

TES STANDAR (*STANDARDIZED TEST*)

Oleh :
Dindin Abdul Muiz Lidinillah

PENDAHULUAN

Evaluasi merupakan bagian penting dalam siklus pendidikan. Hasil evaluasi sangat berpengaruh dalam pembuatan keputusan oleh pihak yang terkait seperti guru. Oleh karena itu, evaluasi merupakan salah satu kegiatan utama yang harus dilakukan oleh seorang guru dalam kegiatan pembelajaran. Dengan evaluasi, guru akan mengetahui perkembangan hasil belajar, intelegensi, bakat khusus, minat, hubungan sosial, sikap dan kepribadian siswa atau peserta didik serta secara umum dapat mengetahui berhasil dan tidaknya program pembelajaran.

Agar evaluasi dapat berhasil dengan baik diperlukan alat evaluasi yang tepat dan telah teruji dengan baik. Alat evaluasi harus juga dapat menghasilkan data yang diperlukan sesuai dengan tujuan evaluasi. Dalam pembelajaran, guru membutuhkan data yang berkaitan dengan perkembangan belajar siswa, oleh karena itu guru melakukan serangkaian pengukuran sesuai dengan jenis penilaian.

Untuk keperluan evaluasi diperlukan alat evaluasi yang bermacam-macam, seperti kuesioner, tes, skala, format observasi, dan lain-lain. Dari sekian banyak alat evaluasi, secara umum dapat dikelompokkan menjadi dua, yakni alat tes dan nontes. Khusus untuk evaluasi hasil pembelajaran alat evaluasi yang paling banyak digunakan adalah tes.

Dari berbagai bentuk jenis tes, ada suatu jenis tes yang instrumennya telah distandarisasi. Artinya, jenis tes tersebut telah diuji validitas dan reliabilitasnya secara ketat serta telah diuji kelayakan penggunaannya secara praktis. Jenis tes ini disebut tes standar (*standardized test*). Berbeda dengan tes yang dibuat oleh guru yang hanya dapat mengukur kemampuan siswanya sendiri, tes standar dapat digunakan oleh siswa-siswa lintas sekolah selama siswa tersebut masuk dalam kriteria yang telah ditetapkan. Hasil tes itu menunjukkan tingkat kemampuan siswa pada bidang tertentu yang diakui secara normatif. Salah satu tes standar yang terkenal adalah Tes Binet untuk mengukur tingkat IQ seseorang.

Dari hal-hal di atas, penulis tertarik untuk membahas tentang tes standar meliputi bagaimana membuat, menggunakan dan menilai suatu tes standar yang telah dibuat. Oleh karena itu, makalah ini akan memaparkan secara singkat tapi jelas tentang tes standar tersebut.

KONSEP DASAR TES STANDAR

Istilah tes secara bahasa berasal dari bahasa Prancis kuno “testum” artinya piring untuk menyisihkan logam-logam mulia. Tes adalah sekumpulan pertanyaan atau latihan atau alat lain yang digunakan untuk mengukur keterampilan, pengetahuan, kecerdasan, kemampuan, atau bakat yang dimiliki oleh seseorang atau kelompok.

Oleh karena itu, tes merupakan alat ukur yang berbentuk pertanyaan atau latihan, dipergunakan untuk mengukur kemampuan yang ada pada seseorang atau sekelompok orang. Sebagai alat ukur dalam bentuk pertanyaan, maka tes harus dapat memberikan informasi mengenai pengetahuan dan kemampuan obyek yang diukur. Sedangkan sebagai alat ukur berupa latihan, maka tes harus dapat mengungkap keterampilan dan bakat seseorang atau sekelompok orang. Tes merupakan alat ukur yang standar dan obyektif sehingga dapat digunakan secara meluas untuk mengukur dan membandingkan keadaan psikis atau tingkah laku individu. Dengan demikian berarti sudah dapat dipastikan akan mampu memberikan informasi yang tepat dan obyektif tentang obyek yang hendak diukur baik berupa psikis maupun tingkah lakunya, sekaligus dapat membandingkan antara seseorang dengan orang lain.

