

## BAB VII

### STRATEGI EVALUASI PENDIDIKAN

#### A. Pengertian

Strategi evaluasi pendidikan adalah cara yang digunakan guru untuk mengukur dan menilai tingkat kemampuan anak berkaitan materi agama yang telah diberikan (kognitif, afektif dan psikomotorik).

#### B. Macam-macam Strategi

1. Dilihat dari cara dalam melaksanakan evaluasi dapat dibedakan menjadi tiga, yaitu
  - a. Evaluasi secara lisan, dalam hal ini berupa :
    - 1). Tes lisan untuk menyaring aspek kognitif
    - 2). Wawancara untuk menyaring aspek afektif
    - 3). Perintah melafalkan untuk menyaring aspek psikomotorik
  - b. Evaluasi secara tertulis untuk menguji kemampuan menyimpulkan/mengemukakan gagasan secara tertulis (kognitif), hasil belajar aspek afektif dan psikomotorik.
  - c. Evaluasi secara praktek, untuk mengetahui kecakapan anak (ketrampilan anak) berkaitan dengan Shalat, membaca al-Qur'an, Hadits, Wudlu, Zakat, Haji, dll.
2. Sedangkan dilihat dari cara memberi skor evaluasi dapat dibedakan menjadi dua yaitu:
  - a. Cara kualitatif, yaitu memberi skor dengan pertanyaan memuaskan, baik sekali, baik, cukup, sedang, kurang. Cara pemberian nilai demikian hanya dilakukan dalam rangka pembinaan semata, bukan untuk menentukan kenaikan atau kelulusan.
  - b. Cara kuantitatif, yaitu memberikan nilai dengan rentangan angka-angka, antara 0-10 atau 10-100 → kognitif, afektif dan psikomotorik.

#### C. Penyusunan Tes (*Test Construction*)

Terdapat dua belas langkah dalam menyusun tes, yakni:

1. Menetapkan tujuan tes:
  - a. Untuk tujuan apa hasil tes yang akan disusun itu dipergunakan ?

- b. Tujuan tes pencapaian belajar adalah untuk mendapatkan informasi tentang seberapa jauh siswa sudah menyerap isi bahan pengajaran yang disajikan dalam kegiatan belajar mengajar.
2. Analisis Kurikulum
- Untuk menetapkan isi bahan yang akan ditanyakan melalui tes itu.
  - Yang dianalisis : TK, TPU, PB, SPB beserta uraian yang ada dalam GBPP → esensial tidaknya.
3. Analisis buku pelajaran dan sumber materi belajar lainnya :
- Untuk memperinci pokok bahasan dan sub pokok bahasan.
  - Langkah ini (2 dan 3) diperlukan untuk memperkecil kesalahan/ kekeliruan (*error*) dalam memilih sampel bahan untuk soal tes. Kekurangcermatan dalam menganalisis buku dapat menimbulkan/ mengakibatkan kesimpulan/penilaian yang sesat.
4. Menyusun kisi-kisi tes (*blue print, table of test specification, lay out, plan, frame work*). Disusun dalam bentuk matriks dengan komponen-komponen :
- Jenis sekolah/jenjang sekolah
  - Program/jurusan
  - Bidang studi/mata pelajaran
  - Kurikulum yang diacu
  - TPU
  - Pokok bahasan/sub pokok bahasan
  - Bahan kelas : I, II, III = 1 : 2 : 3
  - Uraian materi → dari GBPP
  - Aspek Perilaku (*performance behaviour*)  
 Penetapan aspek kompetensi yang hendak diukur menurut Benyamin Bloom meliputi aspek:
    - Aspek Kognitif, meliputi tingkatan:
      - Ingatan atau pengetahuan faktual (I) atau C1
      - Pemahaman (P) atau C2
      - Aplikasi (penerapan) (A) atau C3
      - Analisis (A) atau C4
      - Sintesis (S) atau C5
      - Evaluasi (E) atau C6
 Proporsinya : I : (P + A) : (A + S + E) = 1 : 2 : 1
    - Aspek Afektif, tingkatan A1, A2, A3, A4, A5
    - Aspek Psikomotorik, tingkatan P1, P2, P3, P4, P5

- j. TPK (Indikator). Indikator adalah rumusan tingkah laku yang dapat diamati untuk digunakan sebagai petunjuk bahwa suatu tujuan pembelajaran tercapai. Pada dasarnya, indikator adalah perwujudan dari TPK.

Rumusan Indikator:

- Memuat ciri-ciri perilaku Tujuan Pembelajaran Umum (TPU) yang dipilih.
- Memuat salah satu kata kerja operasional.
- Berkaitan dengan uraian materi, PB/SPB, dan TPU

- k. Bentuk soal

- 1). Obyektif
- 2). Essay

- l. Tingkat kesukaran (didasarkan oleh pendapat guru)

- Mudah : soal dapat dijawab benar lebih dari 70%  
Sedang : soal dapat dijawab benar 50 – 70%  
Sukar : soal dapat dijawab benar kurang dari 50%

Adapun perbandingannya masing-masing tingkat kesukaran tersebut sebagai berikut 1 : 2 : 1. Tinggi rendahnya tingkat kesukaran bertalian dengan aspek intelektual (jenjang kemampuan).

- m. Bobot Soal

Bobot soal adalah besarnya angka yang ditetapkan pembuat soal berdasarkan tingkat kerumitan suatu soal dibandingkan dengan soal lainnya. Biasanya jumlah bobot soal bentuk uraian 100, sedangkan untuk pilihan ganda 1. Penetapan bobot soal didasarkan pada:

- 1). Tingkat kesukaran jawaban
- 2). Kedalaman materi yang dipertanyakan
- 3). Kepentingan soal
- 4). Kerumitan topik, karena berkaitan dengan lebih dari satu PB atau mata pelajaran lain (Suderadjat;2002:18).

- n. Jumlah soal dan waktu yang diperlukan dipertimbangkan dengan pokok bahasan, jenjang kemampuan yang hendak diukur, tingkat kesukaran soal dan bentuk soal.

5. Menulis TPK/indikator

6. Menulis soal. Adapun kegiatan yang dilakukan adalah :

- a. *Review* soal (menelaah soal)
- b. Seleksi soal
- c. Merakit soal menjadi tes

7. Reproduksi tes terbatas (memperbanyak untuk uji coba)
8. Uji coba (dengan sampel) → untuk pengkajian mutu soal
9. Analisis soal
  - a. Untuk mengetahui mutu butir-butir soal
  - b. Dengan melihat taraf kesukarannya, fungsi *steam* (pokok soal), fungsi *distractor* (pengecoh), penyebaran jawaban pada pengecoh dalam tabel kelompok.
10. Revisi soal → jelek → revisi → uji coba
11. Menentukan soal-soal yang baik
12. Merakit soal menjadi tes → tes standar : hasil perakitan butir-butir soal-soal yang telah dilakukan melalui proses analisis soal (Silverius; 1991:13-15).

# KISI-KISI PENULISAN SOAL TEST HASIL/PENCAPAIAN BELAJAR

Jenis Sekolah/Jenjang Sekolah : .....

Kelas/program/Jurusan : .....

Bidang Studi/Mata Pelajaran : .....

Kurikulum Acuan : .....

Jumlah Soal : .....

Waktu : .....

No Urut	No PB/SPB	PB/SPB	KLS	TK SKR	INGATAN (C1)					PEMAHAMAN (C2)					APLIKASI (C3)					ANALISIS (C4)					SINTESIS (C5)					EVALUASI (C6)					JUMLAH										
					A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	A	B	C	D	E	Mid	Sd	Sk	Soni	%						
1	2	3	4	5	6						7						8						9						10						11						1	2			
JUM	Variasi Soal																																												
LAH	Aspek Intelektual																																												
	Prosentase Aspek Intelektual																																												

Keterangan :

PB = Pokok Bahasan

SPB = Sub Pokok Bahasan

A = Pilihan ganda biasa

B = Analisis kasus

C = Grafik, tabel, gambar, matriks

D = Uraian obyektif

E = Uraian non-obyektif

Tkt Skr = Tingkat kesukaran

Md = Mudah

Sd = Sedang

Sk = Sukar

Format di atas merupakan bagian akhir dari rangkaian format kisi-kisi karena format ini ditutup dengan mencantumkan rekapitulasi jumlah soal menurut bentuk, aspek intelektual dan prosentasenya.

