

Dr. H. Masrukhin, S.Ag., M.Pd.

Pengembangan Sistem Evaluasi Pendidikan

Pengembangan Sistem Evaluasi Pendidikan

Buku Pengembangan Sistem Evaluasi Pendidikan Agama Islam ini antara lain dimaksudkan untuk memberikan informasi-informasi tambahan bagi guru sebagai pengajar, dalam usahanya pengembangan kegiatan proses belajar mengajar atau pembelajaran yang dinamis, kreatif dan inovatif sampai pada sistem evaluasinya, lebih khusus pada materi Pendidikan Agama Islam. Penggunaan sistem evaluasi Pendidikan Agama Islam yang dilakukan secara sistematis baik secara manual maupun berbasis digital, maka pengukuran keberhasilan proses belajar mengajar akan semakin mudah diketahui hasilnya sedini mungkin secara cepat dan tepat. Keterlambatan hasil evaluasi yang dilakukan oleh guru pada setiap siswa yang berbeda-beda kemampuannya, akan mengakibatkan sulitnya guru dalam mencapai keberhasilan pembelajaran secara maksimal pada setiap peserta didik. Hal inilah yang menjadi titik tolak motivasi penulis dalam menulis buku ini.

Penerbit Media Ilmu Press
Jl. Conge Ngembalrejo Bae Kudus
<http://mediailmupress.blogspot.com>
Email: Intannajma@gmail.com

ISBN : 978-602-99125-7-9



9 789786 021409

Pengembangan Sistem Evaluasi Pendidikan

Dr. H. Masrukhin, S.Ag., M.Pd.



MEDIA ILMU PRESS

Dr. H. Masrukhin, S.Ag., M.Pd.

**PENGEMBANGAN
SISTEM EVALUASI
PENDIDIKAN**



MEDIA ILMU PRESS

Judul Buku : Pengembangan Sistem Evaluasi Pendidikan

Penulis : Dr. H. MASRUKHIN, S.Ag., M.Pd.

ISBN : 978-602-99125-7-9

Penerbit : Mibarda Publishing

Pencetak : MEDIA ILMU Press

HAK CIPTA PADA PENULIS

Edisi Pertama

Cetakan Pertama, Agustus 2013

KATA PENGANTAR

Segala puji bagi Allah Swt yang selalu memberi berbagai macam kenikmatan baik berupa lahir maupun batin, melindungi, menunjukkan dan membimbing hamba-Nya yang selalu menjadikan dirinya beriman.

Shalawat dan salam senantiasa menghiasi Rasulullah SAW sebagai utusan yang memiliki berbagai macam kecerdasan dalam memberikan pencerahan kehidupan manusia. Tidak lupa pula doa kepada para sahabat beliau serta umat muslim pada umumnya yang masih setia melestarikan pentingnya pendidikan pada setiap generasi untuk kemaslahatan umat di masa yang akan datang.

Guru merupakan sumber daya manusia yang menjadi salah satu kata kunci keberhasilan pendidikan. Oleh karena itu segala upaya untuk membantu guru dalam meningkatkan kualitas, kompetensi mengajar, menambah wawasan serta memperluas cakrawala di kegiatan proses belajar mengajar sampai pada sistem evaluasinya merupakan suatu hal yang tidak dapat ditawar-tawar lagi, kehadiran buku dengan judul *Pengembangan Sistem Evaluasi Pendidikan* ini kiranya akan dapat memberikan kontribusi positif untuk perbaikan kualitas guru dalam kegiatan pembelajaran yang ada secara terus-menerus dan berkesinambungan untuk mencapai keberhasilan proses pembelajaran kepada peserta didik.

Buku *Pengembangan Sistem Evaluasi Pendidikan* ini antara lain dimaksudkan untuk memberikan informasi-informasi tambahan bagi guru sebagai pengajar, dalam usahanya pengembangan kegiatan proses belajar mengajar atau pembelajaran yang dinamis, kreatif dan inovatif sampai pada sistem evaluasinya, lebih khusus pada materi Pendidikan Agama Islam. Penggunaan sistem evaluasi Pendidikan Agama Islam yang dilakukan secara sistematis baik secara manual maupun berbasis *digital*, maka pengukuran keberhasilan proses belajar mengajar akan semakin mudah diketahui hasilnya sedini mungkin secara cepat dan tepat. Keterlambatan hasil

evaluasi yang dilakukan oleh guru pada setiap siswa yang berbeda-beda kemampuannya, akan mengakibatkan sulitnya guru dalam mencapai keberhasilan pembelajaran secara maksimal pada setiap peserta didik. Hal inilah yang menjadi titik tolak motivasi penulis dalam menulis buku ini.

Kepada para dosen, calon dosen, guru, calon guru dan mahasiswa keguruan akan berkenan membaca buku ini insya Allah akan memperoleh tetesan manfaat. Saran dan kritik yang konstruktif dari pembaca sangat penulis harapkan demi perbaikan penerbitan selanjutnya. Dan khususnya tidak lupa pula penulis sampaikan ucapan terimakasih kepada Penerbit : **Media Ilmu Press** yang berkenan menerbitkan buku ini. Semoga bermanfaat, Amin.

Kudus, Agustus 2013

Penulis

DAFTAR ISI

	Halaman
Halaman	i
Judul.....	iii
Kata Pengantar	v
Daftar Isi.....	v
..	
BAB I PENDAHULUAN.....	1
...	
A. Pengertian Evaluasi	1
B. Tujuan dan Fungsi Evaluasi Pendidikan.....	3
C. Mengapa dalam Pendidikan Memerlukan Evaluasi.....	6
D. Karakteristik Evaluasi Pendidikan	8
E. Ruang Lingkup Evaluasi Pendidikan.....	9
 BAB II EVALUASI HASIL BELAJAR MENGAJAR....	 12
A. Prosedur Evaluasi	12
B. Prinsip-Prinsip Dasar Evaluasi Pendidikan.	19
C. Pemilihan Atau Penyusunan Tes Hasil Belajar.....	21
 BAB III. STATISTIK ELEMENTER	 25
A. Pendahuluan.....	25
B. Distribusi Frekuensi.....	25
C. Ukuran Gejala Pusat.....	31
D. Variabilitas.....	35
E. Teknik Korelasi.....	47
 BAB IV. JENIS DAN SISTEM EVALUASI PENDIDIKAN.....	 51
A. Jenis Evaluasi Pendidikan	51
B. Sistem Evaluasi Pendidikan.....	53
C. Cara Penilaian.....	66
D. Pembelajaran dan Penilaian KTSP.....	67

BAB	V.	ALAT-ALAT EVALUASI PENDIDIKAN.....	72
	A.	Pendahuluan.....	72
	B.	Alat-Alat Evaluasi Pendidikan.....	74
	C.	Penyusunan Soal.....	79
	D.	Pengolahan Nilai.....	86
	E.	Cara Pengolahan Skor.....	89
BAB	VI.	PROSEDUR EVALUASI PENDIDIKAN.....	97
	A.	Prosedur Pelaksanaan Tes Hasil Belajar.....	97
	B.	Prosedur Pelaksanaan Sistem Evaluasi	103
BAB	VII.	UJI VALIDITAS DAN RELIABILITAS SUATU KONSTRUK ATAU KONSEP.....	106
	A.	Pendahuluan.....	106
	B.	Uji Reliabilitas.....	109
	C.	Uji Validitas.....	113
	D.	Degrees of Deficulty dan Discriminating Power.....	123
BAB	VIII	ANALISA HASIL TES HASIL TES BELAJAR.....	133
	A.	Skala Penilaian.....	133
	B.	Mengubah Skor Mentah Skor Standar.....	135
	C.	Distribusi Nilai.....	166
BAB	IX	MENILAI INTEGENSI ATAU BAKAT SEKOLAH.....	173
	A.	Definisi Intelegensi.....	174
	B.	Sejarah Pengukuran Intelegensi.....	177
	C.	Metode Intelegensi.....	181
BAB	X	MENILAI BAKAT-BAKAT KHUSUS.....	186
	A	Pengertian Bakat (Aptitude).....	186
	B	Metode Penilaian.....	187
	C	Analisa Tes Bakat.....	188
BAB	XI	MENILAI SIKAP.....	191
	A	Pengertian Sikap.....	191
	B	Metode Pengukuran Sikap.....	191
	C	Analisa Tes Bakat.....	188

BAB XII	MENILAI KEPERIBADIAN.....	200
A	Pengertian Sikap.....	200
B	Metode Pengukuran Kepribadian.....	202
	Daftar Pustaka.....	221
	Lampiran.....	224



HAKEKAT EVALUASI PENDIDIKAN AGAMA ISLAM (PAI)

A. Pengertian Evaluasi PAI

Ada berbagai macam ragam pengertian evaluasi yang dapat dilihat dari sudut pandang bahasa dan istilah. Kata evaluasi berasal dari bahasa Inggris *evaluation* yang berarti penilaian atau penaksiran (John M. Echols dan Hasan Shadily, 1983:220). Dalam buku *Essentials of Educational Evaluation* karangan Edwind Wand dan Gerald W. Brown dikatakan bahwa : *Evaluation refer to the act or process to determining the value of something* (1975:1). Jadi evaluasi merupakan suatu tindakan atau suatu proses untuk menentukan nilai daripada sesuatu. Sedangkan menurut istilah evaluasi adalah kegiatan yang terencana untuk mengetahui keadaan sesuatu obyek dengan menggunakan instrument dan hasilnya dibandingkan dengan tolok ukur untuk memperoleh kesimpulan.

Anne Anastasi mengartikan evaluasi sebagai “*A systematic achieved by pupils*” (Anne Anastasi, 1978:6). Evaluasi bukan sekedar menilai suatu aktivitas secara spontan dan incidental, melainkan merupakan kegiatan untuk menilai sesuatu secara terencana, sistemik, dan terarah berdasarkan atas tujuan yang jelas. Dengan demikian evaluasi merupakan suatu kegiatan yang dikelola secara sistematis, mulai dari awal perencanaan kegiatan sampai pada akhir kegiatan.

Dalam perkembangannya istilah evaluasi berkembang menjadi beberapa istilah yang sering kita gunakan dalam istilah evaluasi, yaitu *measurement* (pengukuran), *assesment* (penaksiran), dan *test*. Ketiga istilah tersebut sering dianggap sama dengan evaluasi. Padahal sebenarnya ketiga istilah tersebut mempunyai perbedaan satu sama lain.

- (1). *Measurement* (pengukuran) adalah “*Measurement means the act or proses of axestaining the extent or quantity of something*” (Wondt, Edwin and G. W. Brown, 1957:1). Pengukuran diartikan sebagai proses untuk menentukan luas atau kuantitas sesuatu dengan pengertian lain pengukuran. adalah suatu usaha untuk mengetahui keadaan sesuatu seperti adanya yang dapat dikuantitaskan, hal ini dapat diperoleh dengan jalan tes atau cara lain. Hasil suatu pengukuran belum banyak memiliki arti sebelum ditafsirkan dengan jalan membandingkan hasil-pengukuran dengan standar atau patokan yang telah ditentukan sebelumnya. Dalam penilaian pendidikan patokan itu dapat berupa batas minimal kompetensi materi pelajaran yang harus dikuasai, atau rata-rata nilai yang diperoleh oleh kelompok. Sebagai contoh siswa yang memperoleh skor tujuh, dapat berarti memiliki nilai rendah apabila dibandingkan dengan rata-rata kelompok yang mencapai skor delapan, tetapi nilai tersebut dapat berarti tinggi apabila dibandingkan dengan batas lulus yang hanya dibutuhkan angka lima misalnya. Dengan demikian dapat dikatakan bahwa ada perbedaan mendasar antara pengukuran dan penilaian. Pengukuran lebih menekankan pada pertanyaan “How much”, sedangkan penilaian akan memberikan jawaban terhadap pertanyaan “What value”. Namun demikian hubungan antara pengukuran dan penilaian tidak dapat dipisahkan, karena pengukuran dalam bentuk apapun tidak akan memberikan arti apa-apa kalau tidak kita hubungkan dengan penilaian.
- (2). *Test* lebih ditekankan pada penggunaan alat pengukuran. Cronbach memberikan batasan tes sebagai berikut: ... *a systematic procedure for observing a person's behavior and discribing it with the aid of a numerical scale or a category system* (Cronbach, 1970:26). Lebih lanjut Sumadi Suryabrata memberikan pengertian sebagai berikut: "Tes adalah pertanyaan-pertanyaan yang harus dijawab dan atau perintah-perintah yang harus dijalankan, yang

mendasarkan harus bagaimana *testee* menjawab pertanyaan-pertanyaan atau melakukan perintah-perintah itu menyelidik mengambil kesimpulan dengan cara membandingkannya dengan standar atau *testee* yang lain" (Sumadi Suryabrata, 1984:22). Dilihat dari pengertian kedua, tes memiliki arti yang sama dengan evaluasi dan memiliki pengertian yang lebih luas dibandingkan dengan pengukuran.

- (3). *Assessment* tidak sampai ke taraf evaluasi, melainkan sekedar mengukur dan mengadakan estimasi terhadap hasil pengukuran. Dalam melakukan estimasi tersebut seringkali dilakukan dalam kegiatan penelitian. Dengan demikian penilaian memiliki pengertian yang hampir sama dengan kegiatan *research*. Keduanya sama-sama merupakan kegiatan untuk memperoleh gambaran tentang keadaan suatu obyek melalui proses penelaahan secara logik dan sistematis, membutuhkan data empirik untuk membuat kesimpulan, dan menuntut syarat keahlian tertentu bagi pelakunya. Perbedaannya, penelitian hampir selalu dimulai dari kesadaran tentang adanya problem, bertujuan untuk mengembangkan prinsip-prinsip baru melalui proses generalisasi, dan dengan mengadakan analisis hubungan antar variabel, tetapi dalam penilaian perhatian utamanya tidak dimulai dari adanya kesadaran terhadap problema kependidikan, melainkan karena adanya proses pendidikan. Analisis yang dikembangkan tidak sekedar mencari hubungan antar variabel, melainkan mencari koherensi antara tujuan, proses, dan pencapaian tujuan pada setiap program pendidikan. Penilaian juga tidak berkepentingan terhadap generalisasi, namun memperhatikan aspek prediktif dari hasil evaluasi. Penelitian memiliki pengertian yang lebih luas daripada evaluasi. Oleh karena itu evaluasi pendidikan dapat berfungsi sebagai bagian dari penelitian yang sering disebut dengan *action research* atau *participation action research* yaitu suatu proses penelitian yang hasil-hasilnya selalu dipakai untuk memperbaiki

pelaksanaan proses, kemudian diadakan penelitian ulang, yang hasilnya dipakai menyempurnakan lagi kegiatan tersebut, begitu seterusnya. Contoh yaitu Evaluasi Program Bantuan Operasional Manajemen Mutu (BOMM) di Sekolah SD Se-Kabupaten Kudus.

Berdasarkan berbagai macam pengertian diatas, maka evaluasi pendidikan agama Islam adalah suatu kegiatan pendidikan agama Islam yang dikelola secara sistematis, mulai dari awal perencanaan kegiatan sampai pada akhir kegiatan untuk mengetahui keadaan sesuatu obyek atau peserta didik dengan menggunakan instrument dan hasilnya dibandingkan dengan tolok ukur keberhasilan pembelajaran sehingga memperoleh kesimpulan.

B. Tujuan dan Fungsi Evaluasi PAI

Pada dasarnya evaluasi PAI bertujuan dan berfungsi untuk mengetahui keberhasilan pembelajaran PAI secara keseluruhan, baik pada aspek kognitif, afeksi dan psikomotor pada setiap peserta didik yang telah dilakukan dengan menggunakan serangkaian proses pembelajaran yang berdasarkan silabus yang telah ditetapkan dan diharapkan dapat berfungsi bagi peserta didik, orang tua, guru, sekolah dan masyarakat untuk perbaikan mutu pembelajaran secara terus menerus dan berkesinambungan.

Tujuan dilakukannya evaluasi menurut Muchtar Buchori ada dua yaitu: (1) untuk mengetahui kemajuan belajar peserta didik setelah ia menyadari pendidikan selama jangka waktu tertentu, dan (2) untuk mengetahui tingkat efisien metode-metode pendidikan yang dipergunakan pendidikan selama jangka waktu tertentu tadi (Muchtar Buchori, 1980:6).

Adapun tujuan evaluasi secara umum dapat dikaitkan dengan fungsi evaluasi dalam pendidikan. Julian C. Stanley dan Kenneth mengklasifikasikan tujuan evaluasi dalam

pendidikan. dalam tiga fungsi yang saling terkait satu dengan yang lainnya, yaitu: (1) *instructional*, (2) *administrative*, dan (3) *guidance* (Stanley dan Kenneth, 1978:7).

Thorndike dan Hagen, merinci tujuan evaluasi pendidikan dalam delapan bidang yang tidak jauh berbeda dengan pembagian fungsi menurut Stanley, kedelapan bidang tersebut adalah:

- a. bidang pengajaran,
- b. hasil belajar,
- c. diagnosis dan usaha perbaikan,
- d. fungsi penempatan,
- e. fungsi seleksi,
- f. bimbingan dan penyuluhan,
- g. kurikulum dan
- h. penilaian kelembagaan (Thorndike R.L. dan Hagen E.P, 1977:1-24).

Evaluasi dalam bidang pendidikan dan pengajaran secara keseluruhan mempunyai beberapa fungsi sebagai berikut :

1. Untuk mengetahui taraf kesiapan daripada anak-anak untuk menempuh suatu pendidikan tertentu;
2. Untuk mengetahui seberapa jauh hasil yang telah dicapai dalam proses pendidikan yang dilaksanakan.
3. Untuk mengetahui apakah suatu mata pelajaran yang kita ajarkan dapat kita lanjutkan pada materi baru atau mengulangi materi yang telah diajarkan.
4. Untuk mendapatkan bahan-bahan informasi dalam memberikan bimbingan tentang jenis pendidikan pada peserta didik.

5. Untuk mendapatkan bahan-bahan informasi apakah anak dapat dinaikkan ke kelas yang lebih tinggi atau mengulang di kelas semula.
6. Untuk membandingkan apakah prestasi yang dicapai oleh anak-anak sudah sesuai dengan kapasitasnya atau belum.
7. Untuk menafsirkan apakah seorang anak telah cukup matang untuk dilepas ke masyarakat atau masih perlu dididik dan dilatih.
8. Untuk mengadakan seleksi.
9. Untuk mengetahui efisiensi dan efektifitas metode yang telah digunakan dalam pembelajaran.
10. Untuk mengetahui tercapai tidaknya tujuan pengajaran, dalam hal ini adalah tujuan instruksional khusus.

Adapun fungsi evaluasi pendidikan bila dilihat dari kepentingan masing-masing pihak, dapat dirinci sebagai berikut

Fungsi evaluasi pendidikan *bagi guru*, adalah untuk:

1. mengetahui kemajuan belajar peserta didik;
2. mengetahui kedudukan masing-masing individu peserta didik dalam kelompoknya;
3. mengetahui kelemahan-kelemahan dalam cara belajar-mengajar dalam PBM;
4. memperbaiki proses belajar-mengajar, dan
5. menentukan kelulusan peserta didik.

Bagi *peserta didik*, evaluasi pendidikan berfungsi:

1. mengetahui kemampuan dan hasil belajar,
2. memperbaiki cara belajar, dan
3. menumbuhkan motivasi dalam belajar.

Bagi *sekolah*, evaluasi pendidikan berfungsi:

1. mengukur mutu hasil pendidikan;
2. mengetahui kemajuan dan kemunduran sekolah;
3. membuat keputusan kepada peserta didik, dan
4. mengadakan perbaikan kurikulum.

Bagi orang tua peserta didik, fungsi evaluasi pendidikan adalah untuk:

1. mengetahui hasil belajar anaknya;
2. meningkatkan pengawasan dan bimbingan serta bantuan kepada anaknya dalam usaha belajar, dan
3. mengarahkan pemilihan jurusan, atau jenis sekolah pendidikan lanjutan bagi anaknya.

Adapun fungsi evaluasi pendidikan bagi *masyarakat dan pemakai* jasa pendidikan, adalah untuk:

1. mengetahui kemajuan sekolah;
2. ikut mengadakan kritik dan saran perbaikan bagi kurikulum pendidikan pada sekolah tersebut, dan
3. lebih meningkatkan partisipasi masyarakat dalam usahanya membantu lembaga pendidikan.

C. Karakteristik Evaluasi Pendidikan Agama Islam

Untuk dapat melaksanakan kegiatan evaluasi pendidikan agama Islam dengan baik, maka seorang evaluator harus memahami karakteristik pokok-pokok evaluasi pendidikan, sebagaimana dijelaskan oleh H.G. Shane (1958: 55-57) maupun Suharsimi Arikunto (1981:9), sebagai berikut:

1. Penilaian dalam pendidikan itu dilakukan secara *tidak langsung*. Obyek pengukuran dan penilaian dalam pendidikan adalah peserta didik, tidak dilihat dari sosok fisiknya, seperti berat dan tinggi badannya, melainkan aspek psikologiknya,

seperti sikap, minat, bakat, inteligensia dan hasil belajar. Aspek-aspek tersebut tidak dapat diukur secara langsung. Sebagai contoh untuk mengukur kepandaian peserta didik yang dapat dilakukan hanyalah mengukur hasil belajar dengan jalan menjawab atau mengerjakan soal-soal tes. Jawaban terhadap soal tes tersebut yang dipakai untuk menggambarkan kepandaian peserta didik.

2. Penggunaan ukuran kuantitatif, karena penilaian selalu dimulai dari pengukuran, maka hasil pengukuran akan menggunakan satuan-satuan secara kuantitatif. Penggunaan satuan *kuantitatif* ini untuk mendapatkan hasil pengukuran yang obyektif, dan pasti setelah itu dapat diolah dan ditafsirkan ke dalam satuan *kualitatif*.
3. Penilaian pendidikan itu menggunakan unit satuan yang tetap. Obyek pengukuran hendaknya menggunakan satuan yang tetap. Sebab apabila penggunaan satuan pengukuran tidak tetap, akan berakibat hasil evaluasi tidak memiliki nilai kejelasan, prediksinya menjadi rendah.
4. Penilaian pendidikan bersifat relatif, artinya hasil penilaian itu kendatipun sudah menggunakan satuan yang tetap, hasilnya tidaklah selalu sama dari waktu ke waktu. Sebab hasil penilaian tidak semata-mata ditentukan oleh alat ukur yang valid, namun juga dipengaruhi oleh keadaan obyek yang selalu berkembang, serta keadaan lingkungan yang mempengaruhi pelaksanaan pendidikan tersebut, apalagi dalam evaluasi pendidikan tidak dapat dilaksanakan secara langsung sebagaimana dijelaskan di atas. Kendatipun demikian, relatifitas hasil-hasil penilaian itu harus tetap dalam batas-batas obyektivitas.
5. Penilaian pendidikan tidak mungkin terhindar dari kesalahan. Kesalahan tersebut dapat diakibatkan alat ukur yang kurang valid, atau sikap subyektif penilai, maupun kesalahan dalam penghitungan, keadaan fisik dan psikis siswa yang dinilai, serta situasi tempat pelaksanaan penilaian itu dilakukan.

BAB II

RUANG LINGKUP EVALUASI PENDIDIKAN AGAMA ISLAM (PAI)

A. Ruang Lingkup Evaluasi Pendidikan Agama Islam

Evaluasi pendidikan yang akan penulis bicarakan dalam buku ini adalah evaluasi dalam pendidikan dan pengajaran Agama Islam. Ruang lingkup evaluasi pendidikan dan pengajaran agama Islam meliputi : hasil belajar, intelegensi, bakat khusus, minat, hubungan sosial, sikap dan kepribadian. Namun dalam perkembangannya ruang lingkup Evaluasi Pendidikan Agama Islam dapat dilihat dari : (1) Domain Hasil Belajar, (2) Sistem Pembelajaran, (3) Penilaian Proses dan Hasil Belajar, (4) Penilaian Berbasis Kelas.

(1) Domain Hasil Belajar

Menurut Benyamin S. Bloom, dkk (1956) hasil belajar dapat dikelompokkan menjadi tiga domain yaitu kognitif, afektif, dan psikomotor. Pada setiap domain disusun menjadi beberapa jenjang kemampuan, mulai dari hal yang sederhana sampai dengan hal yang kompleks, mulai dari yang mudah sampai yang sukar, dan mulai dari hal konkrit sampai dengan hal yang abstrak.

Adapun rincian domain Kognitif (*cognitive domain*) yaitu :

- a. Pengetahuan (*knowledge*) yaitu jenjang kemampuan yang menuntut peserta didik untuk dapat mengenali atau mengetahui adanya konsep, prinsip, fakta atau istilah tanpa harus mengerti atau dapat menggunakannya. Kata operasional yang dipakai yaitu: *mendefinisikan, memberikan, mengidentifikasi, memberi nama, menyusun daftar, mencocokkan, menyebutkan, membuat garis besar, menyatakan kembali, memilih, dan menyatakan.*

- b. Pemahaman (*Comprehension*) yaitu jenjang kemampuan yang menuntut peserta didik untuk memahami atau mengerti tentang materi pelajaran yang disampaikan guru dan dapat memanfaatkannya tanpa harus menghubungkannya dengan hal-hal lain. Kemampuan ini dijabarkan menjadi tiga, yaitu: *menterjemahkan, menafsirkan, dan mengekstrapolasi*. Kata kerja operasional yang dapat digunakan adalah *mengubah, mempertahankan, membedakan, memprakirakan, menjelaskan, menyatakan secara luas, menyimpulkan, memberi contoh, melukiskan kata-kata sendiri, meramalkan, menuliskan kembali, dan meningkatkan*.
- c. Penerapan (*Application*) yaitu jenjang kemampuan yang menuntut peserta didik untuk menggunakan ide-ide umum, tata cara ataupun metode, prinsip teori-teori dalam situasi baru dan konkrit. Kata kerja operasional yang digunakan adalah : *mengubah, menghitung, mendemonstrasikan, mengungkapkan, mengerjakan dengan teliti, menjalankan, memanipulasikan, menghubungkan, menunjukkan, memecahkan, dan menggunakan*.
- d. Analisis (*Analysis*) yaitu jenjang kemampuan yang menuntut peserta didik untuk menguraikan suatu situasi atau keadaan tertentu ke dalam unsur-unsur atau komponen-komponen pembentuknya. Kemampuan analisis dapat dikelompokkan menjadi tiga yaitu : analisis unsur, analisis hubungan, dan analisis prinsip-prinsip yang terorganisasi. Kata kerja operasional yaitu : *mengurai, membuat diagram, memisahkan, menggambarkan kesimpulan, membuat garis besar, menghubungkan dan merinci*.
- e. Sintesis (*Synthesis*) yaitu jenjang kemampuan yang menuntut peserta didik untuk menghasilkan sesuatu yang baru dengan cara menggabungkan berbagai faktor. Hasil yang diperoleh dapat berupa tulisan, rencana atau mekanisme. Kata kerja operasional yaitu: *menggolongkan, menggabungkan, memodifikasi, menghimpun, menciptakan, merencanakan,*

merekonstruksikan, menyusun, membangkitkan, mengorganisir, merevisi, menyimpulkan dan menceritakan.

- f. Evaluasi (*Evaluation*) yaitu jenjang kemampuan yang menuntut peserta didik untuk dapat mengevaluasi suatu situasi, keadaan, pernyataan, atau konsep berdasarkan kriteria tertentu. Kata kerja operasional yang dapat dipakai adalah : *menilai, membandingkan, mempertentangkan, mengkritik, membeda-bedakan, mempertimbangkan kebenaran, menyokong, menafsirkan, menduga.*

Adapun rincian domain afektif (*affective domain*) yaitu : internalisasi sikap yang menunjuk ke arah pertumbuhan batiniah dan terjadi bila peserta didik menjadi sadar tentang nilai yang diterima, kemudian mengambil sikap sehingga menjadi bagian dari dirinya dalam membentuk nilai dan menentukan tingkah laku. Domain afektif terdiri atas beberapa jenjang kemampuan, yaitu:

- a. Kemauan menerima (*receiving*), yaitu jenjang kemampuan yang menuntut peserta didik untuk peka terhadap eksistensi fenomena atau rangsangan tertentu. Kepekaan ini diawali dengan penyadaran kemampuan untuk menerima dan memperhatikan. Kata kerja operasional yang dapat digunakan yaitu: *menanyakan, memilih, menggambarkan, mengikuti, memberikan, berpegang teguh, menjawab, menggunakan.*
- b. Kemampuan menanggapi/menjawab (*responding*), yaitu jenjang kemampuan yang menuntut peserta didik untuk tidak hanya peka pada suatu fenomena tetapi juga bersaksi terhadap salah satu cara. Penekannya pada kemauan peserta didik untuk menjawab secara sukarela, membaca tanpa ditugaskan. Kata kerja operasional yang dapat digunakan diantaranya : *menjawab, membantu, memperbincangkan, memberi nama, menunjukkan, mempraktikkan, mengemukakan, membaca, melaporkan, menuliskan, memberi-tahu, dan mendiskusikan.*

- c. Menilai (*valuing*), yaitu jenjang kemampuan yang menuntut peserta didik untuk menilai objek, fenomena atau tingkah laku tertentu secara konsisten. Kata kerja operasional yang digunakan diantaranya: *melengkapi, menerangkan, membentuk, mengusulkan, mengambil bagian, memilih dan mengikuti.*
- d. Organisasi (*organization*), yaitu jenjang kemampuan yang menuntut peserta didik untuk menyatukan nilai-nilai yang berbeda, memecahkan masalah, membentuk suatu sistem nilai. Kata kerja operasional : *mengubah, mengatur, menggabungkan, membandingkan, mempertahankan, menggeneralisasikan, dan memodifikasi.*
- e. Karakteristik kehidupan (*Characterization*), yaitu menggunakan nilai-nilai sebagai pandangan hidup (*worldview*), mempertahankan nilai-nilai yang sudah diyakini.

Domain psikomotor (*Psychomotor Domain*), yaitu kemampuan peserta didik yang berkaitan dengan gerakan tubuh atau bagian-bagiannya, mulai dari gerakan yang sederhana sampai dengan gerakan yang kompleks. Perubahan pola gerakan memakan waktu sekurang-kurangnya 30 menit. Kata kerja operasional yang dikerjakan adalah :

- a. *Muscular or motor skill*, yang meliputi : *mempertontonkan gerak, menunjukkan hasil, melompat, menggerakkan, dan menampilkan*
- b. *Manipulations of materials or objects*, yang meliputi : *mereparasi, menyusun, membersihkan, menggeser, memindahkan, dan membentuk.*
- c. *Neuromuscular coordination*, yang meliputi : *mengamati, menerapkan, menghubungkan, menggandeng, memadukan, memasang memotong, menarik, dan menggunakan.*

Berdasarkan taksonomi Bloom di atas, maka kemampuan peserta didik dapat diklasifikasikan menjadi dua yaitu tingkat tinggi dan tingkat rendah. Kemampuan tingkat rendah terdiri atas pengetahuan, pemahaman, dan aplikasi. Sedangkan kemampuan

tingkat tinggi meliputi analisis, sintesis, evaluasi dan kreatifitas. Kegiatan peserta didik dalam menghafal merupakan kemampuan tingkat rendah, sedangkan kemampuan berfikir kritis dan kreatif merupakan kemampuan tingkat tinggi. Berfikir kreatif adalah kemampuan melakukan generalisasi dengan menggabungkan, merubah atau mengulang kembali keberadaan ide-ide tersebut. Sedangkan kemampuan berfikir kritis merupakan kemampuan memberikan rasionalisasi sesuatu dan mampu memberikan penilaian terhadap sesuatu tersebut. Rendahnya kemampuan peserta didik dalam berfikir sebagai akibat dari kebiasaan guru yang memberikan ruang, waktu dan gerak dalam mengajar lebih terfokus pada kemampuan menghafal dan sistem evaluasi yang hanya menggunakan *paper and pencil test*.

(2) Sistem Pembelajaran

Ruang lingkup evaluasi pembelajaran hendaknya bertitik tolak dari tujuan evaluasi pembelajaran itu sendiri. Hal ini disebabkan tujuan evaluasi pembelajaran itu sendiri adalah untuk mengetahui sedini mungkin efektifitas dan efisiensi sistem pembelajaran yang telah dilaksanakan berkaitan dengan tujuan, materi, metode, media, sumber belajar, lingkungan guru dan peserta didik, serta sistem evaluasi.

Dengan demikian ruang lingkup evaluasi pembelajaran adalah:

- a. Program pembelajaran, yang berkaitan dengan :
 - 1) Tujuan pembelajaran umum atau kompetensi dasar, yaitu target yang harus dikuasai peserta didik dalam setiap pokok bahasan/topik. Kriteria yang digunakan untuk mengevaluasi tujuan pembelajaran umum atau kompetensi dasar ini adalah keterkaitannya dengan tujuan kurikuler atau standar kompetensi dari setiap bidang studi dan tujuan kelembagaan, kejelasan rumusan kompetensi dasar, kesesuaiannya dengan tingkat perkembangan peserta didik, pengembangannya dalam bentuk hasil belajar dan indikator, penggunaan kata kerja operasional

dalam indikator, dan unsur-unsur penting dalam kompetensi dasar, hasil belajar dan indikator.

- 2) Isi/materi pembelajaran, yaitu isi kurikulum yang berupa topik/pokok bahasan dan sub topik/sub pokok bahasan beserta rinciannya dalam setiap bidang studi atau mata pelajaran. Isi kurikulum meliputi 3 unsur yaitu logika (pengetahuan benar salah, berdasarkan prosedur keilmuan), etika (baik buruk), dan estetika (keindahan). Materi pembelajaran dapat dikelompokkan menjadi enam jenis yaitu fakta, konsep/teori, prinsip, proses, nilai dan ketrampilan. Kriteria yang digunakan antara lain yaitu : kesesuaiannya dengan kompetensi dasar dan hasil belajar, ruang lingkup materi, urutan logis materi, kesesuaiannya dengan tingkat perkembangan peserta didik, dll.
- 3) Metode pembelajaran, yaitu cara guru menyampaikan materi pelajaran, seperti metode ceramah, tanya jawab, diskusi, pemecahan masalah, dan sebagainya. Kriteria yang digunakan, antara lain : kesesuaiannya dengan kompetensi dasar dan hasil belajar, kesesuaiannya dengan kondisi kelas/sekolah, kesesuaiannya dengan tingkat perkembangan peserta didik, kemampuan guru dalam menggunakan metode, waktu yang tersedia, dan sebagainya.
- 4) Media pembelajaran, yaitu alat-alat yang membantu untuk mempermudah guru dalam menyampaikan isi/meteri pelajaran. Media dapat dibagi tiga kelompok, yaitu media audio, media visual, dan media audio-visual. Kriteria yang digunakan seperti pada komponen metode.
- 5) Sumber belajar, yang meliputi : pesan, orang, bahan, alat, teknik, dan latar. Sumber belajar dapat dibedakan menjadi dua jenis, yaitu sumber belajar yang dirancang (*resources by design*) dan sumber belajar yang digunakan (*resources by utilization*). Kriteria yang digunakan sama seperti komponen metode.

- 6) Lingkungan, yaitu lingkungan masyarakat, sekolah dan keluarga. Kriteria yang digunakan adalah hubungan peserta didik dengan masyarakat, teman sekelas/sekolah maupun di luar sekolah, guru orang tua; kondisi keluarga dan lain sebagainya.
 - 7) Penilaian proses dan hasil belajar, baik menggunakan tes maupun non tes. Kreterian yang digunakan antara lain: kesesuaiannya dengan kompetensi dasar, hasil belajar, dan indikator; kesesuaiannya dengan tujuan dan fungsi penilaian, unsur-unsur penting dalam penilaian, aspek-aspek yang dinilai, kesesuaiannya dengan tingkat perkembangan peserta didik, jenis dan alat penilaian.
- b. Proses pelaksanaan pembelajaran
- 1) Kegiatan, yang meliputi : jenis kegiatan, prosedur pelaksanaan setiap jenis kegiatan, sarana pendukung, efektifitas dan efisiensi, dll.
 - 2) Guru, terutama dalam hal : menyampaikan materi, kesulitan-kesulitan guru, menciptakan suasana pembelajaran yang kondusif, menyiapkan alat-alat dan perlengkapan yang diperlukan, membimbing peserta didik, menggunakan teknik penilaian, menerapkan disiplin kelas, dan sebagainya.
 - 3) Peserta didik, terutama dalam hal: peranserta didik dalam kegiatan belajar dan bimbingan, memahami jenis kegiatan, mengerjakan tugas-tugas, perhatian, keaktifan, motivasi, sikap, minat, umpan balik, kesempatan melaksanakan praktik dalam situasi yang nyata, kesulitan belajar, waktu belajar, istirahat dan sebagainya.
- c. Hasil pembelajaran, baik untuk jangka pendek (sesuai dengan pencapaian indikator), jangka menengah (sesuai dengan target untuk setiap bidang studi/mata pelajaran), dan jangka panjang (setelah peserta didik terjun ke masyarakat).

(3) Penilaian Proses dan Hasil Belajar

Dalam kurikulum berbasis kompetensi 2004 terdapat empat komponen pokok, yaitu kurikulum dan hasil belajar, penilaian berbasis kelas, kegiatan belajar-mengajar, dan pengelolaan kurikulum berbasis sekolah. Dalam komponen kurikulum dan hasil belajar, setiap mata pelajaran terdapat tiga komponen penting, yaitu kompetensi dasar, hasil belajar, dan indikator pencapaian hasil belajar berkaitan dengan sikap, pengetahuan dan pemahaman peserta didik terhadap bahan pelajaran, kecerdasan peserta didik, perkembangan jasmani/kesehatan, dan ketrampilan. Dengan demikian, hasil belajar merupakan gambaran tentang apa yang harus digali, dipahami, dan dikerjakan peserta didik.

Ruang lingkup evaluasi pembelajaran dan penilaian hasil belajar di atas merupakan aspek-aspek minimal yang harus dievaluasi oleh guru dalam pembelajaran. Aspek-aspek tersebut masih bersifat umum dan global, karena itu perlu dirinci lagi sampai pada tingkat operasional dan spesifik, sehingga aspek-aspek itu betul-betul dapat diukur (*measurable*) dan dapat diamati (*observable*). Untuk mengukur aspek-aspek tersebut, guru harus membuat instrumen evaluasi atau penilaian secara bervariasi, baik tes maupun non tes.

(4) Penilaian Berbasis Kelas.