Dalam kaitan dengan rumusan tersebut, sebagai alat evaluasi hasil belajar, tes minimal mempunyai dua fungsi, yaitu: untuk mengukur tingkat penguasaan terhadap seperangkat materi atau tingkat pencapaian terhadap seperangkat tujuan tertentu ; dan untuk menentukan kedudukan atau peringkat siswa dalam kelompok, tentang penguasaan materi atau pencapaian tujuan pembelajaran tertentu.

Jenis-jenis tes dari segi tujuannya dalam bidang pendidikan ada beberapa jenis, yaitu : (1) Tes Kecepatan (*Speed Test*), (2) Tes Kemampuan (*Power Test*), (3) Tes Hasil Belajar (*Achievement Test*), (4) Tes Kemajuan Belajar (*Gains/Achievement Test*), (5) Tes Diagnostik (*Diagnostic Test*), (6) Tes Formatif, dan (7) Tes Sumatif.

Dilihat dari jawaban siswa yang dituntut dalam menjawab atau memecahkan persoalan yang dihadapinya, maka tes hasil belajar dapat dibagi menjadi 3 jenis : (1) Tes lisan (*oral test*), (2) Tes tertulis (*written test*), dan (3) Tes tindakan atau perbuatan (*performance test*).

Pengertian Tes Standar

Jenis tes dapat dibedakan juga menjadi 2 jenis dilihat dari ruang lingkup perbandingan hasil penilaian, yaitu : tes standar dan tes buatan guru. Tes Buatan Guru (*Teacher-made Test*) dikembangkan oleh guru untuk keperluan penilaian di kelasnya sendiri. Efektivitas jenis tes ini bergantung kepada ketrampilan dan kemampuan guru dalam merancang satu tes. Sementara Tes Standar (*Standardized Test*) dirancang oleh ahli tes yang bekerja sebagai ahli kurikulum sekaligus sebagai guru. Tes tersebut distandarisasi dalam arti pengelolaan dan penyekoran yang dilakukan berdasarkan standar dan asumsi kondisi yang seragam sehingga hasil dari penilaian dapat dibandingkan untuk kelas atau sekolah yang berbeda.

Untuk keperluan penilaian, guru dapat menggunakan kedua jenis tes tersebut. Sementara tes buatan guru bermanfaat untuk menilai dan membandingkan siswa dalam suatu kelas, tes standar dapat digunakan untuk menilai dan membandingkan kemampuan atau kemajuan belajar siswa dengan siswa di kelas atau kelas lain. Di sisi lain, tes standar juga dapat digunakan untuk mengetahui kemampuan siswa terhadap jenis-jenis kemampuan yang secara normatif telah menjadi standar, seperti tes kemampuan intelektual atau bakat.

Menurut Arikunto (1988), Tes Standar dapat dibedakan dengan Tes Buatan Guru dengan perbandingan sebagai berikut :

Tabel 1
Perbandingan Tes Standar dengan Tes Buatan Guru (Arikunto, 1988)

Tes Standar	Tes buatan guru
1) Didasarkan atas bahan dan tujuan umum dari sekolah-sekolah di seluruh negara	1) Didasarkan atas bahan dan tujuan khusus yang dirumuskan oleh guru untuk kelasnya sendiri
2) Mencakup aspek yang luas dan pengetahuan atau keterampilan dengan hanya sedikit butir tes untuk suatu keterampilan atau topik	2) Dapat terjadi hanya mencakup pengetahuan atau keterampilan yang sempit
3) Disusun dengan kelengkapan staf: profesor, pembahas, editor butir tes	3) Biasanya disusun sendiri oleh guru dengan sedikit atau tanpa bantuan orang lain/tenaga ahli
4) Menggunakan butir-butir tes yang sudah diujicobakan (try out), dianalisa dan direvisi sebelum menjadi sebuah tes	4) Jarang-jarang menggunakan butir-butir tes yang sudah diujicobakan dianalisa dan direvisi
5) Mempunyai realibilitas yang tinggi	5) Mempunyai realibilitas sedang dan rendah
6) Dimungkinkan menggunakan norma untuk seluruh negara	6) Norma kelompok terbatas kelas tertentu

Agar lebih memahami perbedaan tes standar dengan tes buatan guru, berikut ini perbandingannya menurut Sax (1980).