Komponen-komponen yang terdapat dalam matriks berikut cara pengisiannya adalah sebagai berikut :

No. Kolom	Komponen	Cara Pengisian
1	Nomor Urut	Cukup jelas
2	Nomor PB/SPB	Diisi berdasarkan nomor pokok bahasan/sub pokok bahasan terpilih
3	Pokok Bahasan/Sub Pokok Bahasan	Diisi berdasarkan PB/SPB terpilih
4	Kelas	Diisi tingkat kelas yang pokok bahasannya terpilih
5	TKt. Skr	Memuat tingkat kesukaran terdiri dari: Mudah (MD), Sedang (SD), Sukar (SK), yang akan diacu apabila menetapkan jenis soal menurut bentuk dan aspek yang diukur
6-11	Aspek intelektual	Diisi dengan angka pada salah satu bentuk soal (A-E) yang dipakai sesuai dengan aspek yang akan diukur. Angka tersebut menunjukkan jumlah butir soal yang ditulis sesuai dengan PB/SPB, kelas, tingkat kesukaran, aspek intelektual dan bentuk soal yang dipakai
12	Jumlah	Diisi dengan angka pada masing-masing sub kolom MD, SD, SD, yang menunjukkan jumlah soal yang ditulis dengan tingkat kesukaran tersebut. Sub kolom soal, diisi jumlah soal yang akan ditulis mengenai PB/SPB tertentu. Sedangkan sub kolom % menunjukkan prosentase jumlah setiap PB/SPB dari total jumlah soal seluruh tes

Tiga hal yang perlu dipertimbangkan pada pemberian nomor urut soal adalah:

1. Pokok bahasan/sub pokok bahasan yang saling berhubungan dituliskan berurutan dengan bagian yang merupakan prasarat ditempatkan di atas sebelum bagian yang membutuhkan prasarat itu. Yang dimaksud dengan prasarat disini adalah bagian yang harus dikuasai terlebih dahulu sebelum mengikuti atau menguasai bagian lain.

2. Nomor soal untuk bentuk soal yang sama harus berurutan.
3. Nomor soal dengan bentuk yang sama untuk pokok bahasan/sub pokok bahasan yang sama diurutkan menurut tingkat kesukarannya, mulai dari yang mudah ke yang sukar.

Kisi-kisi merupakan suatu pedoman umum untuk menulis soal dalam keseluruhan suatu tes. Jadi, kisi-kisi merupakan kerangka umum suatu tes. Berdasarkan kerangka umum inilah dibuatkan butir-butir soal tes tersebut. Untuk dapat menulis soal yang mengacu pada kisi-kisi tersebut dibutuhkan suatu format penulisan soal yang khusus.

### KISI-KISI PENULISAN SOAL UJIAN

Jenis Sekolah : .....

Bidang Studi/Mata Pelajaran : .....

Program Studi/Jurusan : .....

Kurikulum Acuan : .....

Alokasi Waktu: : .....

Jumlah Soal : .....

Bentuk Soal : .....

No. Urut	TIU (TPU)	PB/SPB	Jumlah soal per PB/SPB	Bahan kelas	Uraian Materi	Indikator	Nomor Soal
1	2	3	4	5	6	7	8
Jumlah							

Sumber : Pedoman penulisan soal SMTA, Departemen Pendidikan dan Kebudayaan Badan Penelitian dan Pengembangan Pendidikan dan Kebudayaan, Pusat Penelitian dan Pengembangan Sistem Pengujian, Jakarta, 1989.

#### **D. Perbandingan Antara Tes Diagnostik, Formatif dan Tes Sumatif**

Untuk memperoleh gambaran mengenai tes diagnostik, tes formatif dan tes sumatif secara lebih mendalam, berikut ini akan disajikan perbandingan antara ketiganya, agar dapat diketahui persamaan dan perbedaannya. Dalam membandingkan, akan ditinjau dari sembilan aspek, yaitu : fungsi, waktu, titik berat dan tekanannya, alat evaluasi, cara memilih tujuan yang dievaluasi, tingkat kesulitan soal-soal tes, cara menyekor tingkat pencapaian dan metode menuliskan hasil tes.

1. Ditinjau dari fungsinya, test dapat diklasifikasikan menjadi tiga, yaitu :
  - a. Tes Diagnostik
    - 1). Menentukan apakah bahan prasarat telah dikuasai atau belum.
    - 2). Menentukan tingkat penguasaan siswa terhadap bahan yang dipelajari.
    - 3). Memisah-misahkan, mengelompokkan siswa berdasarkan kemampuan dalam menerima pelajaran yang akan dipelajari.
    - 4). Menentukan kesulitan-kesulitan belajar yang akan dialami untuk menentukan cara yang khusus untuk mengatasi atau memberikan bimbingan.
  - b. Tes Formatif  
Sebagai umpan balik bagi siswa, guru, maupun program untuk menilai pelaksanaan satu unit program.
  - c. Tes Sumatif  
Untuk memberikan tanda kepada siswa bahwa telah mengikuti suatu program, serta menentukan posisi kemampuan siswa dibandingkan dengan kawannya dalam kelompok.
2. Ditinjau dari waktu
  - a. Tes diagnostik, diberikan:
    - 1). Pada waktu penyaringan calon siswa
    - 2). Pada waktu membagi kelas atau permulaan memberikan pelajaran
    - 3). Selama pelajaran berlangsung bila guru akan memberikan bantuan kepada siswa
  - b. Tes Formatif  
Dilaksanakan selama pelajaran berlangsung, untuk mengetahui kekurangannya agar pembelajaran dapat berlangsung sebaik-baiknya.
  - c. Tes Sumatif

Dilaksanakan pada akhir semester, dengan maksud untuk menentukan tingkat keberhasilan siswa dalam belajar selama satu semester.

3. Ditinjau dari titik berat penilaian, tes dibagi menjadi tiga, yaitu:
  - a. Tes Diagnostik, menekankan pada aspek:
    - 1). Tingkah laku kognitif, afektif dan psikomotorik
    - 2). Faktor-faktor fisik, psikologis dan lingkungan
  - b. Tes Formatif, menekankan pada tingkah laku kognitif.
  - c. Tes Sumatif  
Pada umumnya menekankan pada tingkah laku kognitif, tetapi ada kalanya pada tingkah laku psikomotorik dan kadang-kadang pada afektif. Akan tetapi walaupun menekankan pada tingkah laku kognitif, yang diukur adalah tingkatan yang lebih tinggi, bukan sekedar ingatan atau hafalan saja.
4. Ditinjau dari alat evaluasi
  - a. Tes diagnostik, terdiri dari:
    - 1). Tes prestasi belajar yang sudah distandarisasikan
    - 2). Tes diagnostik yang sudah distandarisasikan
    - 3). Tes buatan guru
    - 4). Pengamatan dan daftar cocok *check list*
  - b. Tes Formatif, dilakukan setiap akhir pokok bahasan dan dimaksudkan untuk mengetahui keberhasilan siswa dalam belajar dan keberhasilan guru dalam mengajar. Tes formatif juga dimaksudkan untuk perbaikan proses belajar mengajar.
  - c. Tes Sumatif, dilakukan pada akhir semester dan dimaksudkan untuk menentukan nilai raport siswa.
  - d. Tes ujian akhir, dilakukan pada akhir tahun, untuk menentukan lulus tidaknya /tamat tidaknya siswa dari suatu lembaga pendidikan.
5. Ditinjau dari cara memilih tujuan yang dievaluasi
  - a. Tes Diagnostik
    - 1). Memilih tiap-tiap ketrampilan prasarat
    - 2). Memilih tujuan setiap program pelajaran secara berimbang
    - 3). Memilih yang berhubungan dengan tingkah laku fisik, mental dan perasaan
  - b. Tes Formatif  
Mengukur semua tujuan instruksional khusus.

