan diberikan setelah pembelajaran; hal tersebut dapat berupa berbagai bentuk pengukuran yang dapat membantu menjawab pertanyaan yang ingin diketahui. Misalnya mengamati siswa berenang sepanjang kolam dapat menjadi pretest yang baik pada pelajaran renang.

3. Pembelajaran

Metode pembelajaran yang baik dibutuhkan untuk mencapai tujuan yang ingin didapat. Prosedur pembelajaran berbeda mungkin diperlukan untuk memenuhi kebutuhan siswa.

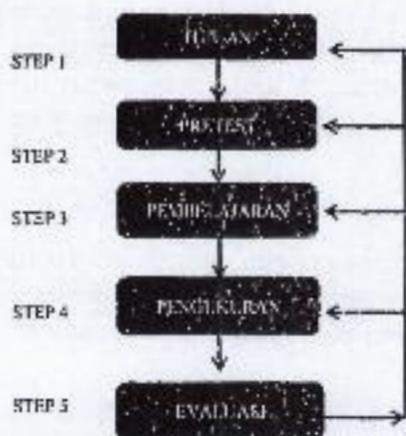
4. Pengukuran

Pengukuran melibatkan pemilihan atau pengembangan tes untuk mengukur pencapaian tujuan. Merupakan hal krusial bahwa tes didesain untuk mengukur perilaku yang dirinci dalam tujuan. Tujuan tersebut dapat berupa kognitif, afektif, psikomotor, atau terkait kebugaran. Elemen kuncinya adalah memilih atau mengembangkan tes yang mengukur tujuan yang ingin diukur. Seringkali, guru perlu mengembangkan tesnya sendiri karena tes-tes standar yang ada tidak sejalan dengan tujuan pembelajarannya. Dalam hal ini, validitas isi dicapai ketika tes yang ada kongruen

dengan tujuan pembelajaran. Ini adalah metode yang umum digunakan untuk menyediakan bukti validitas tes pada setting pendidikan.

5. Evaluasi

Ketika fase pembelajaran telah usai dan kecakapan telah diukur, maka hasil tes akan dievaluasi untuk menyidik apakah perubahan yang diinginkan tercapai seperti tujuan yang dinyatakan. Bagaimana jika siswa tidak mencapai tujuan yang diinginkan? Gambar 1 menyediakan gambaran siklus umpan balik dari evaluasi kembali pada tiap komponen model.



Gambar 1. Model sistematis evaluasi yang sesuai untuk setting pendidikan di sekolah

Sumber: Baumgartner et al. (2007: 11)

Kegagalan untuk mencapai tujuan dapat dikarenakan tiap segmen dari model. Pertama, dapat ditemui bahwa tujuan tidak sesuai atau perlu diubah. Pembelajaran mungkin tidak sesuai dengan kelompok siswa atau tes yang dipilih mungkin

tidak sesuai. Jadi, model evaluasi pendidikan adalah suatu hal yang dinamis dan menyediakan informasi yang dibutuhkan untuk mengubah tiap aspek dalam proses pendidikan.

E. Faktor-Faktor yang Perlu Dipertimbangkan dalam Evaluasi/Penilaian

Beberapa faktor yang lazim dipergunakan sebagai bahan pertimbangan untuk menentukan nilai seorang siswa menurut Miller (2002: 83-86) ada tiga, yaitu: (1) faktor afektif, terdiri atas: sportivitas, kehadiran, partisipasi, dan sikap, (2) faktor kognitif, dan (3) faktor psikomotor, terdiri atas: aktivitas, penampilan permainan, dan kebugaran. Sedangkan menurut Saifit (1986: 460-462) faktor-faktor yang dipakai sebagai dasar untuk memberikan penilaian kepada siswa adalah: (1) kehadiran, (2) usaha, (3) pakaian yang digunakan dan penampilan, (4) sportivitas, dan (5) memberikan reportase/ulasan yang akan dijelaskan sebagai berikut.

1. Kehadiran di kelas

Merupakan salah satu faktor yang dapat dipakai sebagai bahan pertimbangan oleh guru pendidikan jasmani dalam memberikan penilaian. Pada umumnya kehadiran siswa di sekolah/kelas disesuaikan dengan tujuan dan kebijakan sekolah yang bersangkutan. Siswa yang hadir di kelas kurang dari batas minimal kehadiran (misalnya 75% kehadiran) akan dikenai sanksi oleh guru dan biasanya akan mengakibatkan dikenai hukuman atau pengurangan nilai.

2. Faktor usaha

Usaha yang telah dilakukan oleh siswa juga perlu mendapatkan pertimbangan penilaian. Sekalipun siswa mendapatkan nilai hasil belajar yang jelek, namun apabila siswa

tersebut sudah berusaha semaksimal mungkin (misalnya: rajin mengikuti pelajaran, tekun belajar, rajin latihan/mengikuti kegiatan ekstra kurikuler) maka selayaknya kepada siswa tersebut diberikan nilai penunjang sebagai penghargaan atas jerih payah yang telah dilakukannya. Sebaliknya bagi siswa yang memiliki nilai hasil belajar rendah tetapi tanpa ada usaha yang sungguh-sungguh untuk memperbaiki nilainya (sering membolos sekolah, malas latihan/malas mengikuti kegiatan ekstra kurikuler, malas belajar) maka cukup beralasan jika guru pendidikan jasmani memberikan penilaian sesuai apa adanya.

3. Tata cara berpakaian

Cara siswa berpakaian dan berpenampilan di kelas merupakan salah satu aspek yang juga perlu dipertimbangkan dalam penilaian. Menggunakan pakaian olahraga pada saat mengikuti senam maupun aktivitas pendidikan jasmani pada umumnya merupakan persyaratan yang mutlak harus dipatuhi oleh siswa karena hal ini berkaitan dengan kenyamanan pada saat melakukan aktivitas tersebut. Jika pakaian olahraga yang dikenakan oleh siswa masih kotor saat masuk gedung olahraga/*sport hall* (seperti: kaos masih basah, sepatu olahraga kotor), maka keadaan tersebut dapat mengurangi penilaian siswa. Hal ini sangat berbeda apabila saat siswa pada waktu mengikuti pelajaran pendidikan jasmani dalam keadaan berpakaian olahraga rapi, kaos dan sepatu tidak basah/kotor.

4. Sportivitas

Sportivitas merupakan unsur yang tidak terlupakan dari setiap penyelenggaraan pengajaran pendidikan jasmani. Setiap guru pendidikan jasmani seyogyanya berupaya untuk

menanamkan jiwa sportivitas di kalangan anak didik pada proses pembelajaran, walaupun pada kenyataannya mengukur nilai sportivitas siswa bukan merupakan hal mudah karena berkaitan dengan sikap seseorang. Apabila guru pendidikan jasmani ingin menjadikan sportivitas sebagai salah satu unsur penilaian siswa, maka guru tersebut harus selalu memantau dan mencatat dengan cermat perubahan perilaku yang ditunjukkan oleh setiap siswa secara terus-menerus dalam rentang waktu yang lama.

5. Penyajian ulasan/reportase

Penyajian ulasan/reportase dapat juga digunakan sebagai salah satu faktor penilaian pendidikan jasmani. Model pendekatan penilaian ini banyak digunakan oleh sebagian besar guru pendidikan jasmani pada tahun 1982 sebagai dasar penilaian guru pendidikan jasmani terhadap siswanya. Pendekatan penilaian dengan model ini menuntut siswa agar bisa menyajikan informasi kegiatan pendidikan jasmani/olahraga dalam bentuk uraian/reportase seperti seorang presenter (Miller, 2002: 82)



BAB 2

TUJUAN/ALASAN MELAKUKAN TES, PENGUKURAN DAN EVALUASI

Guru pendidikan jasmani akan dihadapkan pada evaluasi pembelajaran sebagai tindak lanjut atas pembelajaran yang dilakukan. Evaluasi yang dilakukan berdasarkan hasil pengamatan berguna untuk memotivasi, mendiagnosa, menyusun peringkat, dan untuk penelitian. Selain itu, data yang diperoleh dari pengamatan dapat digunakan untuk memprediksi kemampuan siswa, mengelompokkan siswa berdasarkan kemampuan /keterampilan (*placement*), ataupun sebagai bahan laporan ke sekolah dan orang tua.

Supaya bermakna, harus diingat bahwa proses pengamatan harus dilaksanakan dengan tujuan tertentu yang jelas. Proses pengamatan akan kurang bermakna jika dilakukan hanya untuk memenuhi tuntutan otoritas tertentu seperti sekolah, meskipun memang hal tersebut adalah alasan yang umum. Bab ini akan membahas *dasar atau alasan-alasan* yang lebih baik untuk melakukan tes dan pengukuran pada pendidikan jasmani dan menyediakan saran mengenai bagaimana pengamatan keterampilan dapat digunakan untuk meningkatkan proses pendidikan.¹

¹ Swand, Bradford N. & Wilson, R. (1993). *Assessing Sport Skills* hal. 3-7. Champaign: Human Kinetics Publishers.

A. Pengukuran Hasil Belajar Psikomotor

Banyak kalangan profesional yang percaya bahwa pencapaian keterampilan psikomotor adalah tujuan nomor satu dari pendidikan jasmani. Jika memang demikian, maka guru perlu menunjukkan bahwa siswanya mengembangkan keterampilan psikomotornya. Banyak siswa yang secara sengaja-sengaja dijadikan subjek tes yang tidak dikembangkan secara bagus, sehingga hasilnya kurang mencerminkan keterampilan siswa sesungguhnya. Oleh karena itu, guru perlu berhati-hati dalam memilih tes keterampilan.

Hasil pengujian keterampilan psikomotor ini dapat digunakan guru untuk menentukan kelebihan dan kekurangan siswa. Selain itu, hasil tes menyediakan bahan masukan mengenai keterampilan apa yang perlu diajarkan serta drill apa yang harus diimplementasikan pada siswa.

Setelah suatu keterampilan diperkenalkan dan dikembangkan, guru dapat memberikan 1 item tes sebagai suatu alat evaluasi *formative* pendek (*short formative evaluation*). Evaluasi *formative* pada 1 item tes memang tidak menyeluruh seperti pada *full test*, namun hal tersebut menyediakan ide cepat bagi siswa untuk mengetahui seberapa baik mereka berkembang.

Banyak guru yang menggunakan item tes keterampilan sebagai kegiatan dalam pembelajaran. Pada dasarnya, menggunakan item tes sebagai drill untuk pengembangan keterampilan meningkatkan validitas dan reliabilitas tes karena siswa telah mempraktikkan item tes tersebut.

Pada akhir unie, anda dapat menggunakan tes keterampilan lengkap sebagai alat evaluasi *summative*. Jika anda memberikan evaluasi *formative* pada tiap keterampilan, kemudian anda menerapkan drill pada item tes, dan jika anda

membantu siswa meningkatkan keterampilannya, maka skor mereka pada tes keterampilan final akan bagus.

B. Sebagai Motivasi

Motivasi adalah daya yang mendorong individu untuk meraih pencapaian puncak. Seringkali guru pendidikan jasmani harus memotivasi siswa yang tidak menyukai pendidikan jasmani. Melalui administrasi tes yang baik (misal mengikuti petunjuk tes secara eksplisit dan memastikan tiap siswa dites dalam kondisi yang sama) dan penanganan yang cermat (seperti menjaga kerahasiaan hasil tes keterampilan), anda dapat memotivasi siswa untuk mencapai level penampilan pendidikan jasmani yang lebih baik.

Jika teradministrasi dengan baik, tes keterampilan dapat memberi siswa pengalaman positif. Namun, tes yang diadministrasi dengan tidak tepat akan mengakibatkan efek negatif pada siswa. Pertimbangkan bahwa banyak anak menerima penghinaan selama tes pendidikan jasmani yang dilakukan oleh guru yang tidak perhatian atau kurang informasi yang merendahkan siswa dengan menanyakan, **"Kamu hanya bisa memasukkan bola basket 1 kali?"** Sayangnya, pada profesi pendidikan jasmani, beberapa siswa yang terhina saat penjas tersebut di masa selanjutnya menjadi pengawas sekolah, kepala sekolah, dan anggota dewan sekolah yang mengontrol pendanaan dan pengambilan keputusan kurikulum. Pada akhirnya, kejadian penghinaan yang didapat pada pembelajaran penjas semata sekolah dapat berimplikasi buruk pada pelaksanaan pendidikan jasmani sendiri.

Ketika dilaksanakan secara tepat, tes keterampilan dapat menyalakan spirit kompetitif pada banyak siswa, namun guru pendidikan jasmani harus memperhatikan kompetisi ini. Ketika seluruh prosedur pengujian berdasarkan kompetisi,

beberapa siswa akan gagal karena mereka tidak peduli atau tidak tertarik untuk berkompetisi dengan siswa lain. Ketika hal ini terjadi, maka *self-competition* dalam bentuk tes peningkatan keterampilan direkomendasikan.

Selanjutnya, motivator terbesar adalah kesuksesan. Bagi siswa, supaya tetap termotivasi, mereka harus melihat bahwa perkembangan keterampilannya meningkat. Tes keterampilan menyediakan *self-competition* dan jalan bagi siswa untuk mengukur kemajuannya. Siswa harus belajar untuk menetapkan tujuan yang realistis dan dapat dicapai yang spesifik sesuai dengan situasi dirinya. Ketika tujuan telah ditetapkan, pengetesan keterampilan dapat membantu siswa untuk mengevaluasi kemajuannya, menetapkan tujuan baru, mengembangkan spirit, dan utamanya mendorong peningkatan fisik dan aktualisasi diri.

Guru harus mendorong siswa secara konsisten untuk mencapai kemampuan terbaiknya baik dalam tes, praktek, maupun *game*. Mengetahui bahwa dirinya telah berkembang adalah motivator besar bagi siswa yang mengembangkan keterampilan baru. Kebanyakan siswa ingin meningkatkan keterampilan, kematangan sosial, atau pengetahuan kognitif. Peningkatan kontinyu akan menginspirasi siswa dan memotivasinya untuk menghadiri kelas, berperan aktif di kelas, dan merasa antusias dengan pelajaran penjas berikutnya.

Banyak siswa senang membandingkan penampilannya dengan siswa lainnya. Membandingkan penampilan memang memotivasi beberapa siswa, namun hal tersebut dapat berdampak negatif bagi lebih banyak siswa lainnya, khususnya bagi individu yang baru saja mempelajari keterampilan baru. Jadi, guru perlu mempertimbangkan dan memperingatkan siswa mengenai kegiatan membanding-bandingkan ini.

Berikutnya, guru perlu menginformasikan pada siswa mengenai level keterampilannya pada suatu unit terkait dengan tujuan unit. Ketika siswa tahu apa yang diharapkan dan menyadari level keterampilannya, banyak siswa akan mengembangkan ketertarikan dan rasa antusias yang cukup untuk meningkatkan level keterampilannya guna memenuhi tujuan pembelajaran. Ketika siswa dapat membandingkan penampilan dirinya dengan tujuan/standar pembelajaran, maka siswa tampak lebih mampu untuk mencapai level tersebut.

Semua siswa berharap dan mencari umpan balik baik dari temannya, atau guru, atau keduanya. Dengan mengingat hal ini, guru dapat menggunakan tes keterampilan yang membuatnya dapat memberikan umpan balik secara positif pada situasi yang tidak mengancam. *Feedback* dapat memotivasi siswa untuk meningkatkan keterampilannya, bukan untuk menyenangkan guru tetapi untuk realisasi diri.

C. Penentuan Peringkat/Grading

Penentuan peringkat memiliki banyak tujuan. Secara utama, penentuan peringkat membuat siswa tahu bagaimana kemajuan mereka terhadap tujuan belajar dan posisinya di antara siswa lainnya. Peringkat dapat mempengaruhi siswa terhadap pendidikan jasmani. Peringkat memberi informasi bagi orang tua mengenai kemajuan dan kecakapan anaknya; memberi informasi bagi guru mengenai kemajuan siswa dan apakah pendampingan tambahan diperlukan; memberi informasi administrator mengenai kesuksesan program dan menyediakan informasi guna pengambilan keputusan administratif.

Terlalu banyak guru pendidikan jasmani yang menentukan nilai atau peringkat berdasarkan pada kesesuaian

pakaian, kebersihan diri, sikap, dan partisipasi. Faktor-faktor tersebut memang penting pada administrasi suatu unit pembelajaran, tetapi kurang berkaitan dengan seberapa banyak siswa belajar atau mencapai tujuan. Ketika digunakan dengan tepat, tes keterampilan membantu guru membuat tujuan daripada mengevaluasi siswa secara subjektif. Penentuan peringkat kemudian lebih mengindikasikan prestasi siswa terkait tujuan mata pelajaran.

Pada intinya, penilaian atau penentuan peringkat harus dilakukan secara objektif, terkait dengan tujuan mata pelajaran, dan harus merefleksikan kemajuan, prestasi, dan peningkatan siswa. Ketika siswa tahu bahwa mereka akan dievaluasi secara objektif berdasarkan prestasi atas tujuan pelajaran yang ditetapkan, mereka akan belajar lebih keras dan akan lebih menerima nilainya, mengetahui bahwa mereka bertanggungjawab atas apa yang mereka capai. Namun, ketika tujuan pelajaran tidak pernah dinyatakan dan siswa tidak pernah dites, maka penentuan peringkat atau nilai menjadi kurang bermakna.

D. Untuk Diagnosis

Diagnosis adalah penilaian berdasarkan fakta-fakta yang diamati. Merupakan fungsi guru untuk secara kontinyu dan objektif mengamati kelemahan, kekuatan, dan kemajuan siswa. Informasi yang diperoleh dari pengamatan keterampilan membantu guru dan siswa mengetahui kekurangan. Merupakan hal yang sulit untuk melihat apakah siswa dapat memegang bolabasket dengan benar jika mereka hanya mendapat satu atau dua kali kesempatan memegang bola selama permainan. Tetapi, tes keterampilan bolabasket dapat digunakan untuk mengidentifikasi kekurangan tertentu yang membuat guru dapat menyusun latihan perkembangan

spesifik untuk siswa yang memerlukan bantuan. Ketika tes keterampilan bolabasket tambahan diberikan, kemajuan siswa dapat dievaluasi kembali.

Guru juga dapat menggunakan hasil tes keterampilan untuk mengevaluasi kembali tujuan lama dan memformulasikan tujuan dan metode baru. Jika hasil tes menunjukkan bahwa siswa tidak mencapai tujuan, guru harus mengubah tujuan supaya sejalan dengan apa yang dapat dicapai siswa atau mengganti pendekatan pembelajaran sehingga lebih banyak waktu dialokasikan untuk pengembangan keterampilan dan utamanya penampilan yang lebih baik dari siswa pada tes keterampilan tercapai.

E. Untuk Prediksi

Prediksi adalah tindakan meramalkan suatu perkiraan atas penampilan masa depan berdasarkan pada fakta-fakta terkini. Siswa mungkin tidak yakin dengan kemampuannya, khususnya jika mereka mengalami perkembangan yang telat. Dari tes keterampilan dan observasi, guru dapat menasihati dan mendorong siswa untuk mengejar aktivitas atau olahraga tertentu. Seorang siswa kelas 7 yang kikuk dengan beberapa keterampilan atletik mungkin menjadi atlet bagus pada saat kelas 11 jika dia memilih untuk tetap menekuni suatu olahraga tertentu. Seringkali, siswa kelas 7 yang kikuk yang tidak menyadari potensinya menjadi berkecil hati dan berhenti.

Melalui tes dan pengukuran, guru dapat memprediksi potensi siswa untuk kesuksesan masa depan pada aktivitas tertentu. Namun, supaya perkembangan terjadi, siswa harus menyadari potensinya untuk sukses. Tes keterampilan menyediakan siswa informasi yang perlu mereka percayai bahwa mereka memiliki potensi pada suatu aktivitas tertentu.

Hasil tes juga membuat siswa percaya diri untuk berpartisipasi pada kompetisi olahraga.

Tes keterampilan juga dapat digunakan untuk memprediksi kesuksesan pada cabang olahraga lainnya yang memiliki karakteristik serupa. Misalnya, jika siswa dapat melakukan tes keterampilan melempar bola sambil berlari dengan baik, maka guru dapat merekomendasikan olahraga atau aktivitas yang memerlukan keterampilan serupa. Hal ini mendorong siswa untuk mengejar jalan baru ketika mereka mengembangkan keterampilan baru.

F. Penempatan (*Placement*)

Dalam usaha meningkatkan pembelajaran, guru dapat mengklasifikasikan siswa pada grup-grup tertentu. Meskipun hal tersebut menuai pro dan kontra, dipercaya bahwa terkadang kegiatan klasifikasi siswa membantu dalam pembelajaran. Siswa dengan kemampuan superior seharusnya tidak dipaksa untuk belajar bersama dengan teman yang kurang terampil dan sebaliknya, siswa dengan keterampilan kurang harus mampu untuk merasakan kesuksesan. Oleh karenanya, pengtesan keterampilan menyediakan metode yang cepat, efektif, dan objektif untuk mengklasifikasikan siswa sesuai kelompok kemampuan. Ketika keterampilannya berkembang, siswa dapat dengan mudah berpindah dari suatu level ke level yang lain.

Ketika siswa ditempatkan pada grup-grup berdasarkan level keterampilan, guru dapat mengindividualisasi pengajaran. Misalnya pada unit sepakbola, siswa berketerampilan lebih dapat bermain game yang melibatkan bertahan dan menyerang, sementara siswa pemula bermain 2 vs 2 untuk mengembangkan keterampilan *passing* dan *catching* ketika lari dan menghindar.

Selain itu, pengelompokan siswa berdasarkan keterampilan dapat meningkatkan situasi pembelajaran. Ketika siswa ditempatkan dengan siswa yang memiliki keterampilan seimbang dalam pembelajaran, praktik, ataupun game, mereka merasa lebih tidak terintimidasi, tidak terhambat, lebih nyaman, lebih aktif, belajar lebih, dan mendapat kesenangan lebih dari pengalaman pembelajaran. Ketika siswa dengan keterampilan serupa dikelompokkan, tujuan pembelajaran dapat disusun sesuai kebutuhan.

G. Untuk Evaluasi Program

Evaluasi program harus menjadi hal kontinyu dari sistem sekolah, karena tanpa evaluasi maka sulit untuk mengetahui jika, kapan, pada bagian mana, dan bagaimana perubahan harus dibuat. Program harus dievaluasi berdasarkan seberapa bagus program memenuhi tujuan yang dinyatakan.

Tes keterampilan membantu evaluasi program pada beberapa cara. Pertama, jika tes menunjukkan bahwa siswa telah berkembang banyak pada keterampilan tertentu, anda mungkin ingin menaikkan tujuan dengan menambah aktivitas serupa atau mencari aktivitas yang mengembangkan keterampilan baru. Kedua, jika keterampilan siswa tidak menunjukkan perubahan atau bahkan menurun, anda mungkin memerlukan aktivitas tambahan untuk mengembangkan keterampilan, atau mungkin memutuskan untuk mengeluarkan aktivitas tersebut dari kurikulum. Ketiga, tes keterampilan menyediakan perbandingan outcome program dengan norma lokal, wilayah, dan negara. Hasilnya dapat digunakan untuk memodifikasi unit, peralatan, fasilitas, melakukan proyek awal, dan mengembangkan teknologi baru. Keempat, hasil tes keterampilan dapat digunakan dalam mengadopsi standar yang tepat untuk program evaluasi.

H. Untuk Evaluasi Guru

Guru harus dievaluasi pada variabel yang dapat mereka kontrol, seperti penampilan dan prestasi siswa terkait dengan tujuan yang dinyatakan. Tes keterampilan dapat membantu evaluasi guru dengan menyediakan informasi apakah tujuan pembelajaran dapat dicapai. Jika beberapa tujuan tidak dapat tercapai, perubahan pada teknik pengajaran dan metodologi disarankan. Tes keterampilan juga dapat mengidentifikasi kelebihan dan kekurangan guru.

Evaluasi juga dapat menunjukkan masalah peralatan. Ketika hanya ada 1 bola basket untuk 35 siswa, sangat sulit bagi guru untuk menyesuaikan pembelajaran guna mencapai tujuan. Hal ini dapat menjadi masukan terkait pengambilan keputusan pada pelaksanaan administrasi sekolah, misalnya terkait alokasi dana untuk pembelian peralatan ataupun kebijakan jumlah siswa per kelas.

I. Untuk Penelitian

Penelitian adalah suatu investigasi sistematis guna meningkatkan pengetahuan. Kualitas data yang dikumpulkan pada penelitian salah satunya bergantung pada presisi dan akurasi instrument pengukuran, teknik pengukuran, dan kesesuaian tes.

Penggunaan tes keterampilan pada penelitian pendidikan jasmani membantu administrator memilih aktivitas yang dimasukkan kurikulum. Misalnya, jika tes keterampilan menunjukkan bahwa siswa memiliki keterampilan melempar di bawah rata-rata, maka aktivitas melempar dapat ditekankan. Berdasarkan hasil tes, guru dapat melihat kebutuhan untuk mengganti metode pembelajaran.

Skor rendah pada suatu item tes dapat mengindikasikan bahwa guru tidak menghabiskan waktu cukup pada pengembangan keterampilan, tidak menjelaskan teknik se jelas mungkin, atau tidak menyediakan waktu latihan yang cukup.

Penelitian menggunakan tes keterampilan dapat menemukan solusi dari masalah kelas. Perubahan kurikulum dan evaluasi program juga dapat dilakukan secara cerdas setelah penelitian tes keterampilan.



BAB 3

KRITERIA TES YANG BAIK

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 penyusun 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. Sesuaikan 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 ini.



Gambar 2. Kriteria tes yang baik

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. Safric (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 Korelasi	Validitas
0.85 – 0.99	Excellent (sangat bagus sekali)
0.80 – 0.84	Very good (sangat bagus)
0.70 – 0.79	Acceptable (dapat diterima)
0.60 – 0.69	Questionable (dipertanyakan)

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 Jeffry 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 Korelasi Reliabilitas

Koefisien Korelasi	Validitas
0.95 - 0.99	Excellent (sangat bagus sekali)
0.90 - 0.94	Very good (sangat bagus)
0.80 - 0.89	Acceptable (dapat diterima)
0.70 - 0.79	Poor (buruk)
0.60 - 0.79	Questionable (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 utit 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 pengesanan, siswa harus belajar sesuatu mengenai dirinya dan variabel yang diteskan.

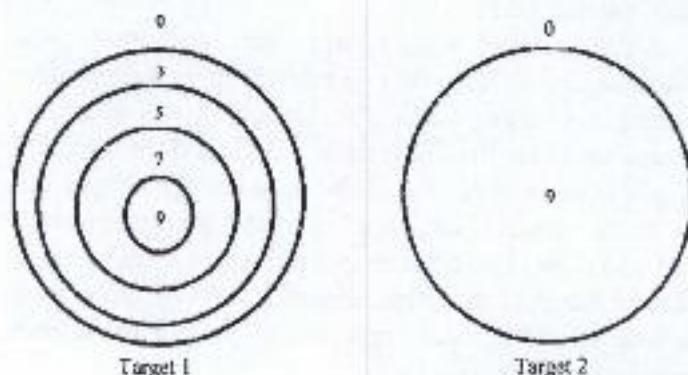
Atribut afektif yang didapat dari partisipasi pada pendidikan jasmani dan olahraga mencakup pengembangan kejujuran, sportivitas, harga diri, integritas, dan keadilan. Beberapa situasi tes membuat siswa mengetes dirinya sendiri atau menghitung temannya. Selama situasi tes khusus ini, guru harus menekankan 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 pengesanan 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

Sebelum menggunakan tes apapun, kajilah apakah tes terlalu memforsir siswa atau dapat mengakibatkan cedera. Sebelum memulai tes fisik, siswa harus diizinkan 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 dilibat, 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 perjas 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 repetisi 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 (Baurgartner & Jackson, 1982). Namun, Safrit (1990) memperingatkan bahwa kemampuan 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 pertombaan; dalam hal ini banyak individu telah mengembangkan polanya sendiri untuk melakukan keterampilan (Safrit, 1990).