Tabel 2
Perbandingan Tes Standar dengan Tes Buatan Guru (Sax, 1980)

Karakteristik	Tes Standar	Tes Buatan Guru
1) Spesifikasi tujuan	1) Tujuan tes berlaku umum untuk siswa lintas kelas atau sekolah	1) Tujuan tes spesifik untuk keperluan penilaian siswa suatu kelas
2) Isi	2) Butir-butir soal tetap dan tidak dapat dimodifikasi, dan hanya mencakup suatu muatan tertentu dari kurikulum	2) Isi dapat diambil dari berbagai muatan kurikulum. Butir-butir tes dapat ditambah, dikurangi dan dimodifikasi sesuai pertimbangan guru
3) Aturan pengelolaan dan penskoran	3) Aturan bergantung kepada pihak yang membuat tes (<i>publisher</i>), mereka menyajikan aturan dan petunjuk dalam sebuah manual	3) Aturan bergantung kepada guru. Mereka dapat melakukan tes secara seragam untuk seluruh siswa, tetapi dapat juga diadaptasi sesuai dengan kondisi siswa
4) Norma	4) Norma dikembangkan oleh pembuat tes (<i>publisher</i>) untuk seluruh guru untuk	4) Tidak ada norma yang menjadi acuan, tetapi norma itu dapat

5) Penilaian tes	membandingkan kinerja suatu kelas berdasarkan usia dan tingkatan siswa 5) Data yang berupa kualitas dari suatu hasil tes dikeluarkan oleh pembuat tes (<i>publisber</i>).	dikembangkan sendiri oleh guru 5) Kualitas dari tes dapat dinilai sendiri oleh guru
------------------	--	--

Secara umum, ujian yang dilakukan berdasarkan tes buatan guru lebih fleksibel. Guru dapat mengubah dan memperbaikinya bergantung pada tujuan dan kondisi yang terjadi di kelas. Akan tetapi, tes ini tidak mampu membandingkan antar siswa dengan kelas lain di luar sekolah tersebut. Hal ini berbeda sekali dengan tes standar, di mana tes ini dapat mengetahui kualitas siswa berdasarkan standar kualitas yang ditetapkan lebih luas, misalnya suatu Negara atau daerah.

Kegunaan Tes Standar

Menurut Arikunto (1988), guru dapat menggunakan Tes Standar untuk beberapa alasan, diantaranya yaitu :

- (1) Membandingkan prestasi belajar dengan bawaan individu atau kelompok
- (2) Membandingkan tingkat prestasi siswa dalam ketrampilan di berbagai bidang studi untuk individu atau kelompok
- (3) Membandingkan prestasi siswa antara berbagai sekolah atau kelas
- (4) Mempelajari perkembangan siswa dalam suatu periode waktu tertentu.

Sebagai perbandingan, kegunaan Tes Buatan Guru adalah sebagai berikut :

- (1) Untuk menentukan seberapa baik siswa telah menguasai bahan pelajaran yang diberikan dalam waktu tertentu
- (2) Untuk menentukan apakah sesuatu tujuan telah tercapai
- (3) Untuk memperoleh suatu nilai

Selanjutnya baik tes standar dan tes buatan guru dianjurkan dipakai jika hasilnya akan digunakan untuk :

- (1) Mengadakan diagnosis terhadap ketidamampuan siswa
- (2) Menentukan tempat dalam suatu kelas atau kelompok
- (3) Memberikan bimbingan kepada siswa dalam pendidikan dan pemilihan jurusan
- (4) Memilih siswa untuk program-program khusus

KRITERIA UNTUK MEMILIH DAN MENILAI SUATU TES STANDAR

Menurut Sax (1980 : 314), suatu tes standar dapat dinilai baik secara teknik (*technically*) dan praktis (*practically*). Pada bagian ini akan dijelaskan mengenai kedua bentuk penilaian tersebut. Pertama kali, akan dijelaskan kriteria penilaian secara praktis kemudian secara teknik.