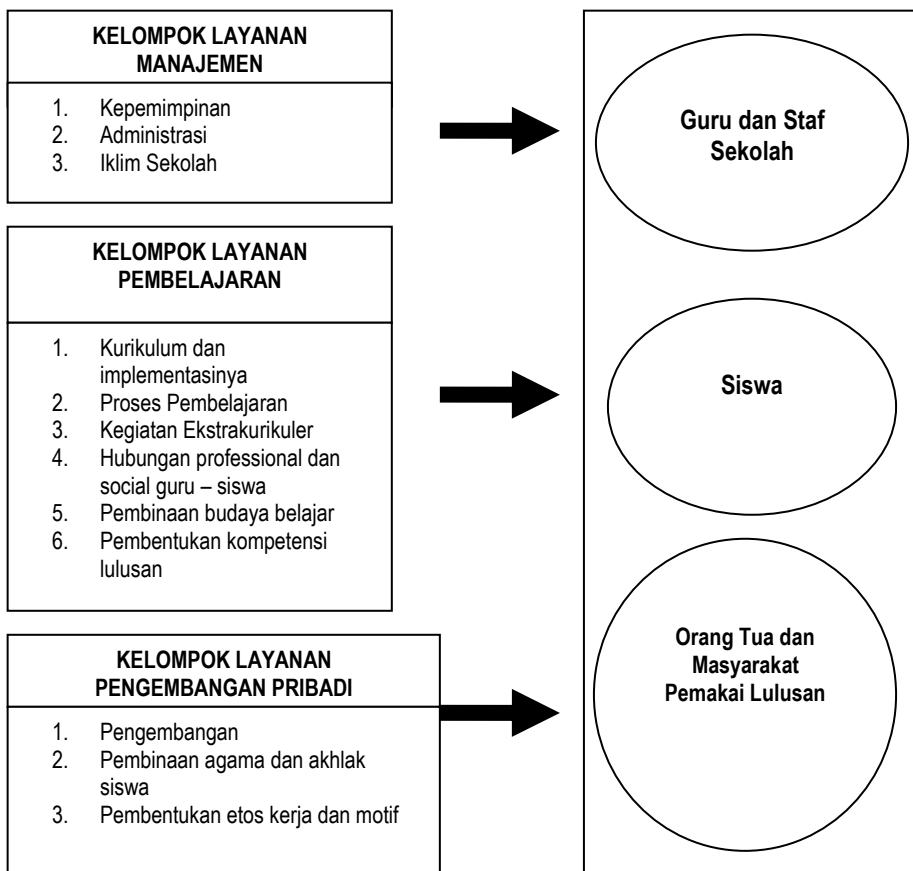
Kompetensi dasar pada hakekatnya adalah pengetahuan, keterampilan, sikap, dan nilai-nilai yang direfleksikan dalam kebiasaan berfikir dan bertindak setelah peserta didik menyelesaikan mata pelajaran tertentu. Sesuai dengan petunjuk pengembangan Kurikulum Berbasis Kompetensi yang dikeluarkan oleh Departemen Pendidikan Nasional (2004), maka ruang lingkup penilaian berbasis kelas adalah sebagai berikut : (1) kompetensi dasar mata pelajaran. Kompetensi dasar pada hakekatnya adalah pengetahuan, keterampilan, sikap, dan nilai-nilai yang direfleksikan dalam kebiasaan berfikir dan bertindak setelah peserta didik menyelesaikan mata pelajaran tertentu. Bertitik

tolak dari materi pelajaran itu, maka dikembangkan alat penilaian. (2) Kompetensi rumpun pelajaran merupakan kumpulan dari mata pelajaran atau disiplin ilmu yang lebih spesifik, (3) Kompetensi lintas kurikulum, merupakan kompetensi yang harus dicapai melalui seluruh rumpun pelajaran dalam kurikulum, (4) Kompetensi Tamatan merupakan pengetahuan, keterampilan, sikap dan nilai-nilai yang direfleksikan dalam kebiasaan berfikir dan bertindak setelah peserta didik menyelesaikan jenjang pendidikan tertentu aspek kognitif, afektif, dan psikomotorik, (5) Pencapaian Keterampilan hidup merupakan penguasaan berbagai kompetensi dasar, kompetensi lintas kurikulum, kompetensi rumpun pelajaran dan kompetensi tamatan melalui berbagai pengalaman belajar dapat memberikan efek positif (*nurturan effects*) dalam bentuk kecakapan hidup (*life skills*). Jenis keterampilan hidup yang perlu dinilai adalah keterampilan diri (keterampilan personal), keterampilan berpikir rasional, keterampilan sosial, keterampilan akademik, dan keterampilan vokasional.

Namun dalam pelaksanaan kegiatan evaluasi ruang-lingkup evaluasi pendidikan, sebenarnya dapat dilihat dari ruang-lingkup proses pendidikan sebagai suatu sistem. Sebagaimana dijelaskan di depan, evaluasi merupakan bagian dari proses pendidikan secara menyeluruh, evaluasi pendidikan bukan hanya sekedar kumpulan teknik-teknik yang diperlukan oleh guru dalam mengukur hasil belajar siswa, melainkan merupakan suatu proses kontinu yang mendasari seluruh proses pendidikan dan pengajaran yang baik. Seperti dijelaskan oleh Anne Anastasi: "*Evaluation includes a number of techniques that are indispensable to the teacher ... however, evaluations is not merely a collection of techniques evaluation is a process, it is a continuous process which underlies all good teaching and learning* (Anne Anastasi, 1978:3).

Apabila kita cermati pada gambar 1 dibawah ini, pendidikan dilihat secara sempit, paling tidak ada beberapa komponen pendidikan yang dapat disederhanakan menjadi tiga, yaitu

pendidik, peserta didik dan bahan pengajaran, jika dilihat dari segi prosesnya terdiri dari tiga aktivitas yaitu *tujuan pendidikan*, *proses interaksional* dan *prosedur evaluasi*. Lebih lanjut jika paradigma proses program pendidikan berbasis layanan. Jenis-jenis layanan itu secara keseluruhan diharapkan oleh semua pihak yang berkepentingan dengan sekolah itu; selanjutnya diidentifikasi kondisi dan jenis-jenis layanan yang sifat-sifatnya diharapkan oleh mereka.



Gambar 1. : Hubungan antara Jenis Layanan dan Pihak-Pihak yang Berkepentingan (Depdiknas 2003 :3)

PARADIGMA EVALUASI PENDIDIKAN AGAMA ISLAM (PAI)

A. Pengertian Paradigma dalam Evaluasi PAI

Dalam kamus bahasa Inggris *The Oxford English Dictionary* paradigma didefinisikan sebagai *the basic meaning of the term paradigm as "a pattern or model, an exemplar"*. Paradigma yaitu pola atau contoh, exemplar. Dalam sejarah ilmu pengetahuan sebagaimana Thomas Kuhn seorang ilmuwan kontemporer dalam bukunya *The Structure of Scientific Revolutions (1962)*. Thomas Kuhn mendefinisikan paradigma ilmiah dengan berbagai macam makna yang berbeda-beda, bahkan sampai 21 makna, yaitu Kuhn defines a scientific paradigm as: "universally recognized scientific achievements that, for a time, provide model problems and solutions for a community of researchers", i.e.,

Paradigma ilmiah sebagai sesuatu yang bersifat universal, diakuinya kemajuan ilmu pengetahuan berlangsung melalui **akumulasi dari temuan dan pembuktian baru**, yang berguna untuk sementara waktu, tersediannya model-model problem-problem dan solusi bagi para komunitas peneliti. Jika terjadi temuan dan pembuktian baru dari hasil-hasil penelitian yang terbaru secara ilmiah, maka temuan dan pembuktian baru dapat digunakan sebagai pengganti temuan terdahulu. Namun perlu diketahui bahwa perubahan mendasar yang sesungguhnya, terjadi karena hasil revolusi ilmiah. **Paradigma** adalah model utama, pola atau metode (untuk meraih beberapa jenis tujuan). Seringkali paradigma merupakan sifat yang paling khas atau dasar dari sebuah teori atau cabang ilmu.

Perspektif evaluasi paradigma merupakan disiplin intelektual adalah cara pandang seseorang terhadap diri dan

lingkungannya yang akan mempengaruhinya dalam berpikir (kognitif), bersikap (afektif), dan bertindak laku (konatif) (Vardiansyah, 2008:27). Paradigma juga dapat berarti seperangkat asumsi, konsep, nilai, dan praktik yang di terapkan dalam memandang realitas dalam sebuah komunitas yang sama, khususnya, dalam disiplin intelektual. Kata paradigma sendiri berasal dari abad pertengahan di Inggris yang merupakan kata serapan dari bahasa Latin ditahun 1483 yaitu paradigma yang berarti suatu model atau pola; bahasa Yunani paradeigma (para+deiknunai) yang berarti untuk "membandingkan", "bersebelahan" (para) dan memperlihatkan (deik).

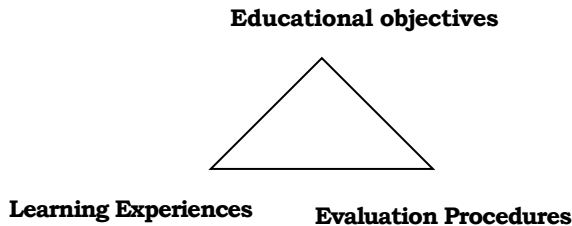
Paradigma evaluasi pendidikan agama Islam merupakan suatu pola atau model yang digunakan untuk membandingkan keberhasilan program pengajaran agama Islam yang dapat dipandang sebagai sesuatu usaha mengubah tingkah laku siswa dengan menggunakan bahan pengajaran agama. Tingkah laku yang diharapkan itu terjadi setelah siswa mempelajari agama dan dinamakan hasil belajar siswa dalam bidang pengajaran agama. Nilai-nilai pengajaran agama Islam itu bermuara pada nilai hakiki atau nilai esensial, yang berbentuk:

- a. *Nilai pembersihan atau pensucian rohani/jiwa*, yang memungkinkan seseorang siap untuk menerima, memahami dan menghayati ajaran agama Islam sebagai pandangan hidupnya.
- b. *Nilai kesempurnaan akhlak*, yang memungkinkan seseorang memiliki akhlakul karimah, yang tercermin pada sifat-sifat Nabi Muhammad saw dan mengamalkan ajaran agama Islam secara sempurna sepanjang hayatnya.
- c. *Nilai peningkatan taqwa kepada Allah swt*, sehingga diri seseorang menjadi semakin akrab kepada-Nya dan dengan penuh gairah serta ketulusan hati menyongsong kehidupan yang hakiki (Zakiah Daradjat, 1995:196).

B. Mengapa dalam Pendidikan Agama Islam Memerlukan Evaluasi

Pendidikan merupakan suatu pelaksanaan program disertai proses kegiatan yang terencana, terkelola, terkoordinatif, terkontrol, dan terevaluasi. Ada tiga alasan utama mengapa dalam kegiatan pendidikan selalu memerlukan evaluasi.

Pertama, apabila dilihat dari *pendekatan proses*, kegiatan pendidikan secara sederhana dapat digambarkan dalam segitiga sebagaimana dikemukakan oleh David McKay yang menggambarkan interaksi ketiga proses tersebut sebagai berikut:



(Julian Stanley dan Kenneth D. Hopkins, 1978:6)

Pada gambar diatas terlihat bahwa interdependensi ketiga komponen tersebut merupakan satu kesatuan. Tujuan pendidikan akan mengarahkan bagaimana pelaksanaan proses belajar-mengajar yang seharusnya dilaksanakan, sekaligus merupakan kerangka acuan untuk melaksanakan kegiatan evaluasi hasil belajar. Pelaksanaan proses belajar-mengajar juga berkepentingan akan adanya perumusan tujuan yang baik, dan prosedur evaluasi haruslah memperhatikan pelaksanaan proses belajar-mengajar. Evaluasi memiliki dua kepentingan, yakni untuk mengetahui apakah tujuan pendidikan sudah tercapai dengan baik, dan kedua untuk

memperbaiki serta mengarahkan pelaksanaan proses belajar-mengajar.

Kedua, kegiatan mengevaluasi terhadap hasil belajar merupakan salah-satu ciri dari pendidik profesional/Satu pekerjaan dipandang memerlukan kemampuan profesional bila pekerjaan tersebut memerlukan pendidikan lanjut (*Advanced Education*) dan latihan khusus (*Special Training*). Pekerjaan pendidik profesional meliputi: menyusun rencana belajar-mengajar, mengorganisasikan, menata mengendalikan, membimbing dan membina terlaksananya proses belajar-mengajar secara relevan, efisien, dan efektif, menilai program dan hasil belajar, dan mendiagnosis faktor-faktor yang mempengaruhi tingkat keberhasilan proses belajar bagi dapat-disempumakannya proses belajar-mengajar selanjutnya (Soediyarto, 1991:7).

Ketiga, bila dilihat dari pendekatan kelembagaan, kegiatan pendidikan adalah merupakan kegiatan manajemen, yang meliputi kegiatan *planning, programming, organizing, actuating, controlling dan evaluating*. Dua hal yang terakhir ini hampir merupakan titik lemah didalam manajemen tradisional yang menganggap bahwa fungsi kontrol dan evaluasi pada setiap proses termasuk pendidikan, dianggap sebagai upaya mengurangi kebebasan dan kemerdekaan para pelaksana kegiatan tersebut. Padahal apabila kedua fungsi manajemen tersebut tidak dilaksanakan dengan baik hampir dapat dipastikan bahwa apabila dalam pelaksanaan program terjadi penyimpangan dan pengorganisasian yang tidak sesuai dengan karakteristik program, maka tujuan tidak akan tercapai. Oleh karena itu berdasarkan tiga alasan utama tersebut di atas evaluasi sangat diperlukan dalam dunia pendidikan, baik ditinjau dari segi profesionalisme tugas kependidikan, proses dan manajemen pendidikan itu sendiri mengharuskan adanya aktivitas evaluasi.

BAB IV

ARTI DAN LETAK TAKSONOMI DALAM PENDIDIKAN

A. Pengertian Taksonomi dalam Pendidikan

Kata **taksonomi** diambil dari bahasa Yunani *tassein* yang berarti *untuk mengelompokkan* dan *nomos* yang berarti *aturan*. Taksonomi dapat diartikan sebagai pengelompokan suatu hal berdasarkan hierarki (tingkatan) tertentu. Di mana taksonomi yang lebih tinggi bersifat lebih umum dan taksonomi yang lebih rendah bersifat lebih spesifik.

Dalam pendidikan, taksonomi dibuat untuk mengklasifikasikan tujuan pendidikan. Dalam hal ini, tujuan pendidikan dibagi menjadi beberapa *domain*, yaitu: kognitif, afektif, dan psikomotor. Dari setiap ranah tersebut dibagi kembali menjadi beberapa kategori dan subkategori yang berurutan secara hirarkis (bertingkat), mulai dari tingkah laku yang sederhana sampai tingkah laku yang paling kompleks. Tingkah laku dalam setiap tingkat diasumsikan menyertakan juga tingkah laku dari tingkat yang lebih rendah.

B. Arti dan Letak Taksonomi dalam Pendidikan

Taksonomi ini pertama kali disusun oleh Benjamin S. Bloom dan kawan-kawan pada tahun 1956, sehingga sering pula disebut sebagai "Taksonomi Bloom". **Taksonomi Bloom** merujuk pada taksonomi yang dibuat untuk tujuan pendidikan. Taksonomi ini pertama kali disoleh Benjamin S. Bloom pada tahun 1956. Dalam hal ini, tujuan pendidikan dibagi menjadi beberapa *domain* (ranah, kawasan) dan setiap domain tersebut dibagi kembali ke dalam pembagian yang lebih rinci berdasarkan hirarkinya.

Tujuan pendidikan dibagi ke dalam tiga domain, yaitu:

1. *Cognitive Domain* (Ranah Kognitif), yang berisi perilaku-perilaku yang menekankan aspek intelektual, seperti pengetahuan, pengertian, dan keterampilan berpikir.
2. *Affective Domain* (Ranah Afektif) berisi perilaku-perilaku yang menekankan aspek perasaan dan emosi, seperti minat, sikap, apresiasi, dan cara penyesuaian diri.
3. *Psychomotor Domain* (Ranah Psikomotor) berisi perilaku-perilaku yang menekankan aspek keterampilan motorik seperti tulisan tangan, mengetik, berenang, dan mengoperasikan mesin.

Beberapa istilah lain yang juga menggambarkan hal yang sama dengan ketiga domain tersebut di antaranya seperti yang diungkapkan oleh Ki Hajar Dewantoro, yaitu: cipta, rasa, dan karsa. Selain itu, juga dikenal istilah: penalaran, penghayatan, dan pengamalan. Dari setiap ranah tersebut dibagi kembali menjadi beberapa kategori dan subkategori yang berurutan secara hirarkis (bertingkat), mulai dari tingkah laku yang sederhana sampai tingkah laku yang paling kompleks. Tingkah laku dalam setiap tingkat diasumsikan menyertakan juga tingkah laku dari tingkat yang lebih rendah, seperti misalnya dalam ranah kognitif, untuk mencapai “pemahaman” yang berada di tingkatan kedua juga diperlukan “pengetahuan” yang ada pada tingkatan pertama.

1. Domain Kognitif

Bloom membagi domain kognisi ke dalam 6 tingkatan. Domain ini terdiri dari dua bagian: Bagian pertama berupa adalah Pengetahuan (kategori 1) dan bagian kedua berupa Kemampuan dan Keterampilan Intelektual (kategori 2-6)

a. Pengetahuan (*Knowledge*)

Berisikan kemampuan untuk mengenali dan mengingat peristilahan, definisi, fakta-fakta, gagasan, pola, urutan, metodologi, prinsip dasar, dsb. Sebagai contoh, ketika diminta menjelaskan manajemen kualitas, orang yg berada di level ini

bisa menguraikan dengan baik definisi dari kualitas, karakteristik produk yang berkualitas, standar kualitas minimum untuk produk.

b. Pemahaman (*Comprehension*)

Berisikan kemampuan yang menuntut peserta didik untuk memahami atau mengerti tentang materi pelajaran yang disampaikan guru dan dapat memanfaatkannya tanpa harus menghubungkannya dengan hal-hal lain. Kemampuan ini dijabarkan menjadi tiga, yaitu: *menterjemahkan*, *menafsirkan*, dan *mengekstrapolasi*.

c. Aplikasi (*Application*)

Di tingkat ini, seseorang memiliki kemampuan untuk menerapkan gagasan, prosedur, metode, rumus, teori, dsb di dalam kondisi kerja. Sebagai contoh, ketika diberi informasi tentang penyebab meningkatnya reject di produksi, seseorang yg berada di tingkat aplikasi akan mampu merangkum dan menggambarkan penyebab turunnya kualitas dalam bentuk fish bone diagram.

d. Analisis (*Analysis*)

Di tingkat analisis, seseorang akan mampu menganalisis informasi yang masuk dan membagi-bagi atau menstrukturkan informasi ke dalam bagian yang lebih kecil untuk mengenali pola atau hubungannya, dan mampu mengenali serta membedakan faktor penyebab dan akibat dari sebuah skenario yg rumit. Sebagai contoh, di level ini seseorang akan mampu memilah-milah penyebab meningkatnya reject, membanding-bandingkan tingkat keparahan dari setiap penyebab, dan menggolongkan setiap penyebab ke dalam tingkat keparahan yg ditimbulkan.

e. Sintesis (*Synthesis*)

Satu tingkat di atas analisis, seseorang di tingkat sintesa akan mampu menjelaskan struktur atau pola dari sebuah skenario yang sebelumnya tidak terlihat, dan mampu mengenali data atau informasi yang harus didapat untuk menghasilkan solusi yg dibutuhkan. Sebagai contoh, di tingkat ini seorang manajer kualitas mampu memberikan solusi untuk menurunkan tingkat reject di produksi berdasarkan pengamatannya terhadap semua penyebab turunnya kualitas produk.

e. Evaluasi (*Evaluation*)

Dikenali dari kemampuan untuk memberikan penilaian terhadap solusi, gagasan, metodologi, dsb dengan menggunakan kriteria yang cocok atau standar yg ada untuk memastikan nilai efektivitas atau manfaatnya. Sebagai contoh, di tingkat ini seorang manajer kualitas harus mampu menilai alternatif solusi yg sesuai untuk dijalankan berdasarkan efektivitas, urgensi, nilai manfaat, nilai ekonomis, dsb.

2. Domain Afektif

Pembagian domain ini disusun Bloom bersama dengan David Krathwol.

a. Penerimaan (*Receiving/Attending*)

Kesediaan untuk menyadari adanya suatu fenomena di lingkungannya. Dalam pengajaran bentuknya berupa mendapatkan perhatian, mempertahankannya, dan mengarahkannya.

b. Tanggapan (*Responding*)

Memberikan reaksi terhadap fenomena yang ada di lingkungannya. Meliputi persetujuan, kesediaan, dan kepuasan dalam memberikan tanggapan.

c. Penghargaan (*Valuing*)

Berkaitan dengan harga atau nilai yang diterapkan pada suatu objek, fenomena, atau tingkah laku. Penilaian berdasar pada internalisasi dari serangkaian nilai tertentu yang diekspresikan ke dalam tingkah laku.

d. Pengorganisasian (*Organization*)

Memadukan nilai-nilai yang berbeda, menyelesaikan konflik di antaranya, dan membentuk suatu sistem nilai yang konsisten.

e. Karakterisasi Berdasarkan Nilai-nilai (*Characterization by a Value or Value Complex*)

Memiliki sistem nilai yang mengendalikan tingkah-lakunya sehingga menjadi karakteristik gaya-hidupnya.

3. Domain Psikomotor

Rincian dalam domain ini tidak dibuat oleh Bloom, tapi oleh ahli lain berdasarkan domain yang dibuat Bloom.

a. Persepsi (*Perception*)

Penggunaan alat indera untuk menjadi pegangan dalam membantu gerakan.

b. Kesiapan (*Set*)

Kesiapan fisik, mental, dan emosional untuk melakukan gerakan.

c. Guided Response (*Respon Terpimpin*)

Tahap awal dalam mempelajari keterampilan yang kompleks, termasuk di dalamnya imitasi dan gerakan coba-coba.

d. Mekanisme (*Mechanism*)

Membiasakan gerakan-gerakan yang telah dipelajari sehingga tampil dengan meyakinkan dan cakap.

e. Respon Tampak yang Kompleks (*Complex Overt Response*)

Gerakan motoris yang terampil yang di dalamnya terdiri dari pola-pola gerakan yang kompleks.

f. Penyesuaian (*Adaptation*)

Keterampilan yang sudah berkembang sehingga dapat disesuaikan dalam berbagai situasi.

g. Penciptaan (*Origination*)

Membuat pola gerakan baru yang disesuaikan dengan situasi atau permasalahan tertentu.



STATISTIK ELEMENTER

A. Pendahuluan

Pemahaman statistik elementer sangat membantu mahasiswa dalam memahami berbagai macam buku-buku yang terkait dengan evaluasi pendidikan. Seorang calon guru hendaknya memahami berbagai macam cara membuat distribusi frekuensi, mencari nilai mean, median, modus, dan variabilitas data yaitu varians dan standar deviasi.

B. Distribusi Frekuensi

Distribusi (*=distribution dalam Bahasa Inggris*), yang berarti “penyaluran/pencaran, pembagian”. Sedangkan frekuensi (*frequency*), yang berarti “kekerapan/ keseringan /jarang-kerapnya”, dalam statistik menunjuk pada angka *berapa kali* suatu variable itu berulang dalam deretan angka. Jadi **distribusi frekuensi** (dalam statistik) adalah suatu keadaan yang menggambarkan bagaimana frekuensi dari gejala atau variabel yang dilambangkan dengan angka itu, telah tersalur, terbagi, atau terpecah”.

1. Pengertian Tabel Distribusi Frekuensi

Tabel adalah alat penyajian data statistik yang berbentuk (dituangkan dalam bentuk) kolom dan lajur. Jadi table Distribusi frekuensi adalah alat penyajian data statistik yang berbentuk kolom dan lajur, yang didalamnya dimuat angka yang dapat melukiskan atau menggambarkan pencaran atau pembagian frekuensi dari variable yang sedang menjadi obyek penelitian. Dengan demikian dalam table distribusi frekuensi

akan memuat : (1) variable, (2) frekuensi, (3) jumlah frekuensi. Sering disebut “tabel frekuensi”.

2. Macam-Macam Tabel Distribusi Frekuensi

Tabel distribusi frekuensi yang dipandang penting yaitu:

- a. Tabel distribusi frekuensi data tunggal.
- b. Tabel Distribusi frekuensi data kelompok.

3. Cara Membuat Tabel Distribusi Frekuensi

Dalam pembahasan ini akan dibicarakan yaitu cara pembuatan tabel distribusi frekuensi data tunggal dan data kelompok.

(1). Cara membuat tabel distribusi data tunggal

Pada tabel distribusi frekuensi ini dapat dibagi menjadi dua yaitu : (a) tabel distribusi frekuensi data tunggal yang semua scornya berfrekuensi satu, (b) tabel distribusi frekuensi data tunggal yang semua scornya berfrekuensi lebih dari satu.

- a. Tabel distribusi frekuensi data tunggal yang semua scornya berfrekuensi satu (data kasar)

Misalnya 10 mahasiswa mengikuti ujian tes Bahasa Arab dengan hasil yaitu :

No.	Nama mahasiswa	Nilai
1	Arifin	6,5
2	Bambang	3
3	Catur	6
4	Dedy	4,5
5	Endang	7,5
6	Farida	4
7	Giyem	7
8	Hanum	5,5
9	Inayah	8
10	Junaidi	5

Selanjutnya kita membuat penyajian tabel distribusi frekuensi sebagai berikut :

Tabel. 1. *Distribusi Frekuensi Nilai Hasil Ujian Mata Kuliah Statistik Pendidikan dari 10 Mahasiswa.*

No.	Nama mahasiswa	Nilai	Frekuensi
1	Inayah	8	1
2	Endang	7,5	1
3	Giyem	7	1
4	Arifin	6,5	1
5	Catur	6	1
6	Hanum	5,5	1
7	Junaidi	5	1
8	Dedy	4,5	1
9	Farida	4	1
10	Bambang	3	1

Namun dalam pembuatan table distribusi frekuensi dewasa ini untuk keterangan frekuensi satu (1) dihilangkan, dengan asumsi semuanya adalah satu (1).

b. Tabel distribusi frekuensi data tunggal yang semua scornya berfrekuensi lebih dari satu.

Missal 40 murid Madrasah Ibtidaiyyah mengikuti ujian pendidikan agama islam sebagai berikut :

5 8 6 4 6 7 9 6 4 5 3 5 8 6 5 4 6 7 7 10
4 6 5 7 8 9 3 5 6 8 10 4 9 5 3 6 8 6 7 6

Langkah yang harus ditempuh yaitu :

1. Mencari nilai tertinggi (H) dan nilai terendah (L) $H = 10$, $L = 3$.

2. Menghitung frekuensi yang ada dengan bantuan jari-jari (tailes).
3. Mengubah jari-jari menjadi angka biasa.

Tabel. 2. *Distribusi Frekuensi dari 40 Murid SD Mengikuti Ujian Pendidikan Agama Islam*

No.	Nilai (X)	Jari-Jari/Tally	Frekuensi
1	10	II	2
2	9	III	3
3	8	IIII	5
4	7	IIII	5
5	6	IIII IIII	10
6	5	IIII II	7
7	4	IIII	5
8	3	III	3
TOTAL			40 = N

(2). Tabel distribusi frekuensi data Bergolong

Misalnya. Data Kasar : Hasil Ujian Statistik Pendidikan dari 15 Mahasiswa STAIN Kudus Kudus Tahun 2004/2005, sebagai berikut :

80 75 73 75 75
80 71 79 77 69
77 77 62 68 72

Untuk mengolah data data tersebut yang perlu kita lakukan adalah :

1. Mencari nilai tertinggi (H) dan nilai terendah (L) H = 80, L = 62.

2. Menentukan luas penyebaran nilai yang ada, dengan formula :

$$R \text{ (total range)} = H - L + 1 \text{ (bilangan konstan)}$$

$$R = 80 - 62 + 1$$

$$R = 19 \text{ butir nilai.}$$

3. Menetapkan jumlah kelas interval yang diperkirakan, dengan menggunakan rumus :

$$K = 1 + 3,3 \log n$$

$$K = 1 + 3,3 \times \log 15$$

$$K = 1 + 3,3 \times 1,176091259$$

$$K = 1 + 3,881101155$$

$$K = 4,881101155 \text{ atau } 5$$

4. mencari interval kelas (I) dengan formula :

$$I = R/K$$

Dimana : I = interval kelas

Range = total range

K = jumlah kelas

$$I = 19/5 = \underline{\mathbf{3,8}} \text{ dapat memilih interval } \mathbf{4}$$

5. Menetapkan bilangan dasar yaitu bilangan batas antara interval yang satu dengan yang lain, dengan memperhatikan bilangan terendah dan tertinggi.
6. Mempersiapkan tabel distribusi frekuensi.

Tabel. 2. *Distribusi Frekuensi dari 15 Mahasiswa STAIN Kudus dalam Mengikuti Ujian Statistik Tahun 2004/2005*

No.	Interval Nilai	Jari-Jari/Tally	Frekuensi
1	62 – 65	I	1
2	66 – 69	II	2
3	70 – 73	III	3
4	74 – 77	IIII I	6
5	78 – 81	III	3
TOTAL			15 = N

4. Grafik Sebagai Alat Penggambaran Distribusi Frekuensi

1.1 Pengertian Grafik dan Susunannya

Grafik atau diagram tidak lain adalah alat penyajian data statistik yang tertuang dalam bentuk lukisan, baik lukisan garis, gambar, maupun lambing. Sebuah grafik yang lengkap pada umumnya terdiri dari 13 bagian, yaitu (1), nomor grafik, (2) judul grafik, (3) sub judul grafik, (4) unit skal grafik, (5) angka skal grafik, (6) tanda skala grafik, (7) ordinat atau ordinal atau sumbu horizontal, (8) koordinat pertolongan/kisi-kisi), (9) abscis (sumbu horizontal), (10) titi nol (titik awal), (11) lukisan grafis, (12) kunci grafik (keterangan grafik), (13) sumber grafik (sumber data).

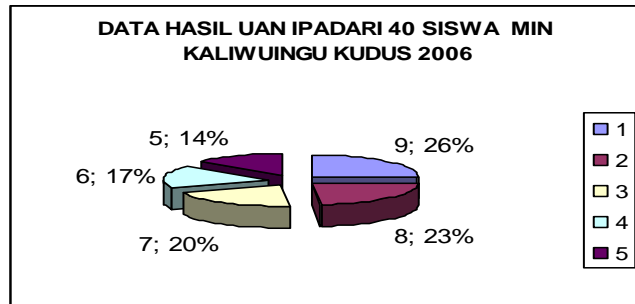
1.2 Macam-Macam Grafik

Berbagai macam grafik: (1) Grafik Balok/garfik batang/ barchart : Grafik balok tunggal, Grafik balok ganda, Grafik balok terbagi, Grafik balok vertical, Grafik balok horizontal, Grafik balok bilateral. (2) Grafik lingkaran (pastel), (3) Grafik gambar (pictograph), (4) Grafik peta (kartogram), (5) Grafik bidang, (6) Grafik volume, (7) Grafik Garis : grafik garis tunggal, garis mejemuk atau ganda dan grafik polygon frekuensi. (8) Grafik ruang atau grafik histogram atau histogram frequency.

1.3 Contoh-Contoh Aplikasi Grafis Distribusi Frekuensi.
4.3.1. Tabel Distribusi Frekuensi Data Tunggal

Tabel 3. Distribusi Frekuensi Nilai Hasil UAN Bidang Studi IPA dari 40 Orang siswa MIN Kaliwungu Kudus 2006

No	Nilai IPA	Frekuensi (f)
1	9	5
2	8	8
3	7	12
4	6	10
5	5	5
Total		N= 40

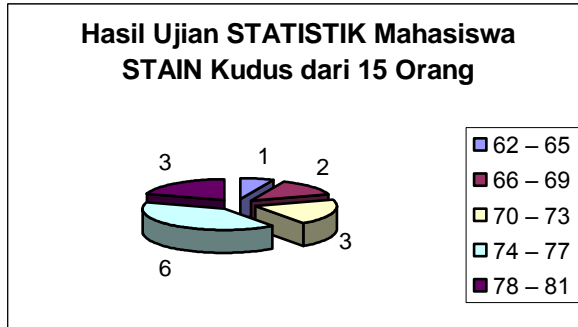


4.3.2. Tabel Distribusi Frekuensi Data Kelompokan

Tabel. 4. *Distribusi Frekuensi dari 15 Mahasiswa STAIN Kudus dalam Mengikuti Ujian Statistik Tahun 2004/2005*

No.	Interval Nilai	Frekuensi
1	62 – 65	1
2	66 – 69	2

3	70 – 73	3
4	74 – 77	6
5	78 – 81	3



C. Ukuran Gejala Pusat (*Central Tendency*)

Apabila kita mencermati dari kajian kita pada Bab yang lalu table distribusi frekuensi dan grafik untuk menganalisa data yang ada belumlah cukup. Bahkan ini merupakan awal dalam kegiatan analisa statistik. Untuk itu perlu melengkapi pisau analisa statistik dengan mempelajari ukuran gejala pusat (*Central Tendency*) dan ukuran letak. Seperti modus, median dan mean. Sehingga dapat menjelaskan kelompok yang telah diobservasi dengan data kuantitatif yang ada. Namun dalam kesempatan kali ini akan kita bahas tentang mencari nilai koefisien mean (nilai rata-rata).

1. Mean

Mean adalah *jumlah dari keseluruhan angka (bilangan) yang ada, dibagi dengan banyaknya angka (bilangan) tersebut.*

Formulanya :

$$M_x = \frac{\sum X_i}{N}$$

Contoh :

Dalam penelitian pada siswa X ditemukan nilai prestasi belajar dari bidang studi PAI, PPKn, Bahasa Indonesia, Bahasa Inggris, Bahasa Daerah, IPS dan IPA sebagaimana pada table berikut ini :

Tabel 4.2. Hasil Prestasi Belajar Siswa X dari 7 Bidang Studi

Nilai (X)	Frekuensi (1)
9	1
8	1
7	1
6	1
5	1
4	1
3	1
42 =	1

Carilah nilai rata-rata siswa X tersebut ?

Jawab :

Diketahui $N = 7$ dan $\Sigma X = 42$, maka :

$$\text{Mex} = \frac{\Sigma X_i}{N}$$

$$\text{Mex} = \frac{42}{7}$$

$$\text{Mex} = \underline{6}$$

Jadi nilai rata-rata siswa X dari 7 bidang studi adalah 6.

2. Menghitung Mean Untuk Data Bergolong

Contoh :

Dari hasil test tentang kemampuan dalam memahami teks Bahasa Arab dari 100 mahasiswa, ditemukan distribusi frekuensi sebagai berikut :

Tabel 4.3. Kemampuan dalam Memahami Teks Bahasa Arab
Dari 100 Mahasiswa

Interval Nilai Kemampuan	Frekuensi/Jumlah
21-30	3
31-40	7
41-50	16
51-60	30
61-70	25
71-80	15
81-90	3
91-100	1
Jumlah	100

Berdasarkan data tersebut di atas hitunglah meannya !.

Jawab :

1. Menghitung Mean (Me)

Untuk distribusi frekuensi data bergolong digunakan rumus :

$$Me = \frac{\sum fiXi}{N}$$

Dimana :

Me = Mean untuk data bergolong

N = Jumlah data/sampel

$f_i X_i$ = produk perkalian antara f_i pada tiap interval data dengan tanda klas (X_i). Tanda klas X_i adalah rata-rata dari batas bawah dan batas pada setiap interval data.

Misalnya : untuk mencari nilai tengah interval

$$21 + 30/2 = \mathbf{25,5}$$
 dikalikan dengan frekuensi 3 = 75,5

Jawab :

Tabel 4.4. Tabel Penolong Kemampuan dalam Memahami Teks Bahasa Arab Dari 100 Mahasiswa

Interval Nilai Kemampuan	Nilai Tengah (X_i)	Frekuensi / Jumlah (f)	$f_i X_i$
21-30	25.5	3	76.5
31-40	35.5	7	248.5
41-50	45.5	16	728
51-60	55.5	30	1665
61-70	65.5	25	1637.5
71-80	75.5	15	1132.5
81-90	85.5	3	256.5
91-100	95.5	1	95.5
Jumlah		100 = N	5840 = $\Sigma f_i X_i$

$$Me = \frac{\Sigma f_i X_i}{N}$$

$$Me = \frac{5840}{100}$$

$$Me = \mathbf{58,4}$$

Jadi nilai rata-rata (mean) kemampuan mahasiswa dalam memahami teks Bahasa Arab adalah 58,4.

D. Variabilitas

Dengan menggunakan ukuran tendensi central, maka kegiatan analisa statistik setapak lebih maju. Namun demikian penganalisaan data statistik dengan mengetahui frekuensinya dan nilai rata-ratanya, masih belum mengetahui bagaimana penyebaran /pemencaran/ variasi/dispersi/variabilitas yang sebenarnya. Contoh : hasil studi Bahasa Arab yang dilakukan oleh siswa MTs tersebar antara nilai 40 sampai 60, sedangkan siswa SMP tersebar diantara nilai 0 sampai 100. apabila kita cermati, maka kita akan menemukan perbedaan kualitas hasil belajar.

Ukuran penyebaran data adalah *berbagai macam ukuran statistik yang dapat dipergunakan untuk mengetahui luas penyebaran data, variasi data, homogenitas data, atau stabilitas data.*

1. Macam-Macam Variabilitas Data

Macam-macam variabilitas data terdiri dari : (1) Range; (2) Deviasi; (3) Variance; (4) Ukuran penyebaran relatif. Namun dalam penggunaannya khususnya di dunia pendidikan yaitu 1,2 dan 3.

2. Rentang Data

Adalah salah satu ukuran statistik yang menunjukkan jarak penyebaran antara skor nilai terendah (*lowest score*) sampai skor nilai tertinggi (*highest score*) dengan singkat dapat dirumuskan :

$$R = H - L$$

Contoh :

Pendapatan dari kotak amal di Masjid X, pendapatan setiap Jum'at sekali dibuka. Setelah dihitung ternyata pendapatan selama 8 sepekan dalam satuan ribu rupiah yaitu :

50, 100, 75, 90, 45, 100, 65, 56

data terkecil dari kelompok itu = 45

data terbesar = 100

jadi rentang (R) = 100 – 45 = **55**

Jadi rentang pendapatan kotak amal 8 sepekan adalah Rp. 55 ribu. Rentang data inilah yang menunjukkan tingkat variasi kelompok.

3. Varians dan Standar Deviasi

Salah satu teknik yang digunakan untuk menjelaskan homogenitas kelompok adalah dengan varians. Varians populasi (σ^2) atau varians sample (s^2) merupakan jumlah kuadrat semua deviasi nilai-nilai individual terhadap rata-rata kelompok. Akar varians disebut standar deviasi atau simpangan baku.

Contoh :

Dilakukan penelitian oleh seorang guru untuk mengetahui variasi dari prestasi siswa pada mata pelajaran PAI dari 15 siswa di Sekolah X. Setelah dilakukan observasi ditemukan data angka prestasi sebagai berikut :

50	60	70	60	78
65	50	80	85	70
50	80	70	75	90

Carilah varian dan simpangan baku atau standar deviasinya sample dan populasi ?