BAB 4

LANGKAH-LANGKAH MEMILIH DAN MENGKONSTRUKSI TES

Banyak guru penjas yang tidak menggunakan tes yang telah disusun sebelumnya karena menemui bahwa tidak ada tes terpublisk 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 tes.¹

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

¹ Strand, Bradford N. & Wilson, R., (1993). *Assessing Sport Skills* halam 9-22. Champaign: Human Kinetics Publishers.

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 memudahkannya, 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 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

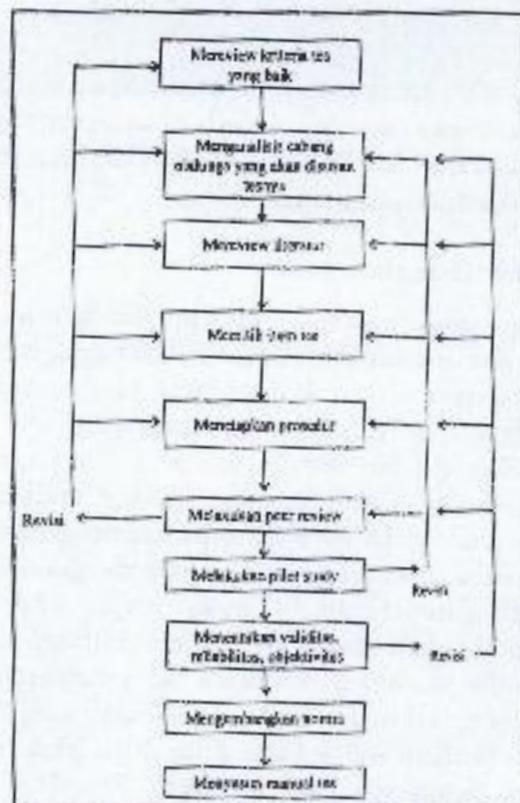
tidak mungkin untuk mengajarkan semua teknik. Maka anda perlu menyortir teknik-teknik dasar terpenting.

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
 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 (Miller, 1988). Kesimpulan dalam tiap aspek akan mengurangi waktu persiapan, waktu administrasi, waktu latihan, dan membuat siswa lebih baik dalam menjalani tes. Pada titik ini, trial dan error terjadi karena guru mencoba untuk menentukan

prosedur terbaik. Revisi dan adaptasi tidak terelakkan dan membantu 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; Saifit, 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 pengujian 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 pengesanan 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 peringkat yang direkomendasikan diukur.

J. Menyusun Manual Tes

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



BAB 5 VALIDITAS TES DAN PENGUJIANNYA

A. Konsep Tentang Validitas Tes

Ketepatan suatu tes dalam mengukur gejala atau bagian gejala sering disebut pula dengan istilah validitas atau kesahihan. Dengan kata lain, validitas suatu tes dapat diartikan derajat ketepatan pengukuran yang besarnya ditentukan oleh statistik korelasi antara skor prediktor dan skor kriterium. Jensen dan Hirst (1980: 22) mengatakan bahwa validitas tes adalah derajat kebenaran sebuah tes mengukur apa sebenarnya yang ingin diukur. Contoh: apabila orang ingin mengukur tinggi badan, alat ukur yang paling tepat adalah dengan menggunakan *stadiometer*. Demikian juga, apabila orang ingin mengukur tingkat kecepatan lari *sprinter*, alat yang paling tepat digunakan adalah *stopwatch*.

Berbicara masalah validitas tes, ada dua aspek yang harus dipertimbangkan, yaitu: (1) relevansi, dan (2) keterandalan. Relevansi adalah tingkat seberapa jauh alat ukur/tes tersebut memiliki kesesuaian dengan apa yang akan diukur dan fungsi yang dimaksud akan diukur oleh tes tersebut. Di sisi lain keterandalan mengacu pada tingkat ketelitian dan ketetapan dari suatu alat ukur (Ngatman, 2002: 8).

Instrumen evaluasi dikatakan memiliki validitas yang baik apabila instrumen tersebut mengukur secara mat-

cermatnya apa yang dimaksudkan akan diukurnya. Apabila ingin mengetahui keterampilan siswa bermain bulutangkis, tes yang valid untuk mengukur keterampilan itu adalah tes keterampilan bermain bulutangkis dan bukan tes kecepatan smash bulutangkis. Khusus mengenai tes keterampilan pendidikan jasmani, tingkat validitas tes tersebut jangan hanya ditentukan oleh statistik saja, namun pertimbangan lain harus dibuat dalam memilih dan membuat tes untuk mengumpulkan data, yaitu: (1) tes harus mengukur kemampuan penting, (2) tes harus melibatkan hanya satu orang, (3) tes harus dapat diskor secara cermat, (4) tes harus mempunyai jumlah percobaan (*trial*) yang memadai, (5) tes harus memadai tingkat kesulitan (Arma Abdoellah, 1985: 6).

Ada beberapa cara untuk menentukan tingkat validitas/kesahihan dari tes yang dibuat. Dalam hal ini, tes yang dibuat oleh guru pendidikan jasmani dibandingkan dengan sebuah kriterium yang dipilih. Sampel diberikan tes yang dibuat tersebut (tes buatan guru pendidikan jasmani) dan tes kriterium yang telah dipilih, sehingga diperoleh dua set skor. Apabila koefisien korelasi antara kedua set skor itu tinggi, tes yang dibuat tersebut dapat dikatakan valid atau sah.

Menurut Barrow dan Mc. Gee (1979: 71), dalam pendidikan jasmani ada beberapa macam kriterium yang biasa dipergunakan oleh para ahli penyusun tes, di antaranya adalah: (1) penilaian yang dilakukan oleh ahli dalam bidangnya/*rating expert judgement*, (2) tes yang sama yang telah diketahui validitasnya, tetapi tes tersebut tidak ekonomis, (3) kedudukan dalam sebuah turnamen, (4) skor gabungan (*composite score*), dan (5) pertimbangan empiris.

Banyak faktor yang mempengaruhi tingkat validitas tes dalam pendidikan jasmani. Seorang guru pendidikan jasmani

harus mengetahui faktor-faktor ini dan memperhatikannya dalam memilih dan membuat tes. Setiap tes hanya valid untuk kelompok tertentu. Sebuah tes yang dirancang untuk mengukur tingkat keterampilan bermain bola voli untuk mahasiswa, mungkin tidak akan valid untuk mengukur tingkat keterampilan bermain bola voli anak-anak usia sekolah dasar. Demikian halnya, sebuah tes keterampilan yang valid digunakan untuk pemula mungkin tidak bernilai apabila dipergunakan untuk kelompok yang lebih tinggi urtuk kerjanya (*performance*). Jadi, tingkat validitas instrumen evaluasi sifatnya tidak umum, namun bersifat spesifik. Secara ekstrim Kirkendall (1980: 78) berpendapat bahwa kebudayaan dan sifat khusus suatu kelompok sering merupakan faktor yang menentukan validitas suatu tes keterampilan.

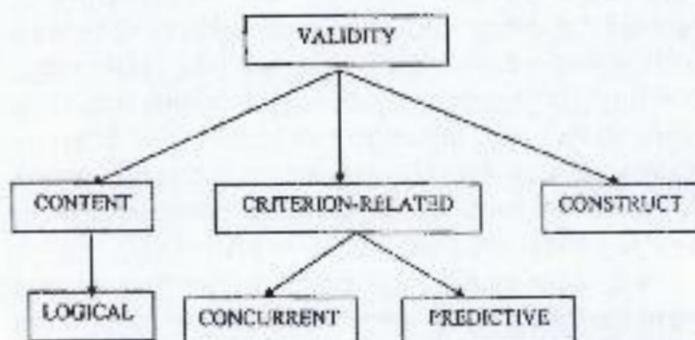
Jadi, dapat ditarik suatu makna bahwa permasalahan yang utama validitas suatu instrumen evaluasi/tes pendidikan jasmani ada dua hal, yaitu:

1. Tingkat kejituan, ketepatan, atau keakuratan instrument atau alat pengukuran. Suatu alat evaluasi dikatakan jitu apabila alat evaluasi tersebut dengan tepat mengenai sasarannya; alat evaluasi yang dapat mengerjakan dengan tepat fungsi yang diserahkan kepadanya.
2. Tingkat ketelitian, keseksamaan atau kecermatan proses pengukuran. Suatu alat evaluasi akan menghasilkan data yang akurat apabila proses pengukuran yang dilakukan dengan alat evaluasi tersebut dilakuakn oleh seseorang yang memiliki kemampuan dan kecermatan serta menguasai instrument dengan baik.

B. Tipe-Tipe Validitas

Dari cara estimasinya yang disesuaikan dengan fungsi setiap tes, tipe-tipe validitas pada umumnya dapat

digolongkan ke dalam tiga kategori, yaitu: (1) validitas isi (*content validity*), (2) validitas konstruk (*construct validity*), dan (3) validitas berdasarkan kriteria (*criterion-related validity*). Dalam pendidikan jasmani, validitas logika merupakan sebuah permasalahan khusus dari validitas muatan (*content validity*) yang sering digunakan. Berkaitan dengan konsep validitas alat evaluasi, ada beberapa tipe validitas seperti tersaji pada gambar 4 berikut ini.



Gambar 4. Tipe-tipe Validitas (Safrit, 1986: 119)

1. Validitas Isi (*Content Validity*)

Validitas isi menunjuk seberapa jauh tes mengukur keseluruhan kawasan pokok bahasan dan perilaku yang hendak diukur. Maksudnya, tes itu harus komprehensif, relevan dan tidak keluar dari batasan tujuan pengukuran. Linn dan Grondlund (1995: 85) mengemukakan esensi tujuan pengujian validitas isi sebagai berikut: *"the goal in consideration of content validation is to determine the extent to which a set of assessment tasks provide a relevant and representative sample of domain of tasks about which interpretation of assessment results are made"*.

Validitas isi adalah validitas yang mempertanyakan bagaimana kesesuaian antara instrument dengan tujuan dan deskripsi bahan yang diajarkan atau deskripsi masalah yang akan diteliti (Nurgiyantoro, dkk., 2012: 339). Lebih lanjut dijelaskan bahwa untuk mengetahui kesesuaian kedua hal itu, penyusunan instrument haruslah mendasarkan diri pada kisi-kisi yang sengaja disiapkan untuk tujuan itu. Pada kisi-kisi itu paling tidak harus terdapat aspek tujuan (umum), deskripsi bahan, indicator, dan jumlah pertanyaan per indicator. Sebelum kisi-kisi dijadikan pedoman penyusunan butir-butir soal instrument penelitian, harus dilakukan telaah terlebih dahulu dan dinyatakan baik. Setelah butir-butir pertanyaan disusun, butir tersebut juga harus ditelaah dengan menggunakan kriteria tertentu di samping disesuaikan dengan kisi-kisi. Penelaah harus dilakukan oleh orang yang berkompeten di bidang yang bersangkutan (*expert judgement*).

Validitas isi merupakan validitas yang diestimasi lewat pengujian terhadap isi tes dengan analisis rasional atau lewat *professional judgement*. Pertanyaan yang dicari jawabannya dalam validitas ini adalah: "sejauh mana *item-item* dalam tes mencakup keseluruhan isi objek yang hendak diukur atau sejauh mana isi tes mencerminkan ciri atribut yang hendak diukur". Pengertian "mencakup keseluruhan isi" tidak saja menunjukkan bahwa tes tersebut harus komprehensif/ menyeluruh isinya tetapi harus pula hanya isi yang relevan dan tidak keluar dari batasan tujuan ukur. Walaupun isinya komprehensif, tetapi apabila suatu tes mengikutsertakan pula *item-item* yang tidak relevan dan berkaitan dengan hal-hal di luar tujuan ukurnya, validitas tes tersebut tidaklah dapat dikatakan memenuhi ciri validitas yang sesungguhnya.

Pengujian validitas isi tidak melalui analisis statistika tetapi menggunakan analisis rasional. Salah satu cara yang

praktis untuk melihat apakah validitas isi telah terpenuhi adalah dengan melihat dan membandingkan apakah *item-item* dalam tes tersebut telah ditulis sesuai dengan kisikisirinya, yaitu kesesuaian dengan batasan domain ukur yang telah ditetapkan sebelumnya dan memeriksa apakah tiap-tiap butir tes telah sesuai dengan indikator perilaku yang hendak diungkapkannya melalui tes tersebut. Misalnya: apabila seorang guru pendidikan jasmani ingin mengetes pengetahuan siswa tentang peraturan permainan bola voli, guru penjas tersebut dapat menyusun tes pengetahuan dan perlu memeriksa kembali cakupan bahan yang menjadi sumbernya, misalnya buku-buku ajar yang relevan. Rancangan soal merupakan cuplikan yang dapat mewakili cakupan seluruh bahan. Apabila hal itu terpenuhi, tes pengetahuan itu disebut memiliki validitas isi. Maksudnya, tes itu benar-benar dimaksudkan dan berfungsi untuk mengukur materi pelajaran yang mesti dikuasai oleh para siswa.

Untuk tes keterampilan/kecakapan dalam pendidikan jasmani, apabila guru pendidikan jasmani ingin mengetahui tingkat kecakapan/keterampilan bermain tenis lapangan siswa-siswanya, sudah sewajarnya apabila salah satu *item* tesnya berupa tes *goundstroke forehand* maupun *goundsroke backhand* karena indikator perilaku yang dominan dari bermain tenis adalah *goundstroke forehand* dan *goundsroke backhand*. Apabila seorang guru pendidikan jasmani ingin mengukur tingkat kecakapan bermain bola basket bagi siswa-siswanya, salah satu butir tesnya berupa tes kecakapan memasukkan bola ke dalam keranjang/*shooting* karena unjuk kerja yang paling dominan dari permainan bola basket adalah kemampuan *shooting*.

Penentuan validitas isi dilakukan berdasarkan proses analisis yang logis. Dengan menentukan butir tes secara

cermat dalam kaitannya dengan tujuan pengajaran, perlu dipertimbangkan tiga hal sebagai syarat validitas isi, yaitu: (1) apakah isi tes paralel dengan tujuan kurikulum, baik dalam hal isi maupun proses, (2) apakah terdapat kesinambungan antara tes dan tekanan dari kurikulum, (3) apakah tes bebas dari suatu keharusan yang tidak relevan (Rusli Lutun, dkk., 2000: 52).

Karena validitas isi didasarkan pada keputusan subjektif, penentuan validitas tipe ini lebih mudah terjadi kesalahan apabila dibandingkan dengan validitas tipe lainnya. Estimasi validitas isi tidak melibatkan perhitungan statistik apa pun melainkan hanya didasarkan pada analisis rasional, maka tidak diharapkan setiap orang akan sama sependapat mengenai seberapa jauh validitas isi suatu tes telah tercapai. Meskipun demikian, pada umumnya validitas isi mendapat perhatian pertama dalam pengembangan semua jenis tes, dan butir-butir tesnya ditulis disesuaikan dengan persyaratan-persyaratan validitas isi.

Validitas isi atau validitas muatan dibagi menjadi dua tipe, yaitu: (1) validitas muka atau *face validity*, dan (2) validitas logika atau *logical validity*. Validitas muka adalah tipe validitas yang paling rendah signifikansinya karena hanya didasarkan pada penilaian terhadap format penampilan (*appearance*) tes. Apabila penampilan tes telah meyakinkan dan memberikan kesan mampu mengungkapkan apa yang hendak diukur dapat dikatakan bahwa validitas muka telah terpenuhi.

Secara teoretik memang validitas muka ini kurang sistematis dalam hal analisis secara logika, namun tes yang memiliki validitas muka yang tinggi (tampak meyakinkan) akan memancing motivasi siswa yang dites untuk menghadapi tes tersebut dengan bersungguh-sungguh. Motivasi ini merupakan aspek penting dalam setiap prosedur pengujian.

Sebaliknya, tes yang tampilannya tidak meyakinkan karena dicetak di kertas murahan (misalnya: tes teori pendidikan jasmani/angket yang dicetak di kertas buram), tentu kurang mendapatkan apresiasi dan respek oleh calon responden apabila dibandingkan jika tes/angket tersebut dicetak pada kertas HVS yang putih. Akibatnya tentu dalam menjawab pun responden cenderung akan asal-asalan sehingga data yang diperoleh kurang valid.

Dengan alasan kepraktisan, banyak tes yang pemakaiannya terbatas, seperti tes prestasi belajar di kelas hanya mengandalkan validitas muka. Tes psikologi, baik yang mengukur aspek abilitas kognitif maupun yang mengukur aspek afektif kepribadian, tidak dapat menyandarkan kualitasnya hanya pada validitas muka. Pada tes yang fungsi pengukurannya bersifat menentukan, seperti tes untuk seleksi atau untuk diagnosis kepribadian, dituntut adanya bukti validitas yang mempunyai landasan yang lebih kuat.

Validitas logika disebut juga sebagai validitas sampling (*sampling validity*). Validitas logika diartikan seberapa jauh suatu tes mengukur komponen-komponen keahlian yang paling penting untuk melaksanakan tugas motorik secara memadai (Safrit, 1986: 124). Prosedur umum yang harus diikuti dalam penyusunan tes keterampilan dalam pendidikan jasmani agar tes tersebut memiliki validitas logika yang baik menurut Safrit (1986: 124) adalah: (1) dengan menentukan secara rinci poin-poin pelaksanaan gerak yang paling bagus, (2) menyusun sebuah tes yang mengukur komponen-komponen keahlian yang penting yang hendak diukur sesuai dengan poin-poin yang telah ditentukan, dan (3) memberikan skor tes sesuai dengan gradasi unjuk kerja yang ditampilkan siswa, sehingga skor terbaik menunjukkan suatu performa yang bagus demikian juga sebaliknya skor jelek mengindikasikan suatu performansi yang tidak bagus.

Contoh: penilaian sebuah pukulan *forehand* dalam permainan tenis lapangan. Guru pendidikan jasmani menjelaskan bahwa suatu pukulan *forehand* yang bagus adalah bola melambung rendah di atas net dan masuk ke dalam lapangan belakang (*baseline*). Guru pendidikan jasmani tersebut memutuskan untuk mengetes siswa-siswanya tentang keahlian ini dengan menggunakan tes keterampilan bermain tenis lapangan, targetnya ditempatkan di lapangan belakang (*baseline*) dan bola harus dipukul melewati seutas tali yang direntangkan pada kedua tiang net. Target tersebut telah dibagi-bagi sehingga daerah-daerah dengan skor yang lebih tinggi lebih dekat dengan garis belakang (*baseline*). Dengan prosedur penilaian yang demikian, validitas logika tes keterampilan bermain tenis tersebut dapat terjamin.

Untuk memperoleh validitas logika yang tinggi, suatu tes harus dirancang sedemikian rupa sehingga benar-benar hanya berisi *item* yang relevan dan perlu menjadi bagian tes secara keseluruhan. Suatu objek ukur yang hendak diungkap oleh tes haruslah dibatasi lebih dahulu kawasan perilakunya secara saksama dan konkret. Batasan perilaku yang kurang jelas akan menyebabkan terikutnya *item-item* yang tidak relevan dan tertinggalnya bagian penting dari objek ukur yang seharusnya masuk sebagai bagian dari tes yang bersangkutan. Pada bentuk-bentuk tes non kognitif, yang mengungkap performa atau atribut kepribadian, pembatasan perilaku itu biasanya tidak mudah untuk dilakukan dan tidak dapat dibuat dengan tegas sekali.

Pada bentuk tes kognitif, yang mengukur performansi maksimal, batasan itu dapat dibuat lebih konkret. Sebagai contoh: dalam pengembangan tes prestasi belajar, kawasan perilaku (tingkatan proses berpikir) yang hendak diukur dapat dikembalikan pada Tujuan Instruksional Umum (TIU)

dan Tujuan Instruksional Khusus (TIK). Dengan bantuan silabus pelajaran, cakupan tes yang akan dibuat pun menjadi lebih jelas.

Validitas logika memang sangat penting peranannya dalam penyusunan tes prestasi (tes teori dan tes praktik) dalam pendidikan jasmani. Dalam perencanaan tes prestasi, penyusun tes memanfaatkan suatu *blue print* yang memuat cakupan isi dan cakupan kompetensi yang hendak diungkap. Penggunaan *blue print* semacam itu sangat membantu tercapainya validitas logika yang sangat baik. Berikut ini disajikan contoh pembuatan kisi-kisi (*blue print*) dalam permainan bola voli.

Tabel 1: Kisi-kisi (*blue print*) teori permainan bola voli

Kemampuan Materi	Pengetahuan 30 %	Pemahaman 60 %	Aplikasi 10 %	Jumlah Soal
a. Keselamatan (5 %)				
b. Teknik (30 %)				
c. Strategi (20 %)				
d. Peraturan Permainan (40 %)				
e. Perawatan (5 %)				

Catatan:

Berkenaan dengan besarnya persentase (%) untuk cakupan materi/ isi dan cakupan kompetensi yang ingin diukur, disesuaikan dengan tingkat kemampuan siswa yang dihadapi.

2. Validitas Konstruk (*Construct Validity*)

Validitas konstruk adalah tipe validitas yang menunjukkan derajat seberapa jauh tes mengungkap suatu *trait* atau atribut tertentu dari kemampuan aded, tetapi secara keseluruhan tidak dapat dijelaskan dengan pengukuran yang

sempurna (Safrit, 1986: 129). Sebagaimana terlukis pada bagan tipe-tipe validitas seperti yang diutarakan oleh Margaret J. Safrit, untuk pengujian validitas konstruk diperlukan analisis statistika yang cukup kompleks seperti: prosedur analisis faktor, analisis regresi majemuk, dan pengujian makna perbedaan skor rata-rata dua kelompok yang ekstrim.

Salah satu prosedur pengujian validitas konstruk yang lebih sederhana adalah dengan pendekatan *multi-trait multi-method* yang dikembangkan oleh Campbell dan Fiske (Allen dan Yen, 1979: 102). Metode ini digunakan apabila dua atau lebih *trait* atau atribut yang hendak diukur dengan dua atau lebih metode. Prosedur ini akan menghasilkan adanya bukti validitas diskriminan dan validitas konvergen. Validitas diskriminan diperlihatkan oleh rendahnya korelasi antara skor tes yang mengukur *trait* yang berbeda terutama apabila digunakan metode yang sama. Di sisi lain validitas konvergen diperlihatkan oleh tingginya korelasi antara skor tes-skor tes yang mengukur *trait* yang sama dengan menggunakan metode yang berbeda.

Konsep validitas konstruk sangat berguna pada tes yang mengukur *trait* atau atribut yang tidak memiliki kriteria eksternal. Saefuddin Azwar (1997: 48) menjelaskan bahwa apabila seseorang ingin melakukan *rating* terhadap sifat agresivitas yang tampak dari seorang karateka, seseorang akan melakukannya setelah melaksanakan observasi terhadap perilaku yang menjadi sasaran penelitian untuk beberapa lamanya. Biasanya, *rating* terhadap perilaku sedemikian ini dapat menjadi indikator yang valid bagi ada-tidaknya sifat agresivitas. Apabila seseorang ingin mengungkapkan atau mengukur adanya agresivitas yang laten (yang belum tampak manifestasinya dalam perilaku karateka tersebut) seseorang dapat melakukannya lewat suatu tes proyektif, tidak lewat observasi.

Permasalahannya adalah, bagaimana seseorang dapat menguji validitas pengungkapan agresivitas laten dari seorang karateka itu? Hal itu dapat diawali dari suatu batasan mengenai variabel yang hendak diukur, yang dalam contoh ini adalah agresivitas seorang karateka, kemudian batasan variabel itu dinyatakan sebagai suatu bentuk konstruk logis menurut konsep-konsep agresivitas yang didasari oleh suatu teori. Dari teori itu kemudian ditarik semacam konsekuensi praktis mengenai hasil tes pada kondisi tertentu. Konsekuensi inilah yang kemudian diuji. Apabila hasilnya sesuai dengan harapan, tes itu dianggap memiliki validitas konstruk yang baik.

Contoh lain berkenaan dengan konsep validitas konstruk adalah masalah **kecemasan** (*anxiety*). Kecemasan sering kali merupakan faktor yang menarik dalam diri atlet. Mengukur banyaknya keringat di telapak tangan, mencatat berapa kali mengalami kecemasan, dan memeriksa denyut jantung dan tekanan darah adalah cara-cara untuk menentukan apakah seseorang tersebut cemas atau tidak. Meskipun demikian, ciri-ciri itu hanyalah indikator-indikator dari kecemasan. Kecemasan sesungguhnya tidak dapat diukur secara langsung. Meskipun sebuah konsep **kecemasan** tidak dapat diukur secara tepat, indikator-indikator perilaku tersebut dapat dipakai untuk menilai tingkat kecemasan seorang atlet.

3. Validitas Berdasarkan Kriteria (*Criterion-Related Validity*)

Validitas berdasarkan kriteria menunjukkan bahwa skor tes secara sistematis berhubungan dengan satu atau lebih kriteria hasil. Menurut Fernandes (1984: 87) validitas kriteria menjawab pertanyaan: "*How well test performance predicts future performance (predictive validity) or estimate current performance on some valued measure other than the test it self (concurrent validity)*".

Prosedur pendekatan validitas berdasarkan kriteria menghendaki tersedianya kriteria eksternal yang dapat dijadikan dasar pengujian skor tes. Suatu kriteria adalah variabel perilaku yang akan diprediksikan oleh skor tes atau berupa suatu ukuran lain yang relevan.

Validitas yang berkaitan dengan kriteria ditunjukkan dengan membandingkan skor-skor tes dengan satu atau lebih variabel eksternal yang dianggap mengarahkan penilaian karakteristik atau perilaku yang ditanyakan. Menurut Safrit (1986: 126) ada dua tipe validitas yang berkaitan dengan kriteria, yaitu: (1) validitas yang terjadi bersamaan atau *concurrent validity*, yang meliputi perbandingan antara penilaian tes yang disusun dengan penilaian tes kriteria, (2) validitas prediktif, yang menyangkut prediksi/ramalan perilaku pada tes kriteria.

Validitas yang berkaitan dengan kriteria ditentukan dengan metode statistika, sedangkan validitas isi/muatan dan validitas logika berperan sebagian dalam tahapan awal penyusunan tes hasil belajar. Validitas yang berkaitan dengan kriteria digunakan apabila sebuah tes diajukan sebagai pengganti tes lain yang dianggap valid. Contoh: mengapa seorang penyusun tes ingin mencari tes pengganti apabila suatu tes yang valid telah tersedia, meskipun validitas tes kriteria telah ditetapkan tes itu tidak bisa dijalankan. Hal ini dikarenakan oleh: (1) tes yang ada memerlukan waktu yang sangat banyak, (2) memerlukan tenaga yang terlatih dengan peralatan yang mahal, (3) tes tersebut merupakan tes laboratorium yang tidak bisa dilakukan karena situasi dan kondisi sekolah yang tidak memadai.

Contoh: telah diketahui bahwa pengujian seseorang dengan tes di *treadmill* memberikan indikator-indikator fungsi kardiorespirasi yang lebih valid daripada tes di lapangan, akan

tetapi tes di *treadmill* diperlukan pakar fisiologi, peralatan laboratorium yang mahal, dan pelaksanaan tesnya satu demi satu sehingga kurang efisien. Oleh karena itu, dalam konteks pendidikan jasmani, tes-tes lapangan yang lebih murah dan praktis untuk mengukur fungsi kardiorespirasi banyak dipergunakan, seperti: *aerobic cooper* dan *multistage test*.

Prosedur validasi tes berdasarkan kriteria menghasilkan dua macam validitas, yaitu: (1) validitas konkuren (*concurrent validity*), dan (2) validitas ramalan (*predictive validity*).

a. Validitas Konkuren (*Concurrent Validity*)

Validitas konkuren adalah tingkat kesesuaian hasil pengukuran dari dua alat pengukur yang dilakukan dalam waktu yang bersamaan atau hampir bersamaan. Validitas tipe ini pengujian validitas tesnya dilakukan berdasarkan data empirik. Disebut demikian, karena yang menjadi kriteria bagi tes yang akan disusun adalah tes yang telah ada yang telah terbukti/diakui validitasnya.