Kriteria Praktis

Ada beberapa kriteria praktis yang harus dipenuhi oleh suatu tes standar, yaitu :

Biaya (cost)

Yang dimaksud biaya dalam hal ini meliputi biaya untuk mengikuti suatu tes bagi pengguna serta biaya yang dilihat berdasarkan penyedia suatu tes mulai dari pembuatan lembar soal dan jawaban, biaya pelaksanaan tes dan biaya pengelolaan hasil tes. Mahal dan tidaknya suatu tes bergantung kepada biaya-biaya tersebut dan besarnya relatif.

Batasan waktu (time limitation)

Waktu yang digunakan untuk tes harus efektif yang bergantung kepada format tes. Tes yang singkat dapat lebih ekonomis dari segi biaya tetapi terkadang kurang terjaga realibilitasnya.

Kemudahan pelaksanaan (easy administration)

Pelaksanaan dari tes akan baik jika manual yang dibuat memuat informasi yang lengkap. Karena tes standar dirancang untuk suatu kondisi tertentu yang seragam, kesalahan penjelasan dari atau pemahaman terhadap manual akan menjadi sumber kurang validnya suatu tes.

Format

Format tes harus mudah untuk dibaca dan tidak membuat bingung. Format tes harus memuat petunjuk pelaksanaan yang dapat terbaca dengan baik dan cepat oleh pihak yang diuji dan penguji. Petunjuk tersebut harus sama dengan manual.

Form alternatif (alternate form)

Form alternatif harus tersedia untuk memudahkan melakukan tes ulang.

Level yang bertingkat (multiple levels)

Ujian atau tes yang bertingkat menyajikan perbandingan norma untuk setiap tingkatan. Tujuannya adalah agar guru dapat membandingkan kemajuan kemampuan siswa sesuai dengan kemajuan norma kelompok tingkatan, misalnya dari grup sekolah dasar kepada grup sekolah menengah.

Lembar jawaban dan prosedur penskoran yang simpel (answer sheets and simple scoring procedures)

Lembar jawaban harus mudah digunakan serta mudah untuk dilakukan penskoran.

Kemudahan untuk diinterpretasi (easy interpretation)

Manual untuk suatu tes harus memuat kasus-kasus yang terjadi atau contoh kesalahan yang sering terjadi dari interpretasi terhadap skor tes. Contoh-contoh yang harus dan jangan dilakukan harus termuat dalam manual.

Kriteria-kriteria di atas dapat dijadikan rujukan layak dan tidak layak suatu tes standar dari segi praktisnya. Jika salah satu kriteria tidak terpenuhi, maka akan mengganggu kegiatan ujian tersebut sehingga dan akan menyebabkan kualitas hasil tes menurun.

Kriteria Teknik

Sementara kriteria praktis dapat menjaga kelayakan suatu tes secara praktis, kriteria teknis lebih menitikberatkan kepada kualitas tes dari isi dan muatannya.

Dalam Sax (1980) diungkapkan bahwa, pada tahun 1974 (saat ini mungkin telah diperbaharui), American Psychological Association menerbitkan *Standards for Educational and Psychological Test* yang merupakan hasil kerjasama dengan American Educational Research Association dan National Council on Measurement in Education. Standar-standar yang dibuat oleh lembaga tersebut berupa kriteria teknis suatu tes standar.

Standar-standar yang dibuat dibagi menjadi dua bagian, yaitu : (1) *standards for test, manuals, and report* (sajian informasi tentang tes, bantuan/petunjuk untuk interpretasi, petunjuk untuk pelaksanaan dan penskoran, norma dan skala, validitas dan realibilitas serta kesalahan-kesalahan pengukuran) ; dan (2) *standards for the uses of test* (kualifikasi pengguna, pilihan dan pengembangan tes atau metode, pelaksanaan dan penskoran, serta interpretasi dari skor). Selanjutnya akan disajikan tentang beberapa hal yang dianggap paling penting dari riteri teknis ini.