Jawab :

1. Mencari nilai rata-rata (Me) prestasi siswa dari 15 orang.

$$Mex = \frac{\sum Xi}{N}$$

$$Mex = 50 + 65 +$$

$$50 + 60 + 50 + 80 + 70 + 80 + 70 + 60 + 85 + 75 + 78 + 70 + 90 / 15$$

$$Mex = 1033 / 15 \quad Mex = \underline{\underline{68,87}}$$

2. Membuat tabel penolong untuk menghitung varians

Tabel 5.1. Cara Menghitung Varian dan Simpangan Baku Sekelompok Prestasi Siswa Pada Matpel PAI di Sekolah X

No.Siswa	Skor (Xi)	Mex	(Xi-Mex)	(Xi-Mex) ²
1	50	68.87	-18.87	356.0769
2	65	68.87	-3.87	14.9769
3	50	68.87	-18.87	356.0769
4	60	68.87	-8.87	78.6769
5	50	68.87	-18.87	356.0769
6	80	68.87	11.13	123.8769
7	70	68.87	1.13	1.2769
8	80	68.87	11.13	123.8769
9	70	68.87	1.13	1.2769
10	60	68.87	-8.87	78.6769
11	85	68.87	16.13	260.1769

12	75	68.87	6.13	37.5769
13	78	68.87	9.13	83.3569
14	70	68.87	1.13	1.2769
15	90	68.87	21.13	446.4769
N=15	1033			2319.734

3. Memasukkan ke dalam rumus varian dan simpangan baku

$$\sigma^2 = \frac{\sum(X_i - M_{ex})^2}{n} \quad (\text{rumus varian populasi}).$$

$$\sigma = \sqrt{\frac{\sum(X_i - M_{ex})^2}{n}} \quad (\text{rumus SD Populasi})$$

$$s^2 = \frac{\sum(X_i - M_{ex})^2}{(n - 1)} \quad (\text{rumus varian sampel}).$$

$$s = \sqrt{\frac{\sum(X_i - M_{ex})^2}{(n - 1)}} \quad (\text{rumus standar deviasi sampel})$$

Dimana :

σ^2 = varian populasi

σ = standar deviasi populasi

s^2 = varian sampel

s = standar deviasi sample
n = jumlah /populasi sample

$$\sigma^2 = \frac{\sum(X_i - M_{ex})^2}{n}$$

$$\sigma^2 = 2319,734 / 15$$

$$\sigma^2 = \underline{\underline{154,6489333}}$$

$$\sigma = \sqrt{\frac{\sum(X_i - M_{ex})^2}{n}}$$

$$\sigma = \sqrt{154,6489333}$$

$$\sigma = \underline{\underline{12,43579243}}$$

$$s^2 = \frac{\sum(X_i - M_{ex})^2}{(n - 1)}$$

$$s^2 = 2319,734 / 15-1$$

$$s^2 = 2319,734 / 14$$

$$s^2 = \underline{\underline{165,6952857}}$$

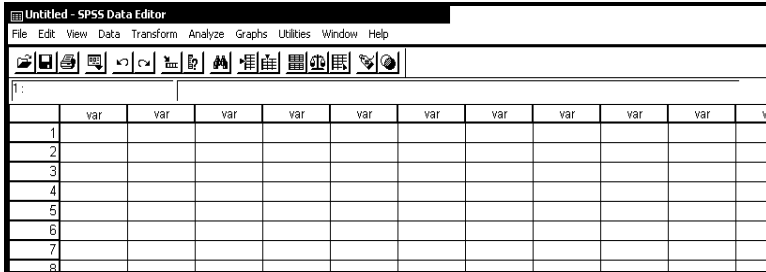
$$s = \sqrt{\frac{\sum(X_i - M_{ex})^2}{(n - 1)}}$$

$$s = \sqrt{165,6952857}$$

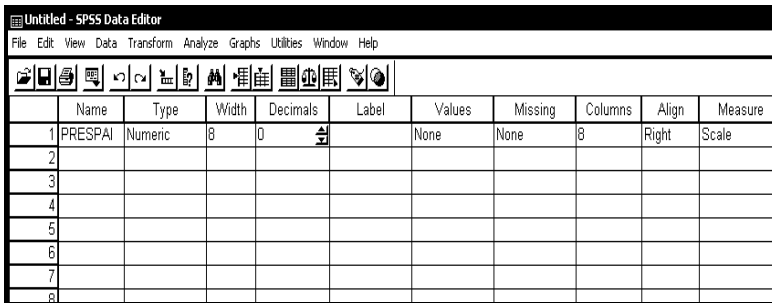
$$s = \underline{\underline{12,87226809}}$$

Langkah-langkah proses analisis data program SPSS:

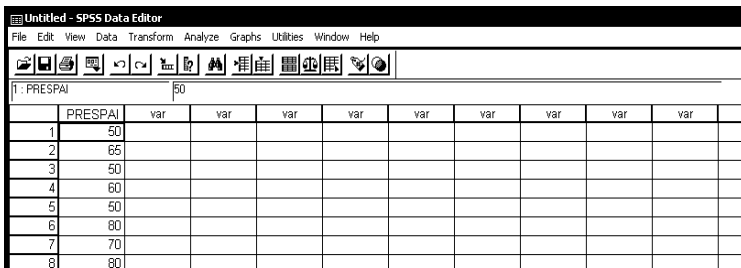
1. Bukalah program SPSS for Windows



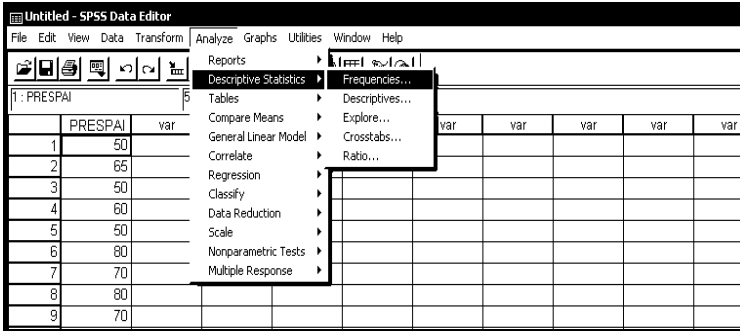
2. Klik variable view dan definisikan variabel (8 karakter) dan desimal ketik 0.



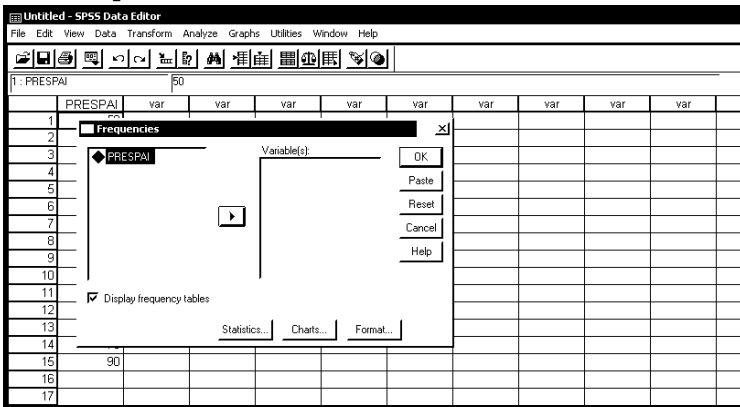
3. Klik data view



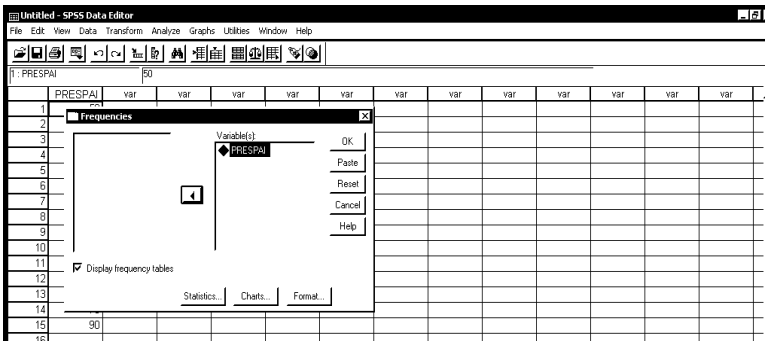
4. Klik analisis deskriptif statistics dan klik frequencies



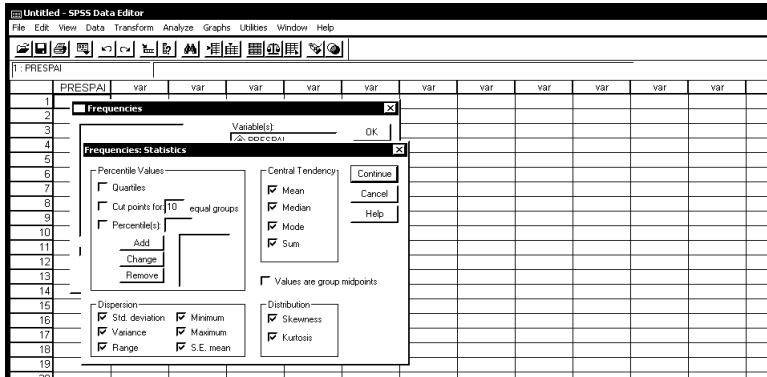
5. Klik frequencies



6. Blok presbel dan klik tanda panah ke dalam kotak variabel



7. Frequencies : Statistics



8. Klik Continue dan Klik OK

Hasil Olah Data SPSS

Frequencies

Statistics

PRESPAI		
N	Valid	15
	Missing	0
Mean		68.87
Std. Deviation		12.872
Variance		165.695
Sum		1033

PRESPAI

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	50	3	20.0	20.0	20.0
	60	2	13.3	13.3	33.3
	65	1	6.7	6.7	40.0
	70	3	20.0	20.0	60.0
	75	1	6.7	6.7	66.7
	78	1	6.7	6.7	73.3
	80	2	13.3	13.3	86.7
	85	1	6.7	6.7	93.3
	90	1	6.7	6.7	100.0
Total		15	100.0	100.0	

4. Menghitung Standard Deviasi Untuk Data Bergolong

$$s = \sqrt{\frac{\sum f_i (X_i - M_{ex})^2}{(n-1)}} \quad \begin{array}{l} \text{(rumus standar deviasi} \\ \text{sampel)} \end{array}$$

Contoh :

Dari hasil test tentang kemampuan dalam memahami teks Bahasa Arab dari 100 mahasiswa diperguruan tinggi X, ditemukan distribusi frekuensi sebagai berikut :

Tabel 5.2. Kemampuan dalam Memahami Teks Bahasa Arab
Dari 100 Mahasiswa di PT. X

Interval Nilai Kemampuan	Frekuensi/Jumlah
21-30	3
31-40	7
41-50	16
51-60	30
61-70	25
71-80	15
81-90	3
91-100	1
Jumlah	100

Berdasarkan data tersebut di atas hitunglah standar deviasi sample pada data bergolong !.

Jawab :

1. Membuat table penolong untuk mencari M_{ex} .
2. Membuat table penolong untuk mencari SD.

3. Memasukan ke dalam rumus.

Langkah-langkah tersebut di atas dapat dilakukan sebagai berikut :

Tabel 5.3. Tabel Penolong Untuk Mencari Mean X Kemampuan dalam Memahami Teks Bahasa Arab Dari 100 Mahasiswa

Interval Nilai Kemampuan	Nilai Tengah (Xi)	Frekuensi / Jumlah (f)	fiXi
21-30	25.5	3	76.5
31-40	35.5	7	248.5
41-50	45.5	16	728
51-60	55.5	30	1665
61-70	65.5	25	1637.5
71-80	75.5	15	1132.5
81-90	85.5	3	256.5
91-100	95.5	1	95.5
Jumlah		100 = N	5840 = $\Sigma fiXi$

$$Me = \frac{\Sigma fiXi}{N}$$

$$Me = \frac{5840}{100}$$

$$Me = \mathbf{58,4}$$

Jadi nilai rata-rata (mean) kemampuan mahasiswa dalam memahami teks Bahasa Arab adalah 58,4.

Tabel 5.4. Tabel Penolong Untuk Mencari SD pada Kemampuan dalam Memahami Teks Bahasa Arab Dari 100 Mahasiswa

Interval Nilai	Nilai Tengah (Xi)	(fi)	Mex	Xi-Mex	(Xi-Mex) ²	fi(Xi-Mex) ²
21-30	25.5	3	58.4	-32.9	1082.41	3247.23
31-40	35.5	7	58.4	-22.9	524.41	3670.87
41-50	45.5	16	58.4	-12.9	166.41	2662.56
51-60	55.5	30	58.4	-2.9	8.41	252.3
61-70	65.5	25	58.4	7.1	50.41	1260.25
71-80	75.5	15	58.4	17.1	292.41	4386.15
81-90	85.5	3	58.4	27.1	734.41	2203.23
91-100	95.5	1	58.4	37.1	1376.41	1376.41
-	-	100	-	-	-	19059

Berdasarkan rumus untuk menghitung standar deviasi data bergolong diatas, maka dapat dikerjakan sebagai berikut:

$$s = \sqrt{\frac{\sum fi (Xi - Mex)^2}{(n - 1)}}$$

$$s = \sqrt{\frac{19059}{100-1}}$$

$$s = \sqrt{\frac{19059}{99}}$$

$$s = \sqrt{192,5151515}$$

$$s = 13,87498294$$

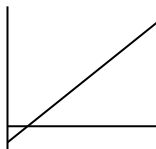
$$s = \underline{\underline{13,875}}$$

Jadi standar deviasi nilai kemampuan memahami teks bahasa Arab dari 100 mahasiswa di perguruan tinggi X sebesar **13,875**.

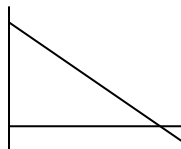
E. Teknik Korelasi

Penggunaan teknik analisis korelasi yang sering dipakai adalah teknik analisis Korelasi Product Moment. Jika variable yang dikorelasikan berbentuk data kontinu, sample bersifat homogen dan regresi linier. Lambangnya “ r_{xy} ” atau R.

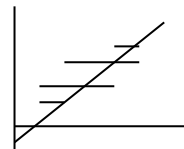
Teknik analisa korelasional adalah teknik analisa statistik mengenai hubungan antara dua variabel ataulebih. Tujuannya adalah mencari bukti ada atau tidak ada hubungan, menjawab pertanyaan (lemah, cukup, kuat), memperoleh kejelasan dan kepastian (signifikan atau tidak). Tandanya : (+) korelasi searah, (-) korelasi berlawanan arah(bukan tanda aljabar). Sifatnya relatif tidak eksak, seperti mengukur panjang sebuah mistar. Arah korelasi dapat dilihat pada tabel dibawah ini :



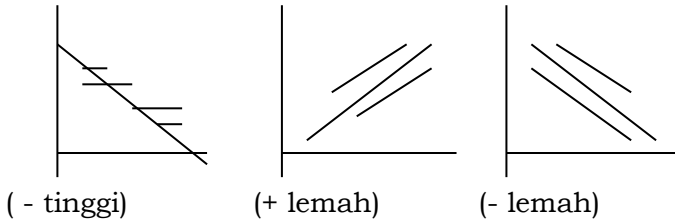
(+ maksimal)



(- maksimal)



(+ tinggi)



Contoh :

Dilakukan pengujian untuk menemukan hubungan antara hasil ujian pelaksanaan ujian mata pelajaran bahasa arab dan nahwu shorof dari 10 siswa. Setelah dilakukan ujian diketemukan angka sebagai berikut :

No.	Bahasa Arab	Nahwu Shorof
1	7	8
2	8	8
3	9	9
4	7	7
5	6	7
6	7	7
7	7	8
8	6	7
9	6	7
10	7	8

Adapun langkah-langkah dalam mencari nilai hubungan yaitu :

1. Bukalah program SPSS fro Windows dan definisikan variabel

Untitled - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	X	Numeric	8	0	Bahasa Arab	None	None	8	Right	Scale
2	Y	Numeric	8	0	Nahwu Shorof	None	None	8	Right	Scale
3										
4										
5										
6										
7										

2. Masukkan data kedalam variabel

Untitled - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

	X	Y	var	var	var	var
1	7	8				
2	8	8				
3	9	9				
4	7	7				
5	6	7				
6	7	7				
7	7	8				
8	6	7				
9	6	7				
10	7	8				
11						

3. Lakukan analisis dengan memilih correlate dan bivariate

Untitled - SPSS Data Editor

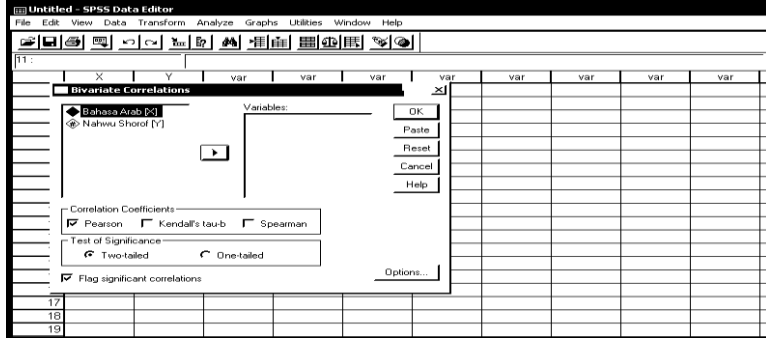
File Edit View Data Transform Analyze Graphs Utilities Window Help

Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Correlate
Regression
Classify
Data Reduction
Scale
Nonparametric Tests
Multiple Response

Bivariate...
Partial...
Distances...

	X	Y	var	var	var	var
1	7					
2	8					
3	9					
4	7					
5	6					
6	7					
7	7					
8	6	7				
9	6	7				
10	7	8				
11						
12						
13						
14						

4. Masukkan variabel ke dalam kotak variabel



5. Hasil olah data SPSS

Correlations

		Bahasa Arab	Nahwu Shorof
Bahasa Arab	Pearson Correlation	1	,843**
	Sig. (2-tailed)	.	,002
	N	10	10
Nahwu Shorof	Pearson Correlation	,843**	1
	Sig. (2-tailed)	,002	.
	N	10	10

** . Correlation is significant at the 0.01 level (2-tailed).

Pada tabel tersebut diatas terlihat bahwa hasil korelasinya diketemukan 0,843 dengan angka signifikansi 0,002. Adapun kriteria penerimaan adalah:

1. Jika angka signifikan $\leq 0,05$ maka dapat dikatakan $H_a =$ diterima dan $H_o =$ ditolak.
2. Jika angka signifikan $> 0,05$ maka dapat dikatakan $H_a =$ ditolak dan $H_o =$ diterima.

Dengan melihat kriteria diatas, maka dapat disimpulkan nilai koefisien korelasi 0,843 dengan angka signifikan 0,002 diatas, termasuk dalam kategori 1. Dengan demikian dapat disimpulkan bahwa ada korelasi antara nilai bahasa arab dan nahwu shorof.



EVALUASI HASIL BELAJAR MENGAJAR

Pelaksanaan kegiatan program pendidikan yang telah dilakukan, hendaknya dilakukan penilaian (evaluasi). Evaluasi dapat menentukan tercapai tidaknya tujuan pendidikan dan pengajaran yang telah ditetapkan. Penilaian atau evaluasi pada dasarnya adalah memberikan pertimbangan atau harga atau nilai berdasarkan kriteria tertentu. Kegiatan pembelajaran merupakan kegiatan yang berproses dan bertujuan. Tujuan tersebut dinyatakan dalam bentuk rumusan tingkah laku yang diharapkan dimiliki siswa setelah mengalami pembelajaran.

Hasil yang diperoleh dari penilaian dinyatakan dalam bentuk hasil belajar. Oleh karena itu tindakan atau kegiatan tersebut dinamakan penilaian hasil belajar.

Agar dalam kegiatan proses evaluasi belajar mengajar dalam dilakukan secara efektif, maka perlu memperhatikan hal-hal sebagai berikut :

A. Prosedur Evaluasi

Kegiatan evaluasi hasil pengajaran pada umumnya dilakukan, baik dalam bentuk formatif maupun sumatif, telah dilaksanakan oleh guru. Melalui pertanyaan secara lisan atau tulisan pada akhir pengajaran guru menilai keberhasilan pengajaran (tes formatif). Demikian juga tes sumatif yang dilakukan pada akhir program seperti akhir kuartal atau akhir semester, penilaian diberikan terhadap para siswa untuk menentukan kemajuan belajarnya. Tes tertulis, baik jenis tes uraian (*essay test*) maupun tes ob-jektif, dilakukan oleh guru dalam penilaian sumatif tersebut. Penilaian hasil belajar bertujuan melihat kemajuan belajar para siswa. Dalam hal

penguasaan materi pengajaran yang telah dipelajarinya sesuai dengan tujuan-tujuan yang telah ditetapkan.

Agar dalam kegiatan evaluasi dapat berjalan secara efektif, maka diperlukan beberapa prosedur evaluasi. Menurut Yulien Standley mengatakan bahwa : “langkah-langkah evaluasi itu terdiri dari : menetapkan tujuan program, memilih alat yang layak, pelaksanaan pengukuran, memberikan skor, menganalisa dan menginterpretasikan skor, membuat catatan yang baik, dan menggunakan hasil-hasil pengukuran (Yulien Stanley, 1964:299).

Menurut Mochtar Buchari langkah-langkah pokok dalam evaluasi terdiri dari perencanaan, pengumpulan data, verifikasi data, analisa data dan penafsiran data (1972:24). Adapun prosedurnya adalah :

1. *Langkah Perencanaan* dalam melakukan langkah ini evaluator harus merencanakan hal-hal sebagai berikut :
 - a. *Merumuskan tujuan evaluasi*, yang hendak dilaksanakan dalam suatu proses pendidikan didasarkan atas tujuan yang hendak dicapai dalam program pendidikan tersebut, sesuai dengan tujuan program yang telah ditetapkan. Evaluasi yang dilakukan konselor pendidikan tentu berbeda dengan evaluasi tes seleksi. Evaluasi yang dilakukan oleh konselor pendidikan bertujuan untuk mendapatkan informasi sebanyak-banyaknya tentang peserta didik yang berguna dalam proses pembimbingan, sedangkan evaluasi tes seleksi bertujuan untuk mengetahui potensi-potensi yang ada pada para calon untuk dapat memilih calon yang memiliki kemampuan sesuai dengan persyaratan yang telah ditetapkan.
 - b. *Sasaran evaluasi*. Sasaran atau objek evaluasi hasil belajar adalah perubahan tingkah laku yang mencakup *bidang kognitif, afektif, dan psikomotor secara seimbang*. Masing-masing bidang terdiri dari sejumlah aspek. Aspek-aspek

tersebut sebaiknya dapat diungkapkan melalui penilaian tersebut. Dengan demikian dapat diketahui tingkah laku mana yang sudah dikuasainya oleh siswa dan mana yang belum sebagai bahan bagi perbaikan dan penyempurnaan program pengajaran selanjutnya.

- c. *Menetapkan aspek-aspek yang harus dinilai.* Penentuan aspek-aspek yang akan dinilai ditentukan oleh tujuan evaluasi yang dilaksanakan.
- d. *Menentukan metode evaluasi yang akan digunakan.* Metode evaluasi yang akan dipergunakan. Metode evaluasi yang akan dipergunakan ditentukan oleh jenis aspek yang akan dinilai. Misalnya untuk menilai motivasi belajar dapat digunakan metode checklist, interviu, dan observasi. Untuk menilai hasil belajar dapat dipergunakan metode observasi dan tes.
- e. *Memilih atau menyusun alat-alat evaluasi yang akan dipergunakan.* Alat-alat evaluasi yang dipergunakan ditentukan oleh metode evaluasi yang kita pergunakan. Penggunaan alat evaluasi hendaknya komprehensif meliputi tes dan bukan tes sehingga diperoleh gambaran hasil belajar yang objektif. Demikian juga penggunaan tes sebagai alat evaluasi tidak hanya membiasakan diri dengan tes objektif, tetapi juga tes esey. Tes esey sangat bermanfaat dalam melatih anak berpikir dan berbahasa tulisan. Dengan demikian kelemahan tes objektif dapat diimbangi dengan tes esey. Sebaliknya kelemahan tes esey dapat ditutupi dengan tes objektif. Penilaian hasil belajar hendaknya dilakukan secara berkesinambungan agar diperoleh hasil yang menggambarkan kemampuan anak yang sebenarnya di samping sebagai alat untuk meningkatkan motivasi belajarnya.
- f. *Menentukan kriteria yang akan dipergunakan.* Misalnya dalam evaluasi hasil belajar dapat dipergunakan skala

lima, skala sebelas, atau skala seratus. Evaluasi hasil belajar biasanya menggunakan dua kriteria, yakni kriteria mutlak atau penilaian acuan patokan (PAP) dan kriteria relatif atau penilaian acuan norma (PAN). Penilaian acuan patokan (PAP) sangat bermanfaat dalam upaya meningkatkan kualitas hasil belajar sebab siswa dipaksa untuk mencapai standar yang telah ditentukan. Dengan demikian hasil belajar siswa dapat diketahui derajat pencapaiannya, namun, resikonya bisa terjadi, yaitu melemahkan semangat belajar siswa apabila hasil yang diperolehnya relatif rendah atau di bawah standar yang diinginkan. Sebaliknya dalam penggunaan penilaian acuan norma, keberhasilan siswa ditentukan berdasarkan rata-rata kelas. Artinya, hasil yang dicapai oleh siswa dibandingkan dengan hasil belajar siswa lainnya. Dengan demikian dapat diketahui prestasi siswa dibandingkan dengan kelompoknya. Ada tiga kategori keberhasilan siswa, yakni di atas rata-rata kelasnya. Rata-rata kelas tidak berarti standar mutlak sebab ditentukan oleh prestasi semua anak. Oleh sebab itu, dianjurkan agar kedua kriteria penilaian di atas dilaksanakan secara seimbang.

g. *Menetapkan frekuensi evaluasi.* Artinya berapa kalikah evaluasi dalam suatu periode tertentu. Hal ini tergantung kepada tujuan yang hendak dicapai.

2. *Langkah Pengumpulan Data*

Dalam melaksanakan langkah pengumpulan data dapat dilakukan langkah-langkah sebagai berikut : pelaksanaan evaluasi, memeriksa hasil-hasil evaluasi, dan memberi kode atau skor.

Data yang kita peroleh dalam pengumpulan data masih merupakan data mentah yang belum dapat memberikan gambaran yang jelas kepada kita. Agar kita dapat gambaran

yang jelas dari evaluasi yang kita laksanakan, maka kode atau skor yang diperoleh harus dianalisa lebih lanjut.

3. *Langkah Verifikasi Data*

Agar data yang telah diperoleh benar-benar objektif, maka perlu dilakukan verifikasi data. Apakah data sesuai dengan realitas, atau tidak. Apakah data yang diperoleh itu sudah diyakini kebenarannya, maka data tersebut masih diperkirakan mengandung kesalahan-kesalahan, maka perlu diadakan verifikasi data terlebih dahulu. Setelah verifikasi data dilanjutkan dengan seleksi data. Adapun langkah verifikasi data yang perlu diperhatikan adalah :

a. Tanda-Tanda Kesestatan

Untuk mengetahui ada tidaknya kesalahan pada sekumpulan data yang kita dapat berpedoman pada ada tidaknya tanda-tanda kesestatan tersebut. Adapun tanda-tanda kesestata sekumpulan data adalah (1) Data yang kita hadapi memberikan gambaran yang sangat berbeda dengan gambaran-gambaran yang telah kita peroleh berdasarkan hasil-hasil evaluasi yang telah kita lakukan pada masa-masa lampau, (2) Distribusi data yang kita peroleh jauh sekali menyimpangnya dari distribusi normal. Khusus yang memakai teknik analisis secara kuantitatif perlu diuji prasyarat data yang diperoleh, yang digunakan untuk menentukan teknik analisis apa yang dipakai. Apakah pengolahan data memakai statistik parametris atau non parametris.

b. Kemungkinan Sumber-Sumber Kesestatan

Setelah kita melihat adanya tanda-tanda kesestatan yang terdapat pada data yang kita hadapi, maka langkah selanjutnya ialah menetapkan cara-cara yang ditempuh untuk menemukan dimana sumber-sumber kesestatan itu. Menurut Mochtar Buchori mengatakan bahwa ada beberapa hal yang memungkinkan timbulnya kesalahan-kesalahan dalam

sekumpulan data yaitu: (1) Kesalahan-kesalahan yang mungkin ditimbulkan karena kurang sempurnanya alat-alat evaluasi, (2) Kesalahan-kesalahan yang mungkin ditimbulkan oleh kurang sempurnanya prosedur pelaksanaan evaluasi yang dilakukan, (3) Kesalahan-kesalahan yang mungkin ditimbulkan oleh kurang sempurnanya cara pencatatan hasil-hasil evaluasi.

c. Prosedur verifikasi

Dalam kegiatan proses verifikasi yang perlu diperhatikan adalah : (1) kita periksa item-item tes yang digunakan dalam pengumpulan data tentang hasil belajar yang kita hadapi, (2) kita tinjau kembali bagaimana evaluasi itu dilaksanakan, (3) kita periksa kembali pencatatan skor yang telah dilakukan. Mengadakan pemeriksaan pada ketiga sumber yang mungkin menimbulkan kesalahan, dan ternyata data yang kita peroleh memang sudah benar adanya, maka data tersebut langsung dapat kita olah.

4. *Langkah Analisa Data*

Teknik pengolahan data atau analisa data biasanya diklasifikasi menjadi dua macam yaitu pengolahan statistik (statistical analysis) dan pengolahan bukan secara statistik (non statistical analysis).

5. *Langkah Penafsiran Data*

Apabila kita melakukan pengolahan terhadap sekumpulan data yang kita peroleh dalam kegiatan evaluasi, maka setelah itu adalah penafsiran data yang telah kita olah tersebut.

Dengan demikian langkah pengolahan data tidak bisa kita pisahkan dari langkah pengolahan data. Sebagaimana yang dikemukakan oleh C. Stanley mengatakan bahwa : “For analysis is worthless without interpretation, and interpretation is impossible without analysis” (1964:319). Justru dengan melakukan penafsiran data seorang evaluator akan mampu memberikan kategori dalam penilaian tersebut berdasarkan

kriteria yang telah ditetapkan. Norma ini dapat ditetapkan atau disiapkan terlebih dahulu secara rasional sebelum suatu evaluasi dilaksanakan, tetapi dapat pula dibuat berdasarkan hasil-hasil yang diperoleh dalam pelaksanaan evaluasi. Apakah kategori sangat berhasil, berhasil, kurang berhasil, dan tidak berhasil.

Penafsiran yang dapat kita lakukan terhadap data yang telah kita olah, dapat kita bedakan atas dua jenis penafsiran yaitu : (1) penafsiran yang bersifat individual, dan (2) penafsiran yang bersifat klasik.

(1). Penafsiran individual dalam praktek evaluasi pendidikan pada umumnya kita bedakan tiga jenis penafsiran individual :

- a. Penafsiran tentang tingkat kesiapan adalah menafsirkan apakah seseorang anak sudah mencapai taraf kesiapan yang cukup memadai sehingga ia bisa dinaikkan ke dalam kelas yang berikutnya atau tetap tinggal di kelas yang sama?
- b. Penafsiran tentang kelemahan-kelemahan adalah penafsiran tentang pada sub-sub tes mana dari suatu mata pelajaran anak itu menunjukkan kelemahan?. Sehingga seorang pendidik akan dapat memberikan penguatan pembelajaran pada bagian pembahasan pada suatu mata pelajaran tertentu.
- c. Penafsiran tentang pertumbuhan adalah suatu penafsiran tentang kemajuan yang telah dicapai oleh anak dalam suatu periode pendidikan. Kemajuan apakah yang telah dicapai oleh seorang dalam mata pelajaran tertentu?

(2). Penafsiran klasikal yang dimaksudkan disini ialah penafsiran terhadap kelas secara keseluruhan tentang hasil

yang mereka capai dalam tes yang kita berikan. Penafsiran klasikal ini dapat kita bedakan atas :

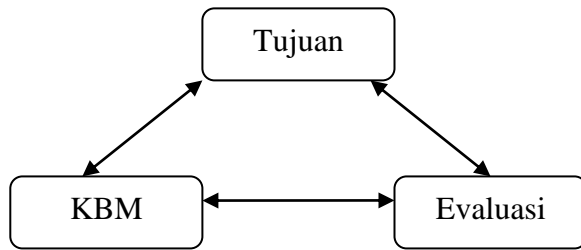
- a. Penafsiran tentang kelemahan-kelemahan kelas bertujuan untuk menafsirkan pada bagian-bagaimanakah dan seluruh bagian-bagaimanakah pembahasan dari suatu mata pelajaran anak menunjukkan kelemahan dan
- b. Penafsiran tentang prestasi kelas bertujuan untuk mengetahui prestasi anak pada kelas tertentu, bagaimanakah prestasi anak secara klasikal terhadap evaluasi yang kita berikan?. Untuk mengetahui perbedaan prestasi antara siswa satu dengan siswa lainnya dapat dilihat pada nilai indeks prestasi kelas
- c. Penafsiran tentang perbandingan antar kelas bertujuan untuk mengetahui kelas mana yang mencapai hasil yang terbaik pada tes yang kita adakan, dengan menggunakan indeks prestasi kelas.
- d. Penafsiran tentang susunan kelas bertujuan untuk mengetahui apakah kelas yang dihadapi merupakan kelas yang heterogen, normal atau homogen.

B. PRINSIP –PRINSIP DASAR EVALUASI PENDIDIKAN

Seorang evaluator dalam melakukan kegiatan evaluasi pendidikan hendaknya memahami satu prinsip umum dan penting dalam kegiatan evaluasi, yaitu adanya triangulasi atau hubungan erat tiga komponen yaitu:

1. Tujuan pembelajaran.
2. Kegiatan pembelajaran atau KBM
3. Evaluasi.

Triangulasi tersebut dapat digambarkan dalam bagan sebagai berikut :



Gambar 1. Trianggulasi dalam Evaluasi Pendidikan

(Suharsimi Arikunto, 2005: 20)

Dengan melihat bagan diatas terlihat adanya hubungan yang saling ketergantungan satu komponen dengan komponen lainnya (interpededensi). Kegiatan belajar-mengajar yang dirancang dalam bentuk rencana mengajar disusun oleh guru dengan mengacu pada tujuan yang hendak dicapai. Hubungan antara tujuan dengan evaluasi adalah kegiatan pengumpulan data untuk mengukur sejauhmana tujuan sudah tercapai.

Seorang evaluator dalam melaksanakan evaluasi pendidikan hendaknya memperhatikan berbagai macam prinsip dasar evaluasi pendidikan sebagaimana dikemukakan oleh Sudjono (1996: 31-33) sebagai berikut : (1) Prinsip Keseluruhan, (2) Prinsip Kesenambungan, dan (3) Prinsip Obyektivitas.

1. Prinsip Keseluruhan (al-Kamal= الكمال)

Prinsip keseluruhan atau prinsip menyeluruh juga dikenal dengan istilah prinsip komprehensif (*comprehensive*). Dengan prinsip komprehensif dimaksudkan di sini bahwa evaluasi hasil belajar dapat dikatakan terlaksana dengan baik apabila evaluasi tersebut dilaksanakan secara bulat, utuh atau menyeluruh tidak parsial. Dengan kata lain, evaluasi hasil

belajar harus dapat mencakup berbagai aspek yang dapat menggambarkan perkembangan atau perubahan tingkah laku yang terjadi pada peserta didik. Adapun aspek yang perlu diungkap adalah aspek proses berpikir (*cognitive domain*), aspek kejiwaan (*affective domain*) dan aspek keterampilan (*psychomotor domain*) yang melekat pada diri masing-masing individu peserta didik. Jika dikaitkan dengan proses pembelajaran Pendidikan Bahasa Arab (PBA) dan Pendidikan Agama Islam, maka evaluasi hasil belajar dalam mata pelajaran Pendidikan Bahasa Arab dan Pendidikan Agama Islam itu hendaknya bukan hanya mengungkap pemahaman peserta didik terhadap penguasaan kemampuan berbahasa arab dan ajaran-ajaran agama Islam, melainkan juga harus dapat mengungkap sudah sejauh mana peserta didik dapat menghayati dan mengamalkan kemampuan berbahasa arab dan ajaran-ajaran Islam tersebut dalam kehidupan mereka sehari-hari. Pada dasarnya tujuan dilakukannya evaluasi hasil belajar secara komprehensif adalah untuk memperoleh bahan-bahan keterangan dan informasi yang lengkap mengenai keadaan dan perkembangan subyek didik yang sedang dijadikan sasaran evaluasi.

2. Prinsip Kesenambungan (Istimrar = استمرار)

Untuk meningkatkan kualitas dalam proses pembelajaran. Seorang evaluator harus melaksanakan prinsip kesinambungan juga dikenal dengan istilah prinsip kontinuitas (*continuity*). Dengan prinsip kesinambungan dimaksudkan di sini bahwa evaluasi hasil belajar yang baik adalah evaluasi hasil belajar yang dilaksanakan secara teratur dan sambung-menyambung dari waktu ke waktu. Dengan demikian pelaksanaan evaluasi hasil belajar yang telah dilaksanakan secara teratur, terencana dan terjadwal. Sehingga dapat diperoleh informasi yang dapat memberikan gambaran mengenai kemajuan atau perkembangan peserta

didik, sejak dari awal mula mengikuti program pendidikan sampai pada saat-saat mereka mengakhiri program pendidikan yang mereka tempuh itu. Apakah telah mencapai tujuan yang telah ditetapkan atau belum, jika belum peneliti dapat memberikan perbaikan dalam kegiatan program pendidikan dan sebaliknya jika sudah mencapai tujuan pembelajaran dan program pendidikan yang telah ditetapkan, maka tugas civitas akademika adalah mempertahankan dan meningkatkan kualitas yang jauh lebih baik dari apa yang telah dicapai sekarang.

3. Prinsip Obyektivitas (Maudluiyyah = موضوعية)

Dalam pengertian sehari-hari prinsip obyektivitas (*objectivity*) telah dengan cepat dapat diketahui bahwa bersikap objektif berarti tidak adanya unsur pribadi bersifat subyektif yang mempengaruhi dalam kegiatan evaluasi pendidikan. Sehubungan dengan itu, dalam pelaksanaan evaluasi hasil belajar, seorang evaluator harus senantiasa berpikir dan bertindak wajar, menurut realitas yang ada, tidak dicampuri oleh kepentingan-kepentingan yang bersifat subyektif yang dapat menodai dalam kegiatan evaluasi pendidikan.

C. PEMILIHAN ATAU PENYUSUNAN TES HASIL BELAJAR

Apabila untuk suatu tindakan evaluasi tertentu telah tersedia alat-alatnya, maka kita tinggal memilih saja alat-alat evaluasi yang sesuai dengan apa yang akan ingin kita evaluasi, baik tes maupun nontes. Tetapi jika alat-alat evaluasi yang akan kita pergunakan belum tersedia, maka kita harus menyusun atau membuat sendiri alat-alat yang akan dipergunakan.

Dalam penyusunan tes hasil belajar ada beberapa langkah yang ditempuh seperti yang akan dibicarakan di bawah ini :

(1). Menyusun *lay out*

Suatu tes hasil belajar baru dapat dikatakan tes baik, jika materi yang tercantum dalam item-item tes tersebut merupakan pilihan yang cukup representatif terhadap materi pelajaran yang diberikan di kelas yang bersangkutan. Dengan analisa rasional dalam “blue print” atau “layout” tentang pokok-pokok yang akan kita kemukakan dalam tes. Dalam *lay out* ini kita cantumkan beberapa hal yang penting yaitu :

- a. Menentukan ruang lingkup (skope) dari pengetahuan yang akan diukur sesuai dengan rencana pelajaran yang telah ditetapkan dalam kurikulum atau dalam program evaluasi.
- b. Menentukan nilai proporsi jumlah item daripada tiap-tiap sub materi. Proporsi jumlah item untuk tiap-tiap sub materi hendaknya sesuai dengan proporsi daripada luas masing-masing sub materi. Misalnya apabila dalam suatu materi terdiri dari tiga sub dengan proporsi 15 % : 40 % : 45 %, maka jumlah itemnya pun harus mengikuti proporsi 15 % : 40 % : 45 %.
- c. Menetapkan jenis pengetahuan atau aspek proses mental yang hendak diukur. Misalnya pengetahuan faktual, pengertian dan aplikasi.
- d. Memilih bentuk tipe tes yang akan dipergunakan lebih dari satu bentuk/tipe test. Misalnya : *multiple-choice*, *matching/essay* dengan *completion*.