Beberapa kriteria yang biasanya sering dipakai dalam penyusunan tes pendidikan jasmani/olahraga menurut Rusli Lutan dkk., (2000: 54) di antaranya adalah:

1) Pendapat Para Ahli

Sekelompok pakar yang ahli dalam bidangnya menilai secara subjektif kemampuan siswa yang bersangkutan dan kemudian hasilnya dikorelasikan dengan skor yang diperoleh dari pengukuran dengan menggunakan tes yang baru disusun. Koefisien korelasinya menunjukkan seberapa besar koefisien validitas konkuren.

2) Hasil pertandingan dalam cabang atau nomor yang bersangkutan,

Skor yang diperoleh dari tes yang baru disusun dikorelasikan dengan kedudukan dalam turnamen. Misalnya skor dari tes bulutangkis yang sedang disusun

dikorelasikan dengan urutan kedudukan hasil turnamen setengah kompetisi. Koefisien korelasinya menunjukkan seberapa besar koefisien validitas konkuren.

3) Skor tes lama yang telah dibakukan dan valid.

Skor tes lama yang baku dan terbukti valid dikorelasikan dengan skor tes yang baru disusun. Misalnya tes braddy volley yang valid dikorelasikan dengan tes keterampilan bola voli yang sedang disusun. Koefisien korelasinya menunjukkan seberapa besar koefisien validitas konkuren.

b. Validitas Ramalan (*Predictive validity*)

Validitas prediktif sangat penting artinya apabila tes dimaksudkan untuk berfungsi sebagai prediktor bagi performa di waktu yang akan datang. Validitas prediktif diartikan sebagai derajat seberapa jauh perilaku atau sifat kriteria dapat diramalkan dengan menggunakan skor tes dari tes prediktor. Dalam pendidikan jasmani/olahraga kegunaan validitas prediktif menurut Safrit (1986: 127) adalah: (1) sebagai suatu deskripsi dari status individu saat ini atas sifat kriteria, (2) sebagai prediktor sifat kriteria di waktu mendatang, yaitu kriteria yang mengacu pada standar kegunaan tes yang telah ditentukan. Prediktor-prediktor itu adalah tes-tes atau variabel-variabel yang memprediksi sifat kriteria.

Penerapan validitas prediktif dalam meramalkan perilaku atau sifat di masa mendatang memakai satu atau lebih variabel prediktor untuk memprediksi sifat kriteria pada beberapa hal di masa mendatang. Dalam konteks ini, skor-skor atas variabel prediktor dan kriteria tidak didapat secara bersamaan. Meskipun pada awalnya nampak ganjil, ini adalah satu-satunya yang logis untuk dilakukan. Jika sifat di masa mendatang diprediksi, fokus perhatiannya adalah pada

penampilan seseorang di masa mendatang bukan penampilan pada saat ini.

Suatu tes memiliki kualitas validitas prediktif apabila tes itu mampu menggambarkan seberapa cocok prediksi hasil yang nyata. Misalnya: soal-soal ujian masuk perguruan tinggi (misalnya ke FIK) akan disebut memiliki validitas prediktif yang baik apabila nilai hasil ujian masuk mahasiswa yang bersangkutan memiliki korelasi yang tinggi dengan nilai hasil-hasil perkuliahan di FIK.

Dengan kata lain, sampai seberapa jauh skor tes itu berkaitan dengan keberhasilan seseorang di perguruan tinggi menggambarkan "seberapa jauh kemampuan tes itu untuk memenuhi fungsinya guna meramalkan prestasi belajar seseorang". Kecermatan prediksi itu biasanya tergambar dari besarnya koefisien korelasi antara skor tes dan kriteria. Tabel berikut akan menyajikan ilustrasi pengujian validitas prediktif.

Tabel 2. Ilustrasi pengujian validitas prediktif

No.	Skor Tes Ujian Masuk PT (ke FIK)	Hasil Nilai Rata-Rata Perkuliahan Di FIK
1.	114	9
2.	109	9
3.	100	7
4.	99	4
5.	113	10
6.	107	8
7.	107	7
8.	100	6
9.	103	5
10.	110	8

Apabila korelasi antara distribusi skor tes ujian masuk PT (ke FIK) dan hasil nilai rata-rata perkuliahan di FIK dalam tabel tersebut dihitung, diperoleh koefisien korelasi $r_{xy} = 0,864$. Koefisien korelasi inilah yang merupakan koefisien validitas prediktif tes ujian masuk PT (ke FIK) selaku prediktor keberhasilan perkuliahan di FIK. Dalam contoh ini koefisien validitas tes ujian masuk PT sangat tinggi, artinya tes tersebut dapat dikatakan valid sebagai prediktor tingkat pencapaian nilai perkuliahan di FIK.

C. Teknik Pengujian Validitas Tes Penjasorkes

Analisis terhadap validitas tes hasil belajar pendidikan jasmani/olahraga sebagai suatu totalitas pada dasarnya dapat dilakukan dengan dua cara. **Pertama**, analisis yang dilakukan dengan jalan berpikir secara rasional atau analisis dengan menggunakan logika (*logical analysis*). **Kedua**, analisis yang dilakukan dengan mendasarkan diri pada kenyataan empiris (*empirical analysis*).

1. Pendekatan Berdasarkan Pertimbangan Rasional

Pendekatan ini dilakukan terhadap semua aspek yang terintegrasi dalam suatu masalah yang diukur. Pada umumnya, tes hasil belajar setelah dilakukan analisis rasional ternyata memiliki ketepatan mengukur atau dengan kata lain tes hasil belajar tersebut telah memiliki validitas logika atau validitas *das sollen* yang cukup memadai (Anas Sudijono, 2005: 164). Lebih lanjut Anas Sudijono menyatakan bahwa validitas rasional adalah validitas yang diperoleh atas dasar hasil pemikiran atau validitas tes hasil belajar yang diperoleh dengan berpikir secara logis. Dengan demikian, suatu tes hasil belajar dapat dikatakan telah memiliki validitas rasional apabila setelah dilakukan penganalisisan secara rasional

ternyata tes hasil belajar itu memang dengan tepat telah dapat mengukur apa yang seharusnya diukur.

Untuk dapat menentukan apakah tes hasil belajar sudah memiliki validitas rasional atau belum, dapat dilakukan penelusuran dari dua segi, yaitu dari segi isinya (*content*) dan dari segi susunan atau konstruksinya (*construct*). Oleh karena itu validitas isi (*content validity*) dan validitas konstruk (*construct validity*) ditentukan berdasarkan pendekatan analisis rasional ini.

2. Pendekatan Berdasarkan Empirik-Statistik

Pendekatan empirik-statistik ini dilakukan untuk mengetahui hubungan yang ada antara alat ukur yang sedang dipelajari (prediktor) dan pengukuran lain atau kenyataan-kenyataan lain yang sudah dibakukan (kriterium). Pendekatan ini tingkat validitasnya diperoleh atas dasar pengamatan di lapangan (bukti empirik).

Beritik tolak dari konsep ini, maka tes hasil belajar dapat dikatakan telah memiliki validitas empirik apabila berdasarkan hasil analisis yang dilakukan terhadap data hasil pengamatan di lapangan, terbukti bahwa tes hasil belajar itu dengan secara tepat telah dapat mengukur hasil belajar yang seharusnya diungkap atau diukur lewat tes hasil belajar tersebut. Untuk dapat menentukan apakah tes hasil belajar sudah memiliki validitas empirik ataukah belum, dapat dilakukan penelusuran dari dua segi, yaitu dari segi daya ketepatan bandingannya (*concurrent validity*) dan daya ketepatan meramalnya (*predictive validity*).

D. Cara Mengetahui Validitas Alat Ukur

1. Cara Mengetahui Validitas Dengan Kriteria Expert Judgement

Sebuah tes dikatakan memiliki validitas baik jika hasilnya sesuai dengan kriteria, dalam arti memiliki kesejajaran antara hasil tes yang sedang disusun dengan kriteria. Teknik yang dapat digunakan untuk mengetahui kesejajaran adalah teknik korelasi *Pearson Product Moment*. Rumusnya yaitu sebagai berikut.

$$r_{xy} = \frac{N \cdot \sum XY - (\sum X) (\sum Y)}{\sqrt{\{N (\sum X^2) - (\sum X)^2\} \{N (\sum Y^2) - (\sum Y)^2\}}}$$

r_{xy} = koefisien korelasi antara variabel X dan variabel Y, dua variabel yang dikorelasikan.

Contoh Soal

Seorang guru penjas sedang menyusun tes keterampilan sepakbola. Untuk menguji validitasnya, guru tersebut menggunakan kriteria penilaian ahli (*expert judgement*). Uji coba tes dilakukan pada 10 siswa dengan data sebagai berikut.

Subjek	X	Y
1	9	8
2	8	7
3	7	6
4	8	7
5	6	6
6	5	6
7	8	7
8	9	8
9	8	7
10	7	6

X = data skor tes keterampilan sepakbola yang sedang disusun

Y = nilai dari *expert judgement*

Pertanyaan!

- Hitunglah koefisien korelasi antara tes keterampilan sepakbola yang sedang disusun dengan nilai dari *expert judgement* (r_{xy})!
- Bagaimana kesimpulan mengenai tes yang sedang disusun tersebut?

Jawab

Subjek	X	Y	X ²	Y ²	X.Y
1	9	8	81	64	72
2	8	7	64	49	56
3	7	6	49	36	42
4	8	7	64	49	56
5	6	6	36	36	36
6	5	6	25	36	30
7	8	7	64	49	56
8	9	8	81	64	72
9	8	7	64	49	56
10	7	6	49	36	42
Jumlah	75	68	577	458	518

Dimasukkan ke dalam rumus

$$r_{xy} = \frac{N \cdot \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N(\sum X^2) - (\sum X)^2\} \{N(\sum Y^2) - (\sum Y)^2\}}}$$
$$r_{xy} = \frac{10 \cdot 518 - 75 \cdot 68}{\sqrt{\{10 \cdot 577 - 5625\} \{10 \cdot 468 - 4624\}}}$$

$$r_{xy} = \frac{5180 - 5100}{\sqrt{\{5770 - 5625\} \{4680 - 4624\}}}$$

$$r_{xy} = \frac{80}{\sqrt{145.56}}$$

$$r_{xy} = \frac{80}{\sqrt{8120}}$$

$$r_{xy} = 0,88779$$

- Koefisien korelasi antara tes keterampilan sepakbola yang sedang disusun dengan nilai dari expert judgement = 0,88779.
- Kesimpulan: tes keterampilan sepakbola yang sedang disusun memiliki validitas *excellent* (sangat bagus sekali).

Keterangan:

Tabel. Standard untuk Menginterpretasikan Koefisien Korelasi Validitas

Koefisien Korelasi	Validitas
0.85 - 0.99	Excellent (sangat bagus sekali)
0.80 - 0.84	Very good (sangat bagus)
0.70 - 0.79	Acceptable (dapat diterima)
0.60 - 0.69	Questionable (dipertanyakan)

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

1. Korelasi positif (+) menunjukkan adanya hubungan sejajar atau linear antara dua hal. Misalnya hal pertama nilainya naik maka hal kedua ikut naik. Sebaliknya jika hal pertama turun, yang kedua ikut turun.
2. Korelasi negatif (-) menunjukkan adanya hubungan kebalikan antara dua hal. Misalnya, hal pertama nilainya naik, justru yang kedua turun. Sebaliknya jika yang pertama turun, yang kedua naik.
3. Koefisien korelasi selalu terdapat antara -1,00 sampai +1,00.
4. Penafsiran harga koefisien korelasi ada 2 cara yaitu:
 - a. Dengan melihat harga r dan diinterpretasikan misalnya korelasi bagus, dapat diterima, dan sebagainya.
 - b. Dengan berkonsultasi ke tabel harga kritik r *product moment* sehingga dapat diketahui signifikan tidaknya korelasi tersebut. Jika harga r lebih kecil dari harga kritik dalam tabel, maka korelasi tersebut tidak signifikan. Begitu juga arti sebaliknya.

Cara Menghitung Menggunakan Kalkulator fx-350 MS

- a. Bersihkan memori kalkulator dengan cara tekan **on**, tekan **shift**, tekan **mode**, pilih 3 (**All**), akan muncul tulisan **Reset All** di layar, lalu ketik = diikuti **AC**.
- b. Masukkan data. Tekan **mode**, pilih 3 (**Reg**), pilih 1 (**Lin**).
 - Masukkan data pertama, Tekan 9 (**X1**), Tekan tanda (.) yang ada di samping **M+**, Tekan 8 (**Y1**), Tekan **M+**, maka akan muncul $n=1$ di layar.
 - Masukkan data kedua, Tekan 8 (**X2**), Tekan tanda (.) yang ada di samping **M+**, Tekan 7 (**Y2**), Tekan **M+**, maka akan muncul $n=2$ di layar.
 - Masukan terus sampai data ke-10
 - Jika ada data yang salah saat memasukkan maka tekan panah atas/bawah pada tombol **replay** sampai ketemu data tersebut. Masukkan data yang benar, lalu ketik tanda = maka data yang baru akan tersimpan.
- c. Tekan **shift**, lalu angka 2 (**S-VAR**), geser panah ke kanan 2 kali sampai muncul tanda r . Tekan angka 3 (r), tekan = maka akan muncul koefisien korelasi antara X dan Y yaitu 0,88779.
- d. Setiap akan menghitung koefisien korelasi yang baru, bersihkan data dan mode kalkulator dengan cara pada poin "a". Jika hal tersebut tidak dilakukan maka data sebelumnya masih tersimpan.

2. Cara Mengetahui Validitas Dengan Kriteria Tes Terstandar

Tes terstandar adalah tes yang telah dicobakan berkali-kali sehingga dapat dijamin kebaikannya. Tes terstandar biasanya sudah memiliki validitas, reliabilitas, daya pembeda, taraf kesukaran pada kategori baik. Cara menentukan validitas soal yang menggunakan tes terstandar sebagai kriteria

dilakukan dengan mengalikan koefisien validitas yang diperoleh dengan koefisien validitas tes terstandar tersebut.

Contoh:

Diketahui validitas tes terstandar bola voli adalah 0,89. Carilah koefisien tes keterampilan bola voli yang sedang disusun dengan data berikut.

Tabel Persiapan Perhitungan Validitas Tes Keterampilan Bola Voli Dengan Kriteria Tes Terstandar Bola Voli

No.	Nama	X	Y	X ²	Y ²	X.Y
1	Budi	5	7	25	49	35
2	Andi	6	6	36	36	36
3	Nisa	5	6	25	36	30
4	Yanti	6	7	36	49	42
5	Rena	7	7	49	49	49
6	Dina	6	5	36	25	30

X= hasil tes keterampilan bola voli yang sedang disusun

Y= hasil tes standar

$$r_{xy} = \frac{N \cdot \sum XY - (\sum X)(\sum Y)}{\sqrt{\{N(\sum X^2) - (\sum X)^2\} \{N(\sum Y^2) - (\sum Y)^2\}}}$$

$$r_{xy} = \frac{6 \cdot 22 - 35 \cdot 38}{\sqrt{\{6 \cdot 207 - 35^2\} \{6 \cdot 244 - 38^2\}}}$$

$$r_{xy} = \frac{132 - 1330}{\sqrt{\{1242 - 1225\} \{1464 - 1444\}}}$$

$$r_{xy} = \frac{2}{\sqrt{340}} = 0,108$$

Jika dari tes terstandar diketahui validitasnya 0,89, maka bilangan 0,108 ini belum merupakan validitas tes keterampilan bola voli yang dicari. Validitas itu harus dikalikan dengan cara $0,89 \times 0,108 = 0,096$, yang maknanya koefisien validitas tes keterampilan bola voli yang sedang disusun sangat rendah sekali sehingga diperlukan revisi.

3. Cara Menghitung Validitas Butir Soal Atau Validitas Item

Untuk tes yang terdiri dari banyak item, misal tes yang terdiri dari 10 soal, 5 item, dan sebagainya, maka selain validitas soal secara umum atau keseluruhan, perlu juga dicari validitas butir soal atau validitas item. Dengan menghitung validitas item maka akan diketahui item soal mana yang memiliki validitas tinggi dan mana yang memiliki validitas rendah. Untuk soal yang memiliki validitas rendah selanjutnya diperlukan revisi.

Contoh:

Dari ujicoba tes pengetahuan penjas yang sedang disusun diperoleh data sebagai berikut.

Hitunglah validitas item no. 6

No.	Nama	Butir soal/item										Skor total
		1	2	3	4	5	6	7	8	9	10	
1	Indah	1	0	1	0	1	1	1	1	1	1	8
2	Budi	0	0	1	0	1	0	0	1	1	1	5
3	Yoyok	0	1	0	0	0	1	0	1	0	1	4
4	Lina	1	1	0	0	1	1	0	0	1	0	5
5	Reni	1	1	1	1	1	1	0	0	0	0	6
6	Diah	1	0	1	0	1	0	1	0	0	0	4
7	Lintang	1	1	1	1	1	1	1	0	0	0	7
8	Bona	0	1	0	1	1	1	1	1	1	1	8

Jawab

Tabel Persiapan Untuk Menghitung Validitas Item Nomor 6

No.	Nama	X	Y
1	Indah	1	8
2	Budi	0	5
3	Yoyok	1	4
4	Lina	1	5
5	Reni	1	6
6	Diah	0	4
7	Lintang	1	7
8	Bona	1	8

Keterangan:

X = skor item nomor 6

Y = skor total

Data di atas dimasukkan ke dalam rumus korelasi *Pearson Product Moment* sehingga diperoleh hasil koefisien validitas item nomor 6 yaitu 0,421. Kesimpulannya, validitas item nomor 6 rendah sehingga perlu direvisi.

4. Cara Menghitung Validitas Faktor

Selain validitas soal secara keseluruhan dan validitas butir/item, masih ada lagi yang perlu diketahui validitasnya yaitu faktor-faktor atau bagian keseluruhan materi. Setiap keseluruhan materi pelajaran terdiri dari pokok-pokok bahasan atau sekelompok pokok bahasan yang merupakan satu kesatuan.

Contoh:

Guru akan mengevaluasi penguasaan siswa untuk tiga pokok bahasan yaitu: bola voli, bolabasket, dan sepakbola. Untuk keperluan ini guru membuat 12 butir soal, masing-masing

3 butir untuk tiap pokok bahasan. Carilah validitas faktor 3. Adapun datanya sebagai berikut. Contoh tabel analisis butir untuk menghitung validitas butir dan validitas faktor

Butir Subjek	Faktor 1 (bolavoli)				Skor faktor	Faktor 2 (bolabasket)				Skor faktor	Faktor 3 (sepakbola)				Skor faktor	Skor total
	1	2	3	4		1	5	6	7		8	2	9	10		
Amir	1	0	1	1	3	1	1	0	0	2	1	0	1	1	3	8
Indah	1	0	1	1	3	1	1	1	0	3	1	0	1	1	3	9
Dewi	1	1	1	0	3	1	1	0	1	3	1	1	1	0	3	9
Putra	0	1	0	1	2	0	1	0	1	2	0	1	0	0	1	5
Ina	1	1	0	1	3	0	1	1	1	3	0	1	0	0	1	7
Ani	1	1	0	0	2	0	1	1	1	3	0	1	1	0	2	7
Fida	1	0	1	1	3	1	1	0	1	3	0	1	1	1	3	9
Rido	0	1	1	1	3	1	1	0	0	2	1	0	1	1	3	8
Iman	1	1	0	1	3	0	1	1	0	2	1	1	1	1	4	9
Faisal	1	1	0	1	3	0	1	1	1	3	1	1	0	0	2	8

Jawab

Tabel untuk menghitung validitas faktor 3

Butir Subjek	Skor faktor 3 (X)	Skor total (Y)
Amir	3	8
Indah	3	9
Dewi	3	9
Putra	1	5
Ina	1	7
Ani	2	7
Fida	3	9
Rido	3	8
Iman	4	9
Faisal	2	8

Data di atas dimasukkan ke dalam rumus korelasi *Pearson Product Moment* sehingga diperoleh hasil koefisien validitas faktor nomor 3 yaitu 0,844. Kesimpulannya, validitas faktor tersebut sangat baik.



BAB 6

RELIABILITAS TES DAN METODE ESTIMASINYA

A. Konsep Reliabilitas

Reliabilitas merupakan terjemahan dari kata *reliability* yang mempunyai asal kata *rely* dan *ability*. Pengukuran yang memiliki reliabilitas tinggi disebut sebagai pengukuran yang reliabel. Walaupun reliabilitas mempunyai berbagai nama lain seperti kepercayaan, keterandalan, keajegan, kestabilan, dan konsistensi, namun ide pokok yang terkandung dalam konsep reliabilitas adalah seberapa jauh hasil suatu pengukuran dapat dipercaya. Berkenaan dengan konsep reliabilitas ini, Strand (1993: 10) mengemukakan bahwa reliabilitas adalah tingkat keajegan pengukuran oleh sebuah tes dalam mengukur kualitas yang sama dalam setiap pelaksanaan tes tersebut. Reliabilitas suatu tes dikatakan baik ketika semua orang yang menggunakan melakukan suatu tes mendapatkan hasil yang sama atau relatif sama meskipun dilakukan berulang-ulang.

Linn dan Gronlund (1995: 70) menjelaskan, "(1) *reliability refers to the results obtained with an assessment instrument and not to instrument it self*, (2) *an estimate of reliability always refer to particular type of consistency*, (3) *reliability is a necessary but not sufficient condition for validity*, and (4) *reliability is primarily statistical*."

Dari pendapat di atas dapat ditarik suatu makna bahwa reliabilitas sebuah tes menunjuk pada tingkat keajegan atau konsistensi skor-skor yang "relatif" bebas dari kesalahan-kesalahan. Kecenderungan ini mengarah pada ketetapan yang ditunjukkan dengan memberikan ulangan tes pada subjek yang sama. Gejala atau unsur-unsur yang diungkapkan dalam pengukuran pertama ternyata tidak berubah atau sama pada pengukuran kedua dan seterusnya apabila pengukuran dilakukan dengan menggunakan instrumen yang sama.

Pengertian reliabilitas alat ukur dan reliabilitas hasil ukur biasanya dianggap sama, namun penggunaannya masing-masing perlu diperhatikan. Konsep reliabilitas dalam arti alat ukur erat berkaitan dengan masalah eror pengukuran (*error of measurement*). Kesalahan pengukuran sendiri menunjuk seberapa jauh inkonsistensi hasil pengukuran terjadi apabila pengukuran dilakukan ulang pada kelompok subjek yang sama. Konsep reliabilitas dalam arti reliabilitas hasil ukur erat kaitan dengan *error* dalam pengambilan sampel (*sampling error*) yang mengacu pada inkonsistensi hasil ukur apabila pengukuran dilakukan ulang pada kelompok individu yang berbeda.

B. Sumber Kesalahan Pengukuran Reliabilitas

Dalam konsep reliabilitas terkandung kesalahan pengukuran, sehingga skor yang diperoleh adalah skor yang sebenarnya plus kesalahan atau galat (*error*). Dengan kata lain:

$$\text{Obtain Score} = \text{True Score} + \text{Error Score}$$

Setiap pengukuran dalam pendidikan jasmani harus diusahakan memiliki kesalahan yang sekecil mungkin. Untuk itu perlu ditinjau beberapa sumber kesalahan pengukuran.

Rusli Lutan dkk. (2000: 57) mengatakan bahwa dua macam kesalahan (*error*) yang mempengaruhi *error* skor dan variansi yaitu: kesalahan pengukuran itu sendiri dan kesalahan sistematis. Kesalahan pengukuran (*error measurement*) dimaksudkan sebagai perubahan dalam skor perfoma seseorang disebabkan oleh alat/perlengkapan, *error* dalam skor, dan kesalahan dalam pelaksanaan, termasuk membaca skala (misalnya: hasil lompatan Rifki pada nomor lompat tinggi 145,5 cm terbaca 145 cm atau 146 cm).

Kesalahan sistematis (*systematic error*) adalah perubahan dalam penampilan atau perilaku siswa disebabkan oleh faktor biologis. Seseorang akan menampilkan kemampuan yang berbeda dari hari ke hari atau antara usaha yang pertama dan seterusnya, karena kemampuan seseorang dalam satu bidang bukan sebagai perilaku alamiah. Sebagai contoh, tinggi badan si Wibi mungkin memiliki variasi, yaitu 175 cm, 175,5 cm, dan 176 cm. Variasi tersebut mungkin karena stadiometer sebagai alat untuk mengukur tinggi badan berbeda-beda kecermatannya, dan si pengukur itu sendiri melakukan kesalahan dalam membaca skala atau memasang stadiometer. Yang pertama disebut kesalahan pengukuran (*error measurement*) dan yang kedua disebut kesalahan sistematis (*systematic error*). Dalam tes ketetapan olahraga, misalnya, justru kesalahan sistematis inilah yang lebih sulit untuk dikontrol. Namun sebaliknya, dalam tes tertulis (misalnya, tes pengetahuan dalam olahraga/pendidikan jasmani) kesalahan sistematis ini kurang begitu kuat dibandingkan dengan tes keterampilan dalam suatu cabang olahraga.

Menurut Djemari Mardapi (1996: 10) sumber kesalahan pengukuran terletak pada: alat ukur itu sendiri, pada yang menggunakan, yang diukur, dan lingkungan. Sumber kesalahan pada alat ukur dapat dilihat pada seberapa besar

tingkat validitas dan reliabilitas alat ukur. Sumber kesalahan pada yang menggunakan, bisa dilihat pada cara pelaksanaan pengukuran atau pengujian, misalnya penampilan para pengawas ujian. Penampilan pengawas atau pelaksana ujian mempengaruhi tingkat kecemasan (*anxiety*) yang diukur. Semakin tinggi tingkat kecemasan seseorang akan semakin besar kesalahan pengukuran.

Kesalahan yang bersumber pada yang mengukur bisa bersifat spesifik. Spesifik berarti kesalahan pengukuran bersifat positif untuk semua peserta tes atau negatif untuk semua peserta tes. Kesalahan ini sering disebabkan oleh proses penilaian khususnya tes yang nonobjektif dan tes keterampilan (tes unjuk kerja), seperti tes-tes esai dalam teori pendidikan jasmani maupun proses penjurian dalam beberapa cabang olahraga, senam artistik, loncat indah, dan cabang olahraga tidak terukur lainnya. Ada penilai yang murah dan ada penilai yang mahal. Kesalahan yang spesifik ini adil karena berlaku untuk semua peserta, namun tetap menjadi masalah apabila informasi hasil tes/penilaian digunakan untuk membuat kebijakan, karena hasil pengukuran tidak menunjukkan kemampuan yang sebenarnya.

Sumber kesalahan yang terletak pada yang diukur bisa bersifat acak dan bisa spesifik. Kesalahan acak berarti bahwa kesalahan pengukuran pada setiap orang tidak sama, dari waktu ke waktu berubah secara acak. Sumber kesalahan ini bersumber pada variasi fisik, mental, dan emosi dari waktu ke waktu. Menurut teori tes klasik, kesalahan pengukuran yang bersifat acak pada populasi sama dengan nol. Hal ini berdasarkan asumsi bahwa keadaan fisik, mental, dan emosi populasi yang diukur ada yang bersifat positif dan negatif. Positif berarti kondisi fisik, mental, dan emosi yang diukur baik sekali, sedangkan negatif berarti kondisi fisik, mental,

dan emosi yang diukur tidak baik. Kesalahan yang bersifat spesifik terjadi apabila unjuk kerja seseorang cenderung lebih rendah atau lebih tinggi daripada kemampuan sebenarnya pada saat pengukuran atau ujian. Hal ini dimungkinkan oleh adanya faktor kecemasan pada diri yang diukur atau disebabkan oleh alat ukur yang tidak baik.