Sajian informasi (dissemination of information)

Manual harus memuat secara lengkap informasi tentang tes. Manual dari tes harus menjelaskan secara lengkap tentang pengembangan dari tes yang terdiri dari raional, bentuk-bentuk butir soal, prosedur dan hasil analisis butir soal atau hasil-hasil penelitian lain tentang tes tersebut.

Bantuan untuk interpretasi (aids to interpretation)

Tes, manual, laporan hasil tes serta dokumen lain harus dapat dipahami dengan benar oleh pengguna termasuk interpretasi hasil tes sehingga pengguna tidak mengalami kesalahan dalam menggunakannya. Manual harus secara eksplisit mencantumkan tujuan dan maksud dari tes dengan jelas.

Petunjuk untuk pelaksanaan dan penskoran (directions for administration and scoring)

Petunjuk pelaksanaan harus termuat dalam manual. Prosedur penskoran juga harus termuat dalam manual.

Norma-norma dan skala-skala (norms and scales)

Norma-norma harus dipublikasikan dalam tes manual. Norma-norma yang disajikan dalam manual harus didefinisikan dan dijelaskan dengan jelas sesuai dengan karakter populasi atau pengguna. Skala penskoran harus juga termuat dalam manual, sebaiknya skala menggunakan persentil agar dapat terlihat keragaman hasil tes.

Validitas (validity)

Hasil uji validitas atau validitas tes harus termuat dalam manual. Manual juga harus menyajikan contoh kasus atau kejadian yang dapat mengganggu validitas dari tes.

Realibilitas dan kesalahan pengukuran (reliability and measurement error)

Manual harus menunjukkan realibilitas dari tes dapat berupa laporan hasil penelitian atau bentuk penjelasan lain. Manual juga harus menyajikan hal-hal yang dapat mengganggu realibilitas tes. Untuk menunjukkan tingkat realibilitas tes, maka format tes harus dibuat dalam beberapa versi dengan tingkat realibilitas yang tinggi dan sama.

Kualifikasi dan kriteria pengguna (qualification and concern of users)

Manual harus menyajikan kualifikasi pengguna tes karena pengguna tes harus memiliki pengetahuan tentang prinsip-prinsip pengukuran dan mampu memahami tes dengan baik. Kualifikasi dan kriteria pengguna bergantung kepada norma-norma yang diacu.

Pilihan-pilihan dan pengembangan tes atau metode (choice and development of test or metode)

Manual juga harus menunjukkan strategi pengembangan tes serta metode-metode penilaian.

Pelaksanaan/pengelolaan dan penskoran (administration and scoring)

Prosedur pelaksanaan tes harus dimengerti oleh pengguna. Pembuat tes harus selalu memeriksa kelayakan perlengkapan untuk pelaksanaan tes. Pelaksana tes harus selalu menjaga kondisi pelaksanaan tes dengan baik. Penilai hasil tes harus menjaga akurasi penskoran sampai penyajian hasil tes.

Interpretasi dari skor (interpretation of score)

Interpretasi skor harus disajikan dengan baik dalam manual sehingga mudah dipahami oleh pengguna. Hal ini penting untuk penetapan target pencapaian skor.

Untuk mengetahui kualitas serta kelayakan suatu tes standar, maka kriteria teknis dan praktis tersebut harus menjadi acuan. American Psychological Association juga membuat panduan untuk menilai suatu tes standar.

BEBERAPA JENIS TES STANDAR

Dalam Sax (1980) dipaparkan beberapa jenis tes standar berdasarkan jenis normanya, yaitu tes intelegensi, tes bakat, tes prestasi akademik, tes minat dan sikap serta tes kepribadian. Berikut ini penjelasan tiga jenis tes yang pertama.

Tes Intelegensi (Intelligence Test)

Pada tahun 1869, Sir Francis Galton melakukan studi pertama kali tentang perbedaan intelektual. Tetapi, tes intelegensi pertama yang terkenal adalah yang dikembangkan oleh Alfred Binet (1911) tahun 1905. Binet mengembangkan berbagai metode dan pengukuran intelegensi yang kemudian dikenal dengan Tes Binet. Kemudian tes ini direvisi oleh asistennya Theo Simon sehingga terkenal dengan Tes Binet-Simon dan digunakan sampai tahun 1927. Henry Goddard mempublikasikan versi bahasa Inggris (pada awalnya tes Binet berbahasa Prancis) pada tahun 1908 dan 1911. Louis M. Terman dari Stanford University melakukan standarisasi ulang tes Binet pada tahun 1916 dan 1937 sehingga terkenal dengan Tes Stanford-Binet. Terakhir distandarisasi ulang tahun 1960 oleh Maud Merrill.