Untuk memperjelas bentuk *lay out* dapat dilihat pada tabel di bawah ini :

Tabel 1. Bentuk *Lay Out* dalam Tes

Jenis Materi	Proporsi	Jenis Pengetahuan	Bentuk Soal			Σitem
			T. F.	M.C	Mch	
Amanah	15 %	<ul style="list-style-type: none"> • Pengetahuan Faktual • Pengertian • Aplikasi 	1	2	1	4
			1	1	1	3
			1	-	-	1
Tasamuh	40 %	<ul style="list-style-type: none"> • Pengetahuan Faktual • Pengertian • Aplikasi 	4	3	3	10
			5	4	4	13
			5	2	3	10
Adil	45 %	<ul style="list-style-type: none"> • Pengetahuan Faktual • Pengertian • Aplikasi 	5	4	4	13
			6	4	4	14
			5	3	3	11
Jumlah	100%	-	33	23	23	79

- (2). Menulis soal yaitu menuliskan pertanyaan-pertanyaan yang biasanya dilakukan oleh seorang guru setelah persiapan mengajar selesai disusun.
- (3). Menata soal yaitu dalam pengaturan ini kita kelompok-kelompokkan soal-soal itu menurut bentuknya, tingkat kesukaran. Jadi bukan menurut jenis materinya dan bukan pula menurut pola atas jenis pengetahuan yang hendak diukur. Dengan demikian ada kelompok soal multiple-choice, ada kelompok soal *matching* dan sebagainya. Tingkat kesukaran dari yang termudah sampai tersulit atau ringan sedang, berat, sedang, ringan.
- (4). Menetapkan skor yaitu setelah pengaturan soal-soal selesai kita lakukan, langkah selanjutnya adalah menetapkan

besarnya skor yang diberikan untuk setiap item. Artinya kita tetapkan beberapa skor yang akan diberikan untuk setiap jawaban yang diberikan oleh anak-anak. Cara menskor yang banyak dilakukan adalah memberikan skor 1 (satu) untuk setiap jawaban yang betul. Pada intinya pemberian skor bergantung pada penting atau tidaknya suatu pertanyaan disebut pemberian skor atas dasar bobot (*weight score*) (Witherington, 1967:68).

- (5). Reproduksi tes yaitu proses penggandaan soal baik dalam bentuk offset, risho, stensil, foto copy dan lain-lain.
- (6). Analisa empiris terhadap suatu tes hasil belajar yaitu analisa yang dilakukan setelah tes selesai dilaksanakan untuk menjawab pertanyaan-pertanyaan tentang kualitas soal. Misalnya : Bagaimanakah kualitas daripada item-item yang kita gunakan? Apakah item-item tersebut sudah cukup atau belum? Dimana letak kelemahan soal tersebut? Item-item mana yang perlu diubah (direvisi) atau dibuan sama sekali? dan lain-lain.

A. JENIS EVALUASI PENDIDIKAN

Pembelajaran yang efektif menghendaki dipergunakannya alat-alat untuk menentukan apakah suatu hasil belajar yang diinginkan telah benar-benar tercapai, atau sampai dimanakah hasil belajar yang diinginkan tadi telah tercapai sesuai dengan tujuan pembelajaran yang telah ditetapkan. Apabila dilihat dari fungsinya, jenis penilaian ada beberapa macam, yaitu penilaian formatif; penilaian sumatif, penilaian diagnostik, penilaian selektif, dan penilaian penempatan.

1. *Penilaian formatif* adalah penilaian yang dilaksanakan pada akhir program belajar mengajar untuk melihat keberhasilan pembelajaran itu sendiri. Penilaian formatif diharapkan guru dapat memperbaiki program pengajaran dan strategi pelaksanaannya.
2. *Penilaian sumatif* adalah penilaian yang dilaksanakan pada akhir unit program, yaitu akhir catur wulan, semester atau akhir tahun. Tujuannya adalah untuk melihat hasil yang dicapai oleh para siswa, yakni seberapa jauh tujuan-tujuan kurikuler dikuasai oleh para siswa. Penilaian ini berorientasi pada produk bukan pada proses sebagaimana penilaian formatif.
3. *Penilaian diagnostik* adalah penilaian yang bertujuan untuk melihat kelemahan-kelemahan siswa serta faktor penyebabnya. Dengan tes ini diharapkan seorang pendidik mengetahui tentang kesulitan belajar yang dihadapi oleh para siswa. Penilaian ini diharapkan dapat mengetahui

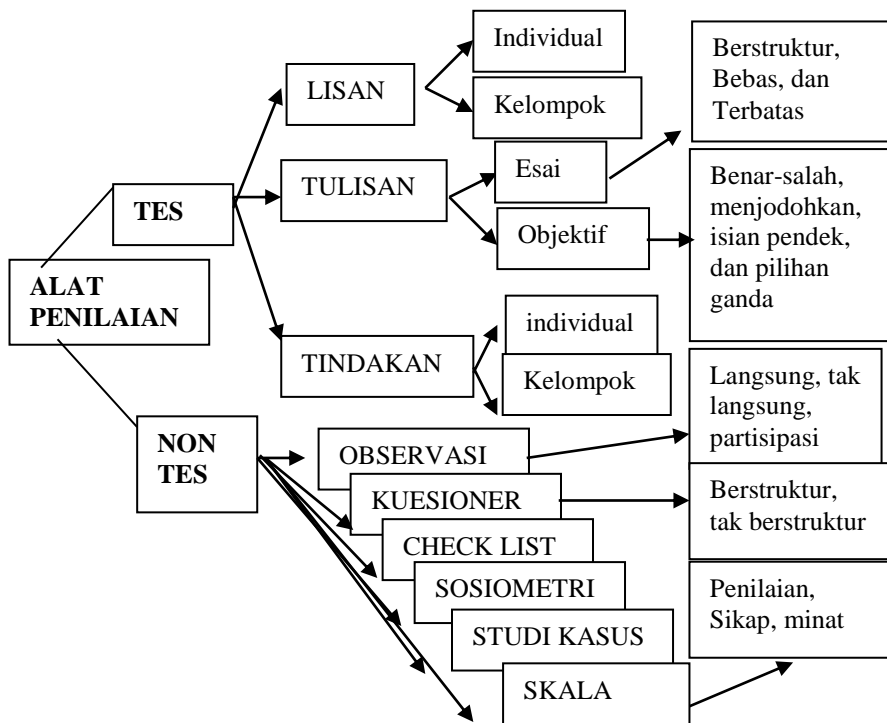
berbagai macam kasus kesulitan belajar dan memberikan bimbingan dan remediasi bagi peserta didik.

4. *Penilaian selektif* yaitu penilai yang bertujuan untuk keperluan seleksi, misalnya ujian saringan masuk ke lembaga pendidikan tertentu.
5. *Penilaian penempatan* adalah penilaian yang ditujukan untuk mengetahui ketrampilan prasyarat yang dipergunakan bagi suatu program belajar dan penguasaan belajar seperti yang diprogramkan sebelum memulai kegiatan belajar untuk program itu.

Ada beberapa alat yang dapat dipergunakan untuk mengetahui kemajuan-kemajuan peserta didik dalam proses kegiatan belajar mengajar. Dilihat dari segi alatnya, penilaian hasil belajar dapat dibedakan menjadi teknik yang dipergunakan adalah *teknik test* dan *non test*. Tes ini ada yang berupa tes tulisan, tes lisan, dan tes tindakan. Untuk mempermudah pembagian tersebut dapat dilihat pada gambar dibawah ini. Sedangkan alat penilaian non test dapat berupa observasi, kuesioner, wawancara, skala, sosiometri, studi kasus dan lain-lain.

Pembuatan alat tes pada dasarnya sudah dibakukan dan tidak dibakukan. Pada umumnya pendidik menggunakan tes buatan oleh guru bidang studi. Hal ini dikarenakan keterbatasan waktu untuk menguji dengan beberapa kali percobaan dan analisis dari segi validitas dan reliabilitasnya.

Alat penilaian hasil belajar mengajar ada yang sifatnya *speed tests* dan *power test*. Tes objektif sebenarnya termasuk dalam kategori *speed tests* (tes yang mengutamakan kecepatan). *Tes essay* (uraian) termasuk dalam kategori *power test* (tes yang mengutamakan kekuatan). Dilihat dari objek yang dinilai atau penyajiannya ada tes yang bersifat individual dan tes yang bersifat kelompok. Untuk mempermudah jenis-jenis tes sebagai alat penilai, termasuk bukan tes (nontes) dapat dilihat pada gambar dibawah ini.



B. SISTEM EVALUASI PENDIDIKAN

Sistem penilaian adalah cara yang digunakan dalam menentukan derajat keberhasilan hasil penilaian sehingga kedudukan siswa dapat diketahui, apakah telah menguasai tujuan instruksional ataukah belum.

Dalam penilaian hasil dan proses belajar dapat digunakan beberapa cara. Cara pertama menggunakan sistem huruf, yakni A, B, C, D, dan G (gagal). Biasanya ukuran yang digunakan adalah A paling tinggi, paling baik, atau sempurna; B baik; C sedang atau cukup; dan D kurang. Cara kedua ialah dengan sistem angka yang menggunakan beberapa standar. Dalam standar empat, angka 4 setara dengan A, angka 3

setara dengan B, angka 2 setara dengan C, dan angka 1 setara dengan D. Ada juga standar sepuluh, yakni menggunakan rentangan angka dari 1-10. Bahkan ada juga yang menggunakan rentangan 1-100. Cara mana yang dipakai tidak jadi masalah asal konsisten dalam pemakaiannya.

Ada dua macam sistem penilaian hasil belajar pada umumnya dibedakan ke dalam dua cara atau dua sistem, yakni penilaian acuan norma (PAN) dan penilaian acuan patokan (PAP).

1. Penilaian Acuan Norma (PAN) atau *Norm Referenced Tes*.

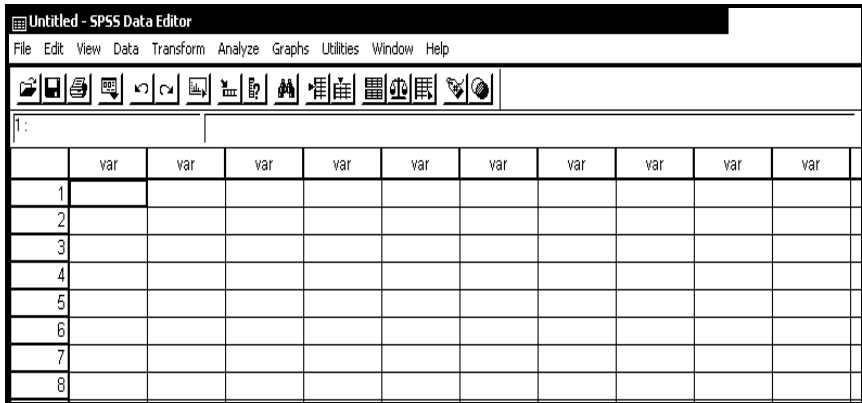
PAN digunakan apabila penilaian hasil belajar siswa ditujukan untuk mengetahui kedudukan siswa dalam kelompoknya. Artinya hasil tes tersebut lebih banyak ditujukan untuk memperoleh gambaran mengenai kedudukan siswa di dalam kelas atau kelompoknya. Hal ini berarti, makna dari angka (nilai) yang diperoleh seseorang siswa, diketemukan dengan cara membandingkan hasil belajarnya dengan hasil belajar siswa lain dalam satu kelas. Dengan cara ini, maka dapat diketahui kedudukan relatif seseorang di dalam kelas. Untuk memperjelas dapat dilihat contoh di bawah ini.

Nama Siswa	Nilai Bahasa Arab	Nilai Nahwu Shorof
Najma	9	9
Luna	[7]	[6]
Khilmi	[6]	[6]
Iman	[6]	[9]
Fuad	[7]	[6]
Isak	8	7
Hafid	[7]	[6]
Imam	[7]	[6]
Junaidi	8	7
Total	65	62

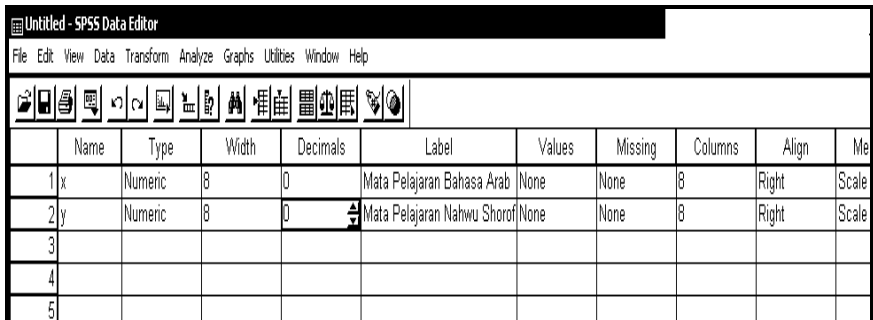
Untuk mempermudah proses penghitungan mean dan standar deviasi, maka dapat digunakan olah data SPSS dengan langkah-langkah sebagai berikut :

Adapun langkah-langkah pengujian dengan menggunakan SPSS adalah :

1. Bukalah program SPSS for Windows



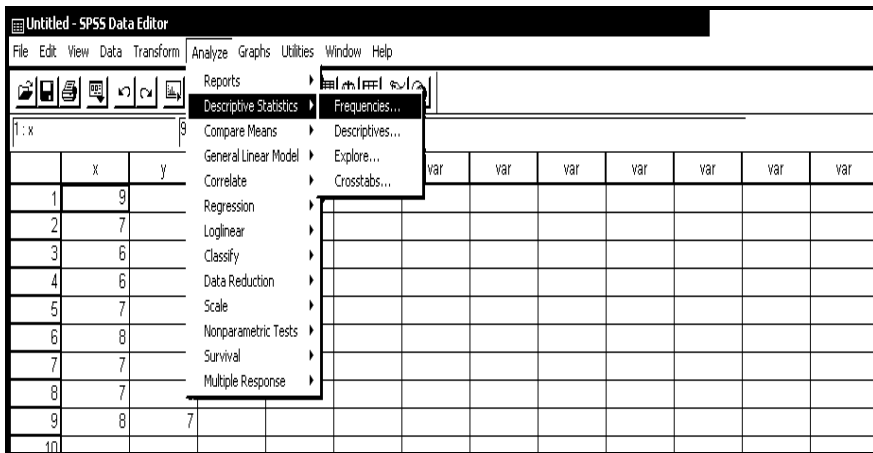
2. Klik variable view dan definisikan variabel (maksimal 8 karakter) dan desimal ketik 0.



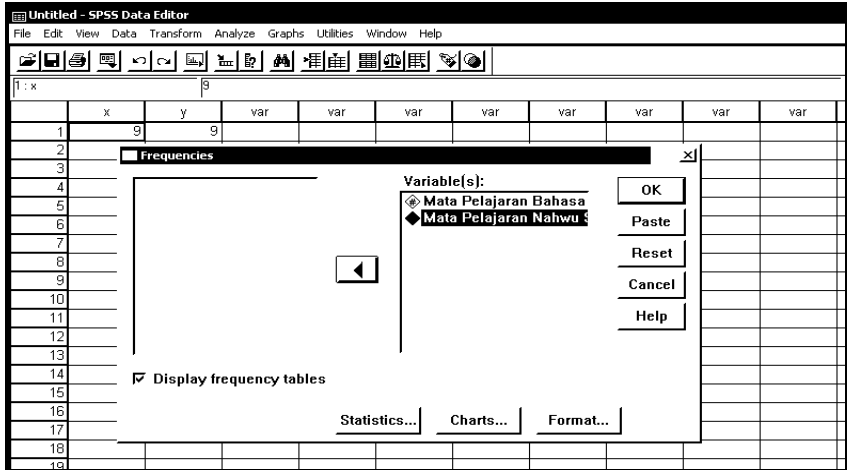
3. Klik data view dan lakukan pengisian data yang akan diuji

	x	y	var	var	var	var	var	var	var	var	var	var
1	9	9										
2	7	6										
3	6	6										
4	6	9										
5	7	6										
6	8	7										
7	7	6										
8	7	6										
9	8	7										
10												

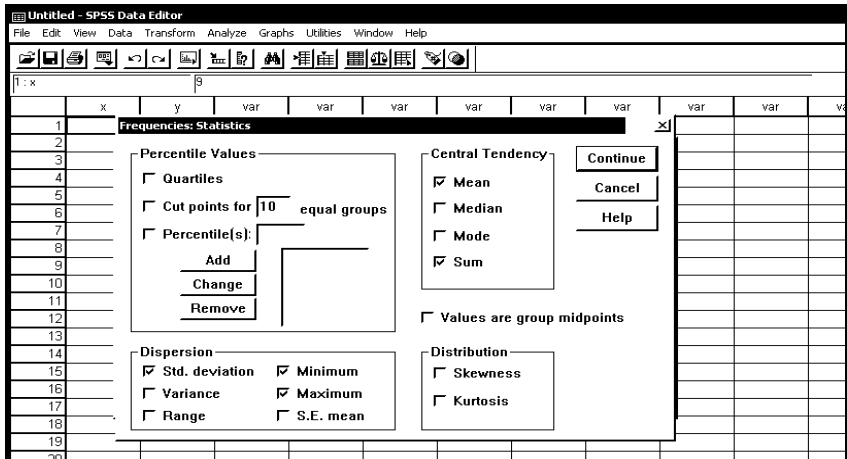
4. Klik Analyze kemudian pilih Descriptive Statistics dan pilih submenu Frequencies



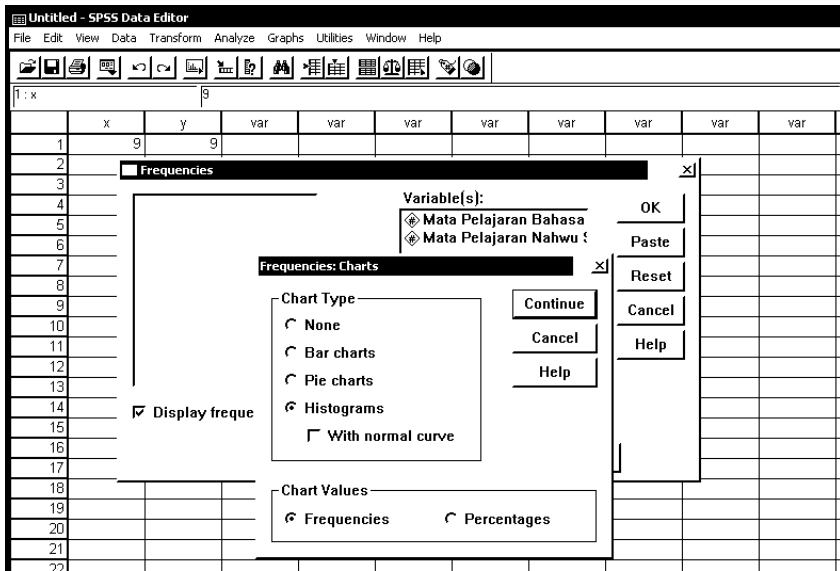
5. Masukkan variabel Mata Pelajaran Bahasa Arab (X) dan Mata Pelajaran Nahwu Shorof (Y) pada kotak Variables



6. Klik tombol Statistics sehingga tampak dilayar windows Frequencies Statistics pada Central Tendency klik mean dan sum, pada Dispersion klik Std. Deviation, Maximum dan Minimum. Lalu Klik Continue.



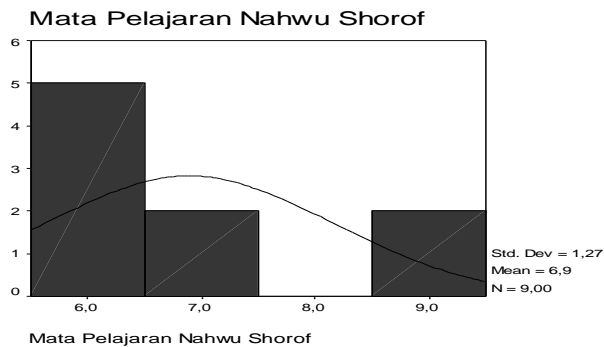
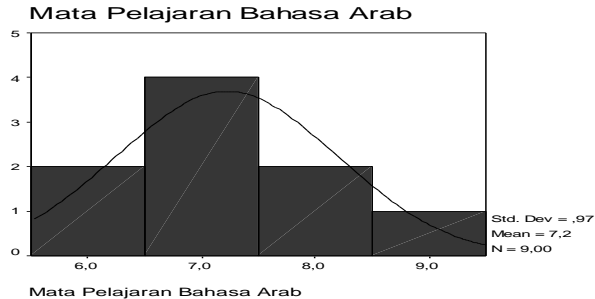
- Klik Charts pilih histograms dan klik with normal curve, klik continue.



- Klik OK. Hasil Oleh Data SPSS

Statistics

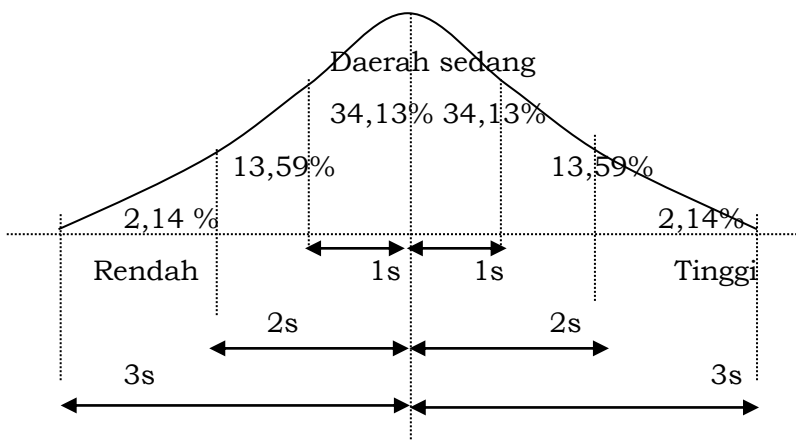
		Mata Pelajaran Bahasa Arab	Mata Pelajaran Nahwu Shorof
N	Valid	9	9
	Missing	0	0
Mean		7,22	6,89
Std. Deviation		,97	1,27
Minimum		6	6
Maximum		9	9
Sum		65	62



Dari olah data SPSS di atas diperoleh nilai rata-rata kelas untuk bidang studi Nilai Bahasa Arab adalah $65/9 = 7,22$. Sedangkan untuk Nahwu Shorof Arab $62/9 = 6,89$. Nilai 7 dan 6 dalam **Mata Pelajaran Bahasa Arab** yang diperoleh Luna, Khilmi, Iman, Fuad, Hafid dan Imam termasuk dalam kategori rendah, karena masih dibawah nilai rata-rata kelas yaitu 7,22. **Nilai Nahwu Shorof** 6 yang diperoleh Luna, Khilmi, Fuad, Hafid, dan Imam termasuk dalam kategori rendah. Namun apabila kita cermati nilai Iman pada Bahasa Arab termasuk dalam kategori rendah yaitu 6 masih dibawah rata-rata 7,22 namun dalam bidang studi Nahwu Shorof Iman termasuk dalam kategori tinggi (diatas rata-rata kelas). Kesimpulannya adalah Iman di

kelasnya termasuk pandai dalam bidang Nahwu Shorof, sedangkan dalam bidang Bahasa Arab ia termasuk dalam kategori rendah.

Dengan melihat deskripsi diatas, maka dapat disimpulkan bahwa ***tujuan dari penilaian acuan norma ialah untuk membedakan siswa atas kelompok-kelompok tingkat kemampuan mulai dari yang terendah sampai pada yang tertinggi (rangking siswa).*** Hasil pembelajaran ideal yang harus diperoleh oleh siswa dapat dilihat pada gambar dibawah ini.



Gambar. **Prosentase Hasil Pembelajaran Yang Ideal Luas Kurve Normal**

Dalam distribusi tersebut dapat dijelaskan bahwa terdapat 9 siswa yang ikut ujian bahasa arab dan nahwu shorof. Nilai rata-rata adalah ujian bahasa arab= 7,2 simpangan baku (standar deviasi) = 0,9718 dan nahwu shorof nilai rata-rata = 6,89 dan standar deviasi =1,2693. Berapa orang yang mendapat nilai 8 ke atas ?

Jawab 1:

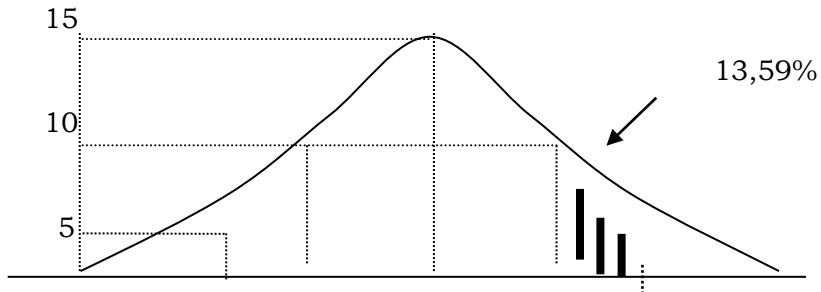
Untuk bahasa arab: Rata-rata kelas (\bar{X}) = 7,2, dan simpangan baku (s) = 0,9718. Dari rumus z dapat dihitung harga z.

$$Z = \frac{(X_i - \bar{X})}{s}$$

$$Z = \frac{(8 - 7,2)}{0,9718}$$

$$Z = 0,8/0,9718 = \underline{0,823}$$

Curve normanya dapat dilihat pada gambar 6.9 di bawah ini.



Gambar 6.9.

Jumlah Mahasiswa yang Mendapat Nilai 8 ke Atas

Dari tabel kurve normal diatas dapat dilihat bahwa daerah 0 sampai dengan 1, luasnya = 34,13%. Ini adalah antara mean (rata-rata) dengan suatu titik yang jauhnya 1 SD di atas mean. Harga ini menunjukkan prosentase jumlah siswa yang mendapat nilai antara 7,2 s/d 8. Dengan demikian prosentase yang mendapat nilai 8 ke atas adalah 50% - 34,13% = 15,87% (50% adalah setengah kurve di atas mean, di mana nilai 8 ke atas berada). Jadi mahasiswa yang mendapat nilai 8 ke atas = 15,87% x 9 = 1,4283 orang atau sekitar 1 atau 2 orang dari 9 jumlah siswa).

Jawab 2:

Untuk bahasa arab: Rata-rata kelas (\bar{Y}) = 6,89, dan simpangan baku (s) = 1,2693. Dari rumus z dapat dihitung harga z.

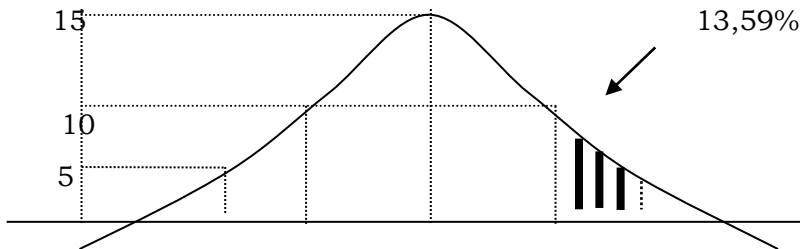
$$Z = \frac{(X_i - M_e)}{S}$$

$$Z = \frac{(8 - 6,89)}{1,2693}$$

$$Z = 1,11/1,2693 = \underline{0,875}$$

Dari tabel kurve normal dapat dilihat bahwa daerah 0 sampai dengan 1, luasnya = 34,13%. Ini adalah antara mean (rata-rata) dengan suatu titik yang jauhnya 1 SD di atas mean. Harga ini menunjukkan prosentase jumlah siswa yang

mendapat nilai antara 6,89 s/d 8. Dengan demikian prosentase yang mendapat nilai 8 ke atas adalah $50\% - 34,13\% = 15,87\%$ (50% adalah setengah kurve di atas mean, di mana nilai 8 ke atas berada). Jadi mahasiswa yang mendapat nilai 8 ke atas = $15,87\% \times 9 = 1,4283$ orang atau sekitar 1 atau 2 orang dari 9 jumlah siswa). Lihat gambar 6.10 di bawah ini.



Gambar 6.9.
Jumlah Mahasiswa yang Mendapat Nilai 8 ke Atas

Sistem penilaian PAN (norm referenced) cocok untuk keperluan seleksi, untuk penempatan siswa, dan untuk tes sumatif. Oleh karena itu bahan tes pada sistem penilaian PAN harus dikembangkan berdasarkan dari bahan yang telah disampaikan kepada siswa. Guru mempunyai kewenangan untuk menentukan bahan mana yang dianggap penting untuk diujikan dengan jumlah soal yang terbatas.

2. Penilaian Acuan Patokan (PAP) atau *Criterion Referenced*

PAP adalah penilaian lebih ditujukan kepada program (penguasaan bahan pelajaran), bukan kedudukan siswa dalam kelas. Oleh sebab itu PAP (*criterion referenced*),

berusaha mengukur tingkat pencapaian tujuan oleh para siswa. Prestasi belajar siswa yang tidak mencapai tujuan yang telah ditetapkan berarti gagal, atau pengajaran yang diberikan oleh seorang guru belum berhasil.

Melihat kasus diatas kelas **Najma** mempunyai nilai rata-rata mata pelajaran bahasa arab sebesar 7,2 dan mata pelajaran nahwu shorof sebesar 6,89. Artinya tingkat keberhasilan kasar kelas Najma pada mata pelajaran bahasa arab sebesar 72% dan mata pelajaran bahwu shorof sebesar 68,9%. Namun jika diukur dengan cara belajar tuntas (mastery learning), hasil yang dicapai dalam bidang studi IPS dan IPA belum mencapai tingkat penguasaan tujuan yang diharapkan, sebab masih kurang dari 80%. Artinya pengajaran mata pelajaran bahasa arab dan nahwu shorof termasuk belum berhasil.

Bagaimana dengan nilai yang didapat oleh Luna pada mata pelajaran bahasa arab mendapat nilai 7 dan mata pelajaran nahwu shorof sebesar 6 ?. Nlai yang didapat Luna dapat dikatakan bahwa nilai Luna baru mencapai 70% dan 60% bila dibandingkan target yang ingin dicapai sebesar 80%, maka nilai Luna belum mencapai target dalam pembelajaran mata pelajaran bahasa arab dan nahwu shorof.

PAP merupakan penilaian ditujukan pada sudah atau belumnya siswa atau kelas mencapai tujuan yang telah ditetapkan. PAP lebih mengutamakan kemampuan siswa dalam menguasai keseluruhan kompetensi yang diinginkan dalam program pembelajaran.

Bagaimana sikap guru jika secara keseluruhan siswa belum mencapai target sasaran 80%? Jika seorang guru mendapatkan kasus yang demikian, maka seorang guru harus memberikan re-teaching (memberikan pembelajaran ulang). Dengan sistem penilaian inilah, maka timbulah konsep belajar tuntas (*mastery learning*).

**PERBEDAAN NORM REFERENCED
DENGAN CRITERION REFERENCED**

No.	Norm Referenced (PAN)	Criterion Referenced (PAP)
1	Befungsi untuk menetapkan kedudukan relative seorang siswa didalam kelasnya	Befungsi menetapkan apakah siswa tela mencapai atau menguasai tujuan atau kemampuan yang ditargetkan
2	Tujuan pengajaran dinyatakan secara umum atau secara khusus	Tujuan pengajaran harus dinyatakan secara khusus (TIK)
3	Belajar tuntas (mastery learning) tidak begitu diutamakan	Sangat mengutamakan adanya belajar tuntas, sehingga perlu dinyatakan standar tingkat keberhasilan tujuan pengajaran
4	Tes (pertanyaan) harus mencakup tingkat kesukaran yang bervariasi dari yang mudah, sedang, sukar	Penyusunan soal lebih mengutamakan pada performance dan kemampuan yang harus dikuasai siswa
5	Skor diolah dengan menggunakan statistika seperti mean, standar deviasi dan lain-lain	Tidak selalu skor harus diolah dengan menggunakan prosedur statistika
6	Tepat dipakai untuk penempatan dan tes sumatif	Tepat dipakai untuk tes diagnostic dan tes formatif
7	Hasil penilaian tepat ditransformasi dala skala angka seperti A, B,C,D, atau dalam standar 0-10	Hasil penilaian tepat dinyatakan dalam bentuk pernyataan: sangat memuaskan, memuaskan, cukup, kurang, dan gagal

C. CARA PENILAIAN

Dalam bagian ini dibahas dua cara penilaian, yaitu (a) Penilaian Kuantitatif, dan (b) Penilaian Kualitatif.

Penilaian kualitatif pada umumnya lebih subyektif daripada penilaian kuantitatif. Penilaian kuantitatif biasanya dinyatakan dengan angka, sedangkan penilaian kualitatif dinyatakan dengan ungkapan, seperti "baik", "memuaskan", "kurang memadai", "kurang sempurna", dan sebagainya.

Dalam penilaian, kedua cara ini harus dilakukan secara seimbang. Ada aspek-aspek individu siswa yang perlu diungkapkan secara kualitatif, di samping secara kuantitatif. Bagaimana mengkombinasikan kedua cara penilaian itu, baik waktu penilaiannya, maupun dalam memperpadukan hasil kedua cara penilaian itu. Kemudian teknik-teknik manakah yang tepat digunakan untuk penilaian yang kuantitatif dan mana pula yang kualitatif, akan diuraikan lebih lanjut pada bagian berikut.

Aspek-aspek tingkah laku siswa dalam bidang kognitif pada umumnya dinilai secara kuantitatif. Aspek-aspek afektif pada umumnya dinilai secara kualitatif, tetapi aspek ketrampilan/psikomotor dapat secara seimbang dinilai secara kuantitatif dan kualitatif. Pada bab ini akan dijelaskan cara penilaian kuantitatif.

Apabila guru akan memperpadukan hasil penilaian kuantitatif dan kualitatif, misalnya untuk menemukan nilai akhir dari suatu bidang pengajaran untuk segala aspek tingkah laku siswa, maka guru harus mengubah salah satu dari kedua cara itu, atau mengubah hasil penilaian kuantitatif ke dalam ungkapan kualitatif, atau sebaliknya mengubah hasil penilaian kualitatif menjadi nilai dalam bentuk angka. Kemudian dijumlah dan dirata-ratakan, dan apabila perlu mempergunakan bobot tertentu sesuai dengan keperluannya. Setelah mendapatkan nilai rata-rata, rata-rata itu dirubah kembali sesuai dengan kebiasaan penilaian akhir di sekolah

yang bersangkutan, dalam bentuk ungkapan atau dalam bentuk angka.

Bagaimanakah kedua cara penilaian itu dilakukan? Penilaian kuantitatif biasanya dilakukan apabila guru ingin memberikan nilai akhir terhadap hasil belajar siswanya. Sedangkan penilaian kualitatif dilakukan apabila guru ingin memperbaiki hasil belajar siswanya. Tetapi perbedaan tersebut jangan dijadikan patokan. Kedua cara penilaian itu dapat saja dilakukan pada kedua kesempatan tadi, sesuai dengan kebutuhannya.

Kedua cara penilaian tersebut membutuhkan teknik pelaksanaannya. Teknik penilaian itu dapat dibagi menjadi dua golongan besar, ialah (a) teknik tes, dan (b) teknik bukan tes atau teknik *non tes*.

a. *Teknik Tes.*

Tes dapat dibedakan menurut materi yang akan dinilai, bentuknya, dan caranya membuat. Menurut materi yang dinilai dibedakan tes hasil belajar, tes cerdasan, tes bakat khusus, tes minat, dan tes kepribadian, tes minat, dll. Bentuk tes dapat dibedakan tes uraian, tes obyektif. Menurut caranya membuat dibedakan tes buatan guru, dan tes baku.

b. *Teknik Non Tes.*

Alat-alat khusus untuk melaksanakan teknik non tes ini dapat dilakukan melalui pengamatan, wawancara, angket, hasil karya/laporan, karangan dan skala sikap.

D. PEMBELAJARAN DAN PENILAIAN BERBASIS KTSP

Definisi pembelajaran dan penilaian berbasis KTSP merupakan suatu proses penerapan ide, konsep, dan kebijakan KTSP dalam suatu aktivitas pembelajaran, sehingga peserta didik menguasai seperangkat kompetensi tertentu, sebagai hasil interaksi dengan lingkungan. Implementasi KTSP juga dapat diartikan sebagai

aktualisasi kurikulum operasional dalam bentuk pembelajaran. Saylor (1981: 227) mengatakan bahwa *"Instruction is thus the implementation of curriculum plan, usually, but not necessarily, involving teaching in the sense of student, teacher interaction in an educational setting"*. Implimentasi perencanaan kurikulum yang ada, tidak harus melibatkan guru dalam menumbuhkan kesadaran peserta didik, guru hanya membuat seting pendidikan yang interaktif.

Pembelajaran berbasis KTSP sedikithya dipengaruhi oleh tiga faktor berikut:

- a. Karakteristik KTSP; yang mencakup ruang lingkup KTSP dan; yaitu strategi yang digunakan dalam pembelajaran, seperti diskusi, pengamatan dan tanya jawab, serta kegiatan lain yang dapat mendorong pembentukkan kompetensi peserta didik.
- b. Karakteristik pengguna kurikulum kejelasannya bagi pengguna di lapangan.
- c. Strategi pembelajaran, yang meliputi pengetahuan, keterampilan, nilai, dan sikap guru terhadap KTSP, serta kemampuannya untuk merealisasikan kurikulum (*curriculum planning*) dalam pembelajaran.

Lebih lanjut Mars (1980) mengemukakan tiga faktor yang mempengaruhi implementasi kurikulum, yaitu dukungan kepala sekolah; dukungan rekan sejawat guru; dan dukungan internal yang datang dari dalam diri guru sendiri. Dari berbagai faktor tersebut guru merupakan faktor penentu di samping faktbr-faktor lain. Dengan kata lain, keberhasilan implementasi KTSP sangat ditentukan oleh faktor guru, karerta bagaimanapun baiknya sarana pendidikan apabila guru tidak melaksanakan tugas dengan baik, maka hasil implementasi kurikulum (pembelajaran) tidak akan memuaskan.

Dalam garis besarnya implementasi KTSP mencakup empat kegiatan pokok, yaitu pengembangan strategi implementasi, pengembangan program, pelaksanaan pembelajaran dan evaluasi.

1. Prinsip-prinsip implementasi KTSP

Ada beberapa prinsip yang harus dijadikan pegangan dalam pelaksanaan KTSP. Adapun prinsip-prinsip tersebut adalah :

- a. Pelaksanaan kurikulum didasarkan pada potensi, perkembangan dan kondisi peserta didik untuk menguasai kompetensi yang berguna bagi dirinya.
- b. Kurikulum KTSP dilaksanakan dengan menegakkan kelima pilar belajar, yaitu: (a) belajar untuk beriman dan ber-takwa kepada Tuhan Yang Maha Esa, (b) belajar untuk memahami dan menghayati, (c) belajar untuk mampu melaksanakan dan berbuat seera efektif, (d) belajar untuk hidup bersama dan berguna bagi orang lain, dan (e) belajar untuk membangun dan menemukan jati diri, melalui proses pembelajaran yang efektif, aktif, kreatif, dan menyenangkan.
- c. Pelaksanaan kurikulum memungkinkan peserta didik mendapat pelayanan yang bersifat perbaikan, pengayaan sesuai dengan potensi yang ingin dikembangkan oleh institusi pendidikan sendiri.
- d. Kurikulum dilaksanakan dalam suasana hubungan peserta didik dan pendidik dengan saling menerima dan menghargai.
- e. Kurikulum dilaksanakan dengan menggunakan pendekatan multistrategi dan multimedia, sumber belajar dan teknologi yang memadai, dan memanfaatkan lingkungan sekitar sebagai sumber belajar.