Kondisi fisik, mental, dan emosi seseorang mempengaruhi ketepatan dan ketetapan hasil pengukuran. Tidak ada seorang siswa pun yang mampu menampilkan unjuk kerja secara optimal pada waktu yang berbeda-beda. Hal ini disebabkan oleh kondisi fisik, mental, dan emosi seseorang selalu berubah. Kondisi fisik, mental, dan emosi ini mempengaruhi peringkat unjuk kerja seseorang. Menurut Feld dan Brennan (1989: 12) faktor-faktor yang mempengaruhi unjuk kerja siswa, di antaranya adalah: variasi efisiensi fisik dan mental yang diukur, variasi tugas yang dikerjakan, dan inkonsistensi para penilai. Kuantifikasi konsistensi untuk kerja seseorang merupakan dasar analisis reliabilitas.

C. Koefisien Reliabilitas Tes Penjasorkes

Secara empirik, tinggi rendahnya reliabilitas ditunjukkan oleh suatu angka yang disebut dengan koefisien reliabilitas. Koefisien reliabilitas adalah angka yang menunjukkan besarnya statistik korelasi antara skor tes pertama dan skor tes kedua (ulangannya). Pada awalnya, tinggi-rendahnya reliabilitas tes dicerminkan oleh koefisien korelasi antara skor pada dua tes yang paralel, yang dikenakan pada sekelompok individu yang sama. Makin tinggi koefisien korelasi termaksud berarti konsistensi antara hasil pengenaan dua tes tersebut makin baik dan hasil ukur kedua tes itu dikatakan makin reliabel. Sebaliknya apabila dua tes yang

dianggap paralel ternyata menghasilkan skor yang satu sama lain berkorelasi rendah, dapat dikatakan bahwa reliabilitas hasil ukur tersebut tidak tinggi.

Koefisien korelasi antara dua variabel dilambangkan oleh huruf r . Apabila skor pada tes pertama diberi lambing X^1 dan skor pada tes yang kedua yang paralel diberi lambing X^2 , koefisien korelasi antarkedua tes tersebut adalah r_{xx^1} . Simbol inilah yang digunakan sebagai simbol koefisien reliabilitas.

Secara teoretik besarnya koefisien reliabilitas berkisar dari 0,0 sampai dengan 1,0, akan tetapi pada kenyataannya koefisien korelasi sebesar 1,0 dan 0,0 tidak pernah dijumpai. Disamping itu, walaupun koefisien korelasi dapat saja bertanda negatif (-), koefisien reliabilitas selalu mengacu pada angka positif (+) dikarenakan angka negatif tidak memiliki makna bagi interpretasi reliabilitas hasil ukur. Koefisien reliabilitas $r_{xx^1} = 1,0$ berarti ada konsistensi yang sempurna pada hasil ukur yang bersangkutan. Konsistensi yang sempurna seperti ini tidak dapat terjadi dalam pengukuran aspek-aspek psikologi dan sosial yang menggunakan manusia sebagai subjek. Hal ini disebabkan oleh terdapatnya berbagai sumber kesalahan (*error*) dalam diri manusia dan dalam pelaksanaan pengukuran yang sangat mudah mempengaruhi kecermatan hasil pengukuran.

Dalam tes pendidikan jasmani, diharapkan koefisien reliabilitas tes agar bisa mencapai 0,90 atau lebih tidaklah masuk akal. Jenis tes ternyata sangat berpengaruh terhadap besarnya koefisien reliabilitas. Menurut Saifit (1986: 136), jika pelaksanaan tes itu memerlukan usaha maksimum, maka reliabilitas tes 0,85 atau lebih sudah cukup masuk akal. Sebagai contoh, sebuah tes kekuatan menggenggam dilakukan dengan menggunakan *hand grip dynamometer*, ini membutuhkan usaha yang maksimal sudah selayaknya menuntut tingkat koefisien reliabilitas yang

setinggi-tingginya. Hasil tes ini bisa sangat reliabel. Hal ini juga berlaku bagi tes-tes lainnya yang memiliki kemampuan sejenis. Makin besar tuntutan kebutuhan tenaga (melempar sekuat mungkin, melompat sejauh mungkin, menarik *back and leg dynamometer* sekuat mungkin), maka tuntutan tingkat koefisien reliabilitas pun juga semakin tinggi.

Namun sebaliknya, bagi tes-tes dalam pendidikan jasmani yang menuntut tingkat keakuratan (*accuracy*) dalam pelaksanaannya, mengharapkan koefisien reliabilitas 0,85 atau lebih tidaklah masuk akal. Pada kenyataannya, suatu tes yang terutama mengukur tingkat keakuratan, bisa diterima koefisien reliabilitas sebesar 0,70 sudah dianggap memuaskan. Adapun jenis-jenis tes yang masuk dalam kategori ini di antaranya: tes melakukan pukulan *chip* dalam permainan golf, servis pendek (*short service*) dalam bulutangkis, tes memanah dalam cabang olahraga panahan, tes menembak ke dalam ring basket, dan tes *placement serve* dalam tenis. Standard untuk menginterpretasikan koefisien korelasi reliabilitas dapat dilihat pada tabel berikut.

Tabel 2. Standard untuk Menginterpretasikan Koefisien Korelasi Reliabilitas

Koefisien Korelasi	Validitas
0.95 - 0.99	Excellent (sangat bagus sekali)
0.90 - 0.94	Very good (sangat bagus)
0.80 - 0.89	Acceptable (dapat diterima)
0.70 - 0.79	Poor (buruk)
0.60 - 0.79	Questionable (dipertanyakan)

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

D. Metode Estimasi Reliabilitas Tes

Metode estimasi terhadap tingginya reliabilitas dapat dilakukan melalui berbagai metode pendekatan. Tiap-tiap metode pendekatan dikembangkan sesuai dengan sifat dan fungsi alat ukur yang bersangkutan dengan mempertimbangkan pula segi-segi praktisnya. Secara tradisional, menurut prosedur yang dilakukan dan sifat koefisien yang dihasilkannya, ada beberapa cara atau metode yang dapat dipergunakan untuk mengestimasi/menaksir reliabilitas tes pendidikan jasmani. Menurut Phillips (1979: 105) metode-metode tersebut di antaranya: (1) metode tes-tes ulang (*tes-retest*), (2) metode bentuk sejajar (*parallel form*), dan (3) metode belah dua (*split-half*). Secara umum, ketiga metode tersebut akan menghasilkan taksiran koefisien reliabilitas yang berbeda-beda.

1. Metode Tes dan Tes Ulang (*Test-retest*)

Pendekatan tes dan tes ulang dilakukan dengan menyajikan tes dua kali pada satu kelompok subjek dengan tenggang waktu di antara kedua penyajian tersebut. Asumsi yang menjadi dasar dalam cara ini adalah bahwa suatu tes yang reliabel tentu akan menghasilkan skor-tampak yang relatif sama apabila dikenakan dua kali pada waktu yang berbeda. Makin besar perbedaan variasi skor subjek antara pengenaan itu berarti makin sulit untuk mempercayai bahwa tes itu memberikan hasil ukur yang konsisten.

Koefisien reliabilitas dengan cara ini diperoleh dengan menghitung koefisien korelasi linear antara distribusi skor subjek pada pemberian tes yang pertama dan distribusi skor subjek pada pemberian tes yang kedua. Koefisien korelasi yang memperlihatkan keceratan variasi skor antara kedua

pemberian tes itu merupakan koefisien reliabilitas tes yang bersangkutan. Formula yang lazim dipergunakan untuk mencari koefisien reliabilitas tes dan tes ulang adalah teknik korelasi *Pearson Product Moment*. Adapun rumusnya adalah sebagai berikut:

$$r_{xy} = \frac{N \times \sum XY - (\sum X)(\sum Y)}{\sqrt{(N(\sum X^2) - (\sum X)^2)(N(\sum Y^2) - (\sum Y)^2)}}$$

Untuk mendapatkan gambaran lebih jelas dan lengkap, berikut akan disajikan aplikasi penggunaan teknik korelasi *Pearson Product Moment* untuk menghitung besarnya koefisien reliabilitas dengan metode tes dan tes ulang.

Contoh: Hasil pelaksanaan tes dan tes ulang kemampuan melakukan tes lempar lembing dari 10 siswa.

Tabel 2: Hasil analisis koefisien korelasi metode tes dan tes ulang

Subjek	Skor Tes Pertama (X ₁)	Skor Tes Kedua (X ₂)
1	20	22
2	19	20
3	22	22
4	17	18
5	24	24
6	17	16
7	20	21
8	15	17
9	24	23
10	19	19
Koefisien Korelasi (r _{x₁,x₂}) = 0,933		

Dari ilustrasi data di atas, dengan menggunakan teknik korelasi *Pearson Product Moment* diperoleh koefisien reliabilitas tes sebesar 0,933. Koefisien reliabilitas setinggi itu pada umumnya dianggap sebagai indikasi adanya tingkat kestabilan pengukuran dari tes lempar lembing tersebut.

Dalam menggunakan metode tes dan tes ulang ini harus diperhatikan pula kemungkinan adanya perubahan kondisi subjek sejalan dengan berbedanya waktu di antara kedua penyajian tes. Perubahan kondisi subjek yang terjadi tidak pada keseluruhan subjek dan tidak searah sedikit banyak akan ada pengaruhnya terhadap koefisien reliabilitas yang diperoleh. Sebagai contoh, apabila siswa tidak bersungguh-sungguh dalam mengerjakan tes, dalam keadaan lelah, atau memang tidak siap ketika dikenai tes pertama kali, lalu dia belajar kemudian siap atau ia bersungguh-sungguh dalam mengerjakan tes tersebut, skor siswa pada kedua pemberian tes akan banyak berbeda. Kalau hal itu terjadi pada sebagian siswa dan perubahan skor itu tidak searah bagi semua siswa, reliabilitas yang ditunjukkan oleh korelasi antara kedua pemberian tes tidak akan tinggi. Koefisien tersebut bukan merupakan estimasi yang benar terhadap reliabilitas tes, melainkan merupakan estimasi yang lebih rendah daripada semestinya (*under estimate*). Itulah salah satu contoh kasus terjadinya *carry over effect* (efek bawaan) yang sering kali menjadi problem serius dalam metode tes dan tes ulang.

Metode tes dan tes ulang dapat dipergunakan untuk menentukan reliabilitas tes khususnya dapat diaplikasikan dalam tes pendidikan jasmani/tes psikomotor. Jika metode tes dan tes ulang ini dipergunakan dalam tes psikomotor, hendaknya perlu diperhatikan beberapa faktor, di antaranya: faktor kelelahan siswa, urutan pelaksanaan tes terlebih jika tes yang diujikan merupakan rangkaian dari beberapa tes (*battery*

test), dan ketersediaan waktu. Namun sebaliknya, metode ini kurang tepat dipergunakan untuk menentukan reliabilitas tes ranah kognitif, seperti: tes teori dalam pendidikan jasmani, karena siswa akan mengingat soal-soal pada pelaksanaan tes yang pertama (Phillips, 1979: 105).

Istilah koefisien reliabilitas tes dan tes ulang dapat pula disebut koefisien stabilitas. Istilah stabilitas dimaksudkan untuk menunjukkan bahwa seseorang siswa dapat berubah dari waktu ke waktu. Hal itu bukan karena perubahan dalam tes, melainkan fluktuasi dalam beberapa aspek karakteristik yang diukur. Variasi yang terjadi di dalam individu itu sendiri dan variasi antarindividu yang diukur. Reliabilitas *test-retest* diperoleh dengan cara melaksanakan pengtesan dua kali terhadap sekelompok subjek dengan memakai tes yang sama.

Hal yang menjadi persoalan dalam pengujian reliabilitas *test-retest* ini adalah berapa lama selang waktu antara tes pertama dan kedua? Keberatan pendekatan tersebut adalah subjek pada pengtesan kedua mungkin meningkat karena dia masih ingat atau hafal akan tugas-tugas atau soal-soal yang harus dikerjakan. Disamping itu, mungkin selang waktu antara tes pertama dan kedua, subjek yang bersangkutan melakukan latihan. Kirkendall dkk. (1987: 59) mengemukakan selang waktu antara tes pertama dan tes ulangnya sebaiknya cukup lama agar subjek yang bersangkutan tidak mengulang kesalahan, atau jangan terlalu lama, sehingga yang bersangkutan ada kesempatan untuk berlatih selama rentang waktu antara tes pertama dan kedua, termasuk pula bagaimana cara melaksanakan atau menyelesaikan tes.

Faktor kematangan dan faktor bias merupakan salah satu kelemahan yang akan berpengaruh terhadap hasil tes ulangnya, jika terjadi interval yang panjang antara tes

pertama dan tes kedua. Jika metode tes dan tes ulangan ini digunakan, laporkan interval waktu pelaksanaan antara tes pertama dan tes ulangannya. Interval waktu yang sangat pendek akan membawa *carry over effect* dalam ingatan testi, sedangkan interval waktu yang lama akan membawa pengaruh perubahan informasi. Ini merupakan kelemahan kedua penggunaan metode tes dan tes ulangan. Perbedaan interval waktu dapat berpengaruh terhadap perbedaan taksiran reliabilitas, kadang-kadang *over-estimate* dan kadang-kadang *under-estimate* terhadap reliabilitas sebenarnya. Kelemahan ketiga penggunaan metode ini adalah sukar mempertahankan kondisi yang selalu sama persis untuk pelaksanaan tes pertama dan tes ulangannya karena banyak variable yang berpengaruh di dalamnya, seperti: kondisi siswa baik fisik maupun psikis pada saat pengujian, dan kondisi lingkungan sekitar pada saat tes dilaksanakan.

2. Metode Tes Sejajar (*Parallel Form*)

Metode tes sejajar kurang tepat digunakan untuk mengukur reliabilitas tes ranah psikomotor, namun sangat bagus dipergunakan untuk mengukur reliabilitas tes ranah kognitif, meskipun dalam pelaksanaannya banyak memakan waktu. Metode tes sejajar dilakukan dengan memberikan sekaligus dua bentuk tes yang paralel/ekuivalen satu sama lain, kepada sekelompok siswa. Agar kualitas soal tes bentuk sejajar dapat terjamin, sebaiknya dikonstruksi menurut *kisi-kisi/blue print* dalam penyusunan soalnya, sehingga cakupan materi dan tingkat kemampuan berpikir yang dikehendaki antara tes pertama dan tes ekuivalennya dapat terjamin. Salah satu persyaratan dalam penyusunan soal tes bentuk paralel di antaranya: (1) tingkat kesulitan soal harus sama, (2) indeks diskriminasi soal harus sama, (3) panjang soal harus sama,

- (4) rerata skor amatan dan simpang baku soal sama, dan
- (5) validitas, reliabilitas dan objektivitas soal harus sama (Ngarman, 2002: 85).

Dalam pelaksanaannya, kedua tes paralel itu dapat digabungkan terlebih dahulu sehingga seakan-akan merupakan satu bentuk tes. Setelah selesai dijawab oleh siswa barulah item pada tiap-tiap tes semula dipisahkan kembali untuk diberi skor masing-masing, sehingga diperoleh dua distribusi skor. Keuntungan cara penggabungan ini adalah tidak memberi kesan yang berat kepada siswa dalam mengerjakan soal dan apabila urutan nomor item gabungan itu diletakkan sedemikian rupa akan dapat mengurangi *carry-over* dari satu bentuk ke bentuk tes yang lain. Kalau penggabungan butir tes dari tiap-tiap tes tidak dimungkinkan, kedua tes tersebut harus diberikan secara berturut-turut dengan tenggang waktu yang hanya sebentar.

Untuk menambah pemahaman tentang metode estimasi reliabilitas tes pendidikan jasmani dengan menggunakan Metode Tes Sejajar (*Parallel Form*), pada tabel 3 disajikan data hasil tes teori pendidikan jasmani dari 10 siswa.

Pada contoh di atas diperlihatkan data skor dari dua bentuk tes yang paralel setelah dikenakan serentak pada sekelompok subjek. Komputasi koefisien korelasi *product moment* antara kedua distribusi skor tersebut menghasilkan $r_{xx'} = 0,907$. Koefisien ini merupakan koefisien reliabilitas tes bentuk sejajar (paralel).

Kerugian metode tes sejajar bagi beberapa tes pendekatan tersebut kurang praktis. Misalnya apabila si pemakai hanya membutuhkan satu bentuk atau tidak untuk mengembangkan bentuk lain. Selain itu, betapa sulit untuk menyusun/mengkonstruksi tes yang isinya sama tetapi butir-butir soalnya nampak berbeda. Karena tes kembar juga

memerlukan waktu pelaksanaan tes sebanyak dua kali (seperti halnya tes dan tes ulangan), skor tes kedua dapat dipengaruhi oleh faktor tertentu, seperti: motivasi siswa, kelelahan, kebiasaan, dan kebosanan. Faktor-faktor tersebut merupakan hambatan bagi penggunaan tes semacam ini.

Tabel 3. Hasil analisis koefisien korelasi tes bentuk paralel

Subjek	Skor Tes Bentuk I (X)	Skor Tes Bentuk II (X')
1	109	104
2	99	99
3	107	109
4	110	109
5	109	111
6	98	100
7	111	113
8	100	102
9	100	100
10	108	108
Koefisien Korelasi ($r_{xx'}$) = 0,907		

Sebenarnya dua tes yang paralel hanya ada secara teoretik, tidak benar-benar paralel secara empirik. Dari segi praktis, dua tes yang memenuhi syarat dan asumsi tertentu dianggap paralel. Untuk membuat dua tes menjadi paralel penyusunannya, haruslah didasarkan pada satu spesifikasi yang sama. Spesifikasi ini meliputi antara lain: tujuan ukur, batasan objek ukur dan operasionalisasinya, indikator-indikator perilakunya, banyaknya butir, format butir, juga kalau perlu meliputi taraf kesukaran butir. Secara empirik, dua tes yang paralel itu haruslah menghasilkan rerata skor dan varians yang setara dan korelasi yang juga tidak berbeda dengan suatu

variabel ketiga. Hanya itulah bukti terpenuhinya sifat paralel antara dua tes yang dapat diperoleh dalam penyusunan tes. Untuk membuktikan bahwa kedua tes menghasilkan dua skor-murni yang sama bagi setiap subjek serta memberikan dua varians eror yang sama sebagaimana dituntut oleh skor-murni klasikal, tidaklah dapat dilakukan.

Apabila telah diperoleh dua tes yang paralel, estimasi reliabilitasnya dilakukan dengan mengenakan dua tes tersebut bersamaan pada sekelompok subjek. Tiap-tiap tes akan menghasilkan distribusi skor-tampak dari seluruh subjek. Dengan menghitung korelasi antara kedua distribusi skor tersebut akan diperoleh koefisien yang menunjukkan seberapa jauh kedua distribusi skor itu bervariasi sama. Variasi bersama itu tercermin dalam kedudukan relatif setiap subjek dalam kedua tes. Makin sama kedudukan setiap subjek pada kedua tes, makin tinggi kecocokan hasil ukurnya. Oleh karena kedua tes telah dianggap paralel, kesamaan kedudukan itu mengandung arti kesamaan skor-tampak pada kedua tes.

Dua tes yang paralel yang menghasilkan skor berkorelasi tinggi satu sama lain disebut tes yang reliabel dan koefisien korelasinya merupakan koefisien reliabilitas tes yang bersangkutan. Sebaliknya, apabila dua tes yang telah dianggap paralel ternyata tidak dapat menghasilkan skor yang berkorelasi tinggi satu sama lain berarti tes tersebut tidak cukup reliabel dan hasil ukurnya tidak dapat dipercaya. Walaupun pendekatan reliabilitas tes bentuk sejajar (*parallel form*) ini dapat menghilangkan masalah penentuan tenggang waktu yang tepat, yang menjadi problem dalam pendekatan tes dan tes ulang (*test-retest*), akan tetapi pendekatan ini tidak menghilangkan sama sekali kemungkinan terjadinya efek bawaan.

3. Metode Belah Dua (*Split-Half*)

Pendekatan dengan menggunakan metode belah dua (*split-half*) yang dipergunakan untuk menaksir besarnya koefisien reliabilitas suatu tes lebih banyak dipergunakan untuk ranah kognitif dan afektif karena prosedur ini dipandang lebih praktis. Pendekatan ini lebih populer disebut dengan istilah pendekatan konsistensi internal. Pendekatan konsistensi internal dilakukan dengan menggunakan satu bentuk tes yang dikenakan hanya sekali saja pada sekelompok subjek (*single-trial administration*). Dengan menyajikan satu tes hanya satu kali, problem yang mungkin timbul pada dua pendekatan reliabilitas terdahulu dapat dihindari. Oleh karena itu, pendekatan ini mempunyai nilai praktis dan efisiensi yang tinggi. Pendekatan reliabilitas konsistensi internal bertujuan untuk melihat konsistensi antarbutir atau antabagian dalam tes itu sendiri. Untuk itu setelah skor setiap butir diperoleh dari sekelompok subjek, tes dibagi menjadi beberapa belahan (Saifudin Azwar, 1997: 42)

Dengan hanya satu kali penguasaan tes akan diperoleh satu distribusi skor tes dari kelompok subjek yang bersangkutan. Untuk itu, prosedur analisis reliabilitasnya diarahkan pada analisis terhadap butir-butir atau terhadap kelompok-kelompok butir dalam tes itu sehingga perlu dilakukan pembelahan tes menjadi beberapa kelompok butir yang disebut bagian atau belahan tes. Setiap bagian atau belahan dapat berisi beberapa item, bahkan dapat berisi hanya satu butir saja. Apabila bagian-bagian tes telah diperoleh, reliabilitas tes diperlihatkan oleh konsistensi di antara butir-butir atau di antara belahan-belahan tes tersebut.

Pembelahan tes dilakukan sedemikian rupa sehingga sedapat mungkin setiap belahan berisi butir dalam jumlah yang sama banyak. Membagi tes kedalam belahan yang berisi

butir dalam jumlah sama banyak tidak perlu untuk dilakukan. Hal ini tidak merupakan masalah lagi, karena sekarang ini telah tersedia rumus baru guna pengujian reliabilitas terhadap tes yang dibelah menjadi bagian-bagian yang berisi butir dalam jumlah yang tidak seimbang.

Contoh: Skor tes 10 kali tembakan dalam permainan bola basket dari 5 siswa.

Tabel 4. Hasil analisis koefisien korelasi tes bentuk ganjil-genap

Subjek	Trial (Percobaan)										X _{ganjil}	X _{genap}
	1	2	3	4	5	6	7	8	9	10		
A	1	1	0	1	1	1	0	0	0	1	2	4
B	1	1	1	1	1	1	1	1	1	1	5	5
C	0	0	1	1	0	1	1	0	1	1	3	3
D	1	0	0	1	1	0	0	1	1	0	3	2
E	1	1	1	0	1	1	1	1	0	1	4	4

Dengan menggunakan formula *Pearson Product Moment* diperoleh koefisien reliabilitas separo butir tes sebesar $r_{ij} = 0,538$. Untuk memperoleh koefisien reliabilitas seluruh tes dipergunakan rumus *Spearman Brown* sebagai berikut:

$$R_{it} = \frac{2 \cdot r_{X_{ganjil} X_{genap}}}{1 + r_{X_{ganjil} X_{genap}}}$$

Dengan demikian, koefisien reliabilitas dari seluruh tes sebesar 0,6996.

Cara pembelahan tes dengan metode belah dua ini juga bergantung pada sifat dan fungsi tes, serta jenis skala

pengukuran yang digunakan dalam tes bersangkutan. Cara pembelahan itu pada gilirannya akan ikut menentukan rumusan atau formula yang harus digunakan dalam menghitung koefisien reliabilitas.

Suatu tes yang hasilnya sebagian ditentukan oleh kecepatan kerja (*speed-test*), misalnya, menghendaki cara pembelahan yang berbeda dari cara pembelahan yang dilakukan terhadap tes yang mengukur kemampuan maksimum (*power-test*). Suatu tes yang berisi butir-butir yang mempunyai taraf kesukaran homogen akan lebih terbuka terhadap berbagai cara pembelahan apabila dibandingkan dengan tes yang berisi butir-butir dengan tingkat kesukaran yang sangat bervariasi. Tidak setiap karakteristik tes menghendaki cara pembelahan khusus, akan tetapi setiap cara pembelahan tes yang digunakan hendaknya ditekankan pada usaha untuk memperoleh bagian-bagian atau belahan-belahan yang relatif setara.

Untuk ranah psikomotor, metode ini tidak direkomendasikan untuk dipergunakan karena berkaitan dengan "terbatasnya item", dan sifat independen sebagian besar butir tes untuk ranah psikomotor (Phillips, 1979: 106). Dalam hal tertentu, metode ini direkomendasikan untuk dipergunakan mengestimasi reliabilitas tes dalam keterampilan olahraga, asalkan tes tersebut bersifat butir tunggal (*single item*), seperti: tes *shooting* dalam bola basket, tes servis dalam permainan tenis, tes *brady* dalam bola voli, dan tes-tes lain yang memiliki sifat sejenis.

Menurut Allen dan Yen (1979: 56) ada tiga cara yang lazim dipergunakan untuk membelah tes menjadi dua bagian, yaitu: (1) Metode ganjil-genap, dengan metode ini butir-butir dikelompokkan berdasarkan butir-butir bernomor ganjil dalam satu kelompok dan butir-butir tes bernomor

genap ke dalam kelompok kedua. Korelasi antara kedua kelompok tersebut *diset-up* dengan rumus Spearman Brown, (2) Metode belah dua sesuai dengan nomor urut. Misalnya, tes terdiri atas 50 butir soal, maka butir 1 sampai dengan butir 25 dikelompokkan dalam satu kelompok, sedangkan butir nomor 26 sampai nomor 50 dikelompokkan ke dalam kelompok lain, dan (3) Metode pembelahan dengan cara random/acak. Membelah tes menjadi dua bagian secara random/acak dapat dilakukan dengan cara undian sederhana guna menentukan butir-butir yang menjadi belahan pertama dan yang menjadi belahan kedua. Pembelahan secara random/acak hanya boleh dilakukan apabila tes yang akan dibelah berisi item-item yang homogen. Pengertian homogen dalam hal ini harus dipandang dari segi isinya (*content homogeneous*) dan juga dari segi taraf kesukarannya apabila tes itu mengukur aspek kognitif. Suatu tes yang berisi butir heterogen apabila dibelah secara random/acak dapat menghasilkan belahan-belahan yang tidak setara satu sama lain, kecuali apabila tes tersebut terdiri atas butir yang berjumlah sangat besar.



BAB 7 OBJEKTIVITAS TES

A. Pengertian Objektivitas

Pengertian objektivitas hampir sama dengan reliabilitas. Perbedaannya terletak pada pendekatan dalam pengujianya. Hal ini disebabkan dalam objektivitas diperlukan sekurang-kurangnya dua atau lebih testor memberikan penilaian terhadap peserta didik dengan prosedur yang sama. Hasil penilaian yang diperoleh dari testor yang satu dikorelasikan dengan hasil penilaian yang diberikan oleh testor lainnya. Hasil korelasi skor/penilaian dari kedua testor inilah yang menunjukkan derajat objektivitas suatu tes. Penilaian dikatakan objektif terkandung makna bahwa penilaian tersebut tidak ada unsur kepentingan pribadi testor yang mempengaruhi hasil pengesanan.

Tingkat keakuratan dalam pemberian nilai/skor suatu tes ke peserta didik sering kali diartikan sebagai objektivitas tes. Tes dikatakan objektif manakala dalam pemberian skoring tes mengandung kesalahan (*error*) yang sangat sedikit atau bahkan tidak terjadi kesalahan sama sekali. Variabilitas di antara para penilai (*skorer*) meningkat apabila penilai memberikan penilaian yang bersifat subjektif (*subjektive rating*), seperti: menilai senam irama, loncat indah, renang indah, peragaan jurus-jurus pencak silat, peragaan nomor-

nomor kata dalam olahraga karate, dan sebagainya. Namun sebaliknya jika penilai ditugaskan untuk memberikan penilaian pada cabang-cabang yang terukur seperti: lari 100 meter, lompat jauh, lompat tinggi, nomor-nomor lempar) maka penilaian yang dihasilkan akan bersifat objektif.