Berikut ini beberapa pengertian dari intelegensi dalam (Sax, 1980 : 395).

- *Intelligence is the ability of an organism to adjust itself adequately to its environment*
- *Intelligence is the heredity capacity to learn*
- *Intelligence is the ability to reason*
- *The intelligency person is one who can solve a wide variety of difficult question rapidly.*

Dapat ditarik kesimpulan bahwa Intelegensi dapat berarti fungsi intelegensi untuk membantu manusia menyesuaikan diri dengan lingkungan; asal muasal

intelegensi apakah sebagai faktor hereditas atau lingkungan; serta struktur dari intelegensi yang terdiri dari berbagai kemampuan.

Secara operasional, intelegensi adalah kemampuan yang tidak dapat diukur secara langsung dengan tes. Ketika tes memuat butir-butir pertanyaan, seorang individu dapat diketahui kadar intelegensinya.

Tes intelegensi dapat digunakan untuk bidang-bidang tertentu karena hasilnya sangat dibutuhkan. Beberapa kaitan tersebut adalah :

- *Intelegensi dengan pendidikan.* Tes intelegensi digunakan untuk mengukur tingkat pencapaian prestasi akademik serta prediksi keberhasilan anak dalam belajar
- *Intelegensi dengan diagnosis.* Para ahli psikologi menggunakan tes intelegensi untuk mengetahui tingkat kenormalan dan keabnormalan seorang anak berkaitan dengan intelegensinya.
- *Intelegensi dengan status sosial ekonomi.* Para ahli meyakini berdasarkan hasil penelitian bahwa tingkat intelegensi bersesuaian dengan status sosial ekonomi.
- *Intelegensi dengan latar belakang etnik.* Walaupun banyak diperdebatkan, perbedaan intelegensi dapat berbeda untuk suatu etnik dengan etnik lainnya.
- *Intelegensi dengan jenis kelamin.* Walaupun antara laki-laki dan perempuan terdapat perbedaan intelegensi tetapi itu hanya berbeda dalam suatu komponen. Misalnya perempuan cenderung lebih baik dalam kemampuan verbal, sementara laki-laki lebih unggul dalam kemampuan numeral.

Kemampuan intelektual didasarkan pada tingkatan usia. Rumus IQ sendiri adalah :
$$IQ = \frac{MA(\text{mental age})}{CA(\text{chronological age})} \times 100.$$
 Baik usia kronologis dengan usia mental dihitung dalam satuan bulan.

Misalnya, anak usia 8 tahun 10 bulan (106 bulan) yang memiliki usia mental (hasil tes IQ) 10-5 tahun (125 bulan), maka IQ nya adalah 118. Berarti anak tersebut memiliki IQ diatas rata-rata.

Adapun komponen dari tes intelegensi, misalkan berdasarkan pada The Wechsler Intelligence Scale for Children-Revised (WISC-R) adalah terdiri dari skala verbal dan unjuk kinerja. Skala verbal terdiri dari : *information, comprehension, arithmetic, similarities, vocabulary, digit span*. Sedangkan skala unjuk kerja terdiri dari : *picture completion, picture arrangement, block design, object assembly, coding dan mazes*.

Tes Bakat (*Aptitude Test*)

Tes bakat adalah pengukuran kognitif yang dirancang untuk memprediksikan sejauh mana individu akan mencapai kesuksesan sebelum mereka dilatih, dipilih dan ditempatkan. Tes bakat yang digunakan untuk memprediksi kesuksesan dalam suatu program khusus disebut tes bakat khusus. Tes bakat sering juga disebut tes bakat skolastik atas tes kecerdasan bakat. Realibilitas dari tes bakat cenderung meningkat seiring usia.