- f. Kurikulum dilaksanakan dengan mendayagunakan kondisi alam, sosial dan budaya dalam menunjang keberhasilan pendidikan dengan muatan seluruh bahan kajian secara optimal.
- g. Kurikulum yang mencakup seluruh komponen-kompetensi mata pelajaran, muatan lokal dan pengembangan diri diselenggarakan dalam keseimbangan, keterikatan dan kesinambungan yang cocok dan memadai antarkelas dan jenis serta jenjang pendidikan.

2. Pengembangan Program

Untuk efektifnya sebuah pelaksanaan pembelajaran KTSP, maka diperlukan pengembangan pembelajaran KTSP yang dilakukan secara terus-menerus. Adapun pengembangan program KTSP mencakup : (1) pengembangan program tahunan, (2) program semester, (3) program modul (pokok bahasan), (4) program mingguan dan harian, (5) program pengayaan dan remedial, dan (6) program bimbingan dan konseling.

3. Penilaian Hasil Belajar

Sistem penilaian hasil belajar dalam KTSP sebagaimana dikatakan oleh Mulyasa (2006 :258-261) dapat dilakukan dengan penilaian kelas, tes kemampuan dasar, penilaian akhir satuan pendidikan dan sertifikasi, *benchmarking*, dan penilaian program. Adapun uraian masing-masing penilaian yaitu :

a. Penilaian Kelas

Penilaian ini dilakukan dalam bentuk ulangan harian, ulangan umum dan ujian akhir.

b. Tes Kemampuan Dasar

Tes kemampuan dasar dilakukan untuk mengetahui kemampuan membaca, menulis, dan berhitung yang diperlukan dalam rangka memperbaiki program pembelajaran.

c. Penilaian Akhir Satuan Pendidikan dan Sertifikasi

Pada setiap akhir semester dan tahunan pelajaran diselenggarakan kegiatan penilaian guna mendapatkan gambaran secara utuh dan menyeluruh mengenai ketuntasan belajar peserta didik dalam satuan waktu tertentu. Tujuan penilaian ini untuk sertifikasi, kinerja dan hasil belajar dalam Surat Tanda Tamat Belajar tidak semata-mata didasarkan atas hasil penilaian pada akhir jenjang sekolah.

d. Benchmarking

Benchmarking merupakan suatu standar untuk mengukur kinerja yang sedang berjalan, proses, dan hasil untuk mencapai suatu keunggulan yang memuaskan. Keunggulan ditentukan di tingkat sekolah daerah dan nasional. Sistem penilaian didasarkan pada prinsip berkesinambungan sehingga peserta didik dapat mencapai satuan satuan tahap keunggulan pembelajaran yang sesuai dengan kemampuan usaha dan keuletannya. Untuk mengetahui informasi ketercapaian *benchmarking* tertentu dapat diadakan penilaian secara nasional yang dilaksanakan pada akhir satuan pendidikan. Hasil penilaian tersebut digunakan untuk : (1) keberhasilan kurikulum dan pendidikan secara keseluruhan, (2) memberikan peringkat kelas tetapi tidak menilai akhir peserta didik, (3) dapat digunakan sebagai salah satu dasar untuk pembinaan guru dan kinerja sekolah.

e. Penilaian Program

Departemen Pendidikan Nasional dan Dinas Pendidikan secara kontinu dan berkesinambungan. Tujuan penilaian program yaitu untuk mengetahui kesesuaian KTSP dengan dasar, fungsi, dan tujuan pendidikan nasional, relevansi perkembangan masyarakat dan kemajuan zaman.

BAB VIII

ALAT – ALAT EVALUASI PENDIDIKAN

A. Pendahuluan

Pengertian secara umum alat adalah sesuatu yang dapat digunakan untuk mempermudah seseorang untuk melaksanakan tugas atau mencapai tujuan secara lebih efektif dan efisien. Kata "alat" biasa disebut juga dengan istilah "instrumen". Dengan demikian maka alat evaluasi juga dikenal dengan instrumen evaluasi.

Untuk memperjelas pengertian "alat" atau "instrumen", terapkan pada dua cara mengupas kelapa, yang satu menggunakan pisau parang, yang satu lagi tidak. Tentu saja hasilnya akan lebih baik dan pekerjaannya berakhir lebih cepat dibandingkan dengan cara yang pertama. Dalam kegiatan evaluasi, fungsi alat juga untuk memperoleh hasil yang lebih baik sesuai dengan kenyataan yang dievaluasi

Pada pelaksanaan evaluasi hasil belajar pengajaran bahasa arab dan agama, anda akan diperkenalkan dengan tiga bentuk evaluasi. Pertama, tes tertulis. kedua, tes lisan, dan ketiga, observasi.

Tidaklah mudah untuk menyusun pertanyaan dalam tes (alat evaluasi), sebab tes/pertanyaan merupakan alat untuk melihat perubahan kemampuan dan tingkah laku siswa setelah ia menerima pengajaran dari guru/pengajaran di sekolah. Alat yang salah dalam penyusunan alat evaluasi akan mengakibatkan hasil pengukuran pada kemampuan siswa yang tidak obyektif.

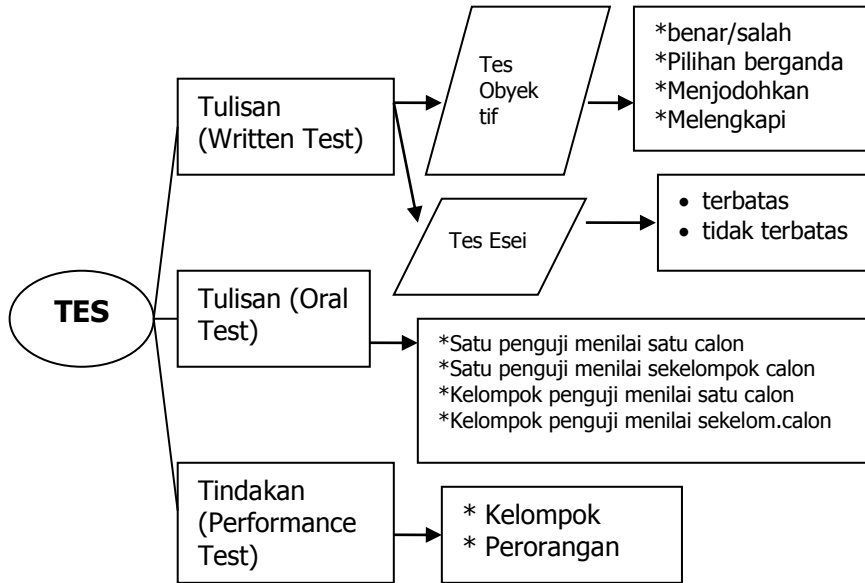
Ada beberapa syarat dan petunjuk yang dapat digunakan oleh evaluator dalam menyusun alat evaluasi sebagaimana dikatakan oleh Sudjana (2000:116-117) ialah :

1. Harus menetapkan dulu segi-segi apa yang akan dinilai, sehingga betul-betul terbatas serta dapat memberi petunjuk bagaimana dan dengan alat apa segi tersebut dapat kita nilai.
2. Harus menetapkan alat evaluasi yang betul-betul valid dan reliabel, artinya taraf ketepatan dan ketetapan tes sesuai dengan aspek yang dinilai.
3. Penilaian harus obyektif, artinya menilai prestasi siswa sebagaimana adanya.
4. Hasil penilaian tersebut harus betul-betul diolah dengan teliti sehingga dapat ditafsirkan berdasarkan kriteria yang berlaku.
5. Alat evaluasi yang dibuat hendaknya mengandung unsur-unsur diagnostis, artinya dapat dijadikan bahan untuk mencari kelemahan siswa belajar dan guru mengajar.

Lebih lanjut dikatakan bahwa beberapa hal yang harus diperhatikan guru atau pengajar dalam melaksanakan penilaian, antara lain :

1. penilaian harus dilaksanakan secara berkelanjutan, artinya setiap saat harus dinilai sehingga akan mendapatkan gambaran kemajuan siswa secara obyektif.
2. dalam proses kegiatan belajar mengajar penilaian dapat dilaksanakan dalam tiga tahap yakni : pre-test, mid-test, dan post-test.
3. penilaian dilaksanakan bukan hanya di dalam kelas tetapi juga diluar kelas. Lebih-lebih penilaian yang berkaitan dengan aspek tingkah laku siswa.
4. untuk memperoleh gambaran yang obyektif penilaian jangan hanya tes tetapi perlu dipergunakan jenis non tes.

Untuk mempermudah gambaran alat-alat evaluasi berdasarkan 3 jenis tes dapat dilihat pada gambar di bawah ini:



Gambar. **Jenis-Jenis TES**

B. Alat-Alat Evaluasi Pendidikan

Ada beberapa bentuk alat evaluasi pendidikan yang ada yaitu:

a. *Tes Tertulis* :

1) Pengertian.

Tes tertulis ialah tes, ujian atau ulangan, yang dialami oleh sejumlah siswa secara serempak dan harus menjawab sejumlah pertanyaan atau soal secara tertulis dalam waktu yang sudah ditentukan.

- 2) Kebaikannya, antara lain:
 - a. Sekaligus dapat menilai sejumlah siswa dalam waktu singkat;
 - b. Bagi siswa terdapat kebebasan memilih dalam menjawab;
 - c. Karena sama, maka skop dan isi pengetahuan yang dinilai ada setiap siswa pun sama pula.
- 3) Kelemahannya antara lain:
 - a. Tidak benar-benar menilai kepribadian siswa.
 - b. Mudah menimbulkan, kecurangan dan kepalsuan jawaban.
 - c. Mudah menimbulkan spekulasi bagi siswa.
- 4) Prosedur pelaksanaan

Dalam penyelenggaraannya perlu diperhatikan hal-hal berikut:

- a. Soal telah tertulis sebelumnya.
- b. Pertanyaan harus mencakup seluruh bahan yang diajarkan.
- c. Menentukan jumlah atau banyaknya pertanyaan atau soal.
- d. Kalimat pertanyaan harus jelas.
- e. Pertanyaan harus mengandung beberapa kemampuan.
- f. Mengandung tingkat kesukaran yang seimbang.
- g. Menyiapkan kunci jawaban.
- h. Menyiapkan norma penilaian.

b. Tes lisan

- 1) Pengertian.

Tes lisan ialah bila sejumlah siswa seorang demi seorang diuji secara lisan oleh seorang penguji atau lebih.

- 2) Kebaikannya, antara lain:
 - a. Lebih dapat menilai kepribadian dan isi pengetahuan siswa, karena dilakukan secara berhadapan-hadapan.
 - b. Jika siswa belum merasa jelas soalnya, penguji dapat mengubah pertanyaan sehingga siswa menjadi faham.
 - c. Dari sikap dan cara menjawab siswa, penguji dapat mengetahui apa yang "tersirat" disamping yang "tersurat".
 - d. Penguji dapat mengoreksi pengetahuan siswa sampai mendetail dan dapat mengetahui bidang mana yang lebih dikuasai atau disenangi siswa.
 - e. Penguji dapat langsung mengetahui hasilnya.
- 3) Kelemahannya, antara lain:
 - a. Jika hubungan antara penguji dengan siswa kurang baik dapat mengganggu objektivitas hasil tes.
 - b. Sifat penggugup pada siswa dapat mengganggu kelancaran jawaban yang dibencannya.
 - c. Pertanyaan yang diberikan tidak dapat senantiasa sama pada setiap siswa.
 - d. Untuk menguji kelas yang besar diperlukan waktu yang lama dan kurang ekonomis.
 - e. Sering tidak terdapat kebebasan bagi siswa
- 4) Prosedur pelaksanaan

Dalam penyelenggaraannya perlu diperhatikan hal-hal berikut:

- a. Penguji mempersiapkan terlebih dahulu pertanyaan yang hendak diajukan secara tertulis.
- b. Pertanyaan hendaknya jelas, sederhana dan santun.
- c. Pertanyaan harus mengandung beberapa kemampuan.
- d. Menentukan jumlah pertanyaan dengan mengingat waktu.
- e. Membuat perencanaan atau penataan aspek pertanyaan.
- f. Membuat kunci jawaban.
- g. Menetapkan norma penilaian.
- h. Memberi skor dan mengolahnya.

c. Observasi

1) Pengertian.

Observasi ialah metode/cara-cara menganalisis dan mengadakan pencatatan secara sistematis mengenai tingkah laku dengan melihat/mengamati siswa atau sekelompok siswa secara langsung. Dalam rangka evaluasi hasil belajar, observasi digunakan sebagai alat evaluasi untuk menilai kegiatan-kegiatan belajar yang bersifat keterampilan atau aspek psikomotor.

2) Kebaikannya, antara lain:

- a. Data observasi diperoleh secara langsung melalui penglihatan/ pengamatan terhadap segala ekspresi siswa dalam situasi atau perangsang tertentu.
- b. Data observasi bersifat lebih objektif dan melukiskan aspek-aspek kepribadian siswa yang sebenarnya.
- c. Di dalam situasi, yang relatif "bebas" dalam arti tanpa tekanan-tekanan, siswa tidak merasa ada yang memperhatikan tetapi juga tidak merasa sendirian. sehingga segala tingkah lakunya, pernyataannya itu dilakukan secara spontan.

- d. Data observasi lebih bersifat menyeluruh, mencakup berbagai aspek kepribadian, aspek hasil belajar, sehingga dalam pengolahannya tidak timpang dan lebih menggambarkan kepribadian siswa secara keseluruhan.
- 3) Kelemahannya, antara lain:
- a. Observasi sebagai suatu teknik evaluasi memerlukan sejumlah keterampilan yang baik, yang dapat dipertanggungjawabkan. Guru harus dapat mengenal perbedaan antara tingkah laku yang terlukiskan (*describing behavior*) dengan tingkah laku yang dievaluasi (*evaluating behavior*). Dengan kata lain, guru harus dapat membedakan apa yang "tersurat" dengan apa yang "tersirat."
 - b. Kepribadian guru sering kali merupakan variabel tambahan. Pengalaman, prasangka-prasangka, nilai-nilai pribadi guru turut terlibat dalam membubuhkan pencatatannya, sehingga sukar dipisahkan secara tegas dari tingkah laku murid yang sedang diamati.
 - c. Tingkah laku yang sama yang diekspresikan oleh beberapa orang siswa, belum tentu mempunyai arti yang sama bagi guru pengamat yang berlainan atau beberapa guru pengamat.
 - d. Data yang diperoleh dari observasi tidak dapat memberikan wawasan (*insight*) yang sama mengenai struktur kepribadian murid.
- 4) Prosedur pelaksanaannya :
- a. Dalam penyelenggaraan observasi sebagai alat penilaian, berupa dengan tes perbuatan, perlu diperhatikan:
 - Menentukan komponen yang akan diamati/dinilai.

- Menentukan aspek setiap komponen.
 - Menetapkan norma penilaian.
 - Menskor, menjumlah dan mengolahnya.
- b. Dalam penyelenggaraan observasi sebagai pengamatan/ pencatatan tingkah laku yang merupakan bantuan yang vital terhadap penilaian, akan dibicarakan tersendiri.

C. Penyusunan Soal

Ketrampilan guru dalam membuat soal merupakan prasyarat mutlak yang harus dimiliki oleh guru. Soal yang valid dan reliabel akan mampu mengukur kompetensi siswa secara obyektif.

1. Tes Esei

Pada tes esei ini disusun pertanyaan dan siswa disuruh menjawab pertanyaan dengan uraian/penjelasan dengan menggunakan kat/kalimat sendiri. Sifat tes ini sangat subyektif, namun demikian bentuk tes ini akan mampu menilai proses mental yang tinggi, berkaitan dengan kecakapan siswa dalam menyusun jawaban, mampu berekspresi, cakap berbahasa, logik dalam berfikir, dan lain-lain.

Ada dua bentuk tes esei ini yaitu :

- Uraian bebas (*free essay*) , dalam bentuk ini siswa akan menjawab secara bebas tentang sesuatu masalah yang ditanyakan. Kelemahannya bentuk ini ialah sukar menentukan standar jawabannya, karena jawaban yang beragama.

Contoh :

Coba jelaskan apa yang kamu ketahui tentang manajemen berbasis sekolah (MBS) ?

- Uraian yang terbatas (*limited essay*), Pada tes bentuk ini jawaban siswa dibatasi dan diarahkan kepada hal yang akan diminta dari pertanyaan tersebut. Aspek yang perlu diungkapkan dalam tes ini adalah : mengenal fakta, membandingkan, mencari sebab dan akibat, mencari contoh, membuat penggolongan, diskusi, kritik dan saran, mengemukakan ide/pendapat baru, dll.

Contoh :

Jelaskan 3 faktor penentu dalam melaksanakan manajemen berbasis sekolah (MBS) ?

2. Tes Obyektif

Tes ini terdiri dari beberapa bentuk tes yaitu:

a. Bentuk benar/salah (*true false*)

Bentuk tes yang ditanya tinggal memilih jawaban diantara benar dan salah, tidak ada pilihan lain.

Bentuk benar/salah ada dua cara yakni :

- **Dengan koreksi** : siswa harus membetulkan bila ia memilih jawaban yang salah.
- **Tanpa koreksi** : artinya meskipun jawaban itu saah tidak perlu siswa membetulkannya.

Petunjuk Soal :

1. Lingkarilah huruf B bila pernyataan itu betul dan lingkarilah huruf S, bila pernyataan itu salah (untuk jenis tanpa koreksi).
2. Lingkarilah huruf B bila pernyataan itu betul, dan lingkarilah huruf S, bila pernyataan itu salah. Bila saudara melingkari huruf S, betulkanlah kata/kalimat yang diberi garis di bawahnya (untuk jenis benar/salah dengan koreksi).

Petunjuk membuat item (soal) bentuk benar-salah

1. Pertanyaan hendaknya jangan meragukan siswa.

Contoh :

Urbanisasi penduduk artinya bahwa penduduk *mungkin* perpindahan dari kota ke desa atau dari desa ke kota.

2. Dalam salah satu pertanyaan diusahakan hanya mengandung satu pengertian yang diminta, jangan memasukkan lebih besar dari satu pengertian yang berbeda.
3. Hindari penggunaan kata-kata yang pasti dan ekstrim. Misalnya : tidak, pasti, sangat tepat, sama sekali, hanya satu-satunya, dan lain-lain.
4. Jawaban soal jangan diatur sistematis sehingga bisa ditebak. Misalnya : no. 1 = S, no. 2= B, no.3=S, no.4=B dst.

CONTOH :

Contoh 1. : tanpa koreksi

B-S Ajaran Islam pada dasarnya adalah cinta damai dan menolak kekerasan.

Contoh 2. : dengan koreksi

B-S mortalitas adalah angka yang menunjukkan tingkat pertumbuhan penduduk untuk setiap 1.000 orang/tahun.

b. Pilihan berganda (*multiple choice*)

Tes bentuk ini menyediakan kemungkinan jawaban lebih dari dua alternatif, biasanya 4 alternatif. Tugas siswa harus memilih satu alternatif yang dianggap benar.

Petunjuk :

”Pilih salah satu jawaban yang paling tepat dengan cara melingkari salah satu huruf yang terdapat dimuka kemungkinan jawaban”.

Petunjuk membuat soal pilihan ganda :

1. Option atau kemungkinan jawaban harus sejenis, dan seolah-olah cocok sebagai jawabannya.
2. Hindari option panjang, tetapi tidak mempunyai arti.
3. Harus ada option yang tepat diantara option-option yang lainnya.
4. Hindari penempatan option sistematis sehingga dapat ditebak siswa.

Tipe pertanyaan bentuk pilihan berganda (*multiple choice*) minimal ada enam jenis yakni :

1. Variasi negatif, artinya setiap pertanyaan mempunyai beberapa kemungkinan jawaban, disediakan satu kemungkinan jawaban yang salah. Tugas siswa memilih kemungkinan jawaban yang salah tersebut.

Contoh :

Pada hekatatnya ajaran Islam adalah :

- a). Menebarkan cinta kasih dan kedamaian
- b). Toleran dengan umat beragama dan antar umat beragama
- c). Suka dengan kekerasan dan pembunuhan
- d). Menebarkan keadilan dan kesejahteraan

2. Variasi berganti artinya pertanyaan memiliki beberapa kemungkinan jawaban yang semuanya betul, tapi ada satu jawaban yang paling betul. Tugas siswa mencari satu jawaban yang betul.

Contoh :

Yang dimaksud dengan urbanisasi adalah :

- a). Perpindahan penduduk dari desa ke kota
 - b). Perpindahan penduduk dari kota ke desa
 - c). Perpindahan penduduk dari pulau ke pulau
 - d). Perpindahan penduduk dari negara satu ke negara lain.
3. Variasi yang tidak lengkap artinya pertanyaan memiliki beberapa kemungkinan jawaban yang kesemuanya belum lengkap, (biasanya untuk perhitungan). Tugas siswa mencari satu kemungkinan dan menyelesaikannya.
 4. Variasi berganda artinya pertanyaan memiliki jawaban yang semuanya betul, tugas siswa memilih jawaban yang paling betul.

Contoh :

Masalah kependudukan di Indonesia ialah :

1. Pertambahan yang cepat
2. Distribusi yang tidak merata
3. Komposisi penduduk muda
4. Tersebar di semua pulau

Pilih :

- a. Jika jawaban 1 + 2 benar
- b. Jika jawaban 1 + 3 benar
- c. Jika semua jawaban benar
- d. Jika hanya jawaban 2 + 4 yang benar

5. Analisis hubungan artinya pertanyaan memiliki sejumlah jawaban siswa diminta menganalisis hubungan yang ada dalam soal tersebut.

Contoh :

Evaluasi dalam pengajaran memegang peranan penting **sebab** evaluasi mengukur keberhasilan pengajaran.

Pilih :

- A. Jika pernyataan benar, alasan benar dan alasan merupakan sebab dari pernyataan
 - B. Jika pernyataan benar, alasan benar tapi alasan bukan merupakan sebab dari pernyataan
 - C. Jika pernyataan benar alasan salah
 - D. Jika pernyataan salah alasan salah
6. Analisis kasus artinya pertanyaan berupa deskripsi kasus/kejadian dan siswa diminta menjawab soal yang diambil dari kasus tersebut.

Contoh :

Seorang guru memberikan tes dalam bentuk objektif tes sebanyak 60 pertanyaan dengan maksimum skor 60. Skor tertinggi yang dicapai oleh siswa adalah 45 dan skor terendah 15. Rata-rata yang dicapai oleh siswa adalah 30.

Batas lulus idealnya adalah :

- a. 30
- b. 32
- c. 35
- d. 45

c. Bentuk pilihan menjodohkan (*Matching*)

Tipe pertanyaan ini terdiri dari 2 kolom. Setiap pertanyaan pada kolom pertama harus dijodohkan dengan urutan pada kolom kedua.

Petunjuk soal:

"Pasangkanlah pernyataan yang ada pada lajur kiri dengan yang ada pada lajur kanan, dengan cara menempatkan (mengisikan) huruf yang terdapat dimuka pernyataan lajur kiri pada titik-titik yang disediakan di lajur kanan".

Syarat umum dalam membuat soal :

- a. Soal jangan terlalu banyak, misalnya 10 sampai 15 pertanyaan sekaligus, kecuali kalau dibagi-bagi menjadi beberapa kelompok.
- b. Jumlah pertanyaan yang ada pada lajur kiri, minimal satu lebih banyak dari yang ada dalam lajur kanan.
- c. Item-item harus homogen artinya mempunyai kategori yang sama.

Contoh :

- | | | | |
|---|--------------|---------|--|
| a | Transmigrasi | 1 | Masuknya penduduk dari negara lain. |
| b | Imigrasi |2 | Pindahnya penduduk ke negara yang lain |
| C | Emigrasi |3 | Pindahnya penduduk dari desa ke kota |
| | |4 | Pindahnya penduduk antar pulau dalam satu negara |

d. Bentuk pertanyaan melengkapi (*completion*)

Sering juga disebut *short-answered* (isian pendek). Tipe pertanyaan ini berupa serangkaian kalimat dengan kata (bagian kalimat) yang penting ditiadakan. Tugas siswa mengisi bagian yang tidak ada tersebut.

Petunjuk Soal :

"Isilah atau lengkapilah bagian yang diberi titik-titik dengan kata yang tepat".

Petunjuk Pembuatan Soal :

1. Hilangkan kata yang penting saja dalam setiap soal.
2. Hindari jangan sampai terlalu banyak bagian yang akan diisi sehingga membingungkan siswa.
3. Jangan meminta isian yang panjang (kalimat)
4. Tempat yang akan diisi (titik-titik) harus sama panjangnya untuk setiap soal.

Contoh :

Orang yang mengelola zakat disebut

D. Pengolahan Nilai

Tes yang telah dilaksanakan hendaknya dilakukan pengolahan untuk mendapatkan skor akhir dari kegiatan belajar mengajar yang dilakukan pada peserta didik.

Cara memeriksa dan menilai tes disesuaikan dengan jenis dan bentuk tes. Dengan mengetahui jenis dan bentuk tes yang berbeda, maka seorang evaluator akan memberikan teknik pemeriksaan dan penilaian yang benar.

Adapun cara memeriksa dan menilai hasil tes belajar siswa yang disesuaikan dengan jenis dan bentuk tes dapat dijelaskan sebagai berikut :

1). Tes Esei

Cara memeriksa tes ini sebaiknya menempuh langkah sebagai berikut :

- a. Tetapkan dulu kunci jawaban standar
- b. Periksa soal demi soal untuk setiap orang. Artinya jangan diperiksa seorang demi seorang semua soal, tetapi satu nomor setiap orang.
- c. Sediakan waktu yang cukup untuk memeriksanya, jangan dipaksakan memeriksa sekaligus, sehingga jika pemeriksa sudah lelah tidak sungguh-sungguh lagi memeriksanya.
- d. Teknik memberikan angka, sebaiknya menggunakan *Weight system*. Artinya memberi angka untuk setiap nomor tidak sama, bergantung kepada tingkat kesukaran yang dimiliki soal tersebut (sistem bobot).

Contoh :

- (1). Apa yang dimaksud dengan istilah *khilafah Islamiyah*?
- (2). Ada beberapa faktor yang menyebabkan amal seseorang tidak diterima, sebutkan dan berilah dalil hadis dan al Qur'annya!
- (3). Kemukakan 5 akibat jika manusia bersikap sombong!

Dengan melihat soal tersebut, maka bobot nilai soal dapat ditentukan dengan mempertimbangkan tingkat kesukaran dan kemudahan.

Misalnya dari soal diatas dapat diberi bobot penilaian sebagai berikut:

Untuk nomor 1 misalnya diberi bobot 2

Untuk nomor 2 misalnya diberi bobot 5

Untuk nomor 3 misalnya diberi bobot 3

2). Tes Obyektif

Dalam melakukan pemeriksaan dan penilaian bentuk tes obyektif, dapat dilakukan dengan prinsip mencocokkan lembar jawab dan kunci jawaban memakai mika transparan.

Sistem penilaian dapat digunakan rumus benar-salah sebagai berikut :

$$S = R - W$$

S = score yang diperoleh (Raw Score)

R = right (jawaban yang benar)

W = Wrong (jawaban yang salah)

3). Tes Pilihan Ganda

Sistem penilaian soal berbentuk pilihan berganda (*multiple choice*) dapat digunakan rumus sebagai berikut :

$$S = R - \left(\frac{W}{N-1} \right)$$

S = Skor yang diperoleh

R = Right (jawaban yang betul)

W = Wrong (jawaban yang salah)

N = Banyak option

I = Bilangan tetap

4). Tes Bentuk Pertanyaan Menjodohkan (*Matching*).

Sistem penilaian soal berbentuk pertanyaan menjodohkan (*Matching*) dapat digunakan rumus sebagai berikut :

$$S = R$$

S = Skor yang diperoleh
R = Right (jawaban yang betul)

5). Tes Tipe Pertanyaan Melengkapi (*Completion*).

Sistem penilaian soal berbentuk pertanyaan melengkapi (*Completion*) dapat digunakan rumus sebagai berikut :

$$S = R$$

S = Skor yang diperoleh
R = Right (jawaban yang betul)

E. Cara Pengolahan Skor

Mengolah skor merupakan kegiatan penting untuk mendapatkan hasil akhir dari proses pelaksanaan kegiatan evaluasi. Pengolahan skor yaitu kegiatan untuk mengubah nilai *raw-score* menjadi *drive score* (skor terjabar). Ada tiga dasar yang harus dilakukan dalam proses mengubah skor dari *raw score* menjadi *drive score* yaitu :

- 1. Mencari Mean**
- 2. Mencari nilai standar deviasi**
- 3. Passing grade (PG) batas lulus.**

Dalam proses pelaksanaan kegiatan skoring dapat dilakukan secara manual maupun memakai bantuan program SPSS.

Contoh :

Dilakukan tes bahasa arab dari 10 siswa dengan hasil sebagai berikut :

No	Nama	Nilai Bahasa Arab
1	Sutarjo	8
2	Sahidin	9
3	Maemunah	7
4	Ngatmin	7
5	Ramijah	5
6	Abdul Ghofur	8
7	Abdul Somad	9
8	Abdul Kohar	7
9	Syamsun	7
10	Rifan	5

Ubahlah nilai tersebut dari *raw score* menjadi *drive score* !

Jawab :

1. Carilah nilai rata-rata

Rumus :

$$\text{Mex} = \frac{\sum X_i}{N}$$

Tabel Nilai Tes Bahasa Arab

No	Nama	Nilai Bahasa Arab
1	Sutarjo	8
2	Sahidin	9
3	Maemunah	7

4	Ngatmin	7
5	Ramijah	5
6	Abdul Ghofur	8
7	Abdul Somad	9
8	Abdul Kohar	7
9	Syamsun	7
10	Rifan	5
	Jumlah	72 (ΣX_i)

$$M_{ex} = 72/10 = 7,2$$

2. Carilah nilai standar deviasi

No.Siswa	Skor X_i	M_{ex}	$(X_i - M_{ex})$	$(X_i - M_{ex})^2$
1	8	7,2	0,800	0,640
2	9	7,2	1,800	3,240
3	7	7,2	-0,200	0,040
4	7	7,2	-0,200	0,040
5	5	7,2	-2,200	4,840
6	8	7,2	0,800	0,640
7	9	7,2	1,800	3,240
8	7	7,2	-0,200	0,040
9	7	7,2	-0,200	0,040
10	5	7,2	-2,200	4,840
	72			17,600

Dari tabel tersebut dapat dimasukkan ke dalam rumus SD :

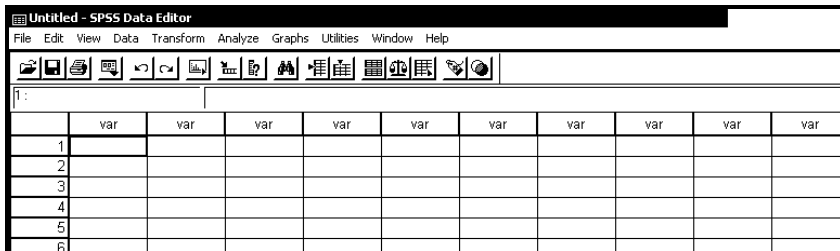
$$SD_x = \sqrt{\frac{\sum(X_i - M_{ex})^2}{N - 1}}$$

$$SD_x = \sqrt{17,600/10-1} = \sqrt{17,600/9} = \sqrt{1,955} = \underline{\underline{1,398}}$$

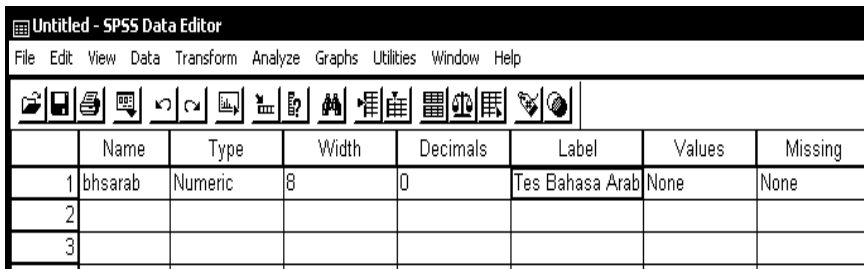
Cara lain untuk mendapatkan nilai mean dan standar deviasi

Program SPSS dapat membantu anda untuk mempercepat memperoleh nilai mean dan standar deviasi dari suatu skor yang diperoleh peserta didik pada pelaksanaan tes. Adapun langkah-langkahnya yaitu :

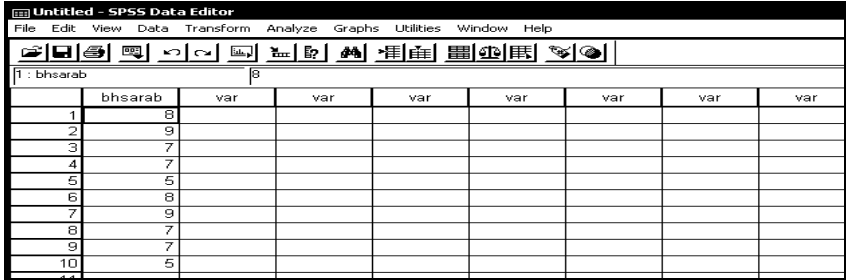
1. Bukalah program SPSS for Windows



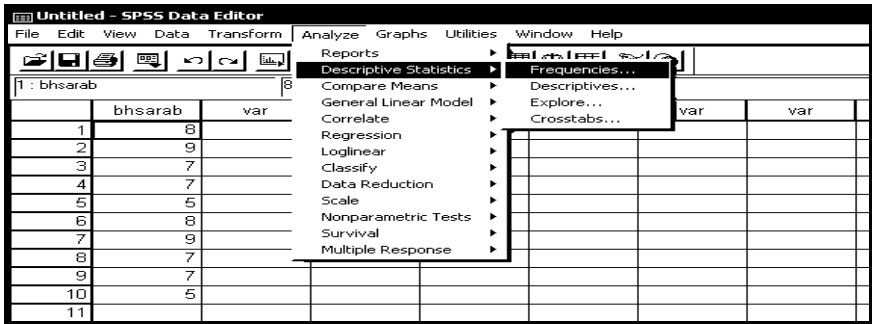
2. Klik variable view dan definisikan variabel (maksimal 8 karakter) dan desimal ketik 0.



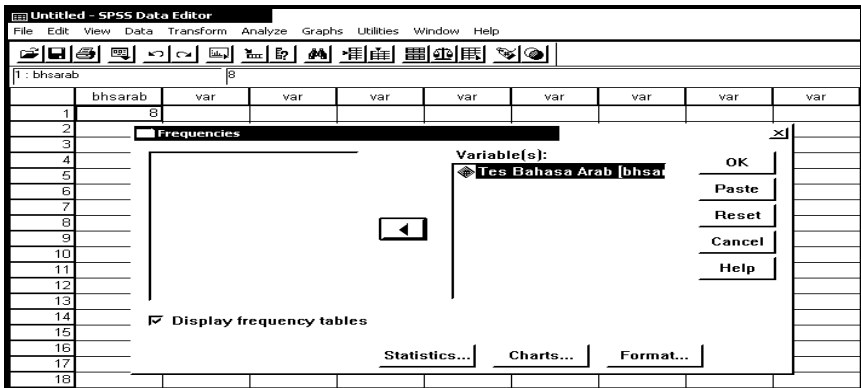
3. Klik data view dan lakukan pengisian data yang akan diuji



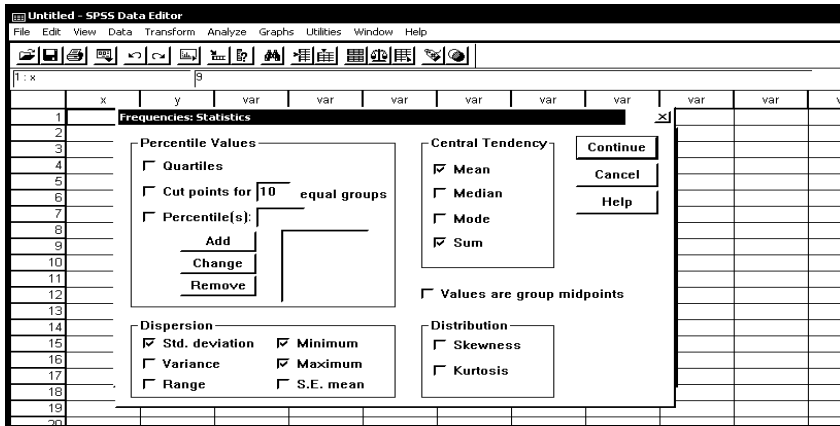
4. Klik Analyze kemudian pilih Descriptive Statistics dan pilih submenu Frequencies



5. Masukkan variabel Tes Bahasa Arab pada kotak Variables



- Klik tombol Statistics sehingga tampak dilayar windows Frequencies Statistics pada Central Tendency klik mean dan sum, pada Dispersion klik Std. Deviation, Maximum dan Minimum. Lalu Klik Continue.



- Klik OK. Hasil Oleh Data SPSS

Statistics

Tes Bahasa Arab

N	Valid	10
	Missing	0
Mean		7,20
Std. Deviation		1,40
Sum		72

Tes Bahasa Arab

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	5	20,0	20,0	20,0
	7	40,0	40,0	60,0
	8	20,0	20,0	80,0
	9	20,0	20,0	100,0
Total	10	100,0	100,0	

3. Membuat batas lulus.

A. Batas lulus aktual

Batas kelulusan aktual didasarkan pada nilai mean dan standar deviasi aktual. Batas lulus yang digunakan di Indonesia ialah angka 6 dalam standar angka 0-10. Angka 6 skala adalah $0,25 \sigma$ (skala sigma). Dengan rumus :

$$\text{Batas kelulusan aktual} = \text{mean} + 0,25 \times \text{SD}_x$$

$$\text{Batas kelulusan aktual} = 7,2 + 0,25 \times 1,398$$

$$\text{Batas kelulusan aktual} = 7,2 + 0,3495$$

$$\text{Batas kelulusan aktual} = \mathbf{7,5495}$$

Skor ini dinyatakan lulus dan memperoleh 6 dalam skala (0-10). Untuk menetapkan nilai 7, 8, 9 dan 10 dapat dikonversikan dengan standar 0-10 dengan menggunakan skala sebagai berikut :

Skala sigma	Skala 0-10	Raw score
-2,25 σ	1	M-2,25 (SD)
-1,75 σ	2	M-1,75 (SD)
-1,25 σ	3	M-1,25 (SD)
-0,75 σ	4	M-0,75 (SD)
-0,25 σ	5	M-0,25 (SD)
+0,25 σ	6	M+0,25 (SD)
+0,75 σ	7	M+0,75 (SD)
+1,25 σ	8	M+1,25 (SD)
+1,75 σ	9	M+1,75 (SD)
+2,25 σ	10	M+2,25 (SD)

B. Batas lulus ideal

Batas kelulusan ini menggunakan Mean Ideal yang besarnya $\frac{1}{2}$ dari skor maksimal, dan standar deviasi ideal yang besarnya $\frac{1}{3}$ dari mean ideal.