Pengertian objektivitas suatu tes dalam penjasorkes pada dasarnya hampir sama dengan pengertian reliabilitas. Kedua-duanya (objektivitas dan reliabilitas) berorientasi pada usaha untuk mendapatkan hasil pengukuran dan penilaian yang sama atau hampir sama. Reliabilitas menunjukkan satu pengambil tes dengan mengadakan ulangan tes terhadap suatu objek kemudian hasilnya dibandingkan antara tes pertama dengan tes kedua terhadap objek yang sama tersebut. Sedangkan dalam objektivitas terdapat dua atau lebih pengambil tes dalam menilai suatu objek yang sama, kemudian dari dua pengambil nilai (*expert judgement*) atau lebih tersebut dibandingkan. Suatu bentuk reliabilitas yang berkenaan dengan testor dalam memberikan penilaian dinamakan "*inter-tester reliability*" atau sering disebut dengan istilah "*objectivity*" (Thomas, 1990: 354).

Berkenaan dengan masalah objektivitas, Phillips (1979: 111) mengatakan: "*objectivity is the degree to which consistent results are obtained from two or more test administrators*". Senada dengan pendapat tersebut, Baumgartner (1982: 110) mengemukakan bahwa: "*objectivity or rater reliability, is an important characteristic of a test or measuring instrument. We can define objectivity as the close agreement between the scores assigned to each subject by two or more judges*".

Dari pendapat di atas dapat ditarik suatu makna bahwa **objektivitas** dapat didefinisikan tingkat konsistensi atau keajegan hasil yang diperoleh dari dua atau lebih pelaksana ukur (testor) terhadap subjek yang sama. Dengan kata lain,

pengukuran itu dikatakan objektif apabila dua orang testor atau lebih memberikan skor atau angka yang sama kepada testi untuk satu tes.

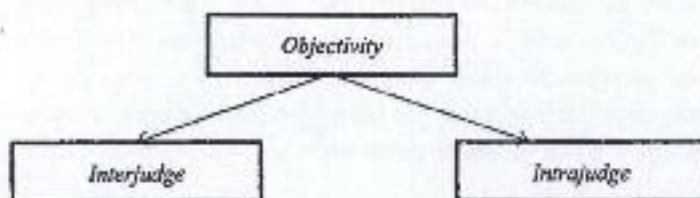
Ada beberapa faktor yang berpengaruh terhadap objektivitas suatu instrumen, di antaranya adalah: (1) Petunjuk pengukuran disusun lengkap, terperinci, dan sejelas mungkin, (2) *Prosedur* pengukuran dibuat sederhana, (3) Menggunakan peralatan elektronik apabila memungkinkan, (4) Memasukkan hasil tes ke dalam skor matematik, (5) Memilih testor yang benar-benar terlatih dan ahli dalam bidangnya, (6) Testor selalu mempertahankan sikap profesional dan sikap ilmiah dalam memberikan nilai, (7) Menggunakan dengan setepat-tepatnya standar prosedur penskoran.

B. Tipe-Tipe Objektivitas

Tipe-tipe *objektivitas* yang lazim dipergunakan dalam pengajaran pendidikan jasmani ada dua, yaitu: (1) Objektivitas intra penilai/juri (*intrajudge*), dan (2) Objektivitas inter penilai/juri (*interjudge*). Objektivitas "*intrajudge*" adalah tingkat konsistensi seorang testor dalam menilai subjek yang sama sebanyak dua kali penilaian atau lebih. Jika tes tertulis sedang dipertimbangkan, maka objektivitas *intrajudge* dapat secara mudah ditentukan. Serangkaian tes dapat dinilai sekali, kemudian disimpan dulu untuk beberapa lama, dan kemudian dinilai lagi oleh orang yang sama. Perbandingan kedua set skor tersebut akan menggambarkan tingkat konsistensi individu dalam skoring tersebut. Dalam tes keterampilan olahraga, pengukuran objektivitas *intrajudge* lebih sulit diperoleh karena penampilan yang sama harus diamati dua kali. Cara ini dapat dipermudah dengan merekam penampilan itu menggunakan *film* atau *video tape*.

Jenis objektivitas kedua adalah objektivitas "*interjudge*", adalah tingkat konsistensi antara dua atau lebih penilai yang bersifat independen dalam menilai penampilan yang sama dari seorang subjek. Objektivitas *interjudge* adalah bagian yang penting dalam merating perlombaan-perlombaan dalam senam, loncat indah, olahraga permainan di mana beberapa juri menilai setiap penampilan individu.

Adapun tipe-tipe objektivitas tersaji pada bagan di bawah ini



Skema 3: Tipe-tipe Objektivitas (Safrit, 1986: 138)

Kedua tipe objektivitas ini menurut Safrit (1986: 139) dapat diestimasi dengan memakai prosedur korelasi yang sama dengan untuk mengestimasi reliabilitas. Pengukuran-pengukurannya diharapkan akan relatif tinggi (0,80 atau > 0,80), karena nilai-nilai yang lebih rendah menunjukkan bahwa juri-juri menilai komponen keahlian secara berbeda.

Secara ringkas, usaha-usaha mengidentifikasi sumber-sumber pembias faktor dalam penilaian pengajaran pendidikan jasmani dan kesehatan sangatlah banyak, namun apabila faktor-faktor tersebut dapat diminimalisir niscaya kegiatan belajar-mengajar akan berlangsung dengan baik serta penilaian yang dilaksanakan akan benar-benar valid, reliabel, dan objektif. Tingkat objektivitas suatu alat ukur

dalam penjasorkes menurut Strand (1993) tergantung pada bagaimana petunjuk tes itu diberikan dengan jelas dan serinci mungkin, serta bagaimana seluruh prosedur tes itu diikuti. Hal senada juga dikemukakan Phillips (1979), bahwa tingkat objektivitas suatu tes atau alat ukur diperoleh apabila:

1. Petunjuk pengukuran disusun dalam susunan kata-kata selengkap dan sejelas mungkin.
2. Prosedur pengukuran dibuat sederhana dan tidak bertele-tele.
3. Bilamana mungkin menggunakan peralatan mekanik, seperti *movie film*.
4. Memasukkan hasil tes ke dalam skor matematik (berupa angka-angka).
5. Memilih testor yang benar-benar terlatih dan ahli/pandai dalam bidangnya.
6. Testor selalu mempertahankan sikap profesional dan sikap ilmiah selama penilaian berlangsung.
7. Pemantauan secara terus-menerus prosedur pelaksanaan pengukuran terhadap administrator tes.
8. Menggunakan dengan setepat-tepatnya standar prosedur pengukuran.

Dengan kata lain untuk menjamin objektivitas penilaian *judge*, maka *judge* harus mengevaluasi testi dengan menggunakan cara dan standard yang sama serta mengikuti petunjuk tes seteliti dan secermat mungkin.

Berapa besarnya koefisien objektivitas suatu tes dalam pendidikan jasmani olahraga dan kesehatan dapat dianalisis dengan menggunakan teknik korelasi. Pada umumnya teknik analisis korelasi yang dipergunakan dapat menggunakan teknik korelasi *Pearson Product Moment*. Korelasi tersebut dihitung berdasarkan hasil-hasil tes yang diperoleh dari masing-masing penilai (testor). Contoh: ada tiga orang

penilaian, yaitu juri A, B, dan C. Untuk mencari tingkat objektivitas penilaian tersebut dapat dilakukan dengan cara:

1. Pertama, melakukan korelasi antara hasil penilaian yang diperoleh dari juri A dan juri B.
2. Kedua, korelasi hasil penilaian antara juri A dan juri C.
3. Ketiga, korelasi hasil penilaian antara juri B dan juri C.

Hasil-hasil analisis koefisien objektivitas suatu tes dituangkan dalam bentuk koefisien korelasi. Sebagai bahan acuan, apakah tes pendidikan jasmani olahraga dan kesehatan mempunyai koefisien objektivitas yang memadai, maka Mathews (1963) memberikan standar untuk menilai besarnya koefisien objektivitas suatu tes berikut disajikan tabel kategori koefisien objektivitas penilaian.

Tabel. Kategori Koefisien Objektivitas

Koefisien Objektivitas	Kategori
0,90 – 0,99	Tinggi
0,80 – 0,89	Baik
0,70 – 0,79	Cukup
0,60 – 0,69	Kurang
< 0,60	Kurang Sekali



BAB 8

PENDEKATAN ACUAN PENILAIAN DAN SISTEM PENILAIAN

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. 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

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.

2. Administrasi

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.

3. Bersifat mewakili

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 perbandingan.

4. Bersifat temporer

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.

5. Memiliki petunjuk/*presentation*,

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.

6. Bisa dibandingkan

Seringkali guru pendidikan jasmani perlu membandingkan skor-skor dari tes-tes yang berbeda untuk mengevaluasi prestasi siswa. Norma-norma pada tes yang terpisah harus komparabel. Dalam rangkaian tes kebugaran jasmani remaja putri dari AAHPER, 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.

Nilai hasil tes pengetahuan umum penjaskes 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 1. 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:

$$\Sigma = 500$$

$$\Sigma (X_i - \bar{X})^2 = 227$$

$$\bar{X} = 25$$

$$\text{Simpangan baku} = 3,45$$

Tabel 2. 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 2 di atas, apabila Si Rifki mendapatkan skor 29, maka nilai Rifki berdasarkan penilaian acuan norma dengan standar 1 - 10 adalah = 8

Tabel 3: Penilaian Acuan Norma dengan 1 - 5 kategori

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

Selain Penilaian Acuan Norma dengan menggunakan standar penilaian 1 – 10, Pendekatan Acuan Norma dapat dikembangkan pula dengan menggunakan 5 kategori, sebagaimana tabel 3.

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.

B. 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 terkontrol dengan baik.

Patokan ini ditetapkan atas dasar pertimbangan legis 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 Saifit (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 3. 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 penjas keskor 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.

C. Sistem Penilaian dalam Pendidikan Jasmani

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. Sebagai contoh, perhatikan tabel 4 berikut ini.

Tabel 4. 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 E 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 5: Jumlah angka bagi setiap tugas dan penentuan nilai

Tugas	Nilai	Skala Penilaian
Tes Keterampilan	90	A = 482 - 530
Metodik I	50	B = 428 - 481
Metodik 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.

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.

Tabel 6. 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

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. Hal 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 kepanjangannya. Namun dapat digunakan sebagai alternatif 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 ketujuh 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:

Tabel. 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		

Tabel. 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 9

PENGUKURAN KESEGERAN JASMANI

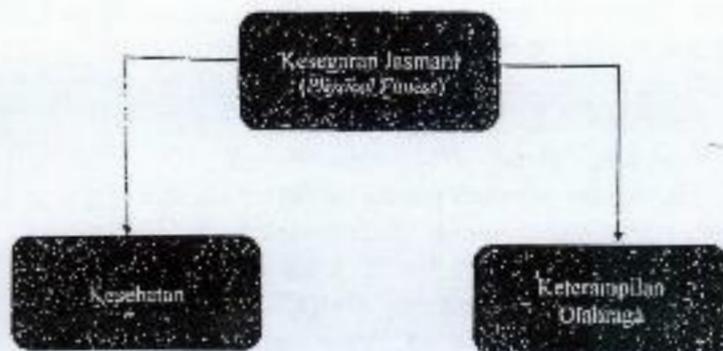
Salah satu sasaran penilaian Penjasorkes yang tercantum dalam PP No. 19 Tahun 2005 tentang Standar Pendidikan Nasional, Permendiknas No. 22 tahun 2006 tentang Standar Isi Pendidikan Nasional, dan Permendiknas No. 23 tahun 2006 Standar Kompetensi Lulusan adalah Kebugaran Jasmani. Tujuan mata pelajaran Penjasorkes adalah untuk meningkatkan kebugaran jasmani peserta didik diamanahkan di semua jenis dan jenjang pendidikan. Ketentuan-ketentuan tersebut juga relevan dengan tuntutan Kurikulum 2013 yang dituangkan dalam Kompetensi Inti (KI) dan Kompetensi Dasar (KD) Mata Pelajaran Pendidikan Jasmani Olahraga dan kesehatan Pembelajaran Penjasorkes bertujuan untuk mengembangkan ranah kognitif, ranah afektif dan ranah psikomotor. Pengembangan ranah psikomotor melalui media pendidikan jasmani olahraga dan kesehatan dapat dilakukan dengan berbagai aktivitas fisik yang dilakukan secara terarah, terprogram, dan sistematis. Pengembangan mata pelajaran penjasorkes di sekolah di arahkan untuk mengembangkan ranah psikomotor, yang terdiri dari: penanaman teknik dasar dan pengembangan keterampilan dalam cabang-cabang olahraga dan pengembangan kebugaran jasmani peserta didik.

Pengembangan kebugaran jasmani peserta didik melalui pembelajaran penjasorkes bertujuan untuk meningkatkan

kondisi fisik peserta didik agar mampu mengikuti aktivitas pembelajaran dengan baik yang diberikan di sekolah sesuai dengan kurikulum yang berlaku. Peserta didik yang memiliki tingkat kesegaran jasmani yang tinggi akan mendukung terhadap aktivitas kegiatan belajarnya, serta mampu untuk melaksanakan aktivitas fisik lainnya, seperti jalan, lari, bersepeda, berolahraga, dan kegiatan mengisi waktu luang lainnya.

Melalui aktivitas olahraga pada pembelajaran penjasorkes maka keterampilan para peserta didik akan meningkat dalam cabang-cabang olahraga yang dipelajarinya. Jenis cabang-cabang olahraga yang diberikan di sekolah mengacu pada kurikulum yang berlaku di sekolah tersebut. Cabang-cabang olahraga yang tercantum dalam Kompetensi Inti dan Kompetensi Dasar Kurikulum di SMA dan SMK antara lain: (1) Permainan dan Olahraga, (2) Senam dan Gerak Berirama, (3) Aktivitas Air/Akuatik, (4) Kesehatan, (5) Aktivitas Pengembangan/Kebugaran Jasmani.

Dalam kehidupan sehari-hari, setiap orang memiliki aktivitas jasmani yang berbeda-beda. Agar dapat melakukan aktivitas jasmani dengan baik diperlukan kualitas kemampuan fungsi organ tubuh sesuai dengan tuntutan tugas gerak yang dilakukannya. Kualitas kemampuan fungsi organ tubuh menunjukkan kualitas tingkat kesegaran jasmani. Dewasa ini konsep kesegaran jasmani berkembang menjadi dua macam, yaitu: kesegaran jasmani yang berhubungan dengan **kesehatan** (*health related fitness*) dan kesegaran jasmani yang berhubungan dengan **keterampilan olahraga** (*skill related fitness*).



Kesegaran jasmani yang berhubungan dengan **kesehatan**, meliputi:

- a. Daya tahan jantung dan paru-paru.
- b. Kekuatan otot
- c. Daya tahan otot
- d. Kelentukan
- e. Komposisi tubuh.

Kesegaran jasmani yang berhubungan dengan **keterampilan olahraga**, meliputi:

- a. Kecepatan.
- b. Kelincahan
- c. Keseimbangan
- d. Koordinasi
- e. Power
- f. Kecepatan reaksi

(Nieman DC, 1993: *Fitness and Your Health*, California: Bull Publishing)

A. Pengertian Kesegaran Jasmani

Pada saat ini kita sering mendengar istilah kesegaran jasmani. Istilah ini merupakan bagian dari *total fitness*. Dalam *total fitness* tercakup beberapa komponen, yaitu:

1. Kesegaran Anatomi (*Anatomical Fitness*)
2. Kesegaran Fisik (*Physiological Fitness*)
3. Kesegaran Psikologi (*Psychological Fitness*)

Kesegaran anatomis (*anatomical fitness*) adalah kesegaran tubuh yang sukar untuk dikembangkan karena untuk pengembangannya harus dimulai sejak masa pertumbuhan anak-anak. Pengembangannya memerlukan jangka waktu yang lama dan hasilnya pun sangat terbatas. Persoalan ini terbentur pada factor keturunan (hereditas) yang tidak dapat kita pengaruhi.

Kesegaran fisiologis (*physiological fitness*) adalah kemampuan tubuh untuk menyesuaikan fungsi fisiologisnya agar dapat mengatasi keadaan lingkungan dan atau tugas fisik yang memerlukan kerja otot secara efisien, tidak mengalami kelelahan yang berlebihan, dan telah memperoleh pemulihan yang sempurna sebelum datangnya tugas-tugas pada hari berikutnya.

Kesegaran psikologis (*Psychological Fitness*) adalah kemampuan tubuh untuk mengatasi masalah-masalah yang berkaitan dengan gangguan emosi dalam menghadapi masalah-masalah setiap hari di lingkungannya.

Kesegaran jasmani adalah merupakan terjemahan dari kata *Physical Fitness* yang dapat diartikan sebagai kondisi jasmani yang menggambarkan kemampuan jasmani, dapat pula diartikan kemampuan seseorang untuk melakukan suatu pekerjaan tertentu dengan cukup baik tanpa mengalami kelelahan. Kesegaran jasmani merupakan aspek fisik dari kesegaran yang menyeluruh yang memberi kesanggupan

kepada seseorang untuk menjalankan hidup produktif dan dapat menyesuaikan diri tiap pembebanan fisik yang layak.

Kesegaran jasmani adalah kondisi jasmani yang menggambarkan potensi dan kemampuan jasmani untuk melakukan tugas-tugas tertentu dengan hasil yang optimal tanpa memperlihatkan keletihan yang berarti. Sedangkan menurut *President's Council on Physical Fitness and Sport* mendefinisikan kesegaran jasmani adalah kemampuan untuk melakukan kegiatan sehari-hari dengan penuh vitalitas dan kewaspadaan tanpa mengalami kelelahan yang berarti, dan masih cukup energi untuk bersantai pada waktu luang dan menghadapi hal-hal yang sifatnya darurat (emergensi). Dari pengertian-pengertian di atas mempunyai kesamaan mengenai kesiapan dan kesanggupan untuk melaksanakan tugas yang memerlukan tenaga fisik secara efisien dan efektif. Hal yang dimaksudkan dengan efisien dan efektif adalah dapat mengatasi dan menyelesaikan tugas tanpa menderita/mengalami kelelahan yang berarti serta dapat melanjutkan tugas-tugas berikutnya atau dengan kata lain masa pemulihannya tidak memerlukan waktu yang lama.

Pentingnya kesegaran jasmani bagi anak usia sekolah antara lain dapat meningkatkan kemampuan organ tubuh, sosial emosional, sportivitas, dan semangat kompetisi. Beberapa ahli juga menyebutkan bahwa kesegaran jasmani memiliki korelasi positif dengan prestasi akademis. Dari sudut pandang pendidikan upaya peningkatan kesegaran jasmani memiliki tujuan antara lain: (1) pembentukan gerak, (2) pembentukan prestasi, (3) pembentukan sosial, dan (4) pertumbuhan badan.

Ada juga tingkatan kebugaran jasmani yang harus dimiliki oleh orang yang bekerja cukup berat seperti militer dan biasa dikategorikan pada *motor fitness* atau para atlet

yang harus memiliki kemampuan fisik melebihi dari tingkat kebugaran jasmani yang baik saja akan tetapi harus memiliki kemampuan fisik secara menyeluruh yang biasa disebut dengan *general motor ability* atau *motor ability*.

B. Pengukuran Komponen Kebugaran Jasmani

— Dalam pembelajaran pendidikan jasmani olahraga dan kesehatan, pengukuran terhadap tingkat kebugaran jasmani peserta didik perlu dilakukan secara berkala. Pengukuran ini berfungsi untuk:

- a. Mengukur kemampuan fisik peserta didik
- b. Menentukan status kondisi fisik peserta didik
- c. Menilai kemampuan fisik peserta didik sebagai salah satu tujuan pengajaran penjasrkes.
- d. Mengetahui perkembangan kemampuan fisik peserta didik
- e. Sebagai bahan untuk memberikan bimbingan dalam meningkatkan kebugaran jasmaninya.
- f. Sebagai salah satu bahan masukan dalam memberikan nilai pelajaran penjasorkes.

1. Daya Tahan Jantung dan Paru (*Cardiorespiratory Endurance*).

Daya tahan jantung dan paru adalah kesanggupan sistem jantung, paru-paru dan pembuluh darah untuk berfungsi secara optimal saat melakukan aktivitas sehari-hari, dalam waktu cukup lama tanpa mengalami kelelahan yang berarti. Daya tahan jantung dan paru sangat penting untuk menunjang kerja otot, yaitu dengan cara mengambil oksigen dan mengeluarkan ke otot yang aktif. Dan pada daya tahan ini ada juga yang dinamakan dengan daya tahan otot lokal.

Daya tahan otot adalah kapasitas otot untuk melakukan kontraksi secara terus menerus pada tingkatan intensitas submaksimal. Pada dasarnya daya tahan kekuatan otot merupakan rentangan antara daya tahan dan kekuatan otot. Daya tahan otot diperlukan untuk mempertahankan kegiatan yang sifatnya didominasi oleh penggunaan otot atau kelompok otot.

Seperti halnya pada komponen lain, daya tahan otot hanya diperlukan sebatas kebutuhan dalam aktivitas otot. Beberapa kegiatan yang dominan memerlukan daya tahan otot pada anak-anak termasuk didalamnya bentuk-bentuk permainan kecil maupun besar (umpamanya bermain tali panjat tebing atau lari lintas alam bagi yang sudah berusia dewasa. Sedangkan pada atlet dapat dicontohkan pada pelaksanaan *pull up*/bergantung angkat badan semampunya.

Item-item tes untuk mengukur daya tahan jantung dan paru dapat dilakukan dengan tes laboratorium (*laboratorium test*) maupun tes lapangan (*field test*). Berdasarkan hasil penelitian, pengukuran daya tahan *cardiorespiracy* yang dilakukan di laboratorium (*laboratory test*) ternyata lebih akurat jika dibandingkan dengan pengukuran yang dilakukan di lapangan, namun biaya yang diperlukan relatif mahal dan memerlukan waktu yang relatif lama. Sedangkan pengukuran daya tahan *cardiorespiracy* yang dilakukan di lapangan (*field test*) lebih efisien dari waktu dan biaya namun tingkat keakuratan data tidak secermat jika pengukuran dilakukan di laboratorium.

Berikut disajikan item-item tes untuk mengukur daya tahan jantung dan paru-paru.

Item tes daya tahan jantung dan paru-paru
(Cardiorespiratory)

1. *Harvard Step Test*
2. *Balke Test* (Jalan-lari 15 menit)
3. *Malty Fitness Test (MFT)*
4. *Katch Pulse Recovery Test*
5. *Aerobic Test* (lari 12 menit, lari 2,4 km, 4,8 km jalan)
6. *Shirky Test*

2. Kekuatan Otot/*Strength*

Secara fisiologis kekuatan otot adalah kemampuan otot atau sekelompok otot untuk melakukan satu kali kontraksi secara maksimal melawan tahanan atau beban. Secara mekanis kekuatan otot didefinisikan sebagai gaya (*force*) yang dapat dihasilkan oleh otot atau sekelompok otot dalam suatu satu kontraksi maksimal. Kekuatan otot merupakan hal penting untuk setiap orang.

Terdapat beberapa macam tipe kekuatan yang harus diketahui, yaitu: kekuatan umum, kekuatan khusus, kekuatan maksimum, daya tahan kekuatan, kekuatan absolut, dan kekuatan relatif (Bompa, 1993). Dengan mengetahui tipe-tipe kekuatan maka kita dapat melatihnya secara efektif. Pengukuran kekuatan otot tidak mudah dilakukan untuk semua otot atau kelompok otot. Karena kekuatan menjadi sasaran penilaian pengajaran penjasorkes maka perlu dipilih mana yang mungkin dilakukan di satuan pendidikan dengan alat dan prosedur pelaksanaan yang sederhana.

Berikut disajikan item-item tes yang pada umumnya dipergunakan untuk mengukur kekuatan otot.

Item Tes Kekuatan Otot

1. *Sit-Ups* 30 detik
2. *Push-Ups* 30 detik
3. *Back-Ups* 30 detik
4. *Pull-Ups*
5. *Chin-Ups*
6. *Wall Sit Test*
7. *Grip Strength Dynamometer*
8. *Back and Leg Dynamometer*
9. *Push Expanding Dynamometer*
10. *Pull Expanding Dynamometer*
11. *Standing Broad Jump*

3. Kelentukan (*Flexibility*)

Fleksibilitas adalah kemampuan sendi untuk melakukan gerakan dalam ruang gerak sendi secara maksimal. Fleksibilitas menunjukkan besarnya pergerakan sendi secara maksimal sesuai dengan kemungkinan gerakan (*range of movement*). Fleksibilitas bagi anak sangat penting dimiliki terutama untuk kegiatan dalam bermain, karena bermain bagi mereka tidak semata-mata dapat bergerak cepat dan kuat, tetapi juga harus lincah dan dapat mengubah arah dengan cepat (*kelincahan*). Kemampuan yang cepat dan lincah dalam mengubah arah memerlukan fleksibilitas tubuh atau bagian tubuh yang lebih dalam kegiatan tersebut. Melakukan perubahan kecepatan dan arah gerakan, dapat mengakibatkan regangan otot yang terlalu kuat sehingga memungkinkan terjadinya cedera otot (*muscle sprain*) apabila fleksibilitas otot yang dimiliki rendah. Pada usia sampai sampai 10 tahun, umumnya anak-

anak memiliki fleksibilitas yang sangat baik. Bagaimanapun juga latihan untuk meningkatkan fleksibilitas tidak boleh berlebihan, karena dapat berpengaruh tidak baik dan bahkan dapat merusak sikap tubuh itu sendiri.

Ada dua macam kelentukan, yaitu kelentukan statis (*static flexibility*) dan kelentukan dinamis (*dynamic flexibility*). Kelentukan statis adalah kemampuan sendi untuk melakukan gerak dalam ruang yang besar. Misalnya, gerakan split. Kelentukan dinamis adalah kemampuan menggunakan persendian dan otot secara terus menerus dalam ruang gerak yang penuh dengan cepat dan tanpa tahanan gerakan. Misalnya, menendang bola tanpa tahanan atau beban pada otot-otot hamstring dan sendi panggul.

Fungsi kelentukan, di antaranya: (1) mengurangi kemungkinan terjadinya cedera, (2) membantu mengembangkan kecepatan, koordinasi, dan kelincahan, (3) membantu mengembangkan keterampilan teknik, (4) membantu efisiensi gerakan, peserta didik yang kelentukannya tinggi menggunakan energi yang lebih sedikit dengan peserta didik yang kelentukannya rendah, dan (5) membantu memperbaiki sikap tubuh (Morrow, 2000).

Item-item tes yang pada umumnya dipergunakan untuk mengukur kemampuan kelentukan dapat disajikan pada tabel berikut:

Tes-Tes Kelentukan (*flexibility*)

1. Sit and Reach
2. Bridge-Up (Kayang)
3. Front Split
4. Side Split
5. Shoulder and Wrist Elevation
6. Ankle Extension (*Plantar Flexion*)

4. Kecepatan (*Speed*)

Kecepatan adalah kemampuan berpindah dari satu tempat ketempat yang lain dalam waktu yang sesingkat-singkatnya. Kecepatan bersifat lokomotor dan gerakannya bersifat siklik (satu jenis gerakan yang dilakukan berulang-ulang seperti lari dan sebagainya) atau kecepatan gerak bagian tubuh seperti melakukan pukulan. Kecepatan penting tidak saja bagi anak-anak terutama saat mereka bermain di sekolah maupun di rumah juga bagi mereka yang sudah dewasa untuk dapat tetap menjaga mobilitasnya.