Tes bakat sering digunakan untuk proses seleksi dan penempatan. Bakat-bakat yang dapat di tes seperti : bakat menulis, mekanik, musik, seni, kreativitas.

Tes Prestasi Akademik (*Achievment Test*)

Tes prestasi akademik berbeda dengan tes bakat dari segi isi dan metode validasinya. Tujuan tes bakat adalah untuk memprediksi kesuksesan, sementara tes prestasi akademik mengukur kemampuan saat itu. Tes bakat mengandung validitas prediktif sementara tes prestasi akademik mengandung validitas isi.

Prosedur yang dapat digunakan untuk mengembangkan tes prestasi akademik adalah :

- Mengembangkan rasional dan tujuan dari tes
- Menulis butir-butir tes yang sesuai dengan tujuan tes
- Menganalisis butir-butir tes
- Mengkonstruksi form akhir dari tes
- Melakukan standarisasi form final
- Mengecek kesiapan (*battery overlap*)
- Menentukan kelayakan form

Tes prestasi akademik disajikan memuat ukuran prestasi dalam persentil, usia, tingkat, dan skor standar. Tes prestasi akademik standar dapat digunakan untuk keperluan berikut :

- Seleksi dan penempatan
- Diagnosis
- Feedback
- Evaluasi program

PENUTUP

Evaluasi tidak dapat dilepaskan dari kegiatan pendidikan. Evaluasi berperan dalam mengukur tingkat pencapaian program pendidikan secara khusus kegiatan pembelajaran. Hasil evaluasi dapat digunakan untuk kepentingan pengambilan kebijakan bagi pihak yang terkait. Untuk melakukan evaluasi diperlukan alat pengukuran yang tepat sesuai dengan jenis evaluasi atau penilaian yang dipilih. Salah satu alat evaluasi adalah tes. Tes adalah sekumpulan pertanyaan atau latihan atau alat lain yang digunakan untuk mengukur keterampilan, pengetahuan, kecerdasan, kemampuan, atau bakat yang dimiliki oleh seseorang atau kelompok.

Sebagai alat evaluasi hasil belajar, tes minimal mempunyai dua fungsi, yaitu: untuk mengukur tingkat penguasaan terhadap seperangkat materi atau tingkat pencapaian terhadap seperangkat tujuan tertentu ; dan untuk menentukan kedudukan atau peringkat siswa dalam kelompok, tentang penguasaan materi atau pencapaian tujuan pembelajaran tertentu.

Guru dapat mengembangkan tes sendiri untuk digunakan di kelasnya sendiri (*teacher-made test*) atau guru dapat menggunakan tes standar (*standardized test*). Tes standar adalah tes yang berdasarkan kepada standar-standar yang telah ditentukan : standar teknis dan praktis, serta telah diuji kelayakannya oleh para ahli. Sementara tes buatan guru hanya berlaku untuk menilai dan membandingkan siswa di kelasnya sendiri sementara tes standar dapat digunakan untuk menilai dan membandingkan kemampuan siswa lintas kelas dan lintas sekolah. Suatu tes standar dapat berlaku lebih umum misalnya suatu daerah, Negara atau berlaku internasional.

Beberapa tes standar yang telah dikembangkan yaitu : tes intelegensi, tes bakat, tes prestasi akademik, tes minat, dan tes kepribadian.

DAFTAR PUSTAKA

- Arikunto, Suharsimi. (1988). *Dasar-dasar Evaluasi Pendidikan*. Penerbit Bina Jakarta : Aksara.
- Ibrahim, R. dan Ali, M. (2007). “Teori Evaluasi Pendidikan”, dalam *Ilmu dan Aplikasi Pendidikan*. Bandung : Pedagogia Press.
- Sax, Gilbert. (1980). *Principles of Educational and Psychological Measurement and Evaluation*. (2nd ed.). California : Wadsworth Publishing Company
- Wirawan, Setia. (2002). “Implementasi Indeks Taraf Kesukaran dan Indeks Daya Pembeda untuk Membangun aplikasi “Pintar” Tes Hasil Belajar”. *Jurnal Informatika dan Komputer*. 3, (7). 150 -159