- c. Tes Sumatif  
M mengukur tujuan instruksional umum.
6. Ditinjau dari cara memilih tujuan yang dievaluasi
- a. Tes diagnostik  
Untuk tes diagnostik mengukur ketrampilan dasar, diambil banyak soal tes yang mudah, yang tingkat kesulitannya (indeks kesukaran) 0,65 atau lebih.
  - b. Tes Formatif  
Belum dapat ditentukan.
  - c. Tes Sumatif  
Rata-rata mempunyai tingkat kesulitan (indeks kesukaran) antara 0.35 sampai 0,70. Ditambah beberapa soal yang sangat mudah dan beberapa lagi yang sangat sukar.
7. Ditinjau dari *scoring* (cara menskor)
- a. Tes Diagnostik  
Menggunakan standar mutlak dan standar relatif (*criterion referenced* dan *norm referenced*).
  - b. Tes Formatif  
Menggunakan standar mutlak (*criterion-referenced*).
  - c. Tes Sumatif  
Kebanyakan menggunakan standar relatif (*norm referenced*), tetapi dapat pula dipakai standar mutlak (*criterion referenced*).
8. Ditinjau dari pencapaian
- Yang dimaksud dengan tingkat pencapaian adalah skor yang harus dicapai siswa dalam setiap tes. Tingkat pencapaian ini tidaklah sama. Tinggi rendahnya tuntutan terhadap tingkat pencapaian tergantung dari fungsi dan tujuan masing-masing tes.
- a. Tes diagnostik  
Berhubung ada bermacam-macam tes diagnostik, maka tingkat pencapaian yang dituntut juga tidak sama. Untuk tes diagnostik yang sifatnya memonitor kemajuan, tingkat pencapaian yang diperoleh siswa merupakan informasi tentang keberhasilannya. Tindakan guru selanjutnya adalah menyesuaikan dengan hasil tes diagnostik.  
Tes prasarat adalah tes diagnostik yang sifatnya khusus. Fungsinya adalah untuk mengetahui penguasaan bahan prasarat yang sangat penting untuk lanjutan studi bagi pengetahuan berikutnya. Untuk itu maka tingkat penguasaannya dituntut 100%.

b. Tes Formatif

Ditinjau dari tujuan, tes formatif digunakan untuk mengetahui apakah siswa sudah mencapai tujuan instruksional umum yang diuraikan menjadi tujuan instruksional khusus. Dalam sistem pendidikan yang lama tidak ada tuntutan terhadap pencapaian TIK/TPK, namun dalam tahun 1975, dengan keluarnya kurikulum tahun 1975 dan modul, tingkat pencapaian untuk tes formatif adalah 75% dari skor yang diharapkan, sedangkan pada kurikulum 1994 minimal 65% dari skor yang diharapkan, bila tidak, siswa yang bersangkutan diwajibkan menempuh kegiatan perbaikan (*remedial program*) sampai siswa yang bersangkutan lulus dalam tes. Siswa tersebut telah mencapai skor maksimal yang diharapkan. Adapun pada kurikulum 2002 standar pencapaiannya adalah kecakapan tertentu yang diraih siswa.

c. Tes Sumatif

Sesuai dengan fungsi tes sumatif, yaitu memberikan tanda kepada siswa bahwa mereka telah mengikuti suatu program, dan untuk menentukan posisi kemampuan siswa dibandingkan dengan kawan dalam kelompoknya, maka tidak diperlukan suatu tuntutan harus berapa tingkat penguasaan yang dicapai. Namun demikian, tidak berarti bahwa tes sumatif tidak penting. Perlu diingat bahwa tes sumatif ini dilaksanakan pada akhir program, berarti nilainya digunakan untuk menentukan kenaikan kelas atau kelulusan. Secara terpisah, tidak ditentukan tingkat pencapaiannya, tetapi secara keseluruhan akan dikenakan suatu norma tertentu yaitu norma kenaikan kelas atau norma kelulusan dan tamat.

9. Ditinjau dari cara pencatatan hasil

a. Tes diagnostik

Dicatat dan dilaporkan dalam bentuk profil.

b. Tes formatif

Prestasi tiap siswa dilaporkan dalam bentuk catatan berhasil atau gagal dalam menguasai sesuatu tugas.

c. Tes sumatif

Keseluruhan skor atau sebagian skor dan tujuan-tujuan yang dicapai.

Scawia B. Anderson membedakan tes menurut dimensi-dimensi seperti tersebut di bawah ini :

1. Tes ditinjau dari unsur suatu kegiatan dapat dibedakan atas : tes pengukur proses dan tes pengukur hasil.

2. Tes ditinjau dari tujuan penggunaan hasil, dibedakan atas : tes formatif, tes sub sumatif dan tes sumatif .
3. Tes ditinjau dari konstruksi yang diukur, dibedakan atas : tes kepribadian, tes bakat, tes kemampuan, tes minat, tes perhatian, tes sikap.
4. Ditinjau dari isi atau bidang studi dibedakan atas tes PAI, Bahasa, matematik, sejarah, IPA, Olah raga, keterampilan dan sebagainya.
5. Tes ditinjau dari lingkup dari materi yang diungkap, dibedakan atas tes pencapaian dan tes penelusuran. Tes hasil belajar mengungkap materi yang luas sedang tes penelusuran dikenakan pada sebagian kecil bahan agar tester dapat lebih cermat mengamati sesuatu.
6. Tes ditinjau keragaman butir atau tugas dibedakan atas tes homogen dan tes heterogen. Tes yang digunakan untuk mengukur sesuatu aspek misalnya faktor minat, maka testnya terdiri dari butir-butir yang seragam (homogen tes terstandar biasanya terdiri dari buti-butir yang heterogen).
7. Tes ditinjau dari cara tester memberikan respons, dibedakan atas tertulis, tes lisan, tes penampilan, tes pengenalan (benar-salah, pilihan ganda, menjodohkan dan sebagainya).
8. Tes ditinjau dari cara skoring dibedakan atas tes obyektif (dikenal dengan "check point") dan tes subyektif (tes yang memerlukan pertimbangan subyektifitas penilai).
9. Tes ditinjau dari standar dalam menentukan jawaban yakni tes yang menuntut adanya kebenaran mutlak (mengenai benar-salah) dan tes yang dimaksudkan untuk sekedar mengetahui keadaan seseorang misalnya tes untuk sikap atau pendapat seseorang.
10. Tes ditinjau dari cara pengadministrasian dibedakan atas *pre test* (tes awal) yang dilakukan sebelum diberikannya perlakuan, dan *post test* (tes akhir yang dilakukan sesudah adanya perlakuan).
11. Tes ditinjau dari tekanan aspek yang diukur, dibedakan atas "*speed test*". Yakni tes yang digunakan untuk mengukur kecepatan testee bekerja dan "*power test*" yakni tes yang digunakan untuk mengukur kemampuan testee. Pembedaan atas tes berdasarkan aspek ini dijumpai pada tes psikologi seperti halnya mengukur tes kemampuan umum (TKU).
12. Tes ditinjau dari banyaknya testee yang dites, dibedakan atas individual dan tes kelompok. Tes pengukuran intelegensi yang sifatnya klinis, merupakan contoh tes individual, sedangkan tes-tes yang berhubungan dengan pencapaian di lapangan pendidikan, industri atau militer, pada umumnya merupakan tes kelompok.

13. Tes ditinjau dari penyusunnya, dibedakan atas tes buatan guru dan tes yang diperdagangkan, yang dikenal dengan tes terstandar (Arikunto,1991:47-48).



## BAB VIII

### SKALA DAN PROSEDUR PENGUKURAN PROSES EFEKTIF

#### A. Teknik Evaluasi

Di dalam pendidikan aspek kognitif dievaluasi dengan teknik tes, aspek afektif dievaluasi dengan teknik non tes, dan aspek psikomotorik dievaluasi dengan tes perbuatan, begitu juga dengan evaluasi mata pelajaran Pendidikan Agama Islam..

1. Evaluasi aspek Kognitif, diarahkan untuk dapat mengungkapkan :
  - a. Segi penguasaan keluasaan materi pembelajaran. Pada Pendidikan Agama Islam meliputi Keimanan, Ibadah, Al-Qur'an, Hadits, Akhlak, Syariah, Mu'amalah, Tarikh.
  - b. Segi kedalaman materi (jenjang aspek kognitif : ingatan, pemahaman, penerapan, analisa, sintes dan evaluasi)

Langkah-langkah dalam penyusunan kisi-kisi :

- a. Perhatikan :
    - 1). Waktu yang tersedia, misal : 90 menit
    - 2). Tingkat kesukaran soal : 25% mudah, 50% sedang, 25% sukar
  - b. Ragam soal dengan perbandingan 35% essay, 65% obyektif  
 $35/100 \times 90 = 31.5$  menit soal essay,  $65/100 \times 90 = 58.5$  menit soal obyektif
  - c. Memperhitungkan tingkat kesukaran
2. Evaluasi Aspek Afektif → dengan non tes  
Perbedaan dengan tes :
    - a. Waktu pelaksanaan → hasil rekaman kondisi afektif siswa muncul pada waktu yang tidak diprogramkan.
    - b. Kondisi siswa yang dievaluasi kadang-kadang tidak ada → melalui orang ketiga sehingga data yang diperoleh tidak lengkap.Alat yang digunakan :
    - 1). Wawancara yang direncanakan, dapat menghemat waktu, hanya beberapa orang saja yang diwawancarai.
    - 2). Observasi : penafsiran dari pihak pengamatan atas penampilan sikap dan perilaku yang ditafsirkan sebagai gambaran dari sikap aktif/afektif siswa yang diobservasi. Bila hasilnya tidak valid → maka perlu pengulangan.