Kecepatan merupakan salah satu komponen biomotorik yang penting untuk melakukan aktivitas olahraga (Bompa, 1990). Dalam olahraga terdapat dua jenis yang dibutuhkan, yaitu kecepatan reaksi (*reaction speed*) dan kecepatan bergerak (*speed movement*). Kecepatan reaksi adalah kualitas yang memungkinkan memulai suatu jawaban kinetis secepat mungkin setelah menerima suatu rangsang. Kecepatan bergerak adalah kualitas yang memungkinkan orang bergerak atau melaksanakan gerakan-gerakan yang sama atau tidak sama secepat-cepatnya.

Dalam kegiatan olahraga, kebanyakan tes kecepatan yang dilakukan oleh guru penjasorkes di sekolah menggunakan tes yang melibatkan kecepatan rata-rata. Kecepatan rata-rata adalah total jarak dibagi total waktu yang digunakan untuk menempuh jarak tertentu.

Adapun jenis-jenis tes yang dipergunakan untuk mengukur kecepatan disajikan pada tabel berikut:

Item Tes Kecepatan (*Speed*)

- | | |
|------------------------------|--|
| 1. Lari cepat 6 detik | 7. <i>Whole Body Reaction Test</i> |
| 2. Lari 30 meter | 8. <i>Speed Anticipation Reaction Test</i> |
| 3. Lari 40 meter | |
| 4. Lari 50 meter | |
| 5. Lari 60 meter | |
| 6. <i>Hand Reaction Test</i> | |

5. *Power* (Daya Eksplosif)

Power atau sering pula disebut daya ledak merupakan salah satu komponen biomotorik yang penting. Power adalah gabungan antara kekuatan dan kecepatan atau pengerahan gaya otot maksimum dengan kecepatan maksimum. Sesuai dengan sifat anak-anak usia sekolah, gerakan eksplosif kuat dan cepat seringkali digunakan, merupakan ciri khas pola bermain yang dikembangkan untuk anak-anak. Anak membutuhkan komponen tersebut untuk menumbuhkan kemampuan kepada orang lain. Bagi orang dewasa, kemampuan yang kuat dan cepat juga diperlukan terutama bagi tindakan-tindakan membutuhkan kemampuan tenaga secara maksimal misalkan pada saat melakukan teknik smash. Tentu saja setiap komponen penampilan aktivitas gerak akan memiliki bobot yang berbeda sesuai dengan tingkatan usia dan kepentingannya.

Power merupakan suatu rangkaian kerja beberapa unsur gerak otot dan menghasilkan daya ledak apabila dua kekuatan tersebut bekerja secara bersamaan. Power memiliki banyak kegunaan pada pembelajaran penjasorkes di sekolah, seperti: berlari, melempar, memukul, menendang. Pelaksanaan gerak dari objek tersebut akan dicapai dengan sempurna apabila peserta didik dapat menerapkan kekuatannya secara maksimal dengan satuan waktu yang sesingkat-singkatnya.

Pada umumnya untuk mengetahui kemampuan power peserta didik dapat dilakukan dengan menggunakan item-item tes berikut

Item Tes *Power* (Daya Eksplosif)

1. *Vertical Jump* (daya eksplosif vertikal)
2. *Standing Broad Jump* (Daya eksplosif horisontal)
3. *Margarita-Kallemen Power Test*
4. *Force Plate*
5. *Medicine Ball Test*

6. Kelincahan (*Agility*)

Kelincahan merupakan salah satu komponen kebugaran jasmani yang diperlukan pada semua aktivitas yang memerlukan kecepatan perubahan posisi tubuh dan bagian-bagiannya. Kelincahan merupakan salah satu komponen gerak yang sangat diperlukan untuk jenis olahraga yang membutuhkan adaptasi yang tinggi terhadap perubahan-perubahan situasi dalam pertandingan. Kelincahan adalah kemampuan untuk mengubah arah atau posisi tubuh dengan cepat dan tepat yang dilakukan bersama-sama dengan gerakan lainnya (Kirkendall, 1987). Bagi anak, kelincahan merupakan komponen kebugaran jasmani yang harus dimiliki. Tanpa kelincahan, anak dikatakan tidak dalam keadaan normal atau mungkin sedang sakit. Kelincahan bagi mereka adalah sesuatu yang khas sesuai dengan kodratnya. Jadi kelincahan harus menempati prioritas utama dalam melatih kebugaran jasmani setiap peserta didik. Bagi orang dewasa kelincahan

tidak berarti kurang penting, tetapi apabila dilihat dari kebutuhan serta aktivitas yang dilakukan kelincahan terbatas pada cabang olahraga yang dilakukannya.

Kelincahan memiliki karakteristik yang unik, karena dapat memainkan peranan yang khusus terhadap mobilitas fisik. Kelincahan bukan merupakan komponen fisik tunggal, melainkan tersusun dari beberapa komponen biomotorik yang lain, seperti: koordinasi, kekuatan, kelentukan, waktu reaksi, dan power. Komponen-komponen tersebut dalam pelaksanaan untuk membentuk gerak kelincahan saling terintegrasi satu dengan yang lain.

Tabel berikut disajikan item-item tes untuk mengukur komponen biomotorik kelincahan (*agility*).

Item Tes Kelincahan (*Agility*)

- | | |
|-----------------------------|--------------------------------|
| 1. Squat thrust | 6. Illinois Agility Run |
| 2. Dodging run | 7. LSU Agility Obstacle Course |
| 3. Hexagonal Obstacle | 8. Agility T-Test |
| 4. Right Boomerang Run Test | 9. Side-Step Test |
| 5. Shuttle Run | 10. Quick Feet Test |

7. Keseimbangan (*Balance*)

Peningkatan keseimbangan peserta didik perlu diukur sebagai salah satu sasaran evaluasi pengajaran penjasorkes karena keseimbangan penting bagi aktivitas geraknya. Keseimbangan adalah kemampuan mempertahankan sikap dan posisi tubuh secara cepat pada saat berdiri (*static balance*) atau pada saat melakukan gerak (*dynamic balance*). Kemampuan untuk mempertahankan keseimbangan dipengaruhi oleh beberapa faktor antara lain: visual, vestibular,

Terdapat dua macam keseimbangan, yaitu keseimbangan statis dan keseimbangan dinamis. Keseimbangan statis adalah kemampuan tubuh mempertahankan keadaan seimbang dalam posisi diam, misalnya: berdiri satu kaki dengan mata dipejamkan (*stork Stand*), sedangkan keseimbangan dinamis kemampuan tubuh mempertahankan keadaan seimbang dalam posisi bergerak, misalnya: berjalan, berlari, melempar, melompat, dan sebagainya. Keseimbangan statis maupun dinamis merupakan komponen kesegaran jasmani yang sering dilakukan oleh anak-anak maupun dewasa. Cara meniti blok atau tician kayu/besi misalnya, sudah menjadi fungsi keseimbangan untuk dapat mempertahankan posisi normalnya. Komponen biomotorik keseimbangan merupakan kemampuan yang penting karena dipergunakan dalam aktivitas sehari-hari maupun aktivitas berolahraga bagi peserta didik.

Berikut akan disajikan item-item tes yang dapat dipergunakan untuk mengukur komponen biomotorik keseimbangan.

Item Tes Keseimbangan (Balance)

- | | |
|---|----------------------------------|
| 1. <i>Stork Stand</i> (Berdiri satu kaki mata dipejamkan) | 6. <i>Side Learing Rest</i> |
| 2. <i>Barz Stick Test</i> (<i>Lengthwise and Crosswise</i>) | 7. <i>One-Knee Balance</i> |
| 3. <i>Modified Bars Test of Dynamic Balance</i> | 8. <i>One Knee-Head To Floor</i> |
| 4. <i>Standing Balance Test</i> | 9. <i>Standing Balance Test</i> |
| 5. <i>One Foot- Touch Head</i> | |

8. TKJI (Tes Kesegaran Jasmani Indonesia)

Pengukuran terhadap tingkat kesegaran Jasmani peserta didik dapat dilakukan melalui beberapa tes disesuaikan dari tujuan pengukuran yang diharapkan akan di ukur. Salah satu tes lapangan (*field test*) dewasa ini yang cukup populer untuk

mengukur tingkat kebugaran jasmani peserta didik mulai dari tingkat SD, SMP, dan SMA/SMK adalah Tes Kebugaran Jasmani Indonesia (TKJI). Tes TKJI merupakan salah satu tes standar (tes baku) dengan dilengkapi petunjuk pelaksanaan tes dan norma penilaian tingkat kebugaran jasmani yang lengkap untuk berbagai tingkatan usia maupun jenis kelamin. Tes ini merupakan rangkaian tes (*battery test*) yang terdiri dari lima item tes. Dalam pelaksanaan tes TKJI, urutan pelaksanaan item tesnya tidak boleh diacak. Tes TKJI merupakan tes yang mengukur komponen biomotorik secara lengkap dalam satu rangkaian pelaksanaan tes. Adapun rangkaian item tes yang tertuang dalam TKJI adalah: tes kecepatan diukur dengan menggunakan tes lari 30 meter, 40 meter, 50 meter, dan 60 meter (sesuai dengan tingkat usia). Tes kekuatan otot perut diukur dengan menggunakan *sit-up* dan tes daya tahan otot lengan dan bahu diukur dengan bergantung siku tekuk (putri), dan bergantung angkat badan (putra). Tes Power diukur dengan menggunakan lompat vertikal (*vertical jump*), serta Tes Daya tahan Jantung dan paru-paru diukur dengan menggunakan tes lari 600 meter, 800 meter, 1000 meter, dan 1200 meter sesuai dengan tingkat usia dan jenis kelamin peserta didik.

Item-item TKJI selengkapnya dapat dilihat pada sajian bagan berikut.

Item-Item Tes TKJ

B. Usia 6 - 9 Tahun (Putra dan Putri)

1. Lari 30 meter
2. Bergantung siku tekak
3. Baring duduk (sit-up) 30 detik
4. Loncat tegak (vertical jump)
5. Lari 600 meter

Item-Item Tes TKJ

A. Usia 10 - 12 Tahun (Putra dan Putri)

1. Lari 40 meter
2. Bergantung siku tekak
3. Baring duduk (sit-up) 30 detik
4. Loncat tegak (vertical jump)
5. Lari 600 meter

Item-Item Tes TKJ

C. Usia 13 - 15 Tahun (Putra)

1. Lari 50 meter
2. Bergantung angkat tubuh 60 detik
3. Baring duduk (sit-up) 60 detik
4. Loncat tegak (vertical jump)
5. Lari 1000 meter

Item-Item Tes TKJ

D. Usia 13 - 15 Tahun (Putri)

1. Lari 50 meter
2. Bergantung siku tekak
3. Baring duduk (sit-up) 60 detik
4. Loncat tegak (vertical jump)
5. Lari 800 meter

Item-Item Tes TKJ

E. Usia 16 - 19 Tahun (Putra)

1. Lari 60 meter
2. Bergantung angkat tubuh 60 detik
3. Baring duduk (sit-up) 60 detik
4. Loncat tegak (vertical jump)
5. Lari 1200 meter

Item-Item Tes TKJ

F. Usia 16 - 19 Tahun (Putri)

1. Lari 60 meter
2. Bergantung siku tekak
3. Baring duduk (sit-up) 60 detik
4. Loncat tegak (vertical jump)
5. Lari 1000 meter

Gambar. Item tes TKJ



[Faint, illegible text]



BAB 10

TES-TES STANDAR DALAM PENDIDIKAN JASMANI DAN OLAHRAGA

A. Tes Kardiovaskuler (Harvard Step Test)

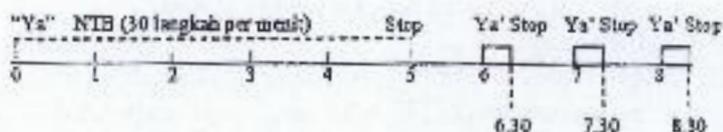
Pengukuran yang paling tua untuk kardiorespirasi adalah Harvard Step-Ups Test (1943) yang dibuat oleh Brouha. Nama Harvard ini diambil karena penelitian tes tersebut dilakukan di Universitas Harvard USA. Item tes Harvard adalah naik turun bangku selama 5 menit.

Pelaksanaan:

- Tinggi bangku 20 feet (50 cm). Irama langkah pada waktu naik turun bangku (NTB) = 30 langkah per menit. Jadi 1 langkah setiap 2 detik.
- Satu langkah terdiri dari 4 gerakan:
 - Hitungan 1: salah satu kaki diangkat (boleh kanan atau kiri terlebih dahulu, tetapi konsisten), kemudian menginjak bangku.
 - Hitungan 2: kaki kiri diangkat lalu berdiri tegak di atas bangku.
 - Hitungan 3: kaki pertama menginjak bangku (pada hitungan 1) diturunkan kembali ke lantai.
 - Hitungan 4: kaki yang lain diturunkan kembali ke lantai untuk berdiri tegak seperti sikap semula.
- Naik turun bangku dilakukan selama 5 menit. Saat aba-aba stop, tubuh harus dalam keadaan tegak. Kemudian duduk di bangku tersebut rileks selama 1 menit.

- Setelah itu hitung denyut nadinya selama 30 detik (dicatat sebagai denyut Nadi ke 1).
- 30 detik kemudian Denyut Nadi (DN) dihitung lagi selama 30 detik (dicatat sebagai Denyut Nadi ke 2).
- 30 detik kemudian Denyut Nadi (DN) dihitung lagi selama 30 detik (dicatat sebagai Denyut Nadi ke 3).
- Supaya irama langkah ajeg, maka digunakanlah **Metronome**.
- Apabila testi tidak kuat melakukan naik turun bangku selama 5 menit, maka waktu lama NTB tersebut dicatat, lalu diukur denyut nadinya sesuai dengan petunjuk pengambilan denyut nadi tersebut.
- Ganti langkah diperbolehkan asalkan tidak lebih dari 3 kali.

Adapun diagram waktu pelaksanaan tes Harvard adalah sebagai berikut:



INDEKS FITNESS (IF)

Long Form Formula = $\frac{\text{Lama waktu NTB dalam detik} \times 100}{2 (DN 1 + DN 2 + DN 3)}$
 (Rumus Panjang)

Short Form Formula = $\frac{\text{Lama waktu NTB dalam detik} \times 100}{(5,5 \times DN 1)}$
 (Rumus Pendek)

Kegunaan tes Harvard adalah untuk mengukur kapasitas umum dalam mengerjakan pekerjaan berat.

Menginterpretasikan tes Harvard ada 2 macam, yaitu:

1. **Rumus Panjang**

Apabila Indeks Fitnessnya	<	55	Jelek
		55 - 64	Rata-rata rendah
		65 - 79	Rata-rata
		80 - 89	Baik
		90 ke atas	Excellent (Baik Sekali)

2. **Rumus Pendek**

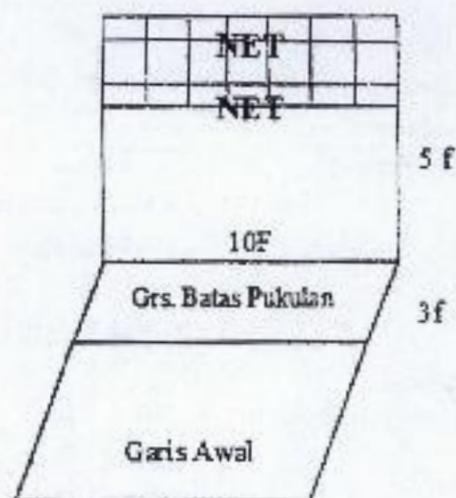
Apabila Indeks Fitnessnya	<	50	Jelek
		50 - 80	Rata-rata
		Di atas 80	Baik

DURASI	PEI (<i>Physical Endurance Index</i>)
Kurang dari 2 menit	25
2 sampai 3 menit	38
3 sampai 3 menit	48
3 sampai 4 menit	52
4 sampai 4 menit	55
4 sampai 5 menit	59

Validitas tes Harvard diperoleh dengan mengkorelasikan skor tes dengan VO₂ max

B. Tes Kecakapan Bermain Bulutangkis (Lockhart - MC Pherson)

Tujuan tes ini yaitu untuk mengukur keterampilan bermain bulutangkis. Tes ini didesain untuk mahasiswa putri (aslinya), namun dapat juga dipakai untuk mahasiswa putra dan untuk pelajar SLA dan SLP. Pelaksanaan tes tersebut sebanyak 3 kali dan setiap kali trial waktunya 30 detik.

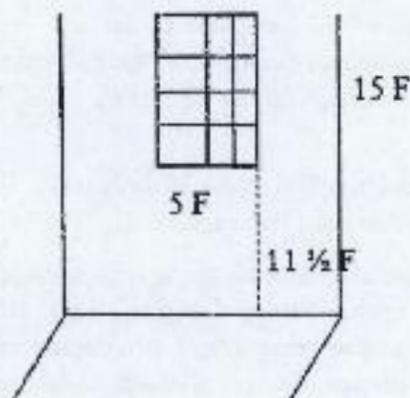


Pelaksanaan:

- Tes dimulai dengan pukulan servis ke arah tembok dilakukan dari belakang garis awal
- Bola yang memantul dari tembok di voli ke daerah sasaran sebanyak-banyaknya selama 30 detik
- Kalau bola atau *shuttlecock* tidak dapat dikuasai, testi mulai dengan pukulan servis dari belakang garis awal untuk melanjutkan tes sampai waktu habis
- Setiap bola yang divoli dari belakang garis batas pukulan dan masuk ke daerah sasaran dan pukulan tersebut sah diberi skor 1
- Skor tes adalah jumlah skor dari 3 kali kesempatan
- Pukulan servis tidak diberi skor
- **Validitas:** - 0,71 dari 3 hakim/judges (68 kasus)
- 0,60 dari pertandingan setengah kompetisi (27 kasus)
- 0,90 dari 3 judges dengan persentase kemenangan ($N = 27$)
- **Reliabilitas** = 0,90 (test - retest)

C. Tes Kecakapan Bermain Bola Voli (Braddy Volley Ball Test, 1945)

Lembar garis: 1 inchi



Tujuan tes ini yaitu untuk mengukur kecakapan umum dalam bermain bolavoli. Tes ini di desain untuk mahasiswa putra dan dilakukan selama 60 detik. Adapun pelaksanaan tes ini yaitu sebagai berikut:

- Testi berdiri menghadap sasaran dengan bolavoli di tangan, setelah ada aba-aba, testi mulai melempar bola ke tembok. Bola yang memantul dari tembok dipukul atau divoli ke daerah sasaran (bukan hanya dilakukan dengan *pass atas*)
- Apabila bola lunas, bola dapat dipegang lalu mulai lagi dengan melempar bola ke tembok untuk dipukul/ divoli sampai waktu habis.

Penilaian:

- Setiap bola yang memantul dari tembok, lalu dipukul/ divoli sah sesuai dengan peraturan permainan, dan bola tersebut termasuk ke daerah sasaran serta mengenai garis

batas daerah sasaran diberi skor 1.

- Skor tes adalah jumlah skor selama 60 detik.
- Tes Braddy hanya dilakukan 1 kali kesempatan.
- Bola yang dilempar ke tembok tidak diberi skor.
- Validitas tes = 0,86 (*Criterion Measure of Subjective Rating*)
- Dilihat pada waktu bertanding.
- Reliabilitas = 0,93 (dengan tes - retest)

Catatan: Tes ini dapat dipergunakan untuk siswa SLTA atau SLTP, tetapi daerah sasarannya dapat diturunkan.

D. Tes Kecakapan Bermain Bola Basket STO (Sekolah Tinggi Olahraga)

Tes ini bertujuan untuk mengukur kecakapan bermain bola basket, untuk menggolongkan siswa sesuai tingkat keterampilannya dan sebagai salah satu dasar pemberian nilai pendidikan olahraga. Tes ini disusun untuk pelajar SLTA putra.

Tes ini merupakan tes battery dan terdiri atas dengan tiga item tes, yaitu:

1. Memantulkan bola ke tembok
2. Menggiring bola
3. Menembak bola ke ring basket selama 1 menit

Validitas tes sebesar 0,804 dicari dengan mengkorelasikan hasil penilaian 5 orang hakim pada waktu bermain tes battery.

Reliabilitas tes sebesar 0,893 yang dicari dengan cara mengkorelasikan hasil tes hari pertama dengan tes hari berikutnya.

Alat-alat dan perlengkapan:

1. Bola basket
2. Kursi makan
3. Dinding tempok yang rata atau papan
4. Stopwatch

5. Kapur
6. Blangko
7. Alat-alat tulis

Keterangan:

1. Memantulkan Bola ke Tembok

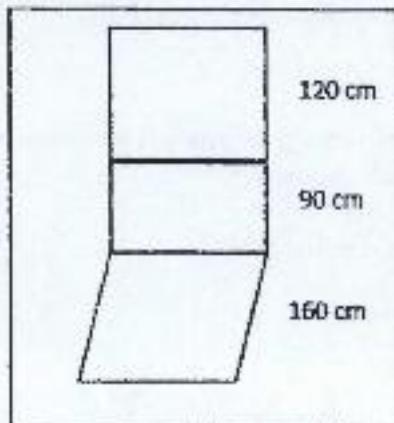
Tujuan: untuk mengukur kecakapan menolak atau melemparkan dan menangkap bola.

Alat-alat:

- Sebuah bola
- Sasaran pada tembok
- Stopwatch atau jam yang ada sekonannya
- Blangko dan alat tulis

Petugas: pengambil waktu dan pencatat

Pelaksanaan:



- Testi melemparkan bola dari belakang garis batas dengan menggunakan lemparan samping atau lemparan atas kepala
- Setiap testi mendapat kesempatan sebanyak 10 kali
- Jika bola jatuh diantara garis Pada aba-aba "Siap", anak

berdiri di belakang garis batas, menghadap ke arah sasaran atau tembok., bola pada ke dua tangan. Pada aba-aba "Ya", bola dipantulkan ke arah sasaran sebanyak-banyaknya selama 15 detik. Bola tidak boleh di voli.

- Pantulan yang sah apabila bola memantul pada garis atau daerah sasaran dan dilakukan pada atau dari belakang garis batas. Dalam memantulkan atau melemparkan bola boleh dengan satu atau dua tangan. Apabila bola tidak terkuasai dan berada jauh dari garis batas, bola segera diambil dan dengan cepat kembali siap di belakang garis batas untuk memulai pantulan berikutnya. Pelaksanaan berhenti setelah aba-aba "Stop".

Penilaian:

Hitunglah jumlah pantulan yang sah selama 15 detik.

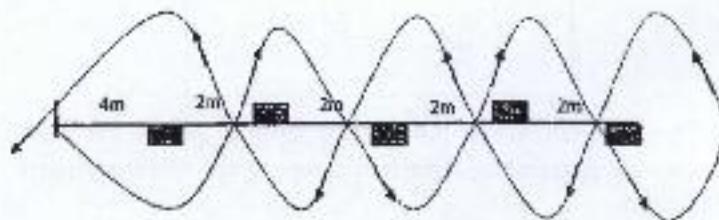
2. Menggiring Bola

Tujuan: mengukur kecepatan dan kecakapan menggiring bola maupun kelincahan mengubah arah.

Alat-alat:

- Bola
- Stopwatch atau jam yang ada sekonannya
- Kursi makan 5 buah
- Peluit
- Blangko dan alat tulis

Pelaksanaan:



- Pada aba-aba "Siap", testi berdiri di belakang garis *start*, letakkan bola di tengah-tengah garis *start*.
- Setelah aba-aba "Ya", testi segera mengambil bola dan menggiringnya sesuai dengan arah atau lintasan yang ditentukan dalam gambar, sampai kembali melewati garis *finish*.
- Menggiring bola boleh dengan berganti tangan, asalkan sesuai dengan peraturan permainan bola basket.
- Setiap kursi harus dilampaui dengan menggiring bola. Pada saat melampaui garis *finish*, bola harus tetap digiring. Garis *start* juga merupakan garis *finish*.
- Apabila saat menggiring bolanya mental jauh, ulangi- lah tes tersebut dengan segera.
- Apabila pada saat menggiring bola tidak memantul atau tidak terkuasai, maka bola boleh dipegang dan segera digiring lagi.

Penilaian:

Kecepatan menggiring dihitung dari saat aba-aba "Ya" sampai melampaui garis *finish*. Kecepatan menggiring dihitung sepersepuluh detik.

3. Menembak Bola Ke Ring Basket Selama Menit

Tujuan: untuk mengukur ketepatan dan ketelitian menembak ke dalam ring basket

Alat – alat:

- Bola basket
- Stopwatch atau jam yang ada sekonannya
- Kursi makan 5 buah
- Jaring Peluit
- Blangko dan alat tulis

Pelaksanaan:

- Pada aba-aba "Siap", testi berdiri bebas di dekat dan ke arah basket df bola di tangan.
- Setelah aba-aba "Ya", segera menembakkan bola ke dalam basket sebanyak-banyaknya selama 1 menit.
- Apabila bola mentol jauh dan tidak terkuasai lagi, bola segera diambil dan dengan lari atau berjalan, kembali secepat-cepatnya ke arah basket untuk kemudian menembakkan lagi ke dalam basket.

Penilaian:

Setiap kali bola masuk ke dalam basket mendapat nilai satu. Jumlah bola yang masuk ke dalam basket selama 1 menit adalah nilai yang diperoleh. Apabila waktu aba-aba "Stop", telah diberikan, sedangkan bola sudah lepas dari tangan dan masuk ke ring tetap mendapatkan satu nilai.

Tabel. Skala T Untuk Tes Kecakapan Bermain Bola Basket SMU Putra

T Score	Memantulkan Bola	Menggiring Bola	Menembak Permenit	T Score
I	II	III	IV	V
80	26	9,0 - 9,1	27	80
79		9,2 - 9,3		79
78		9,4 - 9,5	26	78
77	25	9,6 - 9,7		77
76		9,8 - 9,9	25	76
75		10,0 - 10,1		75
74		10,2 - 10,3	24	7
73	24	10,4 - 10,5		73
72		10,6 - 10,7	23	72
71		10,8 - 10,9		71

70	23	11,0 - 11,1	22	70
69		11,2 - 11,3		69
68		11,4 - 11,5	21	68
67	22	11,6 - 11,7		67
66		11,8 - 11,9	20	66
65		12,0 - 12,1		65
64		12,2 - 12,3	19	64
63	21	12,4 - 12,5		63
62		12,6 - 12,7	18	62
61		12,8 - 12,9		61
60	20	13,0 - 13,1	17	60
59		13,2 - 13,3		59
58		13,4 - 13,5	16	58
57	19	13,6 - 13,7		57
56		13,8 - 13,9	15	56
55		14,0 - 14,1		55
54		14,2 - 14,3	14	54
53	18	14,4 - 14,5		53
52		14,6 - 14,7	13	52
51		14,8 - 14,9		51
50	17	15,0 - 15,1	12	50
49		15,2 - 15,3		49
48		15,4 - 15,5	11	48
47	16	15,6 - 15,7		47
46		15,8 - 15,9		46
45		16,0 - 16,1	10	45
44		16,2 - 16,3		44
43	15	16,4 - 16,5	9	43

42		16,6 – 16,7		42
41		16,8 – 16,9		41
40	14	17,0 – 17,1	8	40
39		17,2 – 17,3		39
38		17,4 – 17,5	7	38
37	13	17,6 – 17,7		37
36		17,8 – 17,9	6	36
35		18,0 – 18,1		35
34		18,2 – 18,3	5	34
33	12	18,4 – 18,5		33
32		18,6 – 18,7	4	32
31		18,8 – 18,9		31
30	11	19,0 – 19,1	3	30
29		19,2 – 19,3		29
28		19,4 – 19,5		28
27	10	19,6 – 19,7	2	27
26		19,8 – 19,9		26
25		20,0 – 20,1		25
24		20,2 – 20,3		24
23	9	20,4 – 20,5		23
22		20,6 – 20,7		22
21		20,8 – 20,9		21
20	8	21,0 – 21,01		20

Tabel. Norma Nilai Kecakapan Bermain Bola Basket

Penggolongan	Jumlah T Score
Baik Sekali	222 ke atas
Baik	193 - 221
Cukup	165 - 192
Sedang	136 - 164
Kurang	107 - 135
Kurang Sekali	79 - 106
Jelek	78 ke bawah

Cara Menilai:

Jumlah bola yang memantul ke tembok, waktu menggiring bola sampai 0,1 detik dan jumlah yang masuk dalam satu menit adalah hasil kasar atau *raw score* dari setiap item tes. Dengan menggunakan tabel tes kecakapan bermain bola basket, setiap hasil kasar, diubah menjadi *T Score*. Tiga macam *T Score* dijumlahkan dan merupakan nilai kecakapan bermain bola basket.