Jadi apabila skor maksimal 9, maka dapat dicari mean ideal = $\frac{1}{2} \times 9 = 4,5$. standar deviasi ideal = $\frac{1}{3} \times 4,5 = 1,5$.

Jadi batas lulusnya = $4,5 + (0,25) \times 15 = 4,5 + 0,375 =$
4,875

skor 4,875 sama dengan standar angka 6 dalam standar 0-10. untuk menghitung nilai-nilai standar 0-10, sama dengan batas kelulusan aktual.

C. Batas lulus purposif

Batas lulus yang ditentukan oleh panitia (guru-guru) berdasarkan pertimbangan-pertimbangan tertentu. Cara ini biasanya didasarkan kepada kebutuhan dan kepentingan tertentu, serta melihat hasil yang dicapai oleh siswa secara keseluruhan.



PROSEDUR EVALUASI PENDIDIKAN

A. PROSEDUR PELAKSANAAN TES HASIL BELAJAR

Pekerjaan mengevaluasi ada prosedur tersendiri, meskipun perlu untuk ditekankan, bahwa pekerjaan mengevaluasi itu lebih tepat untuk dipandang sebagai suatu proses yang kontinu. Suatu continuous proses yang tidak terputus-putus, tetapi ada gunanya juga mengetahui prosedur apa sajakah yang merupakan titik-titik penghubung dari proses yang bersifat kontinu tadi.

Pengetahuan tentang fungsi dalam keseluruhan proses evaluasi akan memungkinkan kita memperoleh gambaran yang cukup jelas tentang sistematis pekerjaan evaluasi pada umumnya. Bayangan yang ada pada diri kita mengenai sistematis dalam rangka pekerjaan evaluasi ini sudah ada pada kita, akan lebih memudahkan bagi kita untuk membangunkan suatu sistem evaluasi yang dapat dipertanggungjawabkan dalam suatu lingkungan pendidikan tertentu.

Dalam praktek, pelaksanaan tes hasil belajar dapat diselenggarakan secara tertulis (tes tertulis), dengan secara lisan (tes lisan) dan dengan tes perbuatan.

Pada tes tertulis, soal-soal tes dituangkan dalam bentuk tertulis dan jawaban tes juga tertulis. Pada tes lisan, soal-soal tes diajukan secara lisan dan dijawab secara lisan pula. Namun demikian dapat juga soal-soal tes diajukan secara lisan dan dalam waktu yang ditentukan, jawaban harus dibuat secara tertulis. Adapun pada tes perbuatan, wujud soal tesnya adalah pemberian perintah atau tugas yang

harus dilaksanakan oleh testee, dan cara penilaiannya dila-kukan terhadap proses penyelesaian tugas dan hasil akhir yang dicapai setelah testee melaksanakan tugas tersebut.

1. Prosedur Pelaksanaan Tes Tertulis

Dalam melaksanakan tes tertulis ada beberapa hal yang perlu mendapat perhatian, yaitu sebagaimana dikemukakan berikut ini.

Pertama, agar dalam mengerjakan soal tes para peserta tes mendapat ketenangan, seyogyanya ruang tempat berlangsungnya tes dipilihkan yang jauh dari keramaian, kebisingan, suara hiruk pikuk dan lalu langnya orang. Adalah sangat bijaksana apabila di luar ruangan tes dipasang papan pemberitahuan.

Kedua, ruangan tes harus cukup longgar, tidak berdesak-desakan, tempat duduk diatur dengan jarak tertentu yang memungkinkan tercegahnya kerja sama yang tidak sehat di antara testee.

Ketiga, ruangan tes sebaiknya memiliki sistem pencahayaan dan pertukaran udara yang baik. Ruangan yang gelap atau remang-remang disamping menyulitkan testee dalam membaca soal dan menuliskan jawabannya, juga akan menyulitkan bagi tester atau pengawas tes dalam menunaikan tugasnya. Ruang tes yang terlalu terang atau terlalu menyilaukan mata, disamping dapat menimbulkan udara panas juga dapat menyebabkan testee cepat menjadi letih.

Keempat, jika dalam ruangan tes tidak tersedia meja tulis atau kursi yang memiliki alas tempat penulis, maka sebelum tes dilaksanakan hendaknya sudah disiapkan alat berupa alas tulis yang terbuat dari triplex, hardboard atau bahan lainnya, sehingga testee tidak harus

menuliskan jawaban soal tes yang diletakkan di atas paha sebagai alas tulisnya.

Kelima, agar testee dapat memulai mengerjakan soal tes secara bersamaan, hendaknya lembar soal-soal tes diletakkan secara terbalik, sehingga tidak memungkinkan bagi testee untuk membaca dan mengerjakan soal lebih awal daripada teman-temannya. Dalam hubungan ini testee harus diberi bahwa mereka baru boleh memulai mengerjakan soal tes setelah tanda waktu mulai bekerja diberikan.

Keenam, dalam mengawasi jalannya tes, pengawas hendaknya berlaku wajar. Artinya jangan terlalu banyak bergerak, terlalu sering berjalan-jalan dalam ruangan tes sehingga mengganggu konsentrasi testee. Sebaliknya, pengawas tes juga jangan selalu duduk di kursi sehingga dapat membuka peluang bagi testee yang tidak jujur untuk bertindak curang (kerja sama dengan testee menyontek). Jika pengawas lebih dari satu sebaiknya berpencar dan jangan bercakap-cakap karena dapat mengganggu ketenangan jalannya tes.

Ketujuh, sebelum berlangsungnya tes, hendaknya sudah ditentukan lebih dahulu sanksi yang dapat dikenakan kepada testee yang berbuat curang. Sanksi itu dapat berupa tindakan mengeluarkan testee dari ruangan tes karenanya tesnya dianggap gugur, atau dengan jalan membuat berita acara tentang terjadinya kecurangan tersebut.

Kedelapan, sebagai bukti mengikuti tes, harus disiapkan daftar hadir yang harus ditandatangani oleh setiap peserta didik. Dalam mengedarkan daftar hadir tes itu hendaknya diusahakan agar tidak mengganggu ketenangan jalannya tes.

Kesembilan, jika waktu yang ditentukan telah habis hendaknya testee diminta untuk menghentikan

pekerjaannya dan secepatnya meninggalkan ruangan tes. Kemudian pengawas tes hendaknya segera mengumpulkan lembar-lembar pekerjaan (jawaban) tes seraya meneliti, jumlah lembar jawaban tes itu sudah sesuai dengan testee yang tercantum dalam daftar hadir tes.

Kesepuluh, untuk mencegah timbulnya berkaitan di kemudian hari, pada Berita Acara Pelaksanaan harus dituliskan secara lengkap, berapa orang testee yang hadir dan siapa yang tidak hadir, dengan menulis jumlahnya dan mengurutkan sesuai dengan nomor urut, nomor induk, nomor ujian, dan lain sebagainya, dan apabila terjadi penyimpangan atau kelainan-kelainan harus dicatat dalam lembar acara pelaksanaan tes tersebut.

2. Prosedur Pelaksanaan Tes Lisan

Beberapa petunjuk praktis berikut ini kiranya akan dapat dipergunakan sebagai pegangan dalam pelaksanaan tes lisan.

Pertama, sebelum tes lisan dilaksanakan, seyogyanya tester sudah melakukan inventarisasi berbagai jenis soal yang akan diajukan kepada testee dalam tes lisan tersebut, sehingga tes lisan dapat diharapkan memiliki validitas yang tinggi, baik dari segi isi maupun konstruksinya.

Kedua, setiap butir soal yang telah ditetapkan untuk diajukan dalam tes lisan itu, juga harus disiapkan sekaligus pedoman atau ancar-ancar jawaban betulnya. Hal ini dimaksudkan agar tester disamping mempunyai kriteria yang pasti dalam memberikan skor atau nilai kepada testee atas jawaban yang mereka berikan dalam tes lisan tersebut, juga tidak akan terpujau atau terkecoh dengan jawaban panjang lebar atau berbelit-belit yang diberikan oleh testee, yang menurut anggapan testee merupakan jawaban betul dan tepat, padahal menurut

kriteria yang telah ditentukan sesungguhnya sudah menyimpang atau tidak ada hubungannya dengan soal yang diajukan kepada testee.

Ketiga, jangan sekali-kali menentukan skor atau nilai hasil tes lisan setelah seluruh testee menjalani tes lisan. Skor atau nilai hasil tes lisan harus sudah dapat ditentukan di saat masing-masing testee selesai dites. Hal ini dimaksudkan agar pemberian skor atau nilai hasil tes lisan yang diberikan kepada testee itu tidak dipengaruhi oleh jawaban yang diberikan oleh testee yang lain.

Keempat, tes hasil belajar yang dilaksanakan secara lisan hendaknya jangan sampai menyimpang atau berubah arah dari evaluasi menjadi diskusi. Tester harus senantiasa menyadari bahwa testee yang ada di hadapannya adalah testee yang sedang "diukur" dan "dinilai" prestasi belajarnya mereka menempuh proses pembelajaran dalam jangka waktu tertentu. Dengan demikian apabila terjadi bahwa jawaban yang diberikan oleh testee yang sekalipun menyimpang dari kriteria yang telah ditentukan, namun sebenarnya tidak dapat disalahkan atau tidak sepenuhnya salah, cukup diberikan skor atau nilai dan tidak perlu disangkal atau diperdebatkan, yang dapat mengakibatkan kegiatan evaluasi berubah menjadi kegiatan diskusi.

Kelima, dalam rangka menegakkan prinsip obyektivitas dan prinsip keadilan, dalam tes yang dilaksanakan secara lisan itu, tester hendaknya jangan sekali-kali "memberikan angin segar" atau "memancing-mancing" dengan kata-kata, kalimat-kalimat atau kode-kode tertentu yang sifatnya menolong testee tertentu alasan "kasihan" atau karena tester menaruh "rasa simpati" kepada testee yang ada dihadapannya itu. Menguji, pada hakikatnya adalah "mengukur" dan bukan "membimbing" testee.

Keenam, tes lisan harus berlangsung secara wajar. Pernyataan tersebut mengandung makna bahwa tes lisan itu jangan sampai menimbulkan rasa takut, gugup atau panik di kalangan testee. Karena itu, dalam mengajukan pertanyaan-pertanyaan kepada testee, tester harus menggunakan kata-kata yang halus, bersifat sabar dan tidak emosional. Penggunaan kalimat-kalimat yang sifatnya "menteror", yang dapat menimbulkan tekanan psikis pada diri testee, haruslah dicegah.

Ketujuh, sekalipun acapkali sulit untuk dapat diwujudkan, namun sebaiknya tester mempunyai pedoman atau ancar-ancar yang pasti, berapa lama atau berapa waktu yang disediakan bagi tiap peserta tes dalam menjawab soal-soal atau pertanyaan-pertanyaan pada tes lisan tersebut. Harus diusahakan terciptanya keseimbangan alokasi waktu, antara testee yang satu dengan testee yang lain.

Kedelapan, pertanyaan-pertanyaan yang diajukan dalam tes lisan hendaknya dibuat bervariasi, dalam arti bahwa sekalipun inti persoalan yang ditanyakan itu sama, namun cara pengajuan pertanyaannya dibuat berlainan atau beragam. Hal ini dimaksudkan agar testee yang dites lebih akhir — (karena sudah memperoleh "informasi" dari testee yang telah dites terdahulu). jangan sampai "memperoleh nasib yang lebih mvijur" ketimbang testee yang dites lebih awal.

Kesembilan, sejauh mungkin dapat diusahakan agar tes lisan itu berlangsung secara individual (satu demi satu). Hal ini dimaksudkan agar tidak mempengaruhi mental testee yang lain. Misalnya, apabila dalam tes lisan itu secara serempak tester berhadapan dengan dua orang testee atau lebih dan pertanyaan yang sedang diajukan kepada testee yang mendapat kesempatan lebih awal tidak mungkin dapat dijawab oleh testee berikutnya, maka mental testee yang belum dites itu akan menjadi menurun

sehingga akan mempengaruhi jawaban-jawaban berikutnya. Kecuali itu hal tersebut di atas juga dimaksudkan agar tidak memberikan "angin segar" kepada testee yang belum dites, sebab mereka mempunyai kesempatan yang lebih luas untuk menyiapkan jawabannya ketimbang testee yang sedang atau sudah selesai dites.

3. Prosedur Pelaksanaan Tes Perbuatan

Tes perbuatan pada umumnya digunakan untuk mengukur taraf kompetensi yang bersifat keterampilan (psikomotorik), di mana penilaiannya dilakukan terhadap proses penyelesaian tugas dan hasil akhir yang dicapai oleh testee setelah melaksanakan tugas tersebut.

Karena tes ini bertujuan ingin mengukur keterampilan, maka sebaiknya tes perbuatan ini dilaksanakan secara individual. Hal ini dimaksudkan agar masing-masing individu yang dites akan dapat diamati dan dinilai secara pasti, sejauh mana kemampuan atau keterampilannya dalam

melaksanakan tugas yang diperintahkan kepada masing-masing individu tersebut.

Dalam melaksanakan tes perbuatan itu, ada beberapa hal yang perlu diperhatikan oleh tester.

Pertama, tester harus mengamati dengan secara teliti, cara yang ditempuh oleh testee dalam menyelesaikan tugas yang telah ditentukan.

Kedua, agar dapat dicapai kadar obyektivitas setinggi mungkin, hendaknya tester jangan berbicara atau berbuat

sesuatu yang dapat mempengaruhi testee yang sedang mengerjakan tugas tersebut.

Ketiga-, dalam mengamati testee yang sedang melaksanakan tugas itu, hendaknya tester telah menyiapkaninstru-menberupa lembar penilaianyangdi dalamnya telah ditentukan hal-hal apa sajakah yang harus diamati dan diberikan penilaian.

B. PROSEDUR PELAKSANAAN SISTEM EVALUASI

Ada beberapa hal yang harus kita lakukan setelah melaksanakan proses kegiatan evaluasi pada peserta didik yaitu penilaian data (verifikasi data), pengolahan data dan penafsiran data (atau interpretasi data).

Pendapat lain langkah-langkah pokok yang telah kita sebutkan tadi kita susun kembali, barangkali urutan langkah-langkah pokok tadi akan terlihat sebagai berikut:

1. Langkah perencanaan (termasuk atau tidak termasuk perumusan kriterium, bergantung keadaan). Tidak akan berlebihan iranya kalau diketahui di sini bahwa, sukses yang akan dapat dicapai oleh suatu program evaluasi telah turut ditentukan oleh memadai atau tidaknya langkah-langkah yang dilaksanakan dalam perencanaan ini.
2. Langkah pengumpulan data.

Pada langkah ini seorang evaluator hendaknya dapat mengumpulkan data dengan berbagai macam teknik. Misalnya memakai teknik tes dan non tes yang dilakukan secara obyektif.

3. Langkah verifikasi data.

Proses penyaringan ini kita sebut penelitian data atau verifikasi data dan maksudnya ialah untuk memisahkan data yang "baik" yang akan dapat memperjelas gambaran yang akan kita peroleh mengenai individu atau sekelompok individu yang

sedang kita evaluasi, dari data yang kurang baik yang hanya akan merusak atau mengaburkan gambaran yang akan kita peroleh apabila turut kita olah juga.

Dalam beberapa situasi kerap kali terjadi, bahwa tidak mungkinlah atau sangatlah sulit bagi evaluator untuk memperoleh data tertentu dari sumber pertama. Ia harus merasa puas dengan sekumpulan data yang telah berasal dari sumber kedua atau bahkan kadang-kadang juga sumber ketiga. Ini tidak semuanya sama baiknya; artinya tingkat persesuaian antara gambaran yang dicerminkan oleh data tadi dengan keadaan si individu "yang sebenarnya" yang dalam kebanyakan hal baru dapat kita cek di kemudian hari tidak sama untuk setiap jenis data yang kita kumpulkan dari sumber kedua atau ketiga tadi.

Di sinilah kita harus bersikap kritis terhadap data yang kita kumpulkan dari sumber pertama ini. Dengan mempergunakan logika kita, pengetahuan yang ada pada kita mengenai. Sifat materi yang tercantum dalam data yang kita hadapi tadi pengetahuan yang dapat kita gali tentang sumber data serta dengan mempergunakan analisis statistik, dapatlah kita mengadakan penelitian atau verifikasi terhadap data yang tidak kita peroleh dari sumber pertama tadi.

Langkah verifikasi data ini dapat memisahkan data yang baik dari data yang kurang baik, data yang "masuk akal" dari data yang tidak dapat diterima akal, data, yang bersih dari data yang kurang bersih.

4. Langkah pengolahan data.

Pada langkah ini seorang evaluator akan melakukan beberapa langkah dalam mengolah hasil tes dari nilai kasar (*raw score*) menjadi nilai (*drive score*). Untuk dapat melakukan proses pengolahan data yang benar, maka seorang evaluator perlu membekali diri dengan

pengetahuan statistik dasar. Misalnya bagaimana mencari nilai mean (nilai rata-rata), standar deviasi, dan korelasi product moment.

5. Langkah penafsiran data.

Pada langkah ini seorang evaluator memberikan penafsiran pada hasil akhir yang diperoleh seorang peserta didik dengan terlebih dahulu merumuskan kriteria kelulusan atau kriteria ketuntasan minimal. Kemudian dari hasil nilai yang ada akan dibandingkan dengan nilai yang didapat oleh peserta didik. Apakah dia termasuk lulus atau tidak lulus.



UJI VALIDITAS DAN RELIABILITAS SUATU KONSTRUK ATAU KONSEP

A. Pendahuluan

Dalam pembicaraan pada Bab ini telah dikemukakan, bahwa validitas adalah salah satu ciri yang menandai tes hasil belajar yang baik. Untuk segi, yaitu: dari segi tes itu sendiri sebagai suatu totalitas, dan dapat menentukan apakah suatu tes hasil belajar telah memiliki validitas atau daya ketepatan mengukur, dapat dilakukan dari dua dari segi itemnya, sebagai bagian tak terpisahkan dari tes tersebut.

Setiap penelitian yang menggunakan pendekatan kuantitatif dalam berbagai macam bidang studi pada umumnya variabel-variabel penelitiannya dirumuskan sebagai variabel laten (atau sering juga disebut faktor atau konstruk) yaitu variabel yang tidak diukur secara langsung, tetapi dibentuk melalui dimensi-dimensi yang diamati atau indikator-indikator yang diamati, dengan menggunakan kuesioner/angket.

Skala yang sering dipakai dalam penyusunan kuesioner adalah skala ordinal atau tipe LIKERT, yaitu skala yang berisi lima tingkat jawaban yang merupakan skala jenis ordinal dengan pilihan sebagai berikut : (1) sangat setuju, (2) setuju (3) ragu-ragu, (4) tidak setuju, (5) sangat tidak setuju.

Instrumen yang valid berarti alat ukur yang digunakan untuk mendapatkan data (mengukur) itu valid. Valid berarti instrumen tersebut dapat digunakan untuk mengukur apa yang hendak diukur. Meteran yang valid dapat digunakan untuk mengukur panjang dengan teliti, karena meteran

memang alat untuk mengukur panjang Meteran tersebut menjadi tidak valid jika digunakan untuk mengukur berat. Instrumen yang reliabel berarti instrumen yang bila digunakan beberapa kali untuk mengukur obyek yang sama, akan menghasilkan data yang sama. Alat ukur panjang dari karet adalah contoh instrumen yang tidak reliabel.

Dengan menggunakan instrumen yang valid dan reliabel dalam pengumpulan data, maka diharapkan hasil penelitian akan menjadi valid dan reliabel. Jadi instrumen yang valid dan reliabel merupakan syarat untuk mendapatkan hasil penelitian yang valid dan reliabel. Hal ini tidak berarti bahwa dengan menggunakan instrumen yang telah teruji validitas dan reliabilitasnya, otomatis hasil (data) penelitian menjadi valid dan reliabel.

Hal ini masih akan dipengaruhi oleh kondisi obyek yang diteliti. Peneliti harus mampu mengendalikan obyek yang diteliti dan meningkatkan kemampuan dan menggunakan instrumen untuk mengukur variabel yang diteliti.

Instrumen yang reliabel belum tentu valid. Meteran yang putus di bagian ujungnya, bila digunakan berkali-kali akan menghasilkan data yang sama (reliabel) tetapi selalu tidak valid. Hal ini disebabkan karena instrumen (meteran) tersebut rusak. Reliabilitas instrumen merupakan syarat untuk pengujiannya validitas instrumen. Oleh karena itu walaupun instrumen yang valid umumnya pasti reliabel, tetapi pengujian reliabilitas instrumen perlu dilakukan.

Pada dasarnya terdapat dua macam instrumen, yaitu instrumen yang berbentuk tes untuk mengukur prestasi belajar dan instrumen yang non tes untuk mengukur sikap. Instrumen yang berupa tes jawabannya adalah "salah atau benar", sedangkan instrumen sikap jawabannya; tidak ada yang "salah atau benar" tetapi bersifat "positif dan negatif.

Pada gambar tersebut ditunjukkan bahwa instrumen yang baik (yang berupa tes maupun non tes), harus valid dan reliabel. Instrumen yang valid harus mempunyai validitas internal dan eksternal. Instrumen yang mempunyai validitas internal dan eksternal. Instrumen yang mempunyai validitas internal atau rasional, bila kriteria yang ada dalam instrumen secara rasional (teoritis) telah mencerminkan apa yang diukur. Jadi kriterianya ada di dalam instrumen itu.

Sedangkan instrumen yang mempunyai validitas eksternal bila kriteria di dalam instrumen disusun berdasarkan luar atau fakta-fakta empiris yang telah ada. Kalau validitas internal instrumen dikembangkan menurut teori yang relevan, maka validitas eksternal instrumen dikembangkan dari fakta empiris. Misalnya akan mengukur kinerja (*performance*) sekelompok pegawai, maka tolak ukur (kriteria) yang digunakan berdasarkan pada tolok ukur yang telah ditetapkan di kepegawaian itu, sedangkan validitas internal dikembangkan dari teori-teori.

Validitas internal instrumen yang berupa tes harus memenuhi *construct validity* (validitas konstruk) dan *content validity* (validitas isi). Sedangkan untuk instrumen yang non tes yang digunakan untuk mengukur sikap cukup memenuhi validitas konstruksi. Sutrisno Hadi (1986) menyamakan *construct validity* dengan *logical validity* dan *validity by definition*. Instrumen yang mempunyai validitas konstruksi, jika instrumen tersebut dapat digunakan untuk mengukur gejala sesuai dengan yang didefinisikannya. Misalnya akan mengukur efektivitas kerja, maka perlu didefinisikan terlebih dahulu apa itu efektivitas kerja. Setelah itu disiapkan instrumen yang digunakan untuk mengukur efektivitas kerja sesuai dengan definisi. Untuk melahirkan definisi, maka diperlukan teori-teori. Dalam hal ini Sutrisno Hadi menyatakan bahwa "bila bangunan teorinya sudah benar maka hasil pengukuran dengan alat ukur (instrumen) yang berbasis pada teori itu sudah dipandang sebagai hasil yang valid".

Instrumen yang harus mempunyai validitas ini adalah instrumen yang digunakan untuk mengukur prestasi belajar (*achievement*) dan mengukur efektivitas pelaksanaan program: dan tujuan. Untuk menyusun instrumen prestasi belajar yang mempunyai validitas isi (*content validity*), maka instrumen harus disusun berdasarkan materi pelajaran yang telah diajarkan. Sedangkan instrumen yang digunakan untuk mengetahui pelaksanaan program, maka instrumen disusun berdasarkan program yang telah direncanakan. Selanjutnya instrumen yang digunakan untuk mengukur tingkat tercapainya tujuan (efektivitas) maka instrumen harus disusun berdasarkan tujuan yang telah dirumuskan.

B. Uji Reliabilitas

Dalam uji reliabilitas sebenarnya adalah alat untuk mengukur suatu kuesioner yang merupakan indikator dari variabel atau konstruk. Suatu kuesioner dikatakan reliabel atau handal, jika jawaban seseorang terhadap kenyataan konsisten atau stabil dari waktu-kewaktu.

Pengukuran reliabilitas dapat dilakukan dengan dua cara yaitu :

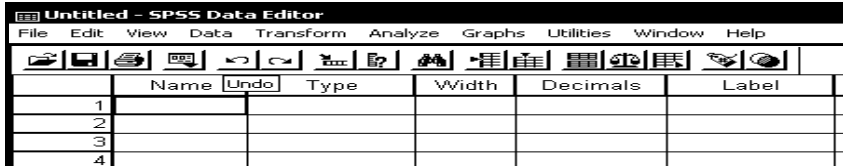
1. *Repeated Measure* atau pengukuran ulang. Disini seseorang akan diberikan pertanyaan yang sama pada waktu yang berbeda, dan dilihat apakah ia tetap konsisten dengan jawabannya.
2. *One Shot* atau pengukuran sekali saja. Pengukuran dilakukan sekali saja dan kemudian hasilnya dibandingkan dengan pertanyaan lain atau mengukur korelasi antar jawaban pertanyaan.

Untuk melakukan uji reliabilitas dapat digunakan program SPSS dengan menggunakan uji statistik Cronbach Alpha. Adapun kriteria bahwa instrumen itu dikatakan reliabel, apabila nilai yang didapat dalam proses pengujian dengan uji

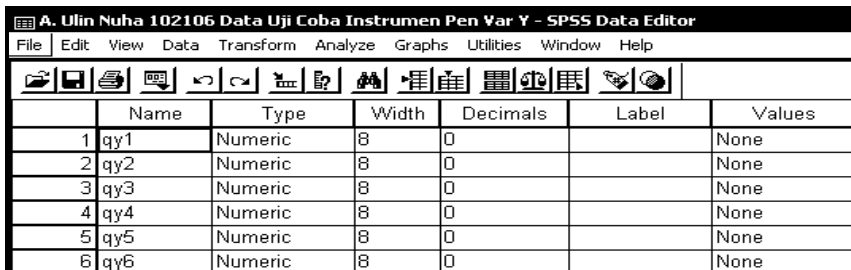
statistik Cronbach Alpha > 0,60 (Nunnally, 1969). Dan sebaliknya jika Cronbach Alpha diketemukan angka koefisien lebih kecil (< 0,60), maka dikatakan tidak reliabel.

Adapun langkah-langkah pengujian dengan menggunakan SPSS adalah :

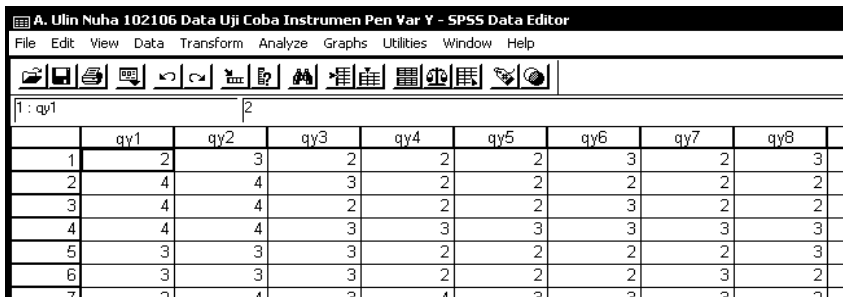
1. Bukalah program SPSS for Windows



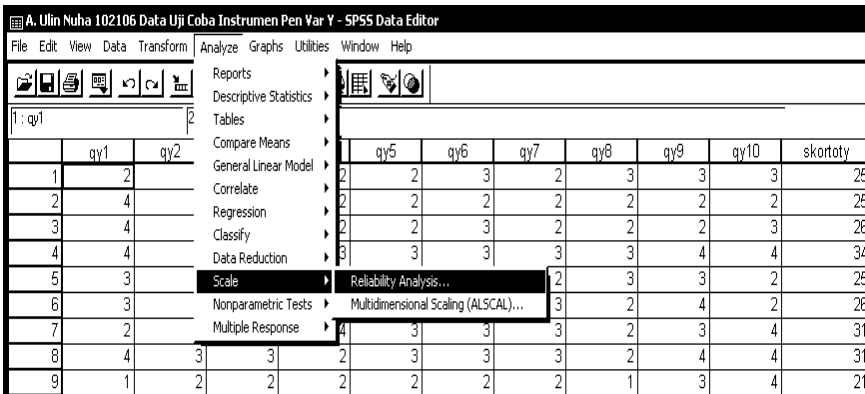
2. Klik variable view dan definisikan variabel (maksimal 8 karakter) dan desimal ketik 0.



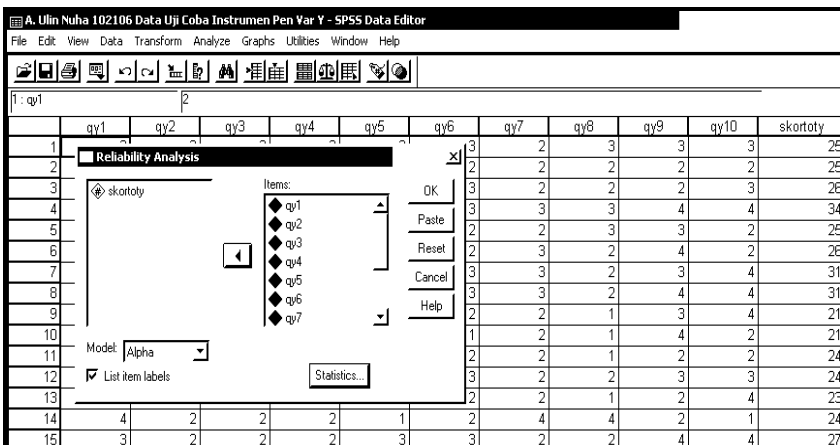
3. Klik data view dan lakukan pengisian data yang akan diuji



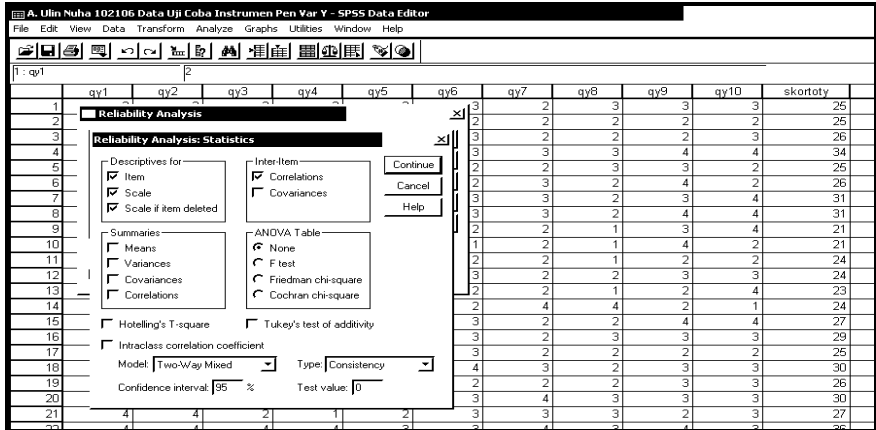
- Klik Analyze kemudian pilih submenu Scale lalu pilih Reliability Analysis



- Masukkan pertanyaan (q) 1 sampai 10 sebagai indikator motivasi kedalam box Items, dan pilih Alpha



- Klik tombol Statistics sehingga tampak dilayar windows Reliability Analysis Statistics dan pada bagian Descriptive pilih Scale if Item deleted.



7. Klik Continue dan OK.

8. Hasil Oleh Data SPSS

Reliability

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,730	,746	10

Item Statistics

	Mean	Std. Deviation	N
qy1	3,07	,907	30
qy2	3,07	,868	30
qy3	2,53	,629	30
qy4	2,53	,776	30
qy5	2,33	,711	30
qy6	2,73	,691	30
qy7	2,60	,724	30
qy8	2,10	,803	30
qy9	2,97	,809	30
qy10	2,73	,828	30

Inter-Item Correlation Matrix

	qy1	qy2	qy3	qy4	qy5	qy6	qy7	qy8	qy9	qy10
qy1	1,000	,213	,117	-,101	,178	,029	,200	,464	-,044	-,067
qy2	,213	1,000	,438	,304	,354	,145	,318	,138	-,046	,265
qy3	,117	,438	1,000	,528	,360	,338	,636	,232	,375	,084
qy4	-,101	,304	,528	1,000	,479	,338	,393	,022	,469	,122
qy5	,178	,354	,360	,479	1,000	,327	,000	-,121	,500	,508
qy6	,029	,145	,338	,338	,327	1,000	,331	,050	,169	,233
qy7	,200	,318	,636	,393	,000	,331	1,000	,486	,212	-,012
qy8	,464	,138	,232	,022	-,121	,050	,486	1,000	-,048	-,062
qy9	-,044	-,046	,375	,469	,500	,169	,212	-,048	1,000	,347
qy10	-,067	,265	,084	,122	,508	,233	-,012	-,062	,347	1,000

The covariance matrix is calculated and used in the analysis.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
qy1	23,60	15,490	,191	,421	,744
qy2	23,60	14,179	,420	,464	,703
qy3	24,13	14,257	,637	,602	,678
qy4	24,13	14,189	,494	,541	,692
qy5	24,33	14,299	,534	,676	,687
qy6	23,93	15,168	,377	,281	,710
qy7	24,07	14,340	,513	,641	,690
qy8	24,57	15,564	,232	,428	,733
qy9	23,70	14,769	,363	,523	,713
qy10	23,93	15,237	,272	,440	,728

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
26,67	17,678	4,205	10

Hasil Cronbach Alpha sebesar 0,730 angka ini jauh diatas 0,60 jadi dapat disimpulkan bahwa reliabilitas dari konstruk atau variable Kedisiplinan Siswa (Y) termasuk dalam kategori tinggi.

C. Uji Validitas

Penganalisisan terhadap tes hasil belajar sebagai suatu totalitas dapat dilakukan dengan dua cara. *Pertama*, peng-analisisan yang dilakukan dengan jalan berpikir secara rasional atau penganalisisan dengan menggunakan logika (*logical analysis*). *Kedua*, penganalisisan yang dilakukan dengan mendasarkan diri kepada kenyataan empiris, di mana penganalisisan dilaksanakan dengan menggunakan *empirical analysis*.

Uji validitas digunakan untuk mengukur sah atau tidaknya suatu kuesioner. Kuesioner dikatakan valid, jika pertanyaan pada kuesioner mampu mengungkapkan sesuatu yang akan diukur. Misalnya kita ingin mengukur Kedisiplinan Siswa (Y) dan diberi 10 (sepuluh) pertanyaan, maka pertanyaan tersebut harus secara tepat mengungkap tingkat Kedisiplinan Siswa.

Pengukuran validitas dapat dilakukan dengan tiga cara :

1. Melakukan korelasi antara skor butir pertanyaan dengan total score konstruk atau variabel, dengan hipotesa : (a) H_a = Skor butir pertanyaan tidak berkorelasi positif dengan total score konstruk, (b) H_0 = Skor butir pertanyaan berkorelasi positif dengan total score konstruk.

Uji signifikansi dilakukan dengan cara membandingkan nilai hitung korelasi dengan nilai hitung r tabel pada $df = n$, dimana $n =$ jumlah sampel. Pada kasus diatas $n = 30$, dengan taraf signifikan 0,05 didapat r tabel 0,361. Jika r hitung (untuk r tiap butir dapat dilihat pada kolom Corrected Item-Total Correlation) lebih besar dari r tabel dan nilai r positif, maka butir atau pertanyaan tersebut dikatakan valid.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
qy1	23,60	15,490	,191	,421	,744
qy2	23,60	14,179	,420	,464	,703
qy3	24,13	14,257	,637	,602	,678
qy4	24,13	14,189	,494	,541	,692
qy5	24,33	14,299	,534	,676	,687
qy6	23,93	15,168	,377	,281	,710
qy7	24,07	14,340	,513	,641	,690
qy8	24,57	15,564	,232	,428	,733
qy9	23,70	14,769	,363	,523	,713
qy10	23,93	15,237	,272	,440	,728

Pada tabel tersebut dapat dilakukan proses penafsiran dengan membuat tabel hasil uji validitas item sebagai berikut :

Tabel Hasil Uji Validitas Item Tentang Kedisiplinan Siswa (Y)

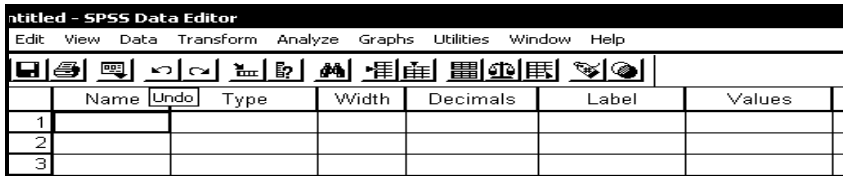
No	Korelasi (r hitung)	R tabel df=28 (5%)	Keterangan
1	0,191	0,361	Tidak Valid
2	0,420	0,361	Valid
3	0,637	0,361	Valid
4	0,494	0,361	Valid
5	0,534	0,361	Valid
6	0,377	0,361	Valid
7	0,513	0,361	Valid
8	0,232	0,361	Tidak Valid
9	0,363	0,361	Valid
10	0,272	0,361	Valid

Terlihat pada tabel diatas terdapat 2 item yang tidak valid, yaitu pada nomor 1 dan 8, sedangkan item 2,3,4,5,6,7,9,dan 10 adalah valid. Pada item nomor 1 dan 8 dapat diubah (direvisi) pertanyaannya dan dapat pula hilangkan.

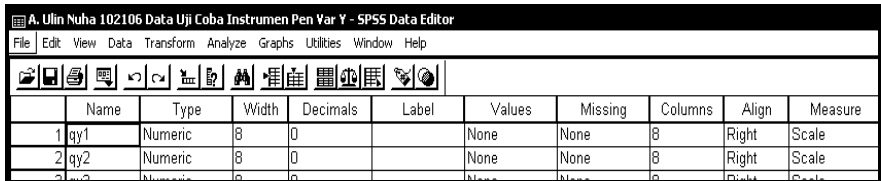
- Menghitung korelasi antara score masing-masing butir pertanyaan dengan total score, dengan menggunakan program SPSS.

Adapun langkah-langkah pengujian dengan menggunakan SPSS adalah :

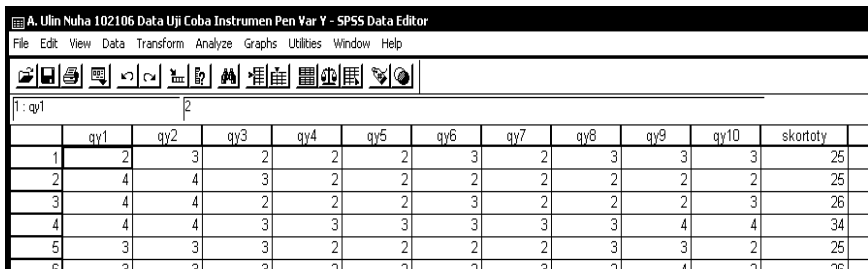
- Bukalah program SPSS for Windows



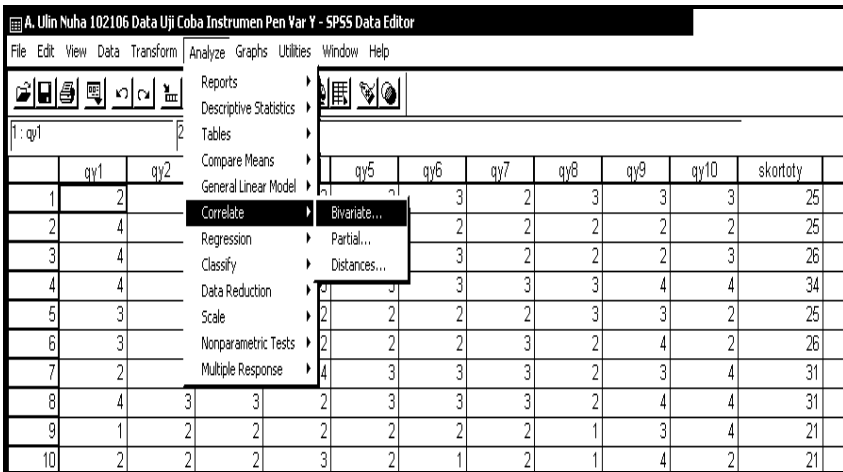
- Klik variable view dan definisikan variabel (maksimal 8 karakter) dan desimal ketik 0.