- 3). Angket : dapat dilaksanakan secara langsung (mengenai minat belajar, gairah beribadah), maupun tidak langsung : ketaatan Shalat.
3. Evaluasi Aspek Psikomotorik
  - a. Yang berkenaan dengan gerakan, misalnya: gerakan Wudhu, Shalat, manasik Haji → dievaluasi praktek gerakan.
  - b. Berkenaan dengan ucapan, misalnya: bacaan Al-Qur'an, bacaan Shalat, dll → dievaluasi dengan praktek lafadz.
  - c. Berkenaan dengan pengamalan, misalnya bidang keagamaan : ketaatan beribadah, kebiasaan dzikir, berdo'a, kegemaran membaca Al-Qur'an, dsb → dievaluasi dengan lembaran observasi

### B. Macam-Macam Skala Penilaian

Terdapat empat skala dalam penilaian Pendidikan, yakni:

1. Skala kelas, yaitu skala yang tidak tetap, ada kalanya memakai skala tertinggi 100, adakalanya 50, 25 dsb, tergantung dari banyak bentuk soal.
2. Skala 1-10, biasanya angka-angka pecahan dibulatkan, katakanlah 5, menjadi 6 dan 5,4 menjadi 5.
3. Skala 1-100. Dalam bilangan ini ditulis puluhan, katakanlah 5,5 menjadi 55, 5,4 menjadi 54.
4. Skala huruf, yakni A, B, C, D dan E

Untuk memberikan nilai rata-rata dari huruf, maka nilai huruf yang harus ditransfer lebih dahulu menjadi nilai angka.

Angka 100	Angka 10	PT	Huruf	Keterangan
80 - 100	8,0 - 10	8,1 - 10	A	Baik sekali
66 - 79	6,6 - 7,9	6,6 - 8,0	B	Baik
56 - 65	5,6 - 6,5	5,6 - 6,5	C	Cukup
50 - 55	4,0 - 5,5	4,1 - 5,5	D	Kurang
30 - 39	3,0 - 3,9	0 - 4,0	E	Gagal

## C. Penskoran dan Penilaian

### 1. Pengertian

Skor ialah hasil pekerjaan menskor yang diperoleh dengan menjumlahkan angka-angka bagi setiap soal tes yang dijawab benar oleh siswa, sedangkan nilai merupakan angka ubahan dari skor, di mana sudah dijadikan satu dengan skor-skor lain serta telah disesuaikan pengaturannya berdasarkan standar tertentu (Suderadjat;2002:19).

Contoh: peruntukan nilai bagi Aisyah

Dari tes yang diikutinya, Aisyah memperoleh nilai 35, sedangkan skor maksimum yang diharapkan adalah 50. Standar skor 100. Ini berarti, Aisyah hanya menguasai  $35/50 \times 100\%$  tujuan pembelajaran khusus atau hanya 70% dari tujuan tersebut. Di dalam daftar nilai dituliskan, Aisyah mendapat nilai 70.

### 2. Prosedur Pemberian Skor

Untuk mengolah skor, digunakan rumus sebagai berikut:

#### a. Bentuk Uraian Objektif (BUO)

##### 1). Prosedur Penyekoran

- Jawaban siswa yang benar diberi skor 1 dan yang salah diberi skor 0.
- Pada setiap soal, lakukan perhitungan skor yang dikalikan bobot dengan rumus:

$$S_i = \frac{a_i}{c} \times b$$

Keterangan:

$S_i$  : Skor yang sudah memperhitungkan bobot

$a_i$  : Skor yang diperoleh pada butir soal

$c$  : Skor maksimal untuk soal itu

$b$  : Bobot soal

Skor akhir didapatkan dengan menjumlahkan skor-skor yang diperoleh siswa dari jawaban-jawaban yang benar.

##### 2). Penentuan Nilai (N) siswa didapat dengan rumus:

$$N = \frac{J_{ss}}{B} \times 100$$

Keterangan:

$N$  : Nilai Siswa

$J_{ss}$  : Jumlah skor semua soal

$B$  : Bobot semua soal

b. Bentuk Uraian Non Objektif (BUNO)

Mengingat jawaban yang dituntut dari bentuk soal seperti ini terbuka dan bebas, tidaklah mungkin menyiapkan kunci jawabannya. Oleh karena itu, biasanya guru menilai setiap jawaban dengan menimbang-nimbang kualitasnya dalam hubungan dengan criteria yang ditetapkan sebelumnya. Jadi, bukan menyekor point demi point dengan kunci jawaban. Untuk itu, biasanya dilakukan dengan mengklasifikasikan jawaban itu ke dalam 5 tingkat. Selanjutnya, diberi nilai 0,1,2,3,4 dan A,B,C,D, dan E.

c. Bentuk Soal Isian dan melengkapi (*Fil in* dan *Completion*)

Terkait dengan penskoran bentuk isian dan melengkapi, ada dua pendapat. *Pertama*, jika suatu tes bentuk isian ada 10 item dan setiap item berisi satu isian atau 3 isian, cara menilainya dihitung menurut jumlah isian yang ada pada seluruh item. Pendapat *kedua*, menyatakan bahwa skor maksimum dihitung menurut jumlah itemnya. Setiap item dinilai satu meskipun jumlah isiannya tidak sama banyaknya.

Rumus Penyekoran:

$$S=R$$

Keterangan:

$S$  = Skor akhir

$R$  = Jumlah isian yang betul

d. *True False* (benar- salah)

1). Dengan denda  $S = R - W$

$S$  = Skor yang diperoleh

$R$  = *Right* (jawaban yang benar)

$W$  = *Wrong* (jawaban yang salah)

2). Tanpa denda  $S = R$

e. *Multiple Choise* (pilihan ganda)

1). Dengan denda  $S R - W/O - 1$



Jadi nilai Aisyah adalah 7,5

- 2). Konversi skor ke dalam nilai pada pendekatan PAN  
Skor yang dicapai *testee* dalam kelompok harus dihitung rata-rata (*mean*), penyebaran skor dan Simpang Bakunya (SD). Kemudian, dengan menggunakan skala actual atau ideal, dihitung nilai bagi setiap *testee*.

Contoh:

Pak Ahmad, guru Fiqh, memberi tes kepada kelas III MAU dengan soal isian, jumlah soal 20. Skor yang diperoleh siswa-siswanya sebagai berikut:

No	Nama/Skor	No	Nama/Skor	No	Nama/Skor	No	Nama/Skor
1	Aan/7	11	Dodi/10	21	Hasan/13	31	Isfi/10
2	Abi/6	12	Endah/15	22	Hasim/9	32	Iftah/5
3	Ahmad/8	13	Endang/11	23	Hakim/4	33	Ifadah/6
4	Bari/10	14	Ema/6	24	Hidayah/11	34	Jannah/12
5	Budi/6	15	Erni/9	25	Husnul/8	35	Junnah/10
6	Candra/8	16	Fajar/9	26	Humairoh/7	36	Khoiro/8
7	Cintia/7	17	Faqih/7	27	Indah/7	37	Khusnul/11
8	Deni/13	18	Fauzi/8	28	Ikhwani/13	38	Khaqim/6
9	Desi/5	19	Ginanjari/5	29	Ismi/9	39	Laila/7
10	Dina/4	20	Gunawan/7	30	Isnaini/10		

Berdasarkan data (skor) Fiqh siswa kelas III MAU diatas, kemudian dihitung MEAN ( $\bar{X}$ ) dan Simpang Baku (SD), dengan rumus Perhitungan

$$\text{MEAN} \quad \bar{X} = \sum xt + i \frac{fd}{n}$$

**Keterangan:**

X = mean/rata-rata hitung

xt = rata-rata hitungan terduga, yaitu titik tengah dari kelas di mana rata-rata hitung terduga terletak (biasanya di tengah-tengah)

i = interval

f = frekuensi

d = simpang terdua

n = jumlah *testee*

$$SD = \sqrt{\frac{fd^2}{n} - \left(\frac{fd}{n}\right)^2}$$

Perhitungan mean ( $\bar{x}$ ) dan simpang baku skor Fiqh siswa Pak Ahmad di kelas III MAU dimulai dengan membuat tabel seperti berikut:

Kelas Interval	f	D	fd	fd 2
4-6	10	-1	10	10
7-9	16	0	0	0
10-13	12	1	12	12
14-16	1	2	2	4
	39		4	26

$$\text{Mean} = X = 8 + 3\left(\frac{4}{39}\right)$$

$$= 8,3$$

$$\text{Simpang Baku} = SD = 3 \cdot \sqrt{\frac{26}{39} - \left(\frac{4}{39}\right)^2}$$

$$= 3 \sqrt{0,66 - 0,01}$$

$$= 2,4$$

Dari hasil perhitungan mean dan simpang baku di atas, selanjutnya tinggal dibuat tabel konversi-skor ke dalam nilai dengan menggunakan skala sigma aktual seperti berikut:

**Tabel Konversi Skor ke Dalam Nilai Standar 10  
Dengan Skala Sigma Aktual (S.SA)**

No	X+SSA (SD)	Interval Skor	St 10
1	8,3 + 2,25 (2,4)	13,7 ke atas	10
2	8,3 + 1,75 (2,4)	12,6-13,6	9
3	8,3 + 1,25 (2,4)	11,3-12,5	8
4	8,3 + 0,75 (2,4)	10,1-11,2	7
5	8,3 + 0,25 (2,4)	8,9-10,0	6
6	8,3 - 0,25 (2,4)	7,7-8,8	5
7	8,3 - 0,75 (2,4)	6,5-7,6	4
8	8,3 - 1,25 (2,4)	5,3-6,4	3
9	8,3 - 1,75 (2,4)	4,1-5,2	2
10	8,3 + -2,25 (2,4)	2,9-4,0	1

Berdasarkan dari tabel di atas, skor setiap siswa diubah ke dalam nilai.

**Contoh:**

- \* Endang dengan skor 15 mendapatkan nilai 10 karena terletak interval skor 13,7 ke atas.
- \* Bari dengan skor 10 mendapat nilai 6 karena terletak di antara 8,9 sampai 10,10.
- \* dan seterusnya.

**4. Menetapkan Batas Kelulusan**

Batas kelulusan yang berorientasi pada PAN, dipakai *batas lulus aktual* atau *batas lulus ideal*, yaitu dengan didasarkan pada nilai rata-rata (*mean*) dan simpang baku (SD), sedangkan batas kelulusan pada sistem yang berorientasi PAP adalah *lulus purposif* (ditentukan berdasarkan kriteria tertentu).

a. Batas Lulus Aktual

Batas lulus ini didasarkan pada rata-rata nilai yang dicapai oleh kelompok siswa (*actual*). Dasar yang diperlukan untuk menentukan batas lulus aktual adalah nilai rata-rata aktual dan simpang baku aktual, biasanya skor yang dinyatakan lulus adalah skor di atas.

$$(\bar{x} + 0,25 \text{ SD})$$

**Contoh:**

Tes Al Qur'an Hadits berbentuk pilihan ganda dengan item 60. Skor untuk jawaban benar 1 sehingga skor maksimum 60. Jika Skor rata-rata kelas 25 dan simpang bakunya 8,0, batas lulusnya:

$$\begin{aligned} \text{Batas lulus} &= (\bar{X} + 0,25 \text{ SD}) && = \text{Skor Rata-Rata} \\ \text{Kelas} & && \\ &= 25 + 0,25 (8,0) && \text{SD} = \text{Simpangan baku/} \\ \text{Deviasi Standar} & && \\ &= 27 && \end{aligned}$$

Jadi, skor di atas 27 dinyatakan lulus, sedangkan di bawah 27 dinyatakan tidak lulus.

b. **Batas Lulus Ideal**

Skor rata-rata dan simpang baku untuk batas lulus ideal digunakan aturan sebagai berikut:

*Nilai rata-rata ideal adalah setengah dari maksimum skor, sedangkan simpang baku ideal sepertiga dari nilai rata-rata ideal*

**Contoh:**

Jika skor maksimum tes Al Qur'an Hadits yang mungkin dicapai siswa 60, rata-rata idealnya  $\times 60$  yaitu 30, sedangkan untuk simpang bakunya  $\times 30$  yaitu 10 sehingga batas lulus tes Al Qur'an Hadits tersebut:  $30 + 0,25 (10) = 32,50$

c. **Batas Lulus Purposif**

Batas lulus purposif mengacu pada penilaian PAP, tidak perlu menghitung nilai rata-rata dan simpang baku. Dalam hal ini, ditentukan kriterianya. Misalnya 75% artinya skor yang dinyatakan lulus adalah skor di atas 75% dari skor maksimum yang mungkin dicapai.

Jadi, jika skor maksimum yang mungkin dicapai siswa adalah 80, batas lulusnya  $75\% \times 80$  yaitu 60. Artinya, skor di atas 60 dinyatakan lulus, sedangkan di bawah 60 dinyatakan gagal.

**5. Kedudukan Siswa dalam Kelompok (Rank)**

Kedudukan siswa dalam kelompok adalah letak seorang siswa di dalam urutan tingkat prestasinya. Dalam istilah umum disebut *ranking*. Ada beberapa cara untuk merencanakan kedudukan siswa di kelasnya, dua diantaranya dengan *ranking sederhana* dan *ranking presentile*.

a. Ranking Sederhana (*Simple Rank*)

Cara ini paling sering dipakai oleh para wali kelas dalam meranking kelas asuhannya mengingat kemudahan pengerjaannya, yaitu dengan cara mengurut skor yang paling tinggi sampai skor terendah dengan urutan ke bawah. Kemudian, tentukan nomor urut dari atas, 1, 2,3, ... dan seterusnya sampai semua siswa mendapat nomor.

**Contoh:**

Skor tes Al Qur'an Hadits bagi 20 siswa adalah:

A=68    F=70    K=40    P=60  
 B=73    G=63    L=82    Q=55  
 C=65    H=75    M=75    R=48  
 D=80    I=35    N=62    S=53  
 E=75    J=60    O=57    T=59

Selanjutnya, data tersebut diurutkan sebagai berikut:

Nama Siswa	Skor	Nomor Urut	Ranking
L	82	1	Ke-1
D	80	2	Ke-2
H	75	3	Ke-4
M	75	4	Ke-4
E	75	5	Ke-4
B	73	6	Ke-6
F	70	7	Ke-7
A	68	8	Ke-8
C	65	9	Ke-9
G	63	10	Ke-10
N	62	11	Ke-11
J	60	12	Ke-13
P	60	13	Ke-13
T	59	14	Ke-14
O	57	15	Ke-15
Q	55	16	Ke-16
S	53	17	Ke-17
R	48	18	Ke-18
K	40	19	Ke-19
I	35	20	Ke-20

**Catatan:**

Siswa yang mempunyai skor sama, juga mempunyai ranking sama, sehingga ada nomor-nomor urut yang tidak digunakan pada contoh di atas, H,M dan E mendapat ranking ke 4 diperoleh dari

$\frac{3+4+5}{3}$  J dan P mendapat ranking ke-13 diperoleh dari

$$\frac{12+13}{2}=12,5 \text{ dibulatkan menjadi } 13. \text{ Ranking ke-3,5 dan 12 tidak}$$

digunakan. Rank terakhir selalu sama dengan nomor urut siswa atau banyak siswa dalam kelompok.

b. Ranking Presentasi (*Percentile Rank*)

Ranking presentasi adalah kedudukan siswa dalam kelompok yang menunjukkan banyaknya prosentase yang berada di bawahnya. Jadi, dalam hal ini siswa dibandingkan dengan siswa lain yang mempunyai skor sama atau lebih kecil dari padanya.

Dengan demikian, prestasi siswa lebih tergambar jika dibandingkan dengan cara perankingan *simple rank* karena pada *percentile rank* ini, angka ranking menunjukkan besarnya prosentase siswa dalam kelompok itu yang berhasil dilampaui.

**Contoh:**

Jika siswa memiliki PR (*percentile rank*) 75, berarti kecakapan siswa tersebut sama atau melebihi 75% dari seluruh kelompok.

Cara menentukan PR (*percentile rank*):

- 1). Tentukan dulu SR (*simple ranknya*)
- 2). Cari banyaknya siswa dalam kelompok itu yang berada di bawahnya.
- 3). Kalikan 100 setelah dibagi dengan kelompok.

**Contoh:**

Dari tabel *simple rank* untuk 20 orang, Zahro menduduki ranking ke-8, maka siswa yang berada di bawahnya berjumlah (20-8) atau

12 orang. Jadi, PR untuk Zahro adalah:  $\frac{12}{20} \times 100$  yaitu 60

Artinya, Zahro itu letaknya dalam kelompok mengalahkan sebanyak 60% untuk prestasi yang bersangkutan.