Jadi nilai kecakapan = Jumlah *T Score*

Hasil akhir kecakapan selanjutnya disesuaikan atas dasar norma yang berlaku.

Contoh:

Seorang pelajar putra SMU kelas II, dari hasil tes ternyata mempunyai hasil kasar sebagai berikut:

Memantulkan bola ke tembok 18

Menggiring bola 12,7 detik; dan

Menembak selama 1 menit 20

Masing-masing hasil kasar diubah menjadi *T Score* hasilnya sebagai berikut:

Menantulkan bola ke tembok	T Score 53
— Menggiring bola 12,7 detik; dan	T Score 62
Menembak selama 1 menit 20	T Score 66
Jumlah	181

Nilai kecakapan = 181, selanjutnya dilihat dalam norma penilaian berlaku, ternyata siswa SMU tersebut termasuk mempunyai keterampilan bermain bola basket dalam kategori **CUKUP**.

Catatan:

1. Waktu yang dibutuhkan guru untuk mengetes sejumlah kurang lebih 40 siswa dengan satu station memantulkan bola, satu station menggiring bola, dan dua station memasukkan bola selama satu menit adalah satu jam pelajaran. Perlu diperhatikan bahwa dalam pengetesan ini guru harus menunjuk para siswa yang dipandang dapat membantunya, untuk itu harus diadakan latihan sebelumnya.
2. Urutan pelaksanaan item-item tes harus sesuai dengan tuntunan, dan tidak boleh dirubah-rubah.
3. Siswa atau testi tidak boleh mengadakan percobaan terlebih dahulu.

D. Johnson Basket Ball Test (1934)

Johnson basket ball test terdiri atas 3 item, yaitu:

1. Menembakkan bola ke basket
2. Melemparkan bola ke arah sasaran
3. Menggiring bola

Tes ini diperuntukkan bagi anak *high school boys* (usia SMP dan SMU).

Alat dan Perlengkapan:

1. Lapangan bola basket yang ada ringnya
2. Stopwatch, bola basket
3. Tembok atau daerah sasaran dengan sasaran 70 x 70 cm, di depan tembok ada ruangan kosong dengan jarak kurang lebih 15 meter dengan bangku (hurdle) banyaknya 5 buah dengan tinggi 6 feet/ 1,80 meter
4. Blangko pencatatan

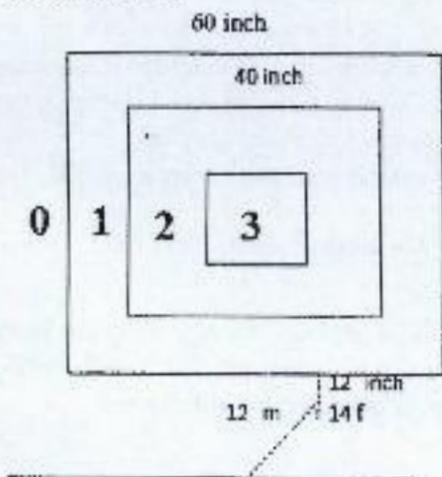
Pelaksanaan:

1. Menembakkan bola ke basket

Testi siap di sembarang tempat dekat ring dengan memegang bola basket. Setelah aba-aba "Ya", testi menembakkan bola ke ring selama 30 detik. Penilaiannya: Jumlah tembakan yang masuk selama 30 detik dicatat sebagai skor tes.

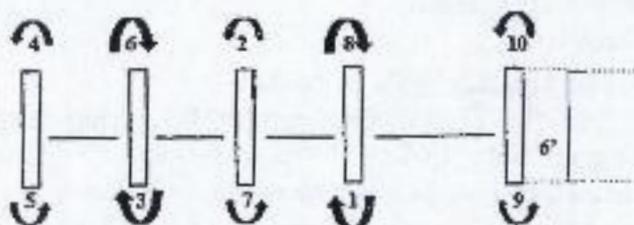
2. Melemparkan bola ke arah sasaran

Pelaksanaan:



- Testi melemparkan bola dari belakang garis batas dengan menggunakan lempatan samping atau lemparan atas kepala
- Setiap testi mendapat kesempatan sebanyak 10 kali
- Jika bola jatuh diantara garis batas, skor diambil yang terbesar
- Skor tertinggi 30 dan terendah 0

3. Menggiring bola



Pelaksanaan:

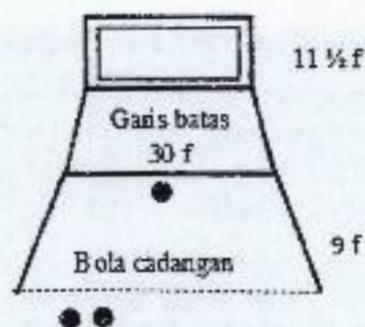
- Tes ini dilakukan selama 30 detik. Jadi berapa banyaknya gawang atau *hurdle* yang dapat dilewati oleh testi
- *Trial* masing-masing testi hanya 1 kali
- **Validitas** tes 0,65 – 0,79 yang diperoleh dari uji coba 180 siswa yang dibagi menjadi 2 kelompok, kelompok baik terdiri dari 50 siswa laki-laki dan kelompok yang kurang baik terdiri atas 130 siswa laki-laki
- **Reliabilitas** tes berkisar antara 0,73 sampai 0,80.

E. Tes Kecakapan Bermain Sepak Bola (Mc. Donalds Soccer Test)

Tujuan: untuk klasifikasi pemain dan dirancang untuk mahasiswa putra. Sebelum melakukan tes, testi boleh melakukan *warming up*, tetapi tidak boleh melakukan tes.

Pelaksanaan:

Gambar lapangan:



- Testi berdiri di belakang garis batas dengan menghadapi bola yang berada di tengah-tengah lapangan tersebut. Setelah ada aba-aba "Ya", testi berusaha menyepak bola ke sasaran.
- Bola pantul disepak secara terus-menerus selama 30 detik. Semua cara menyepak diperbolehkan dan di dalam mengontrol bola boleh memakai seluruh anggota badan termasuk tangan. Pada saat menyepak bola, kaki tumpu harus berada di belakang garis batas, agar mendapat skor. Demikian juga bola pada saat penekanan, harus berada di belakang garis batas. Apabila bola tidak dapat dikontrol, testi boleh mengambil bola cadangan untuk melanjutkan tes sampai waktu habis. Bola dipompa dengan ukuran kekenyalan 13 pound.
- Setiap testi memperoleh kesempatan melakukan tes 4 x¹ (@ 30 detik). Jumlah dari 3 skor trial terbaik menjadi skor tes (dalam buku Donald K Matthews) bola harus pantul harus divoli. Oleh karena itu diharapkan dalam tes ini ada waktu istirahat.

- Validitas tes = 0,85 (dengan *subjective rating of three coach*) saatb testi bertanding dai 53 siswa.
- Reliabilitas tes = tidak dilaporkan.

F. Tes Tenis Meja Dari STO Yogyakarta

Tujuan: untuk mengukur kecakapan memukul bola sebanyak-banyaknya ke papan bagi pelajar SMU putra.

Reliabilitas tes: koefisien reliabilitas 0,738 diperoleh dari hasil tes-retes dari sampel sebanyak 110.

Validitas tes: sebanyak 0,615 diperoleh hasil perbandingan setengah kompetisi 30 pelajar yang diambil secara random dari sampel sebanyak 110 siswa tersebut di atas.

Alat-alat yang dipergunakan terdiri atas:

- Sebuah stopwatch
- Lima buah bola tenis meja
- Sebuah bat
- Sebuah meja tenis meja yang dapat dilipat
- Sebuah kotak karton berukuran 10 x 5 x 3 cm yang dapat ditempelkan dengan pines pada sudut samping kiri belakang meja.
- Dinding atau tiang untuk sandaran bagian meja tenis meja yang didirikan tegak lurus pada bagian meja yang horizontal
- Pita kertas yang lebarnya 2cm
- Blangko dan alat tulis untuk mencatat hasil tes

Tanda garis:

Sebuah garis dari pita selebar 2 cm, dibuat pada bagian meja yang didirikan tegak lurus, sejajar dengan bagian meja yang horizontal dan berjarak 15 cm dari permukaan meja.

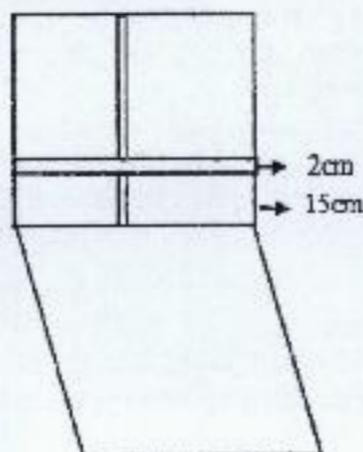
Keterangan Validitas:

Validitas bagi 10 orang kelompok pemain-pemain terbaik menghasilkan validitas $-0,412$, sedangkan validitas bagi 10 orang pemain yang kurang baik adalah $0,782$.

Petugas:

1. Seorang pengambil waktu yang memberikan aba-aba "Ya" dan "Stop"
2. Seorang penghitungan jumlah pantulan yang sah selama 30 detik dan sekaligus mencatat hasilnya.
3. Sekurang-kurangnya seorang pembantu untuk mengambil bola yang tidak dikuasai oleh testi.

Pelaksanaan Tes:



- Testi berdiri di belakang atau lanjutan bagian meja yang horizontal, dengan sebuah bat dan bola di tangan. Pada aba-aba "Ya", testi menjatuhkan bola di atas meja dan kemudian memukul bola ke dalam bagian yang didirikan tegak lurus terhadap bagian meja yang horizontal.

- Testi berusaha memantulkan bola sebanyak-banyaknya dalam waktu 30 detik. Apabila testi tidak dapat menguasai bola, ia dapat mengambil bola yang tersedia di kotak, menjatuhkannya di meja dan melanjutkan usaha memantulkan bola sebanyak-banyaknya dalam sisa waktu yang tersedia. Seorang pembantu mengambil bola yang tidak dikuasai testi dan memasukkannya kembali ke dalam kotak.
- Pantulan dinyatakan tidak sah apabila:
 - a. Bola divoli
 - b. Testi bertelekan dengan tangannya yang bebas pada meja waktu memukul bola
 - c. Bola mengenai bagian meja yang tegak di bawah garis
 - d. Melakukan pukulan servis pada waktu mulai tes
 - e. Memukul bola setelah bola memantul lebih dari satu kali pada meja yang horizontal
 - f. Memukul bola lebih dari satu kali dengan kaki bertumpu di samping meja
- Testor berdiri tegak dengan meja dan menghitung jumlah pukulan yang sah selama 30 detik dan mencatatnya. Kepada testi diberikan kesempatan melakukan tes 3x dengan istirahat 10 detik setiap selesai melakukan tes.

Penilaian:

Skor dari setiap trial adalah jumlah pantulan yang sah selama 30 detik. Skor tes adalah jumlah yang terbanyak dari ketiga trial tersebut.

Catatan: pada waktu aba-aba "Stop", sudah diberikan, tetapi bola sudah dipukul dan pantulan bola sah, maka tetap dihitung.

Tabel. Klasifikasi Kecakapan Bermain Tenis Meja

Skor	Klasifikasi
53 ke atas	Baik
44 - 52	Cukup
26 - 43	Sedang
17 - 25	Kurang
16 ke bawah	Kurang Sekali

H. Tes Keterampilan Bermain Tenis Lapangan (Tes Tenis Dyer, 1938)

Tujuan : Mengukur kecakapan umum dalam bermain tenis.

Kegunaan: Tes ini dapat digunakan untuk pengelompokan dan menilai kemajuan dalam belajar tenis.

Testi : tes ini didesain untuk mahasiswa putri, tetapi ternyata dapat dipakai untuk mengukur kecakapan bermain tenis pelajar SMP, SMA putra maupun putri dan mahasiswa putra.

Alat : raket tenis, satu dosin bola tenis, sebuah stopwatch, pita, pensil, dan blangko untuk mencatat skor.

Arena Tes: tembok atau papan yang permukaannya rata, lebar 15 feet, tinggi 10 feet, dan ruang bebas di depannya. Pada permukaan tembok dibuat garis net sejajar dengan lantai, lebar garis net 3 inci dengan bagian tepi atas berjarak 3 feet dari lantai. Sebuah garis batas pukulan di lantai sejarak 5 feet

dan sejajar tembok.

Pelaksanaan

- Testi berdiri di belakang garis batas pukulan, memegang raket dan dua bola.
- Setelah aba-aba pelaksanaan diberikan, testi memantulkan bola ke lantai kemudian memukulnya ke arah tembok diarahkan ke daerah sasaran di atas garis net. Bola yang memantul dari tembok dipukul kembali ke arah daerah sasaran, demikian dilakukan berulang-ulang selama 30 detik.
- Untuk mempertahankan agar dapat memukul bola, testi boleh melangkah maju melampaui garis batas pukulan dan juga diperbolehkan melakukan pukulan voli.
- Bola yang dipukul dari depan garis batas pukulan tidak dihitung (tidak diskor). Kalau bola tidak dapat dikuasai, testi boleh menggunakan bola cadangan yang disediakan di samping kanan atau kiri arena.
- Penggunaan bola cadangan selama tes tidak dibatasi. Dalam menggunakan bola baru harus dilakukan seperti waktu mulai tes.
- Kesempatan melakukan tes ini 3 kali, setiap kali selama 30 detik.

Skoring: setiap bola yang dipukul dari belakang garis batas pukulan dan masuk ke daerah sasaran atau mengenai garis net diberi skor satu. Jumlah skor dari 3 kali kesempatan merupakan skor tes.

Skala Penilaian: skala penilaian Tes Tennis Dyer disusun berdasarkan skor tes dari 672 mahasiswa putri dan pendidikan jasmani sebagai mata pelajaran utama.

Tabel. T-Scale for New Method Scoring in Dyer Test

T-Scale	Test Score						
100	67	75	50	50	33	25	16
99	66	74	49	49	32	24	15
98		73		48		23	
97	65	72	48	47	31	22	14
96	64	71	47	46	30	21	13
95		70		45		20	
94	63	69	46	44	29	19	12
93	62	68	45	43	28	18	11
92		67	44	42	27	17	10
91	61	66		41		16	
90	60	65	43	40	26	15	9
89	59	64	42	39	25	14	8
88		63		38		13	
87	58	62	41	37	24	12	7
86	57	61	40	36	23	11	6
85		60		35		10	
84	56	59	39	34	22	9	5
83	55	58	38	33	21	8	4
82		57		32		7	
81	54	56	37	31	20	6	3
80	53	55	36	30	19	5	2
79		54		29		4	
78	52	53	35	28	18	3	1
77	51	52	34	27	17	2	
76		51		26		1	

From J. T. Dyer: Revision of backboard test of tennis ability, *Research Quarterly*, 9:25-31, 1938. Courtesy of AAHPER.

Tabel. Ability Ratings For Dyer Tennis Test

RATING*	COLLEGE WOMEN	WOMEN MAJORS PHYSICAL EDUCATION
Superior	46 & up	79 & up
Good	38 - 45	58 - 78
Average	29 - 37	35 - 57
Poor	21 - 28	13 - 34
Inferior	20 & below	12 - below

Based on scores of 672 women students and physical education.

From Miller, Wilma K.: Achievement Levels in Tennis Knowledge and Skill for Woman Physical Education Students, Doctoral dissertation, Boomington Indiana University, 1952

Validitas Tes: Tes Dyer yang asli sebelum direvisi validitasnya antara 0,85 sampai 0,90 dengan cara mengkorelasikan skor tes dengan judges' ratings. Validitas antara 0,85 sampai 0,92 adalah korelasi antara skor tes Dyer yang telah direvisi dengan ranking hasil pertandingan setelah kompetisi

Reliabilitas Tes: Dengan metode tes ulang yang dilakukan terhadap 14 sampai 37 putri sebagai sampel, reliabilitas tes Dyer yang telah direvisi adalah 0,86, 0,87, 0,87 dan 0,92.

Koefisien reliabilitas te Dyer yang asli adalah 0,90.

Keterangan Tambahan

Pada tahun 1935 pada saat pertama kali Dyer memperkenalkan tes Tennis ini tanpa garis batas pukulan. Baru pada tahun 1938 Dyer merevisi tesnya dengan menggunakan garis batas pukulan yang berjarak 5 feet. Juga diadakan perubahan dalam pemberian skor dan penyediaan bola.

Fox menerapkan Tes Tennis Dyer ini pada mahasiswa putri yang sedang mulai belajar tenis. Korelasi antara skor tes dari mahasiswa putri itu dengan ranking keterampilannya dalam *forehand drive*, *backhand drive* dan *serve* adalah 0,53.

Koski menggunakan garis batas pukulan 28 feet dari tembok untuk mahasiswa putra. Korelasi antara skor tes dengan ranking pertandingan adalah 0,51 sampai 0,68. Perubahan tes Dyer terutama pada jarak garis batas pukulan dan lama waktu melakukan pukulan beruntun. Ini adalah pokok isi bahwa tes Dyer membedakan dengan baik kecakapan Tennis pada tingkat *advance* (lanjut). Sebagai akibatnya, perlu diadakan perbedaan jarak garis batas pukulan yang diperbesar, supaya tes itu menjadi lebih sesuai dengan tingkat pemula.

Berdasarkan hasil studi dengan tes Dyer, kelihatannya jarak 20 feet akan lebih sesuai dengan tingkat pemula, dan jarak kurang lebih 30 feet akan lebih berguna untuk mengukur kemampuan tenis bagi pemain tingkat menengah dan *advance*. Suatu masalah dengan garis batas 5 feet yang direkomendasikan dalam revisi yang dilakukan pada tahun 1938 adalah timbulnya teknik pukulan yang buruk. Jarak yang pendek tidak mendukung pemakaian bentuk pukulan yang baik.

I. Modifikasi Tes Keterampilan Bermain Tennis Lapangan Bagi Mahasiswa (NgatmaN, 2000)

1. Tes *Serve*

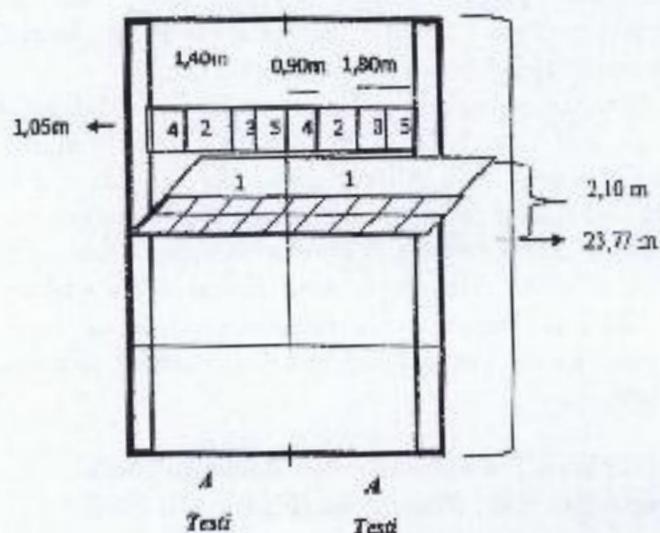
Tujuan : untuk mengukur tingkat keterampilan mahasiswa dalam melakukan *serve* diarahkan pada sasaran tertentu yang dapat menyulitkan lawan dalam permainan tenis.

Alat : raket tenis, 30 buah bola tenis, net, tali, dua buah

tiang setinggi 2,25 meter, pensil, dan blangko untuk mencatat skor.

Testor : seorang pencatat skor tes dan seorang pengawas jatuhnya bola ke sasaran.

Arena Tes: sebuah lapangan tenis, tepat di atas net direntangkan tali yang sejajar dengan net berjarak 2,10 meter dari lantai. Daerah *serve* untuk setiap sisi (sebelah kanan atau sebelah kiri) dibagi menjadi 5 daerah sasaran yang bernilai 1, 2, 3, 4, dan 5 sebagaimana terlihat pada gambar 6. Arena tes serve dapat digambar sebagai berikut:



Gambar. Arena Tes serve

Pelaksanaan

- Testi berdiri di belakang garis belakang (base line) siap untuk melaksanakan tes.

- Sebelum tes dimulai, testi diberi kesempatan melakukan pemanasan secukupnya. Enam kali percobaan latihan diijinkan bagi testi sebelum tes sebenarnya dilakukan (sebaiknya 3x dari sebelah kanan dan 3x dari sebelah kiri).
- Untuk mulai melaksanakan tes *serve*, testi berdiri di belakang base line, melakukan pukulan *serve* 10 kali kesempatan (5x dari sebelah kanan dan 5x dari sebelah kiri) diarahkan ke daerah sasaran di dalam petak servis.
- Bola harus lewat antara net dan tali di atasnya yang berjarak 2,10 meter dari lantai. Lambungan bola yang tidak enak boleh tidak dipukul. Pukulan yang tidak mengenai bola atau bola membentur net sebagai satu pukulan *serve*.

Penskoran: Setiap pukulan *serve* bola lewat di antara net dan tali dan bola tersebut jatuh di daerah sasaran, memperoleh skor sebesar nilai daerah sasaran. Apabila bola lewat di atas tali dan jatuh di daerah sasaran memperoleh skor separuh nilai daerah sasaran. Jika bola jatuh pada garis batas antara dua daerah sasaran diberi skor sebesar nilai daerah sasaran yang besar. Skor akhir tes adalah jumlah nilai dari 10 kali pukulan *serve* (5x dari sebelah kanan dan 5x dari sebelah kiri).

2. Tes *Forehand Groundstrokes*

Tujuan: untuk mengukur tingkat keterampilan mahasiswa dalam melakukan *forehand groundstrokes* yang diarahkan pada sasaran tertentu sehingga dapat menyulitkan lawan dalam permainan tenis.

Alat: raket tenis, 30 buah bola tenis, net, tali, dua buah tiang setinggi 2,25 meter, pensil, dan blangko untuk mencatat skor.

Testor : seorang pencatat skor tes, seorang pengawas jatuhnya bola ke sasaran dan seorang pengumpan bola (*feeder*)

Pelaksanaan

- Tes dimulai dengan testi berdiri di tengah-tengah base line (*centre mark*) menghadap ke arah daerah sasaran yang berada di seberang net siap untuk melakukan tes.
- Sebelum tes dimulai testi diberi kesempatan untuk mencoba melakukan *forehand groundstrokes* terhadap bola yang diumpangkan dari belakang seberang net sebanyak 5 kali.
- Dalam pelaksanaan tes, testi melakukan *forehand groundstrokes* sebanyak 6 kali terhadap bola yang diumpangkan oleh seorang pengumpangan dari garis belakang seberang net dan diusahakan bola yang diumpangkan itu jatuh sedikit di belakang garis *service*.
- Bola yang diumpangkan apabila tidak enak, boleh tidak dipukul. Bola yang dipukul diusahakan lewat lewat di antara net dan tali dan jatuh di dalam daerah sasaran.
- Pukulan yang tidak kena, bola yang menyangkut net atau bola jatuh di luar daerah sasaran dihitung sebagai satu pukulan. Apabila bola menyentuh net dan jatuh di daerah sasaran pukulan diulang.

Penskoran: setiap *forehand groundstrokes* bola lewat di antara net dan tali dan bola tersebut jatuh di daerah sasaran, memperoleh skor sebesar nilai daerah sasaran. Apabila bola lewat di atas tali dan jatuh di daerah sasaran, memperoleh skor separuh bilai daerah sasaran.

Jika bola jatuh pada garis batas dua daerah sasaran memperoleh skor sebesar nilai daerah sasaran yang besar. Skor akhir tes adalah jumlah skor dari 6 kali *forehand groundstrokes*.

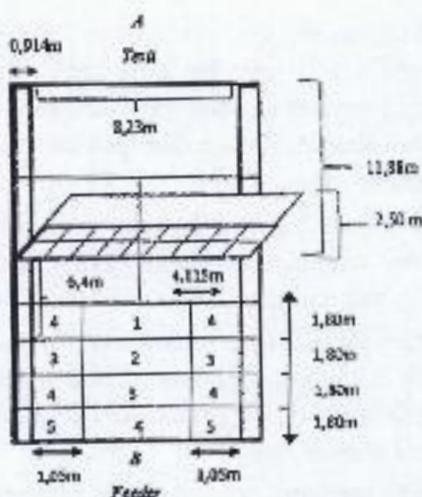
3. Tes *Backhand Groundstrokes*

Tujuan: untuk mengukur tingkat keterampilan mahasiswa dalam melakukan *backhand groundstrokes* yang diarahkan pada sasaran tertentu sehingga dapat menyulitkan lawan dalam permainan tenis.

Alat: raket tenis, 30 bola tenis, net, tali, dua buah tiang setinggi 2,25 meter, pensil dan blangko untuk mencatat skor.

Testor : seorang pencatat skor tes, seorang pengawas jatuhnya bola ke sasaran dan seorang pengumpan bola (*feeder*).

Arena Tes



Gambar 8. Arena tes *Backhand Groundstrokes*

Sebuah lapangan tenis, tepat di atas net direntangkan tali sejajar dengan net yang berjarak 2,50 meter dari lantai. Di dalam lapangan tenis dibuat empat buah garis sejajar dengan garis belakang (*base line*), jarak di antara dua buah garis juga dibuat dengan jarak masing-masing 1,05 meter dari garis permainan tunggal sebelah kanan (*right single side line*)

dan jarak 1,05 meter dari garis permainan tunggal sebelah kiri (*left single side line*). Garis-garis tersebut berpotongan membentuk bidang daerah sasaran yang bernilai 1, 2, 3, 4, dan 5 seperti gambar 8. Arena tes *backhand groundstrokes* dapat dilihat pada gambar di atas.

Pelaksanaan

- Tes dimulai dengan testi berdiri di tengah-tengah base line (centre mark) menghadap ke arah daerah sasaran yang berada di seberang net siap untuk melakukan tes.
- Sebelum tes dimulai testi diberi kesempatan mencoba melakukan *backhand groundstrokes* terhadap bola yang diumpankan dari garis belakang seberang net sebanyak 5 kali.
- Dalam pelaksanaan tes, testi melakukan *backhand groundstrokes* sebanyak 6 kali erhadap bola yang diumpankan oleh seorang pengumpaan (*feeder*) dari garis belakang seberang net dan diusahakan bola yang diumpankan itu jatuh sedikit di belakang garis *service*.
- Bola yang diumpankan apabila tidak enak, boleh tidak dipukul. Bola yang dipukul diusahakan lewat di antara net dan tali dan jatuh di dalam daerah sasaran. Pukulan yang tidak kena, bola menyangkut net atau bola jatuh di luar daerah sasaran dihitung sebagai satu pukulan.
- Apabila bola menyentuh net dan jatuh di daerah sasaran pukulan diulang.

Penskoran: setiap *backhand groundstrokes* bola lewat di antara net dan tali dan bola tersebut jatuh di daerah sasar memperoleh skor sebesar nilai daerah sasaran. Apabila lewat di atas tali dan jatuh di daerah sasaran, maka nilai separuh nilai daerah sasaran.

Jika bola jatuh pada garis batas dua daerah sasaran memperoleh skor sebesar nilai daerah sasaran yang besar. Skor akhir tes adalah jumlah skor dari 6 kali *backhand groundstrokes*.