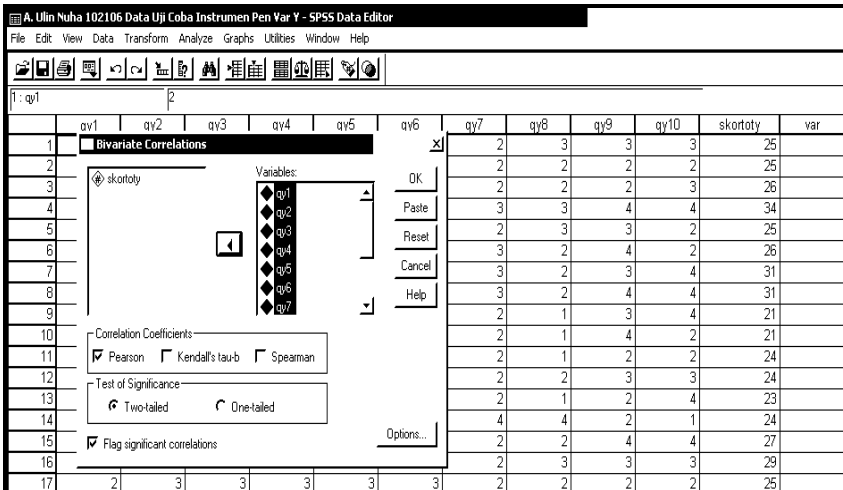
- Klik data view dan lakukan pengisian data yang akan diuji



- Klik Analyze kemudian pilih submenu Correlate, lalu pilih Bivariate.



- Isikan box variabel kesepuluh butir score pertanyaan dan score total y (kedisiplinan siswa), pilih Coefficient Correlation Pearson..



6. Klik Continue dan OK

Correlations

		qy1	qy2	qy3	qy4	qy5	qy6	qy7	qy8	qy9	qy10	skortoty
qy1	Pearson Correlation	1	.213	.117	-.101	.178	.029	.200	.464**	-.044	.067	.395*
	Sig. (2-tailed)	.	.258	.539	.595	.346	.878	.291	.010	.818	.724	.031
	N	30	30	30	30	30	30	30	30	30	30	30
qy2	Pearson Correlation	.213	1	.438*	.304	.354	.145	.318	.138	-.046	.265	.582**
	Sig. (2-tailed)	.258	.	.016	.103	.055	.443	.087	.466	.810	.156	.001
	N	30	30	30	30	30	30	30	30	30	30	30
qy3	Pearson Correlation	.117	.438*	1	.528**	.360	.338	.636**	.232	.375*	.084	.722**
	Sig. (2-tailed)	.539	.016	.	.003	.051	.067	.000	.217	.041	.659	.000
	N	30	30	30	30	30	30	30	30	30	30	30
qy4	Pearson Correlation	-.101	.304	.528**	1	.479**	.338	.393*	.022	.469**	.122	.627**
	Sig. (2-tailed)	.595	.103	.003	.	.007	.067	.032	.908	.009	.522	.000
	N	30	30	30	30	30	30	30	30	30	30	30
qy5	Pearson Correlation	.178	.354	.360	.479**	1	.327	.000	-.121	.500**	.508**	.650**
	Sig. (2-tailed)	.346	.055	.051	.007	.	.078	1.000	.525	.005	.004	.000
	N	30	30	30	30	30	30	30	30	30	30	30
qy6	Pearson Correlation	.029	.145	.338	.327	.327	1	.331	.050	.169	.233	.514**
	Sig. (2-tailed)	.878	.443	.067	.067	.078	.	.074	.794	.373	.215	.004
	N	30	30	30	30	30	30	30	30	30	30	30
qy7	Pearson Correlation	.200	.318	.636**	.393*	.000	.331	1	.486**	.212	-.012	.634**
	Sig. (2-tailed)	.291	.087	.000	.032	1.000	.074	.	.006	.261	.952	.000
	N	30	30	30	30	30	30	30	30	30	30	30
qy8	Pearson Correlation	.464**	.138	.232	.022	-.121	.050	.486**	1	-.048	-.062	.409*
	Sig. (2-tailed)	.010	.466	.217	.908	.525	.794	.006	.	.802	.744	.025
	N	30	30	30	30	30	30	30	30	30	30	30
qy9	Pearson Correlation	-.044	-.046	.375*	.469**	.500**	.169	.212	-.048	1	.347	.524**
	Sig. (2-tailed)	.818	.810	.041	.009	.005	.373	.261	.802	.	.060	.003
	N	30	30	30	30	30	30	30	30	30	30	30
qy10	Pearson Correlation	-.067	.265	.084	.122	.508**	.233	-.012	-.062	.347	1	.449*
	Sig. (2-tailed)	.724	.156	.659	.522	.004	.215	.952	.744	.060	.	.013
	N	30	30	30	30	30	30	30	30	30	30	30
skortoty	Pearson Correlation	.395*	.582**	.722**	.627**	.650**	.514**	.634**	.409*	.524**	.449*	1
	Sig. (2-tailed)	.031	.001	.000	.000	.000	.004	.000	.025	.003	.013	.
	N	30	30	30	30	30	30	30	30	30	30	30

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Terlihat pada tampilan output SPSS korelasi masing-masing butir qy1 sampai qy10 terhadap total score butir-butir pertanyaan kedisiplinan siswa menunjukkan hasil yang signifikan. Jadi dapat disimpulkan bahwa masing-masing butir pertanyaan adalah valid.

3. Uji Analisis Faktor

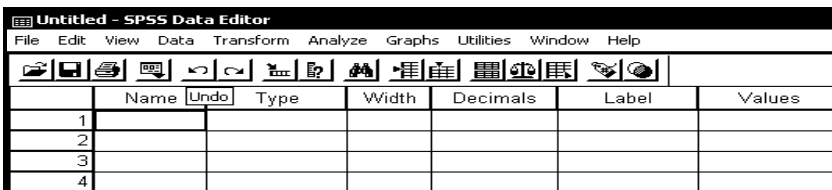
Analisis faktor digunakan untuk menguji validitas item atau indikator q1 sampai q10 yang digunakan dapat mengkonfirmasi sebuah faktor atau konstruk atau variabel (kedisiplinan), dengan ketentuan jika masing-masing item merupakan indikator pengukur kedisiplinan, maka akan memiliki nilai loading factor yang tinggi.

Misalkan kita ingin mengukur Pendidikan Pramuka dan Kedisiplinan Siswa. Pendidikan Pramuka diukur dengan qy1 sampai qy5 dan Kedisiplinan Siswa diukur dengan qy6 sampai qy10.

Faktor analisis dapat kita gunakan untuk menguji apakah butir pertanyaan qy1 sampai qy5 betul-betul merupakan indikator Pendidikan Pramuka, dan qy6 sampai qy10 betul-betul merupakan indikator Kedisiplinan Siswa. Analisis faktor akan mengelompokkan masing-masing pertanyaan qy1 sampai qy5 ke dalam beberapa faktor. Jika qy6 sampai qy10 merupakan indikator Pendidikan Pramuka, maka dengan sendirinya akan mengelompok menjadi satu dengan faktor loading tinggi begitu juga dengan kedisiplinan.

Adapun langkah-langkah pengujian dengan menggunakan SPSS adalah :

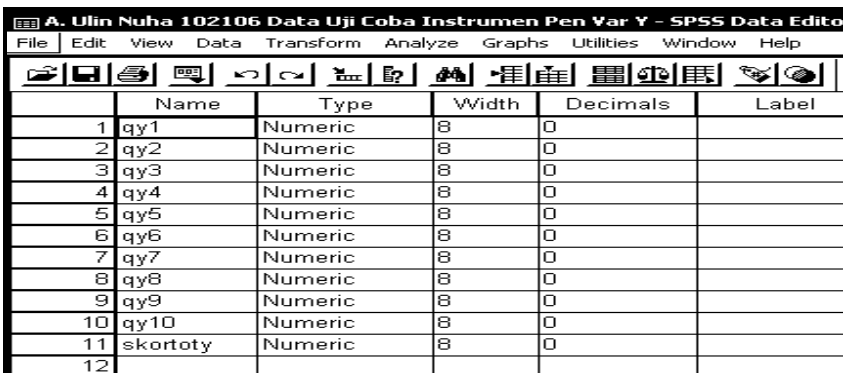
1. Bukalah program SPSS for Windows



The screenshot shows the SPSS Data Editor window titled "Untitled - SPSS Data Editor". The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Window, and Help. Below the menu bar is a toolbar with various icons. The main area contains a table with the following structure:

	Name [Undo]	Type	Width	Decimals	Label	Values
1						
2						
3						
4						

2. Klik variable view dan definisikan variabel (maksimal 8 karakter) dan desimal ketik 0.



The screenshot shows the SPSS Data Editor window titled "A. Ulin Nuha 102106 Data Uji Coba Instrumen Pen Var Y - SPSS Data Editor". The menu bar includes File, Edit, View, Data, Transform, Analyze, Graphs, Utilities, Window, and Help. Below the menu bar is a toolbar. The main area contains a table with the following structure:

	Name	Type	Width	Decimals	Label
1	qy1	Numeric	8	0	
2	qy2	Numeric	8	0	
3	qy3	Numeric	8	0	
4	qy4	Numeric	8	0	
5	qy5	Numeric	8	0	
6	qy6	Numeric	8	0	
7	qy7	Numeric	8	0	
8	qy8	Numeric	8	0	
9	qy9	Numeric	8	0	
10	qy10	Numeric	8	0	
11	skortoty	Numeric	8	0	
12					

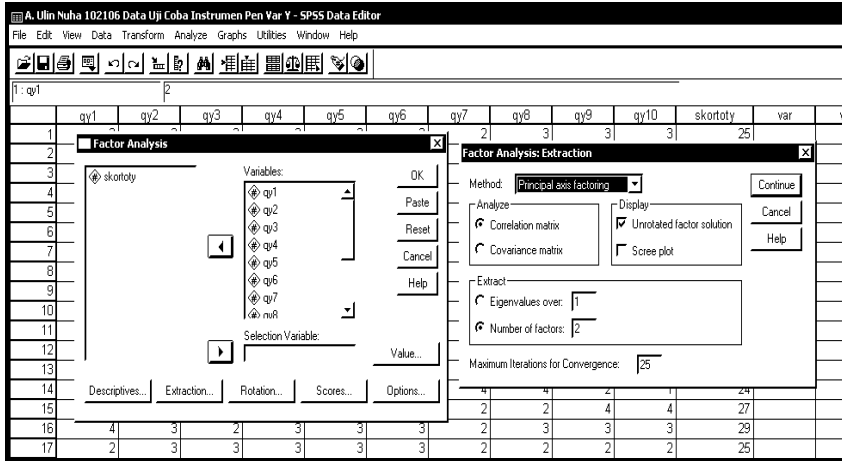
3. Klik data view dan lakukan pengisian data yang akan diuji

	qy1	qy2	qy3	qy4	qy5	qy6	qy7	qy8	qy9	qy10	skortoty	var
1	2	3	2	2	2	3	2	3	3	3	3	25
2	4	4	3	2	2	2	2	2	2	2	2	25
3	4	4	2	2	2	3	2	2	2	3	3	26
4	4	4	3	3	3	3	3	3	4	4	4	34
5	3	3	3	2	2	2	2	2	3	3	2	25
6	3	3	3	2	2	2	3	2	4	2	2	26
7	2	4	3	4	3	3	3	2	3	4	4	31
8	4	3	3	2	3	3	3	2	4	4	4	31
9	1	2	2	2	2	2	2	1	3	4	4	21
10	2	2	2	3	2	1	2	1	4	2	2	21
11	4	4	2	2	3	2	2	1	2	2	2	24

4. Klik Statistics Analyze kemudian pilih Submenu Data Reduction lalu pilih Factor.

	qy1	qy2
1	2	
2	4	
3	4	
4	4	
5	3	
6	3	
7	2	
8	4	3
9	1	2

5. Klik Factor Analysis, pada box variabel, masuk semua butir pertanyaan dari q1 sampai q10, pilih box Extraction, dan aktifkan box Number Factor, isikan 2 lalu continue, pilih box Rotation, dan aktifkan Varimax dan tekan Continue.



6. Klik OK Factor Analysis

Communalities

	Initial	Extraction
qy1	1,000	,354
qy2	1,000	,337
qy3	1,000	,677
qy4	1,000	,568
qy5	1,000	,664
qy6	1,000	,308
qy7	1,000	,701
qy8	1,000	,647
qy9	1,000	,499
qy10	1,000	,409

Extraction Method: Principal Component Analysis.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3,257	32,571	32,571	3,257	32,571	32,571	2,904	29,040	29,040
2	1,907	19,074	51,646	1,907	19,074	51,646	2,261	22,606	51,646
3	1,214	12,144	63,790						
4	,942	9,417	73,207						
5	,823	8,227	81,434						
6	,708	7,083	88,517						
7	,448	4,478	92,995						
8	,305	3,046	96,041						
9	,226	2,263	98,304						
10	,170	1,696	100,000						

Extraction Method: Principal Component Analysis.

Component Matrix

	Component	
	1	2
qy1	,206	,558
qy2	,556	,167
qy3	,797	,206
qy4	,733	-,176
qy5	,673	-,459
qy6	,550	-,070
qy7	,637	,544
qy8	,255	,763
qy9	,578	-,406
qy10	,415	-,486

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

Rotated Component Matrix

	Component	
	1	2
qy1	-,108	,585
qy2	,393	,428
qy3	,579	,585
qy4	,720	,224
qy5	,813	-,050
qy6	,509	,221
qy7	,269	,793
qy8	-,171	,786
qy9	,704	-,053
qy10	,606	-,206

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Component Transformation Matrix

Component	1	2
1	,859	,512
2	-,512	,859

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

Pada tabel diatas terlihat dari ke 10 butir pertanyaan, kita meminta dalam pengolahan program SPSS untuk dikelompokkan menjadi dua faktor (Pendidikan Pramuka dan Kedisiplinan). Tampilan output SPSS terlihat bahwa Faktor 1 mampu menjelaskan variasi sebesar 32,571 % dan faktor 2 mampu menjelaskan 19,074 % atau kedua faktor tersebut mampu menjelaskan variasi sebesar 51,645%.

Hasil dari rotasi varimax menunjukkan bahwa yang masuk ke faktor 1 dengan loading faktor besar adalah qy2, qy3, qy4, qy5, dan khusus qy6, qy7, qy9 sebenarnya merupakan indikator kedisiplinan akan tetapi dalam uji analisis faktor cocok untuk indikator pendidikan pramuka. Faktor 2 dengan loading faktor besar adalah qy8, qy10 dan khusus qy1 merupakan indikator variabel pendidikan pramuka tetapi cocok untuk menjelaskan variabel kedisiplinan. Untuk mempermudah dapat dibuat tabel sebagai berikut :

Tabel Hasil Uji Analisis Faktor

No. Item	Faktor 1	Faktor 2
1		qy1
2	qy2	
3	qy3	
4	qy4	
5	qy5	
6	qy6	
7	qy7	
8		qy8
9	qy9	
10		qy10

D. DEGREES OF DEFICULTY (DERAJAT KESUKARAN) DAN DISCRIMINATING POWER (DAYA BEDA) SUATU TES

Alat berupa tes tidak boleh terlalu mudah, dan juga tidak boleh terlalu sukar. Sebuah item yang terlalu mudah sehingga para siswa mampu menjawab dengan benar oleh semua anak bukanlah merupakan item yang baik. Begitu pula item pada alat tes yang terlalu sulit sehingga para siswa tidak mampu mengerjakan, juga bukan item yang baik. Jadi alat tes yang baik adalah item yang mempunyai derajat kesukaran tertentu.

Adapun kriteria derajat kesukaran (*degrees of difficulty*) adalah derajat kesukaran yang bergerak antara 25% sampai 75%. Item yang mempunyai derajat kesukaran dibawah 25% (< 25%) berarti bahwa item tersebut *terlalu mudah*. Sebaliknya jika ada item yang derajat kesukarannya diatas 75% (>75%), berarti bahwa item tersebut *terlalu sukar*.

Disamping itu alat tes yang kita gunakan juga dimaksudkan untuk memisahkan antara murid-murid yang betul-betul bisa dan murid-murid yang tidak bisa, maka item tes tersebut seharusnya mampu memisahkan dua kelompok murid tersebut. Jadi item disamping memiliki derajat kesukaran juga harus memiliki daya beda.

Adapun Kriteria daya beda (*Discriminating Power*) yaitu daya beda 0,40 ke atas (> 0,40). Namun untuk ulangan-ulangan harian, masih dapat ditolerer daya beda sebesar 0,20.

Contoh :

Murid SDN 1 Megawon mengikuti tes Matematika sebanyak 60 orang. Item tes tersebut berjumlah 6 item. Dimana WL dan WH dapat dilihat pada tebal dibawah ini :

No.Item	WL	WH
1	6	2
2	9	5
3	12	7
4	0	6
5	10	13
6	3	4

Ujilah derajat kesukaran dan daya beda!

Jawab :

Adapun langkah-langkah :

- 1. Susunlah lembar jawaban siswa mulai dari skor tertinggi sampai yang terendah.**
- 2. Ambil 27% lembar jawaban dari atas (kelompok atas), dan ambil 27% lembar jawaban dari bawah (kelompok bawah).**
- 3. Rumus Derajat Kesukaran dan Daya Beda**

$$DK = \frac{WL + WH}{nL + nH} \times 100\%$$

$$DB = \frac{WL - WH}{n}$$

Keterangan :

WL (*Wrong Lower*) = Jumlah individu kelompok bawah (27% dari bawah) yang tidak menjawab atau menjawab salah pada item tertentu.

WH (*Wrong High*) = Jumlah individu kelompok atas (27% dari atas) yang tidak menjawab atau menjawab salah pada item tertentu.

nL = Jumlah kelompok bawah

nH = Jumlah kelompok atas

n = jumlah kelompok atas atau bawah

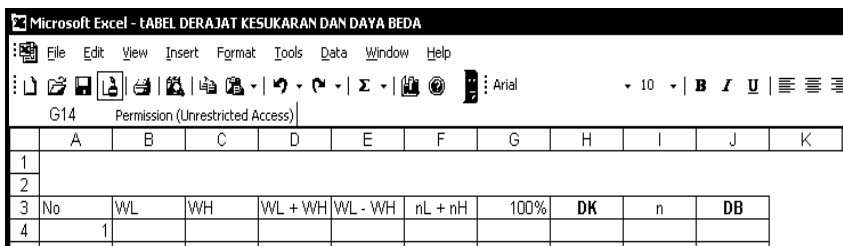
4. Membuat tabel kerja berdasarkan data.

Dalam proses penghitungan dapat digunakan bantuan komputer menggunakan Excel. Hal ini dapat dilihat langkah-langkah penghitungan sebagai gambar berikut ini.

- a. Membuka program Excel dengan melakukan klik dua kali pada shortcut program Excel.



- b. Buatlah tabel kerja berdasarkan unsur-unsur yang dibutuhkan dalam rumus uji derajat kesukaran dan uji daya beda. Sebagaimana gambar di bawah ini



c. Isilah berdasarkan data dan hitunglah !

No	WL	WH	WL + WH	WL - WH	nL + nH	100%	DK	n	DB
1	6	2	8	4	28	100	28,5714	14	0,285714
2	9	5	14	4	28	100	50	14	0,285714
3	12	7	19	5	28	100	67,8571	14	0,357143
4	0	6	6	-6	28	100	21,4286	14	-0,42857
5	10	13	23	-3	28	100	82,1429	14	-0,21429
6	3	4	7	-1	28	100	25	14	-0,07143

5. Interpretasi

Hasil dari proses kalkulasi tersebut dapat dilihat pada tabel di bawah ini.

No.It	WL	WH	WL + WH	WL - WH	nL + nH	100%	DK	n	DB
1	6	2	8	4	28	100	28,5714	14	0,285714
2	9	5	14	4	28	100	50	14	0,285714
3	12	7	19	5	28	100	67,8571	14	0,357143
4	0	6	6	-6	28	100	21,4286	14	-0,42857
5	10	13	23	-3	28	100	82,1429	14	-0,21429
6	3	4	7	-1	28	100	25	14	-0,07143

Dengan menggunakan kriteria derajat kesukaran dan daya beda, maka hasil perhitungan pada tabel diatas dapat diketemukan simpulan pada tiap-tiap item. Adapun kriterianya adalah

a. Kriteria derajat kesukaran (*degrees of deficulty*) adalah derajat kesukaran yang bergerak antara 25% sampai 75%.

Item yang mempunyai derajat kesukaran dibawah 25% (< 25%) berarti bahwa item tersebut *terlalu mudah*. Sebaliknya jika ada item yang derajat kesukarannya diatas 75% (>75%), berarti bahwa item tersebut *terlalu sukar*. Pada tabel diatas terlihat item 4 termasuk dalam kategori item *terlalu mudah* (21,4266%), dan pada item 5 termasuk dalam kategori *terlalu sukar* (82,1429 %).

- b. Kriteria daya beda (*Discriminating Power*) yaitu daya beda 0,40 ke atas (> 0,40). Namun untuk ulangan-ulangan harian, masih dapat ditolerer daya beda sebesar 0,20. Pada tabel diatas terlihat item 6 tidak dapat membedakan murid mana yang pandai dan bodoh hal ini diketemukan angka - 0,07143 masih dibawah 0,20.

Efektifitas Option

Berdasarkan uji derajat kesukaran dan uji daya beda, maka kita dapat melakukan seleksi item-item yang telah kita buat, item-item mana yang memenuhi syarat dan yang tidak. Item-item yang memenuhi syarat kita simpan atau kita gunakan untuk keperluan evaluasi yang akan datang. Item yang tidak memenuhi syarat akan kita lakukan dengan cara membuang atau merevisi item.

Revisi item adalah efektifitas daripada masing-masing option yang digunakan dalam item tersebut. Adapun langkah-langkah yang dilakukan untuk mengetahui suatu option mana yang efektif dan mana yang tidak, dapat dijelaskan sebagai berikut :

1. Ambilah 27% lembar jawaban yang mendapat skor tertinggi dan 27% lembar jawaban yang mendapat skor rendah.
2. Buatlah tabel sejumlah item yang akan diuji efektifitasnya dengan memakai program Excel.

a. Bukalah Program Excel

Microsoft Excel - TABEL DERAJAT KESUKARAN DAN DAYA BEDA

File Edit View Insert Format Tools Data Window Help

A1 Save

	B	C	D	E	F	G	H	I
2								
3								
4								
5								
6								
7								
8								
9								
10								

b. Buatlah tabel

Microsoft Excel - TABEL DERAJAT KESUKARAN DAN DAYA BEDA

File Edit View Insert Format Tools Data Window Help

J11

	A	B	C	D	E	F	G	H
1								
2								
3	Option/Kelompok	a	b	c	d	e		
4	Atas							
5	Bawah							
6								
7								

c. Isikan d distribusi pilihan terhadap option, baik untuk kelompok atas atau bawah.

Microsoft Excel - TABEL DERAJAT KESUKARAN DAN DAYA BEDA

File Edit View Insert Format Tools Data Window Help

A1

	A	B	C	D	E	F	G	H
1								
2								
3								
4	Option/Kelompok	a	b	c	d	e		
5	Atas	1	8	4	0	0		
6	Bawah	8	3	3	1	0		
7								
8								
9								

d. Hitunglah masing-masing option

(1). Option Kunci

Microsoft Excel - tABEL DERAJAT KESUKARAN DAN DAYA BEDA										
File Edit View Insert Format Tools Data Window Help										
D10										
	A	B	C	D	E	F	G	H	I	J
1										
2	Option/Kelompok	a	b	OH + OL	100%	Uji Efektifitas (%)	c	d	e	
3										
4	Atas	1	8				4	0	0	
5	Bawah	8	3			3	1	0		
6	Option Kunci		11	28	100	39,28571429				
7										
8										

(2). Option Pengecoh a, c, d dan e

Microsoft Excel - tABEL DERAJAT KESUKARAN DAN DAYA BEDA									
File Edit View Insert Format Tools Data Window Help									
D20									
	A	B	C	D	E	F	G		
1									
2	Option/Kelompok	a	Frekuensi Kel Atas (OH)	2X jumlah Option Pengecoh	25%	OH + OL	Uji Efektifitas (%)		
3									
4	Atas	1							
5	Bawah	8							
6	Option Pengecoh	9	1	8	0,25	28	0,875		
7									
8									
9	Option/Kelompok	c	Frekuensi Kel Atas (OH)	2X jumlah Option Pengecoh	25%	OH + OL	Uji Efektifitas (%)		
10									
11	Atas	4							
12	Bawah	3							
13	Option Pengecoh	7	4	8	0,25	28	3,5		
14									
15									
16	Option/Kelompok	d	Frekuensi Kel Atas (OH)	2X jumlah Option Pengecoh	25%	OH + OL	Uji Efektifitas (%)		
17									
18	Atas	0							
19	Bawah	1							
20	Option Pengecoh	1	1	8	0,25	28	0,875		
21									
22									
23	Option/Kelompok	a	Frekuensi Kel Atas (OH)	2X jumlah Option Pengecoh	25%	OH + OL	Uji Efektifitas (%)		
24									
25	Atas	0							
26	Bawah	0							
27	Option Pengecoh	0	0	8	0,25	28	0		
28									

e. Interpretasi

Option/ Kelompok	a	b	OH + OL	100%	Uji Efektifitas (%)	c	d	e
Atas	1	8				4	0	0
Bawah	8	3				3	1	0
Option Kunci		11	28	100		39,28571429		

Keterangan:

OH = lembar jawaban skor tertinggi

OL = lembar jawaban skor terendah

Hasil uji efektifitas option diketemukan angka 39,28571429 % atau 39%. Jadi uji efektifitas option kunci (b) **cukup efektif**. Hal ini terlihat 39% berada pada kisaran antara > 25% dan <75%, Disamping itu kelompok pemilih atas (8 orang) lebih besar daripada kelompok pemilih bawah (3 orang).

Option/ Kelompok	a	Frekuensi Kel Atas (OH)	2X jumlah Option Pengecoh	25%	OH + OL	Uji Efektifitas (%)
Atas	1					
Bawah	8					
Option Pengecoh	9	1	8	0,25	28	0,875

Option pengecoh a berfungsi **sangat efektif** sebab jumlah pemilihnya 9 orang. Jadi tidak kurang 0,875 (>0,875). Disamping itu frekuensi pemilih kelompok bawah (8 orang) lebih tinggi daripada frekuensi pemilih atas (1 orang).

Option/ Kelompok	c	Frekuensi Kel Atas (OH)	2X jumlah Option Pengecoh	25%	OH + OL	Uji Efektifitas (%)
Atas	4					3,5
Bawah	3					
Option Pengecoh	7	4	8	0,25	28	

Option pengecoh c berfungsi *tidak efektif* sebab jumlah pemilih kelompok atas (4 orang) lebih tinggi daripada kelompok bawah (3 orang). Disamping itu hasil uji efektifitasnya diketemukan 3,5 diatas **0,875 (>0,875)**.

Option/ Kelompok	d	Frekuensi Kel Atas (OH)	2X jumlah Option Pengecoh	25%	OH + OL	Uji Efektifitas (%)
Atas	0					0,875
Bawah	1					
Option Pengecoh	1	1	8	0,25	28	

Option pengecoh d berfungsi *agak efektif* sebab jumlah pemilih kelompok atas (1 orang) dan kelompok bawah (1 orang). Disamping itu hasil uji efektifitasnya diketemukan 0,875 tidak kurang dari **0,875 (<0,875)**.

Option/ Kelompok	a	Frekuensi Kel Atas (OH)	2X jumlah Option Pengecoh	25%	OH + OL	Uji Efektifitas (%)
Atas	0					0
Bawah	0					
Option Pengecoh	0	0	8	0,25	28	

Option pengecoh d berfungsi **tidak efektif** sebab jumlah pemilih kelompok atas (1 orang) dan kelompok bawah (1 orang). Disamping itu hasil uji efektifitasnya ditemukan 0 dibawah **0,875 (<0,875)**.

Agar dalam menentukan kebijakan item itu proporsional, maka yang perlu diingat adalah apakah item mempunyai daya positif atau tidak, apakah setiap item mengukur hasil belajar yang penting atau tidak?. Jika pertanyaan tersebut dijawab ya, maka item tersebut dapat disimpan dan digunakan pada masa yang akan datang dengan mempertimbangkan kembali analisis item, baik dengan *derajat kesukaran*, *daya beda* dan *efektifitas option*.



ANALISA HASIL TES HASIL TES BELAJAR

A. Skala Penilaian

Kegiatan evaluasi yang perlu dilakukan dalam pengolahan nilai adalah skala penilaian. Kita tahu bahwa dengan mengetahui skala penilaian yang benar diharapkan mampu menghindari persepsi yang salah pada skor yang diperoleh oleh siswa. Adapun beberapa skala penilaian adalah :

1. Skala Bebas

Setiap pelaksanaan ulangan yang dilakukan oleh seorang guru, berkas kertas ujian biasanya diberikan oleh siswa disertai nilainya. Amir, adalah seorang pelajar Madrasah Aliyah yang mendapatkan skor 10 terlihat pada pojok kiri atas kertas ujian. Dengan percaya diri skor 10 yang telah didapat diberitahukan pada teman lainnya, yang menurut dirinya ia telah mendapatkan nilai yang tertinggi. Amir berdiskusi dengan temannya baru pada 4 orang. Ternyata nilai hasil ulangan sebanyak 10 yang didapat oleh Amir termasuk paling rendah diantara 4 orang temannya yang mendapatkan nilai 15,20, 21, dan 25, wajahnya menjadi malu tersipu-sipu.

Mengapa terjadi persepsi yang salah? Hal ini dikarenakan Amir terpancang dengan anggapan bahwa 10 merupakan nilai tertinggi, yang benar adalah penentuan skala penilaian tertinggi dari banyak dan bentuk soal. Jadi angka tertinggi dan skala yang digunakan tidak selalu sama.

2. Skala 0-10

Amir beranggapan bahwa angka 10 adalah angka tertinggi untuk penilaian? Hal ini disebabkan pada umumnya guru-guru di Indonesia mempunyai kebiasaan menggunakan

skala 1-10 atau 1-100 untuk pelaporan hasil belajar siswa dalam rapor. Guru memberikan penilaian berupa skor jarang yang menggunakan pecahan desimal. Misalnya nilai 6,5 kemudian dibulatkan menjadi 7. Padahal siswa yang mendapatkan nilai 7,4 juga dibulatkan menjadi 7. Dengan demikian rentangan angka 6,5 sampai 7,4 (selisih hampir 1), namun dalam rapor nilai tersebut akan keluar skor 7.

3. Skala 0-100

Skala 0-100 dimungkinkan melakukan penilaian yang halus karena terdapat 100 bilangan bulat. Nilai 6,5 dan 7,4 dalam skala 0-10 yang biasanya dibulatkan menjadi 7 dalam skala 0-100 boleh dituliskan dengan 65 dan 74. Penggunaan angka dengan s

4. Skala Huruf

Pemberian nilai dapat dilakukan dengan menggunakan A, B, C, D dan E, bahkan ada juga yang menggunakan huruf sampai abjad G. Penggunaan huruf dalam penilaian akan terasa lebih tepat digunakan karena tidak ditafsirkan sebagai arti perbandingan. Huruf tidak menunjukkan kuantitas, tetapi dapat digunakan sebagai simbol untuk menggambarkan kuantitas. Kelemahannya jika mencari nilai rata-rata prestasi hasil belajar.

Untuk mengatasi hal tersebut dapat digunakan dengan cara menstransfer nilai-nilai huruf itu kedalam angka pada rentangan nilai angka. Sebagai contoh adalah nilai huruf yang terdapat pada *Tabel 18 konversi skor*.

Contoh :

Angka 100	Angka 10	IKIP	Huruf	Keterangan
80-100	8,0-10,0	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
40-55	4,0-5,5	4,1-5,5	D	Kurang

30-39	3,0-3,9	0-4,0	E	Gagal
-------	---------	-------	---	-------

B. Mengubah Skor Mentah Menjadi Skor Standar

Ada beberapa cara dalam mengubah skor mentah menjadi skor standar. Sebagaimana yang diungkapkan oleh Woodworth mengatakan bahwa : *"In general there are two ways of making such comparison. In some situation the individual's score is compared with certain absolute standards... more commonly the individual score is compared directly with the scores of other individual in same performance (Woodworth, 1961, hal 28).*

Ada dua jenis norma yang dapat dipergunakan untuk mengkonversikan nilai dari skor mentah menjadi skor standar), yaitu :

1. Norma Absolut yaitu dengan membandingkan skor yang diperoleh oleh seseorang dengan standar yang absolut.
2. Norma Relatif yaitu dengan jalan membandingkan skor seseorang dengan skor yang diperoleh oleh orang-orang lain dalam tes tersebut

Bahkan dalam perkembangannya ada mengkombinasikan antara norma absolut dan relatif. Untuk lebih jelasnya dapat dilihat pada beberapa penjelasan dari pembagian tiga norma tersebut.

1. Norma Absolut (Mutlak)

Norma absolut ini disebut pula dengan norma aktual atau norma das solen atau Penilaian Acuan Patokan (PAP). Norma absolut merupakan suatu norma yang ditetapkan secara absolut (mutlak) oleh guru atau pembuat tes, berdasarkan atas jumlah soal, bobot masing-masing soal serta prosentase penguasaan yang dipersyaratkan. Dengan demikian nilai atau skor standar yang diperoleh seseorang didasarkan atas norma absolut menunjukkan bahwa penguasaan anak terhadap bahan yang diberikan..

Disamping kita menentukan jenis norma penilaian, juga menentukan jenis skala yang dipergunakan. Adapun jenis skala yang umum dipergunakan ada beberapa jenis yaitu :

a. Skala Lima

Skala lima adalah suatu pembagian tingkatan yang terbagi atas lima kategori. Masing-masing tingkatan dinyatakan dengan huruf A, B, C, D, dan E.

Adapun langkah-langkah yang dapat ditempuh dalam mengkonversi nilai skor mentah menjadi skor standar dengan menggunakan norma absolut skala lima adalah sebagai berikut :

- (1). Mencari skor maksimal ideal (SMI) daripada tes yang diberikan.
- (2). Membuat pedoman konversi yang pada umumnya yaitu :

Tingkat Penguasaan	Skor Standar
90%-100%	A
80%-89%	B
65%-79%	C
55%-64%	D
0%-54%	E

Contoh :

Misalkan suatu tes hasil belajar terdiri dari item-item sebagai berikut :

- 10 item *true-false* masing-masing dengan bobot 1
- 15 item *multiple choice* masing-masing dengan bobot 3
- 15 item *matching* masing-masing dengan bobot 2

3 item *essay* dengan bobot 5

Buatlah pedoman konversi norma absolute dengan skala lima!

Jawab :

a. Skor maksimal ideal (SMI) dari tes:

$$10 \text{ item } \textit{true-false} = 10 \times 1 = 10$$

$$15 \text{ item } \textit{multiple choice} = 15 \times 3 = 45$$

$$15 \text{ item } \textit{matching} = 15 \times 2 = 30$$

$$3 \text{ item } \textit{essay} = \underline{3 \times 5 = 15}$$

$$\text{Jumlah (SMI)} = 100$$

b. Cari skor mentah pada batas-batas bawah interval sebagai kriteria tertentu :

Penguasaan 90% skor mentahnya adalah :

$$90/100 \times 100 = 90$$

Penguasaan 80% skor mentahnya adalah :

$$80/100 \times 100 = 80$$

Penguasaan 65% skor mentahnya adalah :

$$65/100 \times 100 = 65$$

Penguasaan 55% skor mentahnya adalah :

$$55/100 \times 100 = 55$$

Selanjutnya berdasarkan batas-batas bawah interval dapat dibuat interval nilai, dengan demikian kriteria pedoman konversi norma absolut sebagai berikut :

Skor Mentah	Skor Standar
90-100	A
80-89	B
65-79	C
55-64	D
0-54	E

Dengan menggunakan pedoman konversi tersebut, jika siswa mendapat skor mentah 80 akan mendapat skor standar B, skor mentah 55 akan mendapat skor standar D.

b. Skala Sembilan

Skala sembilan adalah suatu pembagian tingkatan yang terbagi atas sembilan kategori. Masing-masing tingkatan terendah sampai tertinggi dinyatakan dengan angka 1 sampai 9.

Adapun langkah-langkah yang dapat ditempuh dalam mengkonversi nilai skor mentah menjadi skor standar dengan menggunakan norma absolut skala lima adalah sebagai berikut :

- (1). Mencari skor maksimal ideal (SMI) daripada tes yang diberikan.
- (2). Membuat pedoman konversi yang pada umumnya yaitu :

Tingkat Penguasaan	Skor Standar
85%-100%	9
75%-84%	8
65%-74%	7
55%-64%	6
45%-54%	5
35%-44%	4
25%-34%	3
15%-24%	2
0%-14%	1

Contoh :

Misalkan suatu tes hasil belajar terdiri dari item-item sebagai berikut :

10 item *true-false* masing-masing dengan bobot 1

15 item *multiple choice* masing-masing dengan bobot 3

15 item *matching* masing-masing dengan bobot 2

3 item *essay* dengan bobot 5

Buatlah pedoman konversi norma absolute dengan skala sembilan!

Jawab :

a. Skor maksimal ideal (SMI) dari tes:

$$10 \text{ item } \textit{true-false} = 10 \times 1 = 10$$

$$15 \text{ item } \textit{multiple choice} = 15 \times 3 = 45$$

$$\begin{array}{rcl}
 15 \text{ item } \textit{matching} & = 15 \times 2 & = 30 \\
 3 \text{ item } \textit{essay} & = 3 \times 5 = 15 & \\
 \text{Jumlah (SMI)} & & = 100
 \end{array}$$

- b. Cari skor mentah pada batas-batas bawah interval sebagai kriteria tertentu :

Penguasaan 85% skor mentahnya adalah :

$$85/100 \times 100 = 85$$

Penguasaan 75% skor mentahnya adalah :

$$75/100 \times 100 = 75$$

Penguasaan 65% skor mentahnya adalah :

$$65/100 \times 100 = 65$$

Penguasaan 55% skor mentahnya adalah :

$$55/100 \times 100 = 55$$

Penguasaan 45% skor mentahnya adalah :

$$45/100 \times 100 = 45$$

Penguasaan 35% skor mentahnya adalah :

$$35/100 \times 100 = 35$$

Penguasaan 25% skor mentahnya adalah :

$$25/100 \times 100 = 25$$

Penguasaan 15% skor mentahnya adalah :

$$15/100 \times 100 = 15$$

Selanjutnya berdasarkan batas-batas bawah interval dapat dibuat interval nilai, dengan demikian kriteria pedoman konversi norma absolut sebagai berikut :

Skor Mentah	Skor Standar
85-100	9
75-84	8
65-74	7
55-64	6
45-54	5
35-44	4
25-34	3
15-24	2
0-14	1

Dengan menggunakan pedoman konversi tersebut, jika siswa mendapat skor mentah 85 akan mendapat skor standar 9, skor mentah 55 akan mendapat skor standar 6.

c. Skala Sebelas

Skala Sebelas adalah suatu pembagian tingkatan yang terbagi atas sembilan kategori. Masing-masing tingkatan terendah sampai tertinggi dinyatakan dengan angka 0 sampai 10.