Rumus untuk menentukan PR:  $PR = \frac{N - SR}{n} \times 100$

**Keterangan:**

N = Jumlah siswa

SR = Ranking Sederhana

n =

## 6. Nilai Akhir

Nilai akhir adalah nilai yang dicantumkan di dalam buku rapor atau Surat Tanda Tamat Belajar (STTB). Bagi seorang siswa, nilai mempunyai arti penting karena merupakan pencerminan dari keberhasilan belajarnya. Tinggi rendahnya nilai dapat memotivasi siswa dalam usaha memperbaiki dan meningkatkan prestasi berikutnya. Oleh karena itu, pemberian nilai akhir hendaknya dilakukan dengan cermat dan terperinci.

### a. Pertimbangan Penentuan Nilai Akhir

Walaupun penetapan nilai di setiap sekolah tidak sama, secara garis besar mempertimbangkan faktor-faktor sebagai berikut:

#### 1). Prestasi

Nilai, apakah itu dalam bentuk angka atau huruf hendaknya hanya merupakan gambaran tentang prestasi (tingkat penguasaan kompetensi) siswa.

#### 2). Usaha

Guru sering beranggapan, bahwa siswa yang memiliki prestasi rendah, rendah pula usahanya. Sebaliknya, yang memiliki prestasi tinggi dianggap tinggi pula usahanya. Padahal belum tentu demikian. Walaupun antara usaha dan prestasi mempunyai hubungan erat, sebaliknya tentang usaha siswa ini dilaporkan secara terpisah dari prestasi.

#### 3). Aspek Pribadi dan Sosial

Yang dinilai di sini adalah respon siswa yang berhubungan langsung dengan proses pembelajaran, misalnya, tata tertib.

#### 4). Kebiasaan Kerja

Yang dinilai adalah hal-hal yang menyangkut kebiasaan melakukan tugas, seperti kerja kelompok, segera mengerjakan tugas, ketelitian kerja, kerapian kerja, dan sebagainya.

### b. Cara Menentukan Nilai Akhir

Penentuan nilai akhir terutama dilakukan pada waktu akan menetapkan nilai untuk dilaporkan pada buku rapor, Surat Tanda Tamat Belajar (STTB), menetapkan kriteria naik kelas dan kelulusan. Penentuan nilai akhir ini, guru mengikuti pedoman yang tidak ditetapkan oleh pemerintah atau kebijakan Yayasan. Berikut ini contoh cara menetapkan pelaporan nilai:

#### 1). Nilai Rapor dan STTB

Untuk nilai rapor digunakan rumus:

$$a). Na = \frac{2p + q + r}{5}$$

$$b). Na = \frac{2r + 3p + 5q}{10}$$

**Keterangan:**

2,3, dan 5 = Bobot

p = Nilai sub sumatif setiap topik/nilai sub kompetensi dan setiap sub diperoleh dari nilai rata-rata.

$$p = \frac{K1 + K2 + K3 \dots + Kn}{n}$$

**Keterangan:**

K1 = Nilai kompetensi

n = Jumlah kompetensi/modul/topik

Sedangkan untuk nilai STTB digunakan rumus:

$$A = \frac{K + 2R}{3}$$

**Keterangan:**

A = prestasi rata-rata

K = Nilai rata-rata semester

R = Nilai hasil Ujian Nasional (UN) Wewenang Sekolah/Ujian Nasional Wewenang Pusat

Pada tahun 2002/2003, penentuan Nilai Akhir/STTB digunakan rumus sebagai berikut:

a). Tes Tulis

$$NT = \frac{\text{JumlahSkorPerolehan}}{\text{JumlahSkorMaksimum}} \times \text{Bobot}$$

**Keterangan:**

NT = Nilai Ujian Tertulis

Skor Perolehan = Skor yang diperoleh siswa dalam ujian

Skor Maksimum = \* Untuk pilihan ganda tiap butir 1

\* Untuk Uraian tiap butir ditentukan oleh guru

Misalnya hasil tes Fiqih terdiri dari 50 soal pilihan ganda 5 soal uraian diperoleh hasil sebagai berikut:

No	Bentuk Soal	Jumlah Soal	Nomor Soal	Skor Maksimum	Skor Perolehan
1	Pilihan Ganda	50	1-50	50	40
2	Uraian	5	51	5	4
			52	3	3
			53	2	2
			54	5	3
			55	2	2
Jumlah		55		67	54

Maka nilai tes tulis Fiqih adalah:

$$NT = \frac{54}{67} \times 10 = 8,06$$

b). Tes Praktek

Aspek yang dinilai:

- Persiapan dengan bobot 2
- Proses dengan bobot 5
- Hasil dengan bobot 3

Rumus Perhitungan:

$$NP = \frac{\text{Skor Perolehan}}{\text{Skor Maksimum}} \times \text{Bobot}$$

**Keterangan:**

NP= Nilai Praktek

Skor Perolehan= Skor yang diperoleh siswa dalam ujian praktek.

Skor Maksimum= Skor tiap aspek yang dinilai dan ditentukan oleh guru

Misalnya hasil ujian Praktek Fiqih diperoleh sebagai berikut:

Aspek Yang Dinilai	Skor Maksimum	Bobot	Skor Perolehan
Persiapan	5	2	4
Proses	11	5	8
Hasil	24	3	20

Maka nilai praktek adalah:

$$\text{Persiapan} = \frac{4}{5} \times 2 = 1,6$$

$$\text{Proses} = \frac{8}{11} \times 5 = 3,8$$

$$\text{Hasil} = \frac{20}{24} \times 3 = 2,5$$

$$\text{Jumlah} = 7,9$$

c). Nilai Akhir

$$NA = \frac{NT + NP}{2}$$

**Keterangan:**

NA = Nilai Akhir Ujian dengan rentang nilai 0-10 dengan 2 digit di belakang koma, misalnya, 4,00 dst.

NT = Nilai Ujian Tulis

NP = Nilai Ujian Praktek

Maka Nilai STTB mata pelajaran Fiqih adalah

$$NA = \frac{8,06 + 7,9}{2} = \frac{15,96}{2} = 7,98 \text{ dibulatkan menjadi } 8$$

2). Penentuan Nilai Kenaikan Kelas

Kenaikan kelas memiliki kriteria sebagai berikut:

- Memiliki nilai setiap pelajaran pada setiap semester
- Nilai NPK (Nilai Penentu Kenaikan) minimum 6 (enam)

$$\text{NPK} = \frac{C1 + 2C2}{3} \geq 6$$

**Keterangan:**

C1 = nilai rata-rata semester 1/3/5

C2 = nilai rata-rata semester 2/4/6

c). Kehadiran kumulatif 95%

3). Penentuan kelulusan

Siswa yang dinyatakan lulus bila telah menyelesaikan keseluruhan program dengan pemilikan kompetensi dasar.

Kriteria kelulusan:

a). Aspek Akademis:

(1). Memiliki nilai seluruh mata pelajaran yang diujikan secara nasional.

(2). Tidak terdapat nilai  $\leq 3,00$

(3). Nilai rata-rata seluruh mata pelajaran yang diujikan secara nasional minimal 6,00

b). Aspek Non Akademis, meliputi: Kelakuan, Kerajinan, dan Kerapian.

## BAB IX

### VALIDITAS DAN RELIABILITAS

Suatu tes dapat dikatakan baik bilamana memiliki ciri sebagai alat ukur yang baik. Kriterianya antara lain: memiliki validitas yang cukup tinggi, reliabilitas yang baik, dan memiliki nilai kepraktisan (Thoha, 1994:109).

#### A. VALIDITAS

##### 1. Pengertian

Validitas adalah kesohehan. Suatu test bisa dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur (*A test is valid if it measures what it purpose to measure*) (Arikunto, 1991:63). Penekanannya bukannya pada tes sendiri, melainkan pada hasil pengetesan atau skornya. Hasil tes sesuai dengan kurikulum, artinya memiliki kesejajaran antara hasil tes dengan kurikulum

##### 2. Macam-Macam Validitas

Terdapat dua macam validitas dalam alat evaluasi, yaitu validitas logis dan validitas empiris.