4. Tes *Volley Forehand*

Tujuan: untuk mengukur tingkat keterampilan mahasiswa dalam melakukan *volley forehand* yang diarahkan pada sasaran tertentu sehingga dapat menyulitkan lawan dalam permainan tenis.

Alat: raket tenis, 30 bola tenis, net, tali, dua buah tiang setinggi 2,25 meter, pensil dan blangko untuk mencatat skor.

Testor : seorang pencatat skor tes, seorang pengawas jatuhnya bola ke sasaran dan seorang pengumpan bola (*feeder*).

Arena Tes: sebuah lapangan tenis, pada lapangan tenis tersebut dibuat empat garis sejajar dengan garis belakang (*base line*), jarak antara dua buah garis yang berdekatan adalah 1,80 meter. Selain itu dua garis juga dibuat dengan jarak masing-masing 1,05 meter: dari garis pemain tunggal sebelah kanan (*right single side line*) dan berjarak 1,05 dari garis permainan tunggal sebelah kiri (*left single side line*). Garis-garis tersebut berpotongan membentuk bidang daerah sasaran yang bernilai 1, 2, 3, 4, dan 5 seperti pada gambar 9.

Arena tes *volley forehand* dapat digambarkan sebagai berikut:

yang di *volley* mengenai net tetapi memantul ke sasaran, percobaan diulang.

Penskoran: setiap bola yang divoli dengan pukulan *forehand* jatuh di dalam daerah sasaran memperoleh skor sebesar nilai daerah sasaran. Apabila bola jatuh pada garis batas daerah sasaran memperoleh skor sebesar nilai daerah sasaran yang besar. Jika bola meyangkut di net atau bola jatuh di luar lapangan diberi skor nol. Skor akhir tes adalah jumlah skor dari 5 kali pukulan *volley forehand*.

5. Tes *Volley Backhand*

Tujuan: untuk mengukur tingkat keterampilan mahasiswa dalam melakukan *volley forehand* yang diarahkan pada sasaran tertentu sehingga dapat menyulitkan lawan dalam permainan tenis.

Alat: raket tenis, 30 bola tenis, net, tali, dua buah tiang setinggi 2,25 meter, pensil dan blangko untuk mencatat skor.

Testor : seorang pencatat skor tes, seorang pengawas jatuhnya bola ke sasaran dan seorang pengumpan bola (*feeder*).

Arena Tes: sebuah lapangan tenis, pada lapangan tenis tersebut dibuat empat garis sejajar dengan garis belakang (*base line*), jarak antara dua buah garis yang berdekatan adalah 1,80 meter. Selain itu dua garis juga dibuat dengan jarak masing-masing 1,05 meter dari garis pemain tunggal sebelah kanan (*right single side line*) dan berjarak 1,05 dari garis permainan tunggal sebelah kiri (*left single side line*). Garis-garis tersebut berpotongan membentuk bidang daerah sasaran yang bernilai 1, 2, 3, 4, dan 5 seperti pada gambar 10. Arena tes *volley backhand* dapat digambarkan sebagai berikut.

- Bola umpan yang tidak enak boleh tidak dipukul. Bola yang di *volley* mengenai net tetapi memantul ke sasaran, percobaan diulang.

Penskoran: setiap bola yang divoli dengan pukulan *backhand* jatuh di dalam daerah sasaran memperoleh skor sebesar nilai daerah sasaran. Apabila bola jatuh pada garis batas daerah sasaran memperoleh skor sebesar nilai daerah sasaran yang besar. Jika bola meyangkut di net atau bola jatuh di luar lapangan diberi skor nol. Skor akhir tes adalah jumlah skor dari 5 kali pukulan *volley backhand*.

J. Tes dan Pengukuran Untuk Daya Tahan Aerobik (Lari 15 menit)

- 1) Tujuan untuk mengukur kapasitas aerobik
- 2) Petunjuk pelaksanaan tes: dengan menggunakan star berdiri, setelah diberi aba-aba oleh petugas, peserta lari menempuh waktu 15 menit sejauh mungkin.
- 3) Penilaian: jarak yang ditempuh selama 15 menit dicatat dalam satuan meter. Kemudian digunakan rumus VO2 max dalam menentukan kapasitas aerobik. Adapun rumusnya ialah:

$$\text{VO}_2 \text{ max} = \left(\frac{\text{X meter}}{15} - 133 \right) \times 0,172 + 33,3$$

Sumber: Depdiknas (2000: 64)

Keterangan:

VO2 max : kapasitas aerobik (ml/kg.BB/menit)

X : jarak yang ditempuh dalam meter

15 : waktu lari 15 menit

Tabel 4
Kesegaran jasmani fungsi kardiorespirasi untuk atlet

No	Jenis Kelamin	Umur			Klasifikasi
		20-29	30-39	40-49	
1	Putri	49 ke atas	45 ke atas	42 ke atas	Baik Sekali
	Putra	49 ke atas	45 ke atas	42 ke atas	Baik Sekali
2	Putri	38-48	34-44	31-41	Baik
	Putra	38-48	34-44	31-41	Baik
3	Putri	31-37	28-33	24-30	Sedang
	Putra	31-37	28-33	24-30	Sedang
4	Putri	24-30	20-27	17-23	Kurang
	Putra	24-30	20-27	17-23	Kurang
5	Putri	23 ke bawah	19 ke bawah	16 ke bawah	Kurang Sekali
	Putra	23 ke bawah	19 ke bawah	16 ke bawah	Kurang Sekali

Sumber: Depdiknas (2000: 64)

K. Tes dan Pengukuran Untuk Daya Tahan Aerobik (Lari 2400 meter)

- 1) Tujuan dari tes: mengukur daya tahan jantung paru
- 2) Petunjuk Pelaksanaan tes: dengan menggunakan start berdiri, setelah diberi aba-aba oleh petugas, peserta lari menempuh jarak 2400 meter dalam waktu secepat mungkin.
- 3) Penilaian: penilaian tes lari 2400 meter ini, dengan mempergunakan waktu tempuh dalam menit dan detik. Untuk mengetahui daya tahan yang dimiliki dapat dilihat pada tabel berikut:

Tabel 5
Norma Tes Lari 2400 Meter

Jenis Kelamin	Usia dan Waktu (menit, detik)			Kategori kesegaran jasmani
	13-19 tahun	20-29 tahun	30-39 tahun	
Putra	< 08,37	< 09,45	< 10,00	Istimewa
Putri	< 11,50	< 12,30	< 13,00	
Putra	08,37-09,40	09,45-10,45	10,00-11,00	Sangat Baik
Putri	11,50-12,29	12,30-13,30	13,00-14,30	
Putra	09,41-10,48	10,46-12,00	11,01-12,30	Baik
Putri	12,30-14,30	13,31-15,54	14,31-16,30	
Putra	10,49-12,10	12,01-14,00	12,31-14,44	Sedang
Putri	14,31-16,54	15,55-18,00	16,31-19,00	
Putra	12,11-15,30	14,01-16,00	14,45-16,30	Kurang
Putri	16,55-18,30	18,31-19,00	19,01-19,30	
Putra	>15,31	>16,01	>16,31	Sangat Kurang
Putri	>18,31	>19,01	>19,31	

Sumber: Kenneth H. Cooper dalam Len Kravitz (2001: 15)

L. Tes dan Pengukuran Untuk Daya Tahan Aerobik (*Multistage Fitness Test*)¹

Pada tes ini, testi berlari bolak-balik sepanjang jalur atau lintasan yang telah diukur sambil mendengarkan serangkaian tanda yang berupa bunyi "tut" yang terekam dalam kaset. Pada mulanya waktu tanda "tut" berdurasi sangat lambat, namun secara bertahap menjadi lebih cepat sehingga makin sulit testi menyamakan kecepatannya dengan kecepatan yang diberikan oleh tanda tersebut. Testi berhenti apabila tidak mampu lagi mempertahankan langkahnya dan tahap ini menunjukkan tingkat konsumsi oksigen maksimal testi tersebut.

¹ Ismaryati. (2009). *Tes dan Pengukuran Olahraga* hal. 80-87. Surakarta: LPP UNS dan UNS Press.

1) Perlengkapan

- a) Halaman, lapangan, atau permukaan datar dan tidak licin, sekurang-kurangnya sepanjang 22 meter.
- b) Mesin pemutar kaset (*tape recorder*)
- c) Kaset audio *multistage fitness test*
- d) Pita pengukur/meteran untuk mengukur jalur sepanjang 20 m
- e) Kurucut sebagai tanda batas jarak
- f) Lebar lintasan kurang lebih 1 hingga 1,5 meter untuk tiap testi
- g) *Stopwatch*

2) Pelaksanaan Tes

- a) Hidupkan *tape recorder*. Pada bagian permulaan pita, jarak antara dua tanda "tut" menandai interval 1 menit. pergunakan saat permulaan untuk memastikan pita kaset belum mengalami peregangan (*molor*). Ketelitian sekitar 0,5 detik dianggap memadai, jika lebih dari itu maka jarak lari perlu diubah untuk koreksi jarak.
- b) Petunjuk kepada testi telah tersedia dalam pita kaset rekaman yang berisi penjelasan pelaksanaan tes dan mengantarkan perhitungan mundur selama 5 detik menjelang pelaksanaan. Setelah itu pita mengeluarkan tanda "tut" tunggal pada beberapa interval yang teratur. Testi diharapkan sampau ke ujung yang berlawanan bertepatan dengan saat "tut" yang pertama berbunyi. Begitu seterusnya.
- c) Tiap interval, kecepatan lari harus makin ditingkatkan. Kecepatan lari pada menit pertama disebut level 1, pada menit kedua disebut level 2, dan seterusnya. Terdapat 21 level. Tiap akhir level ditandai sinyal "tut" tiga kali berturut-turut serta oleh penjelasan dari pita kaset. Pada level 1, testi diberi waktu 9 detik

- untuk menempuh jarak lintasan 20 m.
- d) Testi harus selalu menempatkan satu kaki tepat pada atau dibelakang tanda meter ke 20 tiap lari. Jika sudah sampai sebelum tanda "tut" berikutnya berbunyi, testi berbalik dengan cara bertumpu pada salah satu sumbu putar kaki dan menunggu isyarat bunyi "tut" untuk melanjutkan lari.
 - e) Testi harus berlari selama mungkin sampai tidak mampu lagi mengikuti kecepatan pita kaset dan secara sukarela menarik diri dari tes. Pelatih perlu menghentikan testi yang mulai ketinggalan di belakang dari langkah yang diharapkan. Jika terlambat sekali, testi masih diberi kesempatan 2 kali lari untuk mengejar irama langkah.
 - f) Untuk mengetahui konsumsi oksigen maksimal testi, gunakan tabel. Setelah diketahui konsumsi oksigen maksimal, gunakan tabel berikutnya untuk mengetahui kategori kesegaran jasmaninya.

Tabel 6
Penyesuaian Jarak Lari Bolak-Balik Berdasarkan
Kecepatan Pemutar Kaset

Periode waktu standar (detik)	Jarak lari (meter)	Periode waktu standar (detik)	Jarak lari (meter)
55,0	18,333	60,5	20,166
55,5	18,500	61,0	20,333
56,0	18,666	61,5	20,500
56,5	18,833	62,0	20,688
57,0	19,000	62,5	20,833
57,5	19,166	63,0	21,000
58,0	19,333	63,5	21,166
58,5	19,500	64,0	21,333
59,0	19,666	64,5	21,500
59,5	19,833	65,0	21,666
60,0	20,000		

Perhatian: apabila terdapat kesalahan lebih dari 5 detik pada periode waktu standar, ganti kaset pemutar atau tape recorder dengan yang lain.

Tabel 7
Prediksi Ambilan Konsumsi Oksigen Maksimal Dengan Tes Lari Bolak-Balik

Level	Shuttle	Prediksi VO2Max	Level	Shuttle	Prediksi VO2 Max
4	2	26,8	7	2	37,1
4	4	27,6	7	4	37,8
4	6	28,3	7	6	38,5
4	9	29,5	7	8	39,2
5	2	30,2	7	10	39,9
5	4	31,0	8	2	40,5
5	6	31,8	8	4	41,1
5	9	32,9	8	6	41,8
6	2	33,6	8	8	42,4
6	4	34,3	8	11	43,3
6	6	35,0	9	2	43,9
6	8	35,7	9	4	44,5
6	10	36,4	9	6	45,2
10	2	47,4	9	11	46,8
10	4	48,0	15	2	64,6
10	6	48,7	15	4	65,1
10	8	49,3	15	6	65,6
11	2	50,8	15	8	66,2
11	4	51,4	15	10	66,7
11	6	51,9	15	13	67,5
11	8	52,5	16	2	68,0
11	10	53,1	16	4	68,5
11	12	53,7	16	6	69,0
12	2	54,4	16	8	69,5
12	4	54,8	16	10	69,9
12	6	55,4	16	12	70,5

Level	Shuttle	Prediksi VO2Max	Level	Shuttle	Prediksi VO2 Max
12	8	56,0	16	14	70,9
12	10	56,5	17	2	71,4
12	12	57,1	17	4	71,9
13	2	57,6	17	6	72,4
13	4	58,2	17	8	72,9
13	6	58,7	17	10	73,4
13	8	59,3	17	12	73,9
13	10	59,8	18	2	74,8
13	12	60,6	18	4	75,3
14	2	61,1	18	6	75,8
14	4	61,7	18	8	76,2
14	6	62,6	18	10	76,7
14	8	62,7	18	12	77,2
14	10	63,2	18	15	77,9
14	13	64,0	19	2	78,3
20	2	81,8	19	4	78,8
20	4	82,2	19	6	79,2
20	6	82,6	19	8	79,7
20	8	83,0	19	10	80,2
20	10	83,5	19	12	80,6
20	12	83,9	19	15	81,3
20	14	84,3	21	8	86,5
20	16	84,8	21	10	86,9
21	2	85,2	21	12	87,4
21	4	85,6	21	14	87,8
21	6	86,1	21	16	88,2

Tabel 6. Penilaian Tingkat Kesegaran Jasmani Berdasar Konsumsi Oksigen Maksimal

Kategori	Konsumsi oksigen maksimal (ml/kg bb/menit)			
	<30 tahun	30-39 tahun	40-49 tahun	>50 tahun
Sangat buruk	<25,0	< 25,0	< 25,0	
Buruk	25,0 – 33,7	25,0 – 30,1	25,0 – 26,4	<25,0
Sedang	33,8 – 42,5	30,2 – 39,1	26,5 – 35,4	25,0 – 33,7
Baik	42,6 – 51,5	39,2 – 48	35,5 – 45,0	33,8 – 43,0
Sangat baik	>51,6	>48	>45,1	>43,1

Sumber: Davis, Kimmec, Auty dalam Ismaryati (2009: 84-86)

Formulir Pencatatan Lari Bolak-balik

Level	Nomor Pengulangan															
1	1	2	3	4	5	6	7									
2	1	2	3	4	5	6	7	8								
3	1	2	3	4	5	6	7	8								
4	1	2	3	4	5	6	7	8	9							
5	1	2	3	4	5	6	7	8	9							
6	1	2	3	4	5	6	7	8	9	10						
7	1	2	3	4	5	6	7	8	9	10						
8	1	2	3	4	5	6	7	8	9	10	11					
9	1	2	3	4	5	6	7	8	9	10	11					
10	1	2	3	4	5	6	7	8	9	10	11					
11	1	2	3	4	5	6	7	8	9	10	11	12				
12	1	2	3	4	5	6	7	8	9	10	11	12				
13	1	2	3	4	5	6	7	8	9	10	11	12	13			
14	1	2	3	4	5	6	7	8	9	10	11	12	13			
15	1	2	3	4	5	6	7	8	9	10	11	12	13			
16	1	2	3	4	5	6	7	8	9	10	11	12	13	14		
17	1	2	3	4	5	6	7	8	9	10	11	12	13	14		
18	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
19	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
20	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
21	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

M. Tes dan Pengukuran Keleutukan (*Sit and Reach*)²

- 1) Tujuan: mengukur keleutukan otot punggung ke arah depan dan paha belakang
- 2) Sasaran: laki-laki dan perempuan usia 5 tahun ke atas
- 3) Perlengkapan: box khusus
- 4) Pelaksanaan:
 - a) Testi duduk selunjur tanpa sepatu, lutut lurus, telapak kaki menempel pada sisi box.
 - b) Kedua tangan lurus diletakkan di atas ujung box, telapak tangan menempel di permukaan box.
 - c) Dorong dengan tangan sejauh mungkin, tahan 1 detik, catat hasilnya.
 - d) Dilakukan 4 kali ulangan.
 - e) Pada saat tangan mendorong ke depan, kedua lutut harus tetap lurus.
 - f) Dorongan harus dilakukan dengan dua tangan bersama-sama, bila tidak maka tes harus diulang.
 - g) Sebelum melakukan tes harus pemanasan terlebih dahulu.
- 5) Penilaian: raihan terjauh dari keempat ulangan merupakan nilai keleutukan punggung bawah testi. Angka dicatat sampai mendekati 1 cm.

² Ismaryati. (2009). *Tes dan Pengukuran Olahraga* hal. 101 - 102. Surakarta: LPP UNS dan UNS Press



Sikap Akhir

Sikap Awal

Sumber: Mackenzie, B. (2003)

Tabel 9
Klasifikasi Hasil Tes Sit and Reach

Female (Age)	Very Poor	Poor	Average	Above Avg	Good	Excellent
Under 25	> 14	15-16	17-19	20-21	22	< 23
25-35	> 13	14-17	18-19	20	21-22	< 23
35-45	> 12	13-16	17	18-19	20-21	< 22
45-55	> 10	11-14	15-16	17-18	19-20	< 21
55-65	> 9	10-14	15	16-17	18-19	< 20
65 and over	> 8	10-14	15-16	17	18-19	< 20

Male (Age)	Very Poor	Poor	Average	Above Avg	Good	Excellent
Under 25	> 11	12-13	14-17	18-19	20-21	< 22
25-35	> 9	10-12	13-16	17	18-19	< 20
35-45	> 7	8-11	12-15	16-17	18-19	< 20
45-55	> 6	7-9	10-13	14-15	16-18	< 19
55-65	> 5	6-9	8-11	13	14-16	< 17
65 and over	> 4	5-7	8-11	12-13	14-16	< 17

Sumber: <http://kinesiologists.ca/wp/bcrpa-certification/personal-trainer-store/free-resources/protocol-flexibility-a/>

N. Tes dan Pengukuran Keseimbangan³

a. *Stork Stand*

Tujuan : Mengukur keseimbangan statis

Sasaran : Laki-laki dan perempuan berusia 10 tahun ke atas

Perlengkapan : *Stopwatch*

Pelaksanaan :

- Testi berdiri di atas satu kaki yang dominan, kaki yang lain diletakkan di samping lutut, tangan berada di pinggang.
- Saat diberi aba-aba "ya", testi mengangkat tumitnya dari lantai (jinjit) dan mempertahankan sikap ini selama mungkin tanpa gerakan apapun atau meletakkan tumitnya menyentuh lantai.
- Saat mengangkat tumit dan mempertahankannya, tangan tidak boleh lepas dari pinggang.
- Dilakukan tiga kali ulangan.

Penilaian:

- Waktu yang terlama dalam mempertahankan keseimbangannya merupakan waktu yang digunakan untuk menilai keseimbangan testi
- Waktu dicatat dalam detik, dimulai dari saat testi mengangkat tumit sampai mulai kehilangan keseimbangannya.

³ Yuyun Ari Wibowo & Fidia Dwi Andriyani. (2015). *Pengembangan Elektrokardioler Olahraga Sekolah*. Yogyakarta: UNY Press.



Gambar Stork Stand
Sumber: Mackenzie, B. (2000)

Tabel.
Data Norma Tes *Stork Stand* untuk usia 16-19 tahun
menurut Johnson & Nelson dalam Mackenzie, B. (2000)

Rating	Males	Females
Excellent	>50	>30
Above Average	41-50	23-30
Average	31-40	16-22
Below Average	20-30	10-15
Poor	<20	<10

Tabel
Data Norma Tes Stork Stand untuk laki-laki dan wanita
menurut Schell & Leclerthaepin dalam Mackenzie, B.
(2000)

Rating	Male	Females
Excellent	>50	>27
Above Average	37-50	23-27
Average	15-36	8-22
Below Average	5-14	3-7
Poor	<5	<3

O. Tes Kekuatan dan Daya Tahan Otot⁴

1. Tidur terlentang dengan kedua tangan menyangga kepala atau menyilang di depan dada.
2. Kedua lutut dibengkokkan pada lutut, dengan kedua kaki terletak datar di lantai kurang lebih 15 cm sampai 25 cm dari pantat.
3. Untuk melakukan gerakan "Crunch", bengkokkan badan sehingga tulang belikat terangkat dari lantai (pinggang anda tetap berada di lantai). Lakukan yang mulus.
4. Untuk melakukan tes, hitunglah jumlah "Crunch" yang dapat anda lakukan selama satu menit.

⁴ Kravitz, Len. (2001). *Panduan Lengkap: Bogor Total*. Jakarta: PT Raja Grafindo Persada.



Sumber: Kravitz, Len. (2001)

Hasil Untuk Wanita dan Pria

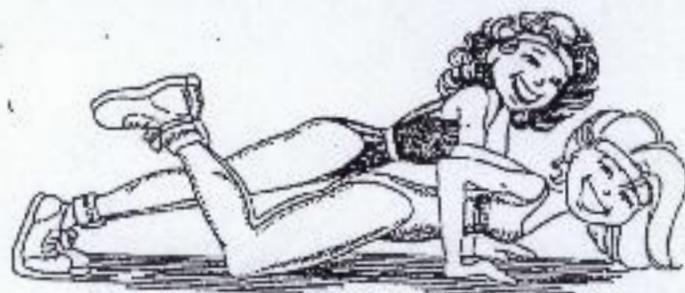
Kategori	Hasil
Sangat Bagus	50-60 "Crunch" atau lebih
Sedang	36-49 "Crunch"
Kurang	Kurang dari 36 "Crunch"

P. Tes Kekuatan dan Daya Tahan Badan Atas⁴

Push Up dapat dilakukan dengan cara standar (dengan badan lurus dan disangga oleh kedua tangan dan kaki) atau dengan cara yang telah dimodifikasi (dengan kedua lutut di lantai).

1. Turunkan dada sehingga hampir menyentuh lantai kemudian luruskan kedua lengan.
2. Untuk melakukan tes, hitunglah jumlah push up yang dapat dilakukan dalam satu menit.
3. Mulailah dengan push up yang sudah dimodifikasi. Sekali anda telah mencapai *tingkatan sangat baik* dengan push up yang dimodifikasi, selanjutnya dengan push up standar.

⁴ Kravitz, Len. (2001). *Panduan Lengkap: Buger Total* hal. 17 Jakarta: PT Raja Grafindo Persada.



Sumber: Kravitz, Len. (2001)

Hasil Untuk Pria Dan Wanita

<i>Push Up yang Telah Dimodifikasi</i>	
Sangat Baik	40-49 push up atau lebih
Sedang	17-39 push up
Kurang	Kurang dari 17 push up

<i>Push Up Standar</i>	
Sangat Baik	40-45 push up atau lebih
Sedang	20-39 push up
Kurang	Kurang dari 20 push up

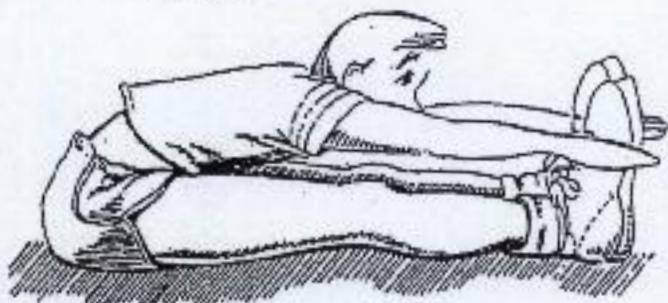
Q. Tes Kelenturan (Tes Duduk dan Meraih)⁶

Kelenturan memiliki sifat yang khas. Ini berarti bahwa tingkat kelenturan di salah satu sendi tidak sama dengan kelenturan pada sendi yang lain pada tubuh. Karena kurangnya kelenturan di pinggang, bagian belakang kaki, dan

⁶ Kravitz, Len. (2001), *Pasukan Lengkap: Sugar Total* hal. 78. Jakarta: PT Raja Grafindo Persada.

pinggul merupakan 80% dari penyebab terjadinya masalah pada pinggang, maka anda dapat memilih tes kelenturan ini.

1. Duduklah dengan kedua kaki direntangkan ke depan anda. Usahakan agar kedua kaki tegak lurus pada lantai. Taruhlah sebuah penggaris di lantai antara kedua kaki.
2. Perlahan-lahan renggangkan ke depan, meraih atau melampaui jari-jari dan tahan (jangan memantul!). Usahakan kedua kaki lurus tetapi tidak terkunci.
3. Lebih baik lakukan ini beberapa kali untuk latihan, perlahan renggangkan lebih banyak sampai titik batasnya.



Sumber: Kravitz, Len. (2001)

Hasil Tes Duduk dan Meraih Bagi Wanita Dan Pria

Kategori	Hasil
Baik Sekali	12,5-20 cm melampaui jari-jari kaki atau lebih
Sedang	5 cm di depan jari-jari kaki sampai 12,5 cm melampaui jari
Kurang	Tidak lebih dari 5 cm di depan jari-jari kaki

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.
- 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.
- Burhan Nurgiyantoro, dkk. (2012). *Statistik Terapan Untuk Penelitian Ilmu-Ilmu Sosial*. Yogyakarta: Gadjah Mada University Press.
- 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.
- 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., & Moed, 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.
- _____. (2001). Petunjuk Praktikum Tes dan Pengukuran. Prodi PJKR FIK UNY.
- 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, Margaret 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.
- Suharsimi Arikunto. (2016). *Dasar-dasar Evaluasi Pendidikan Edisi 2*. Jakarta: Bumi Aksara.
- Yuyun Ari Wibowo & Fitria Dwi Andriyani. (2015). *Pengembangan Ekstrakurikuler Olahraga Sekolah*. Yogyakarta: UNY Press.

TENTANG PENULIS



Drs. Ngatman, M.Pd. berprofesi sebagai dosen di Prodi Pendidikan Jasmani, Kesehatan, dan Rekreasi, Jurusan Pendidikan Olahraga, Fakultas Ilmu Keolahragaan Universitas Negeri Yogyakarta. Ia menempuh gelar S1 dan S2 nya pada Universitas yang sama. Untuk jenjang magister, Ngatman mengambil jurusan Penelitian dan Evaluasi Pendidikan. Bidang keahliannya yaitu Tes, Pengukuran, dan Evaluasi Pendidikan Jasmani.

Ngatman banyak berkiprah dan berprestasi dalam bidang olahraga Tennis. Ia mendampingi para atlet Tennis pada berbagai kejuaraan baik di dalam maupun luar negeri. Ia juga merupakan tutor ITF Level-1 Coaches Course sejak tahun 2013 dan menjadi narasumber pada berbagai kegiatan dengan tema olahraga.

Ngatman menunjukkan kiprahnya dalam dunia pendidikan jasmani dan olahraga melalui berbagai penelitian dan publikasi. Beberapa diantara penelitian dan publikasi yang telah dilakukan yaitu: Evaluasi Standar Kompetensi PPL mahasiswa PPKHB Kabupaten Magelang, Alat Evaluasi Nontes Dalam Pendidikan Jasmani, Pengaruh Metode Circuit Training terhadap Tingkat Kebugaran Jasmani Kelas VII SMP Negeri 1 Jepon, Kabupaten Blora, *Assesing Physical Fitness For Tennis Player*, dan Uji Validitas, Reliabilitas, dan Relevansi Kemp-Vincent Rally Tennis Test terhadap Tingkat Keterampilan Bermain Tennis Mahasiswa FIK UNY.

Anda dapat berkorespondensi dengan Ngatman melalui emailnya yaitu ngatmansoewito@yahoo.com.