- a. Validitas Logis (*logical validity*), yaitu suatu validitas yang diketahui dari hasil pemikiran. Dalam hal ini ada dua macam :
  - 1). Validitas isi (*content validity*), yaitu mampu mengukur tujuan khusus tertentu sejajar dengan materi atau isi pelajaran yang diberikan. Disebut juga validitas kurikuler, karena materi yang diajarkan tertera dalam kurikulum. Pencapaian validitas isi dapat diusahakan sejak penyusunan (tes), dengan cara memperinci kurikulum atau materi pelajaran. Yang menjadi kriteria adalah materi dan tujuan yang ada dalam kurikulum.
  - 2). Validitas konstruksi (*construct validity*), yaitu butir-butir soal yang membangun tes tersebut mengukur setiap aspek berfikir, seperti yang disebutkan dalam TPK. Validitas konstruksi dapat diketahui dengan cara memperinci dan memasang setiap butir soal dengan setiap aspek dalam TPK (yang menjadi kriteria aspek-aspek berfikir).
- b. Validitas Empiris (*Empirical Validity*), yaitu suatu validitas yang diketahui dari hasil pengamatan. Dalam hal ini ada dua macam :

- a). Validitas ada sekarang (*concurrent validity*), yaitu sesuai dengan pengalaman (masa lalu) dengan kehidupan masa kini (yang menjadi alat banding atau kriteria ialah hasil tes masa lalu).
- b). Validitas prediksi (*predictive validity*), yaitu mampu meramalkan apa yang akan terjadi pada masa yang akan datang (yang menjadi alat banding atau kriteria hasil tes nanti). Misalnya : tes masuk perguruan tinggi yang diharapkan mampu meramalkan keberhasilan dibangku kuliah. Jadi yang dijadikan bahan banding adalah prestasi yang diperoleh dibangku kuliah.

Dilihat dari cara mengukur soal, jenis validitas juga ada dua macam, yaitu :

- a. Validitas seluruh soal
- b. Validitas tiap item soal

### 3. Cara mengetahui validitas alat ukur :

Untuk mengetahui validitas alat ukur digunakan teknik korelasi *product moment* yang dikemukakan Pearson (Arikunto, 1991:67-71):

- a. Dengan Simpangan

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

$r_{xy}$  = Koefisien korelasi antara variabel x dan variabel y, dua variabel yang dikorelasikan ( $x = X - \bar{X}$

dan  $y = Y - \bar{Y}$ )

$\sum xy$  = Jumlah perkalian x dengan y

$x^2$  = Kuadrat dari x

$y^2$  = Kuadrat dari y

- b. Dengan angka kasar

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

Koefisien korelasi itu selalu terdapat antara -1,00 sampai +1,00. Namun karena dalam menghitung sering dilakukan

pembulatan angka-angka, maka sangat mungkin diperoleh koefisien lebih dari 1,00. Koefisien negatif menunjukkan hubungan kebalikan, sedangkan koefisien positif menunjukkan adanya kesejajaran.

Besarnya koefisien tersebut dapat diinterpretasikan sebagai berikut :

- 1). sangat tinggi = antara 0,800 – 1,00
- 2). tinggi = antara 0,600 – 0,800
- 3). cukup = antara 0,400 – 0,600
- 4). rendah = antara 0,200 – 0,400
- 5). sangat rendah = antara 0,000 – 0,200

Cara penafsiran harga koefisien korelasi :

- a. Dengan melihat harga  $r$  dan diinterpretasikan, misalnya korelasi tinggi, cukup 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.

#### 4. Pengukuran Validitas Tes Keseluruhan

Langkah-langkah pengukuran validitas-validitas tes keseluruhan :

- a. Mencari *score x* (nilai tes yang akan dicari validitasnya) dan *score y* (kriteria test).
- b. Membuat tabel persiapan
  - 1). Dengan simpangan

No.	Nama	X	Y	x	y	x <sup>2</sup>	y <sup>2</sup>	xy
1.	Ria	6	6	-1,2	-1,2	2,4	2,4	2,4
2.	Sonia	7	6	-0,2	-1,2	0,4	2,4	0,24
3.	Satria	8	7	+0,8	-0,2	0,64	0,4	0,16
4.	Rudi	9	8	-1,7	+0,8	2,89	0,64	1,36
Jumlah		30	27			6,33	5,84	4,16

#### 2). Tabel dengan angka kasar

No.	Nama	X	Y	x <sup>2</sup>	y <sup>2</sup>	xy
1.	Ria	6	6	36	36	36
2.	Sonia	7	6	49	36	42
3.	Satria	8	7	64	49	63
4.	Rudi	9	8	81	64	72
Jumlah		30	27	230	185	213

$$\bar{X} = \frac{\Sigma X}{N} = \frac{30}{4} = 7,2$$

$$\bar{Y} = \frac{\Sigma Y}{N} = \frac{27}{4} = 7$$

$$x = X - \bar{X}$$

$$y = Y - \bar{Y}$$

c. Dimasukkan ke dalam rumus :

1). Dengan Simpangan

$$r_{xy} = \frac{\Sigma xy}{\sqrt{(\Sigma x^2)(\Sigma y^2)}}$$

$$r_{xy} = \frac{4,16}{\sqrt{(6,33)(5,48)}}$$

$$= \frac{4,16}{\sqrt{36,97}}$$

$$= \frac{4,16}{6,08} = 0,68 \text{ (validitas sangat tinggi)}$$

2). Dengan angka kasar

$$r_{xy} = \frac{N \cdot \Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{(N \cdot \Sigma x^2 - (\Sigma x)^2)(N \cdot \Sigma y^2 - (\Sigma y)^2)}}$$

$$= \frac{4 \cdot 213 - (30)(27)}{\sqrt{(4 \times 230 - (30)^2)(4 \times 185 - (27)^2)}}$$

$$= \frac{852 - 810}{\sqrt{(920 - 900)(740 - 729)}}$$

$$= \frac{44}{\sqrt{(20 \times 11)}}$$

$$= \frac{44}{14,83} = 2,83 \text{ (validitas sangat tinggi)}$$

### 5. Perhitungan Tiap Item (Validitas Butir Soal)

Sebuah item dikatakan valid apabila mempunyai dukungan yang besar terhadap skor tabel :

Validitas x adalah skor tiap item

Validitas y adalah skor tabel

Tabel item untuk validitas tes Pendidikan Agama Islam

No	Nama	Butir Soal / Item										Skor Tabel
		1	2	3	4	5	6	7	8	9	10	
1.	Ria	1	1	1	0	1	1	1	0	0	0	6
2.	Sonia	1	0	1	1	1	0	0	1	1	1	7
3.	Satria	0	1	1	1	0	1	1	1	1	1	8
4.	Rudi	1	1	0	1	1	1	1	1	1	1	9

Validitas soal no. 6

No.	Nama	X	Y	$x^2$	$y^2$	xy
1.	Ria	1	6	1	36	6
2.	Sonia	0	7	0	49	0
3.	Satria	1	8	1	64	8
4.	Rudi	1	9	1	81	9
Jumlah		3	30	3	230	23

$$\begin{aligned}
 r_{xy} &= \frac{N \cdot \Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{(N \cdot \Sigma x^2 - (\Sigma x)^2)(N \cdot \Sigma y^2 - (\Sigma y)^2)}} \\
 &= \frac{4 \cdot 23 - (3)(30)}{\sqrt{(4 \cdot 3 - 9)(4 \cdot 230 - 900)}} \\
 &= \frac{92 - 90}{\sqrt{(12 - 9)(920 - 900)}} \\
 &= \frac{2}{\sqrt{(3 \cdot 20)}} \\
 &= \frac{2}{7,75} = 0,62 \text{ (validitas rendah)}
 \end{aligned}$$

atau dengan rumus :

$$Y_{pbi} = \frac{mp - mt \sqrt{p/q}}{St}$$

$Y_{pbi}$  = koefisien korelasi

$mp$  = rerata skor dari subyek yang menjawab betul

$mt$  = rerata skor tabel

$St$  = standar deviasi dari skor tabel

$p$  = proporsi siswa yang menjawab benar

$q$  = proporsi siswa yang menjawab salah ( $q = 1 - p$ )

Bila soal nomor 6 tadi dicari validitasnya :

a. Mencari  $mp = (6 + 8 + 9)/3 = 23/3 = 7,67$

b. Mencari  $mt = (6 + 7 + 9)/4 = 30/4 = 7,5$

c. Mencari standar deviasi

$$\begin{aligned}
 St &= \sqrt{\left(\frac{3}{4}\right) - \left(\frac{3}{4}\right)^2} \\
 &= \sqrt{0,75 - 0,56} \\
 &= \sqrt{19} = 0,44
 \end{aligned}$$

d. Menentukan harga  $p$ , yaitu  $\frac{3}{4} = 0,75$

e. Menentukan harga  $q = \frac{1}{4} = 0,25$  atau  $1 - 0,75 = 0,25$

f. Memasukkan ke dalam rumus :

$$\begin{aligned}
 Y_{pbi} &= \frac{7,67 - 7,65 \sqrt{0,75/0,25}}{0,44} \\
 &= \frac{0,17\sqrt{3}}{0,44} \\
 &= \frac{0,17 \cdot 1,73}{0,44} \\
 &= 0,39
 \end{aligned}$$

## B. RELIABILITAS

### 1. Pengertian :

Suatu tes dikatakan mempunyai taraf kepercayaan tinggi jika tes tersebut memberikan hasil yang tetap. Ketetapan suatu tes apabila diteskan pada subyek yang sama.

### 2. Beberapa hal yang mempengaruhi hasil tes:

- Berhubungan dengan tes sendiri, yaitu panjang tes dan kualitas butir-butir soalnya, semakin panjang tes, maka reliabilitasnya semakin tinggi @ @ penambahan jumlah butir soal.

Rumus :  $nr / 1 + (n-1) r$

$nr$  = besarnya koefisien reliabilitas sesudah tes tersebut ditambah butir soal baru

$n$  = berapa kali butir-butir soal itu ditambah

$r$  = besarnya koefisien reliabilitas sebelum butir-butir soalnya ditambah

Penambahan butir soal tes kadang-kadang tidak berarti, bahkan merugikan hal ini disebabkan :

- Sampai pada suatu batas tertentu, penambahan banyaknya butir soal sudah tidak menambah tinggi reliabel tes.
- Penambahan tingginya reliabilitas tes tidak sebanding nilainya dengan waktu, biaya dan tenaga yang dikeluarkan.

Adapun kualitas butir soal ditentukan oleh :

- Jelas tidaknya rumusan soal
- Baik tidaknya pengenalan soal kepada jawaban sehingga tidak menimbulkan salah jawaban.
- Petunjuknya jelas sehingga mudah dan cepat dikerjakan.

- b. Hal-hal yang berhubungan dengan terdoba (*testee*)
- c. Hal-hal yang berhubungan dengan penyelenggaraan tes (administratif)
  - 1). Petunjuk yang diberikan sebelum tes.
  - 2). Pengawas yang baik
  - 3). Suasana lingkungan dan tempat tes.

### 3. Cara Mencari Besarnya Reliabilitas

Kriteria yang digunakan untuk mengetahui ketetapan ada yang berada di luar tes (*consistency external*) dan pada tes itu sendiri. Berikut merupakan cara mencari besarnya reliabilitas pada tes obyektif. Terdapat tiga cara dalam mencari besarnya reliabilitas, yaitu dengan menggunakan metode Bentuk Paralel, Metode Tes Ulang, dan Metode Belah Dua (Arikunto, 1991:85-88).

#### a. Metode bentuk paralel (*equivalent*) = *double test – double – trial method*

Tes paralel atau tes ekuivalen yaitu dua buah tes yang mempunyai kesamaan tujuan, tingkat kesukaran dan susunan, tetapi butir-butir soalnya berbeda. Dalam bahasa Inggris disebut *Alternative – forms method* (paralel forms).

Dengan metode ini, dua buah yang paralel, misalnya tes Fiqh seri A yang akan dicari reliabilitasnya (x) dan tes seri B (Y) diteskan kepada sekelompok siswa yang sama, kemudian hasilnya dikorelasikan. Koefisien korelasi dari kedua hasil tes inilah yang menunjukkan koefisien reliabilitas tes seri A. Jika koefisiennya tinggi, maka tes tersebut sudah reliabel dan dapat digunakan sebagai alat pengetes yang terandalkan.

Rumus yang digunakan : *Product Moment* :

Angka kasar

$$r_{xy} = \frac{N \cdot \Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{(N \cdot \Sigma x^2 - (\Sigma x)^2)(N \cdot \Sigma y^2 - (\Sigma y)^2)}}$$

Langkah-langkah :

- 1) Mencari skore x dan y
- 2) Membuat tabel persiapan

No.	Nama	X	Y	x <sup>2</sup>	y <sup>2</sup>	xy
1.	Ria	6	6	36	36	36
2.	Sonia	7	6	49	36	42
3.	Satria	8	7	64	49	63
4.	Rudi	9	8	81	64	72
Jumlah		3033	27	230	185	213

3) Memasukkan ke dalam rumus :

$$\begin{aligned}
 & \frac{4.213 - (30)(27)}{\sqrt{(4 \times 30 - 900)(4 \times 185 - 729)}} \\
 &= \frac{852 - 810}{\sqrt{(920 - 900)(7400 - 729)}} \\
 &= \frac{42}{\sqrt{(20 \times 11)}} \\
 &= \frac{42}{14,49} = 2,9
 \end{aligned}$$

Kelebihannya :

Terhindar dari *practice - effect* dan *carry - over effect* (yakni masih ingat soalnya, karena pernah mengerjakan soal tersebut).

Kelemahan :

- 1). Pengetes pekerjaannya berat, karena harus menyusun dua seri tes.
- 2). Membutuhkan waktu lama, karena mencoba dua tes.

**b. Metode Tes Ulang (*Test retest method*) = *Single-test-double-trial method***

Metode ini digunakan untuk menghindari pengukuran dua seri tes. Dengan metode ini, pengetes hanya memiliki satu seri tes tetapi dicobakan dua kali pada siswa yang sama. Untuk tes yang bertujuan untuk mengungkap pengetahuan dan pemahaman, cara

ini kurang tepat, karena si ter coba akan mengingat butir-butir soal yang ada. Oleh karena itu, tenggang waktu antara pemberian tes pertama dengan kedua menjadi persoalan tersendiri. Bila terlalu sempit, siswa masih banyak ingat tentang butir-butir soalnya. Namun bila terlalu lama, kondisi menjadi berbeda dan siswa sendiri barangkali sudah mempelajari sesuatu. Pada umumnya, hasil tes yang kedua hasilnya lebih baik daripada pertama. Hal ini tidak mengapa, karena pengetes harus sadar akan adanya *practice effect* dan *carry over effect*, yang penting adanya ketetapan hasil yang ditunjukkan oleh koefisien yang tinggi. Contoh :

Nama siswa	Test pertama		Test kedua	
	Skor	Ranking	Skor	Rangking
Sanusi	7	3	8	3
Satria	5	1	6	1
Sidi	8	4	9	4
Dia	6	2	7	2

Metode ini juga disebut *self-correlation method* (korelasi diri sendiri), karena mengkorelasikan hasil dari tes yang sama.

Langkah :

- 1). Mencari skor X dan Y
- 2). Membuat tabel persiapan

No.	Nama	X	Y	$x^2$	$y^2$	xy
1.	Dia	6	7	36	49	42
2.	Sanusi	7	8	49	64	56
3.	Satria	5	6	25	36	30
4.	Sidi	8	9	64	81	72
Jumlah		26	30	174	230	200

3). Memasukkan ke dalam rumus

$$\begin{aligned} &= \frac{4 \cdot 200 - (26)(30)}{\sqrt{(4 \times 174 - 676)(4 \times 230 - 900)}} \\ &= \frac{800 - 780}{\sqrt{(696 - 676)(920 - 900)}} \\ &= \frac{20}{\sqrt{(20 \times 20)}} \\ &= \frac{20}{20} = 1,00 \text{ Reliabilitas sangat tinggi} \end{aligned}$$

Kelemahan :

Banyak terjadi *practice effect* dan *carry over effect*

Kelebihan :

- a). Pengetes pekerjaanya ringan
- b). Waktu yang dibutuhkan sedikit

c. **Metode Belah Dua atau *split-half method = single-test-single trial method***

Metode ini untuk mengatasi metode dua tes dua kali percobaan dan satu tes dua kali percobaan. Dengan metode ini, pengetes hanya menggunakan satu tes dan dicobakan sekali. Dalam metode ini, yang dibelah dan dikelompokkan adalah butir soalnya, bukan siswanya, sehingga syaratnya jumlah butir soal harus genap.

Cara membela butir soal :

- 1). Membela item-item genap dan item-item ganjil (belahan ganjil genap)
- 2). Membela item-item awal dan item-item akhir (belahan awal akhir)
- 3). Langkah membela butir soal :
  - a). Mengadakan analisis butir soal (analisis item) item yang dijumlahkan diberi skor 1 dan item yang salah diberi skor 0.
  - b). Memasukkan ke dalam tabel analisis item dan menentukan variabel x dan y.

- c). Memasukkan ke dalam tabel persiapan perhitungan reliabilitas  
- Belah dua genap ganjil

No.	Nama	X	Y	$x^2$	$y^2$	xy
1.	Dia	2	4	4	16	8
2.	Sanusi	2	5	4	25	10
3.	Satria	0	5	0	25	0
4.	Sidi	5	3	25	9	15
Jumlah		9	17	33	75	33

- Belah dua awal akhir

No.	Nama	X	Y	$x^2$	$y^2$	xy
1.	Dia	3	3	9	9	9
2.	Sanusi	3	4	9	16	12
3.	Satria	2	3	4	9	6
4.	Sidi	4	4	16	16	16
Jumlah		12	14	36	50	43