Adapun langkah-langkah yang dapat ditempuh dalam mengkonversi nilai skor mentah menjadi skor standar dengan menggunakan norma absolut skala lima adalah sebagai berikut :

- (1). Mencari skor maksimal ideal (SMI) daripada tes yang diberikan.
- (2). Membuat pedoman konversi yang pada umumnya yaitu :

Tingkat Penguasaan	Skor Standar
95%-100%	10
85%-94%	9
75%-84%	8
65%-74%	7
55%-64%	6
45%-54%	5
35%-44%	4
25%-34%	3
15%-24%	2
5%-14%	1
0%-4%	0

Contoh :

Misalkan suatu tes hasil belajar terdiri dari item-item sebagai berikut :

- 10 item *true-false* masing-masing dengan bobot 1
- 15 item *multiple choice* masing-masing dengan bobot 3
- 15 item *matching* masing-masing dengan bobot 2
- 3 item *essay* dengan bobot 5

Buatlah pedoman konversi norma absolute dengan skala sebelas!

Jawab :

a. Skor maksimal ideal (SMI) dari tes:

$$10 \text{ item } \textit{true-false} \quad = 10 \times 1 \quad = 10$$

$$15 \text{ item } \textit{multiple choice} \quad = 15 \times 3 \quad = 45$$

$$15 \text{ item } \textit{matching} \quad = 15 \times 2 \quad = 30$$

$$3 \text{ item } \textit{essay} \quad = \underline{3 \times 5 = 15}$$

$$\text{Jumlah (SMI)} \quad = 100$$

b. Cari skor mentah pada batas-batas bawah interval sebagai kriteria tertentu :

Penguasaan 95% skor mentahnya adalah :

$$95/100 \times 100 = 95$$

Penguasaan 85% skor mentahnya adalah :

$$85/100 \times 100 = 85$$

Penguasaan 75% skor mentahnya adalah :

$$75/100 \times 100 = 75$$

Penguasaan 65% skor mentahnya adalah :

$$65/100 \times 100 = 65$$

Penguasaan 55% skor mentahnya adalah :

$$55/100 \times 100 = 55$$

Penguasaan 45% skor mentahnya adalah :

$$45/100 \times 100 = 45$$

Penguasaan 35% skor mentahnya adalah :

$$35/100 \times 100 = 35$$

Penguasaan 25% skor mentahnya adalah :

$$25/100 \times 100 = 25$$

Penguasaan 15% skor mentahnya adalah :

$$15/100 \times 100 = 15$$

Penguasaan 5% skor mentahnya adalah :

$$5/100 \times 100 = 5$$

Selanjutnya berdasarkan batas-batas bawah interval dapat dibuat interval nilai, dengan demikian kriteria pedoman konversi norma absolut sebagai berikut :

Skor Mentah	Skor Standar
95-100	10
85-94	9
75-84	8
65-74	7
55-64	6
45-54	5
35-44	4
25-34	3
15-24	2
5-14	1
0-4	0

Dengan menggunakan pedoman konversi tersebut, jika siswa mendapat skor mentah 70 akan mendapat skor standar 7, skor mentah 50 akan mendapat skor standar 5.

d. Skala Seratus

Skala seratus adalah skala yang bergerak antara 0 sampai 100. Untuk mengubah skor mentah menjadi skor standar absolut digunakan rumus T-score. Adapun langkah-langkahnya adalah :

- (1). Mencari skor maksimal ideal
- (2). Mencari angka rata-rata ideal

$$M_i = \frac{1}{2} \times SMI$$

Keterangan : M_i = Mean ideal, SMI = Skor Maksimal Ideal

- (3). Mencari standar deviasi ideal

$$SD_i = \frac{1}{3} \times MI$$

Keterangan =
 SD_i = Standar Deviasi

- (4). Mengkonversi skor mentah menjadi skor standar dengan rumus sebagai berikut :

$$T = 50 + \frac{X-M}{SD} \times 10$$

Contoh :

Misalkan suatu tes hasil belajar terdiri dari item-item sebagai berikut :

10 item *true-false* masing-masing dengan bobot 1

15 item *multiple choice* masing-masing dengan bobot 3

15 item *matching* masing-masing dengan bobot 2

3 item *essay* dengan bobot 5

Jika siswa mendapat nilai 80, maka carilah skor standar anak tersebut !

Jawab :

$$\mathbf{SMI = 100}$$

$$\mathbf{M_i = \frac{1}{2} \times 100 = 50}$$

$$\mathbf{SD_i = 1/3 \times 50 = 0,3333 \times 50 = 16,67}$$

$$\mathbf{T = 50 + X - M / SD \times 10}$$

$$\mathbf{T = 50 + 80 - 50 / 16,67 \times 10}$$

$$\mathbf{T = 50 + 30 / 16,67 \times 10}$$

$$\mathbf{T = 50 + 1,7996 \times 10}$$

$$\mathbf{T = 50 + 17,996}$$

$$\mathbf{T = 67,996}$$

e. **Skala Z skor**

Z- skor adalah suatu ukuran yang menyatakan besarnya penyimpangan suatu skor terhadap angka rata-rata skor dalam kelompok tersebut, dalam satuan deviasi standar. Adapun rumus dalam mencari Z skor adalah :

$$Z = \frac{X-M}{SD}$$

Contoh :

Misalkan suatu tes hasil belajar terdiri dari item-item sebagai berikut :

10 item *true-false* masing-masing dengan bobot 1

15 item *multiple choice* masing-masing dengan bobot 3

15 item *matching* masing-masing dengan bobot 2

3 item *essay* dengan bobot 5

Jika siswa mendapat nilai 30, maka carilah Z skor anak tersebut !

Jawab :

$$Z = \frac{X-M}{SD}$$

$$Z = \frac{30-50}{16,67}$$

$$Z = \frac{-20}{16,67}$$

$$Z = -1,19976$$

2. Norma Relatif

Norma relatif disebut juga norma aktuil atau norma empiris atau norma das sein. Umumnya disebut pulan Penilaian Acuan Norma (PAN). Norma relatif adalah suatu norma yang disusun secara relatif berdasarkan distribusi skor yang

dicapai oleh para pengikut dalam suatu tes. Dengan demikian skor standar yang didapat oleh seseorang dalam tes berdasarkan atas norma relatif ini mencerminkan status individu di dalam kelompok.

Dalam teknik penentuan norma relatif kita juga membutuhkan jenis skala seperti yang ada pada norma absolut. Adapun skala tersebut adalah :

(1).Norma relatif skala lima

Langkah-langkah yang harus dilakukan dalam mengkonversi skor mentah menjadi skor standar dengan norma relatif skala lima adalah :

- a. Mencari angka rata-rata (Mean) aktuul dengan rumus statistik
- b. Mencari Standar Deviasi (SD) aktuul dengan rumus statistik
- c. Membuat pedoman konversi norma relatif skala lima.

	→	A
M + 1,5 SD	→	B
M + 0,5 SD	→	C
M - 0,5 SD	→	D
M - 1,5 SD	→	E

Contoh :

Dilakukan Tes Bahasa Arab pada kelas I MI dari 20 siswa. Setelah dilakukan tes diperoleh skor sebagai berikut :

No	Skor
1	70
2	80
3	90
4	50
5	40
6	30
7	50
8	60
9	60
10	50
11	70
12	70
13	50
14	60
15	65
16	56
17	67
18	50
19	30
20	40

Buatlah standar penilaian dengan menggunakan norma relatif !

Jawab :

- a. Mencari angka rata-rata (Mean) aktual dengan rumus statistik
- b. Mencari Standar Deviasi (SD) aktual dengan rumus statistik

Dengan cara tepat dan cepat dapat data tersebut dilakukan pengolahan data dengan program SPSS for Windos. Dengan langkah-langkah sebagai berikut :

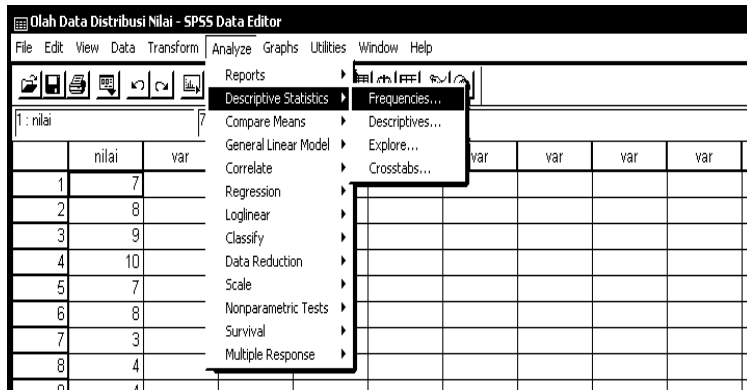
1. Buka Pogram SPSS for Windows dan Klik Variabel View

	Name	Type	Width	Decimals	Label	Values	Missing
1	nilai	Numeric	8	0	Distribusi Nilai	None	None
2							
3							
4							
5							

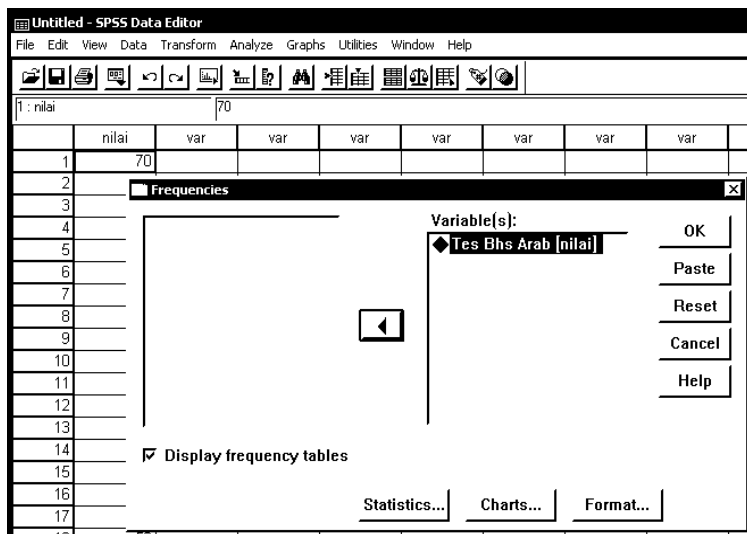
2. Klik data view dan masukkan nilai

	nilai	var	var	var	var	var	var	var	var
1	70								
2	80								
3	90								
4	50								
5	40								

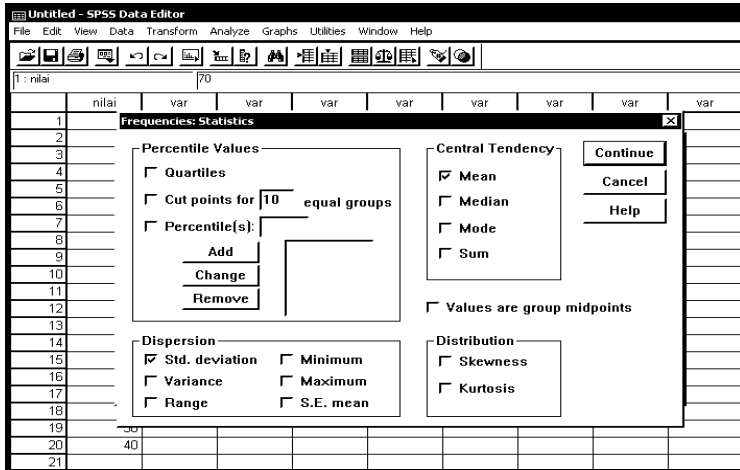
3. Klik Analyze pilih Descriptive Statistics dan pilih Frequencies



4. Klik Frequencies dan pindahkan variable distribusi nilai ke dalam kotak variable.



5. Klik Statistics pilih Central Tendency pilih Mean, dan Dispersion pilih Standar Deviasi klik continue.



6. Klik OK, maka akan muncul hasil olah data SPSS.

Statistics

Tes Bhs Arab

N	Valid	20
	Missing	0
Mean		56,90
Std. Deviation		15,56

Tes Bhs Arab

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 30	2	10,0	10,0	10,0
40	2	10,0	10,0	20,0
50	5	25,0	25,0	45,0
56	1	5,0	5,0	50,0
60	3	15,0	15,0	65,0
65	1	5,0	5,0	70,0
67	1	5,0	5,0	75,0
70	3	15,0	15,0	90,0
80	1	5,0	5,0	95,0
90	1	5,0	5,0	100,0
Total	20	100,0	100,0	

Berdasarkan hasil olah data SPSS ditemukan Mean = **56,90** dan SD = **15,56**

c. Membuat pedoman konversi norma relatif skala lima.

$$M + 1,5 \text{ SD} = 56,90 + 1,5 \times 15,56 = 56,90 + 23,34 = \mathbf{80,24} \quad \longrightarrow \text{A}$$

$$M + 0,5 \text{ SD} = 56,90 + 0,5 \times 15,56 = 56,90 + 7,78 = \mathbf{64,68} \quad \longrightarrow \text{B}$$

$$M - 0,5 \text{ SD} = 56,90 - 0,5 \times 15,56 = 56,90 - 7,78 = \mathbf{49,12} \quad \longrightarrow \text{C}$$

$$M - 1,5 \text{ SD} = 56,90 - 1,5 \times 15,56 = 56,90 - 23,34 = \mathbf{33,56} \quad \longrightarrow \text{D}$$

\(\longrightarrow\) E

Dengan menggunakan pedoman konversi ini, maka jika ada anak mendapat skor mentah 70, maka skor standarnya adalah B.

(2). Norma relatif skala sembilan

Langkah-langkah yang harus dilakukan dalam mengkonversi skor mentah menjadi skor standar dengan norma relatif skala sembilan adalah :

- a. Mencari angka rata-rata (Mean) aktual dengan rumus statistik
- b. Mencari Standar Deviasi (SD) aktual dengan rumus statistik
- c. Membuat pedoman konversi norma relatif skala sembilan.

	→	9
M + 1,75 SD	→	8
M + 1,25 SD	→	7
M + 0,75 SD	→	6
M + 0,25 SD	→	5
M - 0,25 SD	→	4
M - 0,75 SD	→	3
M - 1,25 SD	→	2
M - 1,75 SD	→	1

Contoh :

Misal soal sama

Buatlah standar penilaian norma relatif dengan skala sembilan! **Jawab :**

- Mencari angka rata-rata (Mean) aktuil dengan rumus statistik = 56,90
- Mencari Standar Deviasi (SD) aktuil dengan rumus statistik = 15,56
- Membuat pedoman konversi norma relatif skala sembilan.

$$\begin{aligned} & \longrightarrow 9 \\ M + 1,75 SD &= 56,90 + 1,75 \times 15,56 = 56,90 + 27,23 = \\ & \mathbf{84,13} \end{aligned}$$

$$\begin{aligned} & \longrightarrow 8 \\ M + 1,25 SD &= 56,90 + 1,25 \times 15,56 = 56,90 + 19,45 = \\ & \mathbf{76,35} \end{aligned}$$

$$\begin{aligned} & \longrightarrow 7 \\ M + 0,75 SD &= 56,90 + 0,75 \times 15,56 = 56,90 + 11,67 = \\ & \mathbf{68,57} \end{aligned}$$

$$\begin{aligned} & \longrightarrow 6 \\ M + 0,25 SD &= 56,90 + 0,25 \times 15,56 = 56,90 + 3,89 = \\ & \mathbf{60,79} \end{aligned}$$

$$\begin{aligned} & \longrightarrow 5 \\ M - 0,25 SD &= 56,90 - 0,25 \times 15,56 = 56,90 - 3,89 \\ & = \mathbf{53,01} \end{aligned}$$

$$\begin{aligned} & \longrightarrow 4 \\ M - 0,75 SD &= 56,90 - 0,75 \times 15,56 = 56,90 - 11,67 \\ & = \mathbf{45,23} \end{aligned}$$

$$\begin{aligned} & \longrightarrow 3 \\ M - 1,25 SD &= 56,90 - 1,25 \times 15,56 = 56,90 - 19,45 \\ & = \mathbf{37,45} \end{aligned}$$

$$\begin{aligned} & \longrightarrow 2 \\ M - 1,75 SD &= 56,90 - 1,75 \times 15,56 = 56,90 - 27,23 \\ & = \mathbf{29,67} \end{aligned}$$

→ 1

Dengan menggunakan pedoman konversi ini, maka jika ada anak mendapat skor mentah 30, maka skor standarnya adalah 2

(3). Norma relatif skala sebelas

Langkah-langkah yang harus dilakukan dalam mengkonversi skor mentah menjadi skor standar dengan norma relatif skala sebelas adalah :

- Mencari angka rata-rata (Mean) aktual dengan rumus statistik
- Mencari Standar Deviasi (SD) aktual dengan rumus statistik
- Membuat pedoman konversi norma relatif skala sebelas.

$$M + 2,25 \text{ SD} \rightarrow 10$$

$$M + 1,75 \text{ SD} \rightarrow 9$$

$$M + 1,25 \text{ SD} \rightarrow 8$$

$$M + 0,75 \text{ SD} \rightarrow 7$$

$$M + 0,25 \text{ SD} \rightarrow 6$$

$$M \rightarrow 5$$

$$M - 0,25 \text{ SD} \rightarrow 4$$

$$M - 0,75 \text{ SD} \rightarrow 3$$

$$M - 1,25 \text{ SD} \rightarrow 2$$

$$M - 1,75 \text{ SD}$$

$$\begin{array}{r}
 \longrightarrow 1 \\
 M - 2,25 \text{ SD} \\
 \longrightarrow 0
 \end{array}$$

Contoh :

Misal soal sama

Buatlah standar penilaian norma relatif dengan skala sebelas!

Jawab :

- Mencari angka rata-rata (Mean) aktuil dengan rumus statistik = 56,90
- Mencari Standar Deviasi (SD) aktuil dengan rumus statistik = 15,56
- Membuat pedoman konversi norma relatif skala sembilan.

$$\begin{array}{r}
 \longrightarrow 10 \\
 M + 2,25 \text{ SD} = 56,90 + 2,25 \times 15,56 = 56,90 + 35,01 = \\
 \mathbf{91,91}
 \end{array}$$

$$\begin{array}{r}
 \longrightarrow 9 \\
 M + 1,75 \text{ SD} = 56,90 + 1,75 \times 15,56 = 56,90 + 27,23 = \\
 \mathbf{84,13}
 \end{array}$$

$$\begin{array}{r}
 \longrightarrow 8 \\
 M + 1,25 \text{ SD} = 56,90 + 1,25 \times 15,56 = 56,90 + 19,45 = \\
 \mathbf{76,35}
 \end{array}$$

$$\begin{array}{r}
 \longrightarrow 7 \\
 M + 0,75 \text{ SD} = 56,90 + 0,75 \times 15,56 = 56,90 + 11,67 = \\
 \mathbf{68,57}
 \end{array}$$

$$\begin{array}{r}
 \longrightarrow 6 \\
 M + 0,25 \text{ SD} = 56,90 + 0,25 \times 15,56 = 56,90 + 3,89 = \\
 \mathbf{60,79}
 \end{array}$$

$$\begin{array}{r}
 \longrightarrow 5 \\
 M - 0,25 \text{ SD} = 56,90 - 0,25 \times 15,56 = 56,90 - 3,89 \\
 = \mathbf{53,01}
 \end{array}$$

$$M - 0,75 \text{ SD} = 56,90 - 0,75 \times 15,56 = 56,90 - 11,67 \\ = \mathbf{45,23} \quad \longrightarrow 4$$

$$M - 1,25 \text{ SD} = 56,90 - 1,25 \times 15,56 = 56,90 - 19,45 \\ = \mathbf{37,45} \quad \longrightarrow 3$$

$$M - 1,75 \text{ SD} = 56,90 - 1,75 \times 15,56 = 56,90 - 27,23 \\ = \mathbf{29,67} \quad \longrightarrow 2$$

$$M - 2,25 \text{ SD} = 56,90 - 2,25 \times 15,56 = 56,90 - 35,01 \\ = \mathbf{21,89} \quad \longrightarrow 1$$

$$\longrightarrow 0$$

Dengan menggunakan pedoman konversi ini, maka jika ada anak mendapat skor mentah 40, maka skor standarnya adalah 3.

(4). Norma relatif skala seratus

Skala seratus adalah skala yang bergerak antara 0 sampai 100. Untuk mengubah skor mentah menjadi skor standar absolut digunakan rumus T-score.

Adapun langkah-langkahnya adalah :

(1). Mencari angka rata-rata dengan rumus statistik

$$M_a = \Sigma X_i / N$$

Keterangan : M_a = Mean aktual, ΣX_i = Jumlah Skor, N = Jumlah Kasus

(2). Mencari standar deviasi

$$SD_a = \sqrt{\frac{\sum(X_i - M_{ex})^2}{(n - 1)}} \quad \begin{array}{l} \text{(rumus standar deviasi} \\ \text{sampel)} \end{array}$$

Keterangan =

SD_a = Standar Deviasi Aktuil

(3). Mengkonversi skor mentah menjadi skor standar dengan rumus sebagai berikut :

$$T = 50 + X - M / SD \times 10$$

Contoh :

- a. Mencari angka rata-rata (Mean) aktuil dengan rumus statistik = 56,90
- b. Mencari Standar Deviasi (SD) aktuil dengan rumus statistik = 15,56
- c. Memasukkan nilai mean dan standar deviasi ke dalam T Skor , misal anak siswa X = 60 :

$$T = 50 + X - M / SD \times 10$$

$$T = 50 + 60 - 56,90 / 15,56 \times 10$$

$$T = 50 + 3,1 / 15,56 \times 10$$

$$T = 50 + 0,1992 \times 10$$

$$T = 50 + 1,992$$

$$T = 51,992$$

(5). Norma relatif dengan Z skor

Rumus Z skor norma relatif sama saja dengan rumus Z skor norma absolut. Misalnya apabila distribusi skor yang dicapai oleh anak-anak sama dengan contoh di atas, maka Z skor anak mendapat skor 60, dapat dicari sebagai berikut :

$$Z = X - M / SD$$

$$\begin{aligned} Z &= X - M / SD \\ Z &= 60 - 56,90 / 15,56 \\ Z &= 3,1 / 15,56 \\ Z &= 0,1992 \end{aligned}$$

Jadi Z skor untuk anak yang mendapat skor 60 adalah 0,1992.

3. Kelebihan dan Kelemahan Norma Absolut dan Relatif

Norma absolut dan relatif yang telah dibicarakan diatas mempunyai segi kebaikan dan kelemahan masing-masing. Norma absolut baik dipergunakan apabila derajat kesukaran daripada tes yang dipergunakan betul-betul telah memenuhi syarat tes yang baik. Misalnya pada tes telah distandarisasikan, dan direvisi berdasarkan analisa empiris dan rasio yang memadai. Jika tidak maka penggunaan norma absolut akan menggambarkan kurang tepat. Misanya jika soal itu terlalu sukar, maka siswa yang pandai pun akan mendapat nilai yang rendah, dan sebaliknya jika soal itu terlalu mudah, maka siswa yang bodoh pun akan mendapat nilai yang tinggi.

Norma relatif baik dipergunakan untuk distribusi kecakapan anak-anak dalam tes berdasarkan kurve normal, apabila distribusi kecakapan anak itu tidak mengikuti hukum kurve normal, maka norma relatif tidak akan mampu menggambarkan secara obyektif. Jika soal tes diberikan pada anak-anak yang bodoh, maka anak-anak boddoh akan mendapatkan nilai yang tinggi, sebaliknya anak-anak yang

pandai, maka sebagian anak pandai akan terdapat sebagian anak pandai mendapatkan nilai yang rendah.

4. Norma Kombinasi

Norma kombinasi adalah suatu norma yang didasarkan atas kombinasi antara norma absolut dengan norma relatif. Pada norma kombinasi juga dapat dibedakan menjadi lima jenis skala yaitu skala lima, skala sembilan, skala sebelas, skala seratus, dan z skor. Ancer-ancer yang dipergunakan untuk mengkonversikan skor mentah menjadi skor standar pada tiap jenis skala, adalah kriteria rata-rata antara norma absolut dengan norma relatif. Dalam bab ini penulis akan membicarakan pada skala lima dan T skor.

a. Norma Kombinasi Skala Lima

Prosedur yang harus ditempuh dalam mengkonversikan skor mentah menjadi skor standar dengan norma kombinasi skala lima adalah :

1. Mencari batas-batas bawah interval kriteria norma absolut
2. Mencari batas-batas bawah interval kriteria norma relatif
3. Mencari rata-rata batas bawah interval kriteria norma absolut dengan norma relatif
4. Membuat pedoman konversi dengan menggunakan rata-rata konversi antara norma absolut dan norma relatif.

Contoh:

Misalkan suatu tes bahasa Arab hasil belajar terdiri dari item-item sebagai berikut :

10 item *true-false* masing-masing dengan bobot 1

- 15 item *multiple choice* masing-masing dengan bobot 3
- 15 item *matching* masing-masing dengan bobot 2
- 3 item *essay* dengan bobot 5

Dari soal tersebut dilakukan tes bahasa Arab kelas I MI dari 20 siswa.

Buatlah pedoman konversi norma kombinasi skala lima!

Jawab :

(1). Mencari batas-batas bawah interval kriteria norma absolut

a. Skor maksimal ideal (SMI) dari tes:

10 item <i>true-false</i>	= 10 x 1	= 10
15 item <i>multiple choice</i>	= 15 x 3	= 45
15 item <i>matching</i>	= 15 x 2	= 30
3 item <i>essay</i>	<u>= 3 x 5 = 15</u>	
Jumlah (SMI)		= 100

b. Cari skor mentah pada batas-batas bawah interval sebagai kriteria tertentu :

Penguasaan 90% skor mentahnya adalah :

$$90/100 \times 100 = 90$$

Penguasaan 80% skor mentahnya adalah :

$$80/100 \times 100 = 80$$

Penguasaan 65% skor mentahnya adalah :

$$M + 0,5 SD = 56,90 + 0,5 \times 15,56 = 56,90 + 7,78 = \mathbf{64,68}$$

→ C

$$M + 0,5 SD = 56,90 - 0,5 \times 15,56 = 56,90 - 7,78 = \mathbf{49,12}$$

→ D

$$M - 1,5 SD = 56,90 - 1,5 \times 15,56 = 56,90 - 23,34 = \mathbf{33,56}$$

→ E

(3). Mencari rata-rata batas-batas bawah interval kriteria norma absolut dan relatif.

$$\frac{1}{2} \times (90 + 80,24) = \frac{1}{2} \times 170,24 = 85,12$$

$$\frac{1}{2} \times (80 + 64,68) = \frac{1}{2} \times 144,68 = 72,34$$

$$\frac{1}{2} \times (65 + 49,12) = \frac{1}{2} \times 114,12 = 57,06$$

$$\frac{1}{2} \times (55 + 33,56) = \frac{1}{2} \times 88,56 = 44,28$$

(4). Membuat skala lima norma kombinasi dapat disusun sebagai berikut :

$$85,12 \quad \rightarrow \quad \mathbf{A}$$

$$72,34 \quad \rightarrow \quad \mathbf{B}$$

$$57,06 \quad \rightarrow \quad \mathbf{C}$$

$$44,28 \quad \rightarrow \quad \mathbf{D}$$

44,28

→ E

Dengan menggunakan pedoman konversi ini, maka siswa yang mendapatkan skor mentah 60 berada pada skor standar C.

b. Norma Kombinasi T Skor

Prosedur yang ditempuh untuk mengkonversikan skor mentah menjadi skor standar dengan norma kombinasi dengan T skor adalah sebagai berikut :

- (1). Mencari angka rata-rata kombinasi yaitu rata-rata mean ideal dan aktuil.
- (2). Mencari angka rata-rata kombinasi yaitu rata-rata SD ideal dan aktuil
- (3). Mengkonversi tiap-tiap skor mentah menjadi skor standar dengan rumus T skor.

Contoh : Seperti pada contoh diatas

- (a). Mencari angka rata-rata kombinasi yaitu rata-rata mean ideal dan aktuil.
- (b). Mencari angka rata-rata kombinasi yaitu rata-rata SD ideal dan aktuil.

$$M_i = \frac{1}{2} \times SMI$$

$$SD_i = \frac{1}{3} \times SMI$$

$$\mathbf{SMI = 100}$$

$$\mathbf{M_i = \frac{1}{2} \times 100 = 50}$$

$$\mathbf{SD_i = \frac{1}{3} \times 50 = 0,3333 \times 50 = 16,67}$$

$$M_a = \frac{\sum X_i}{N}$$

$$SD_a = \sqrt{\frac{\sum (X_i - M_{ex})^2}{(n - 1)}}$$

Mencari angka rata-rata (Mean) aktuil dengan rumus statistik = 56,90

Mencari Standar Deviasi (SD) aktuil dengan rumus statistik = 15,56

Dengan hasil angka rata-rata dan standar deviasi diatas, maka dapat dicari angka rata-rata kombinasi yaitu rata-rata SD ideal dan aktuil sebagai berikut :

$$M_k = \frac{1}{2} \times (M_i + M_a)$$

$$M_k = \frac{1}{2} \times (50 + 56,90)$$

$$M_k = \frac{1}{2} \times (50 + 56,90)$$

$$M_k = \frac{1}{2} \times 106,9$$

$$M_k = 53,45$$

$$SD_k = \frac{1}{2} \times (SD_i + SD_a)$$

$$SD_k = \frac{1}{2} \times (16,67 + 15,56)$$

$$SD_k = \frac{1}{2} \times (32,23)$$

$$SD_k = 16,115$$

- (c). Membuat konversi nilai tiap-tiap skor mentah menjadi skor standar, dengan T Skor. Misal: skor mentah 65.

$$T = 50 + X-M/SD \times 10$$

$$T = 50 + X-M/SD \times 10$$

$$T = 50 + 65 - 53,45/16,115 \times 10$$

$$T = 50 + 11,55/16,115 \times 10$$

$$T = 50 + 0,7167 \times 10$$

$$T = 50 + 7,167$$

$$T = 57,167$$

C. Distribusi Nilai

Distribusi nilai yang dimiliki oleh siswa-siswanya dalam suatu kelas berdasarkan pada dua macam standar yaitu :

1. Standar Mutlak

Dengan dasar bahwa hasil belajar siswa dibandingkan dengan sebuah standar mutlak atau dalam hal ini skor tertinggi yang diharapkan, maka tingkat penguasaan siswa akan terlihat dalam berbagai bentuk kurva. Bentuk kurva tersebut sangat dipengaruhi oleh tingkat kesulitan yang dialami oleh siswa dalam mengerjakan soal-soal tes. Jika soal terlalu mudah, maka siswa akan banyak memperoleh nilai antara 8, 9 dan 10. Sebaliknya jika soal terlalu sulit, maka siswa akan banyak memperoleh nilai antara 1, 2, 3 dan 4.

Untuk lebih jelasnya dalam aplikasi tersebut dapat dilihat pada contoh sebagai berikut :

Dilakukan tes pada mata pelajaran bahasa Arab dari 20 siswa. Setelah dilakukan tes diketemukan data sebagai berikut :

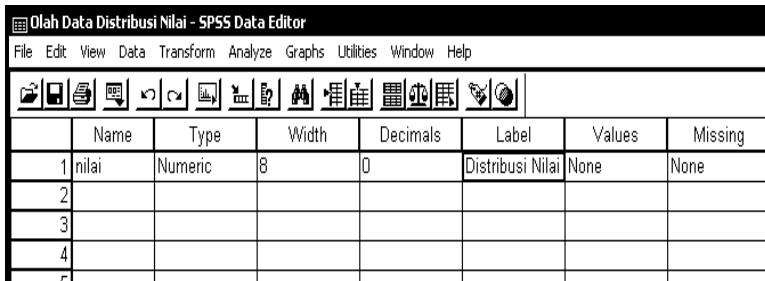
7	8	9	10	7	8	3	4	4	5
6	7	6	7	1	2	6	5	5	7

Buatlah grafis distribusi Nilai !

Jawab :

Langkah-langkahnya yaitu :

- i. Buka Program SPSS for Windows dan Klik Variabel View



- ii. Klik data view dan masukkan nilai

Olah Data Distribusi Nilai - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1.: nilai 7

	nilai	var	var	var	var	var	var	var	var
1	7								
2	8								
3	9								
4	10								
5	7								
6	8								

- iii. Klik Analyze pilih Descriptive Statistics dan pilih Frequencies

Olah Data Distribusi Nilai - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1.: nilai 7

	nilai	var	var	var	var	var	var	var	var
1	7								
2	8								
3	9								
4	10								
5	7								
6	8								
7	3								
8	4								
9	4								

Reports

Descriptive Statistics > **Frequencies...**

Compare Means > Descriptives...

General Linear Model > Explore...

Correlate > Crosstabs...

Regression

Loglinear

Classify

Data Reduction

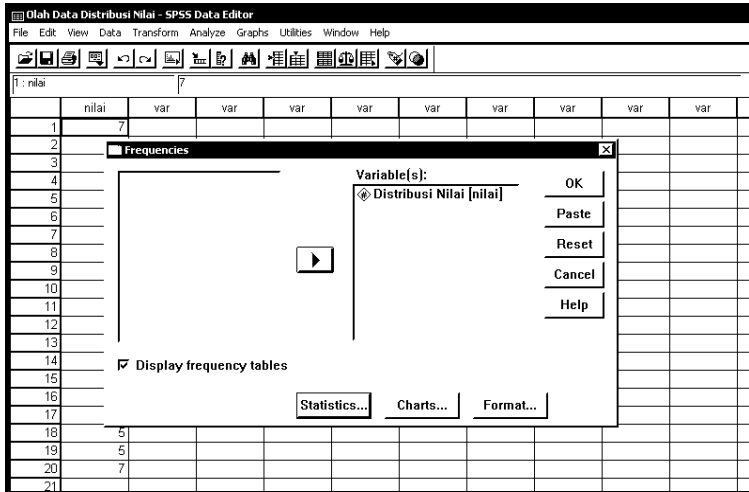
Scale

Nonparametric Tests

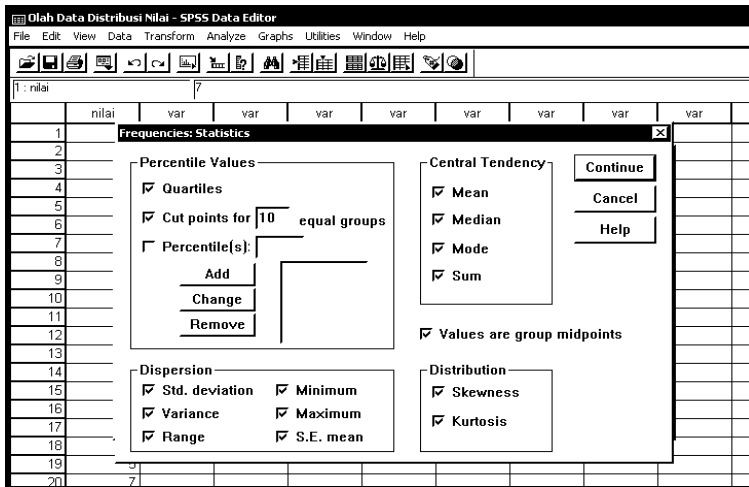
Survival

Multiple Response

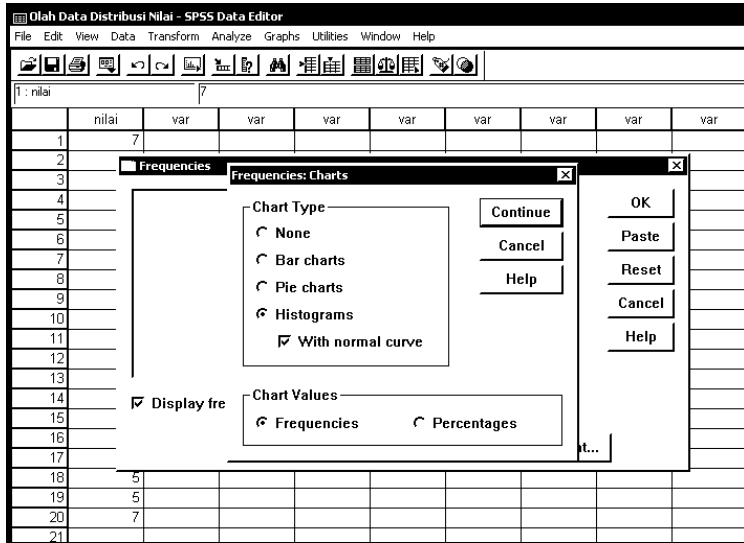
- iv. Klik Frequencies dan pindahkan variable distribusi nilai ke dalam kotak variable.



- v. Klik Statistics pilih Central Tendency, Distribution, Dispersion dan Percentile Values klik continue.



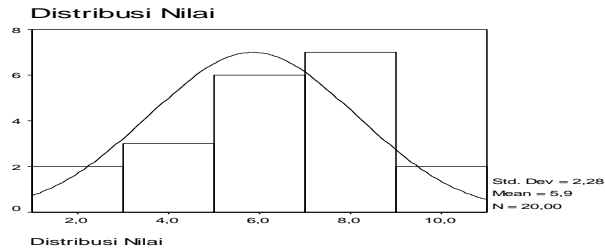
- vi. Klik Charts dan pilih Histograms with normal curve dan klik continue



vii. Klik OK, maka akan muncul hasil olah data SPSS.

Distribusi Nilai

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1	5,0	5,0	5,0
	2	1	5,0	5,0	10,0
	3	1	5,0	5,0	15,0
	4	2	10,0	10,0	25,0
	5	3	15,0	15,0	40,0
	6	3	15,0	15,0	55,0
	7	5	25,0	25,0	80,0
	8	2	10,0	10,0	90,0
	9	1	5,0	5,0	95,0
	10	1	5,0	5,0	100,0
	Total	20	100,0	100,0	



Statistics

Distribusi Nilai		
N	Valid	20
	Missing	0
Mean		5,85
Std. Error of Mean		,51
Median		6,13 ^a
Mode		7
Std. Deviation		2,28
Variance		5,19
Skewness		-,389
Std. Error of Skewness		,512
Kurtosis		-,020
Std. Error of Kurtosis		,992
Range		9
Minimum		1
Maximum		10
Sum		117
Percentiles		
	10	2,50 ^b
	20	4,00
	25	4,40
	30	4,80
	40	5,50
	50	6,13
	60	6,63
	70	7,14
	75	7,43
	80	7,71
	90	8,67

a. Calculated from grouped data.

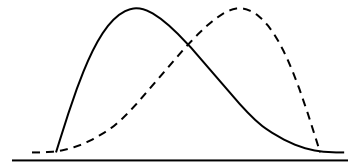
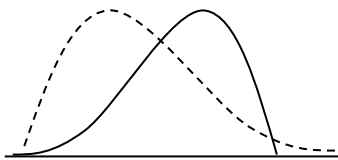
b. Percentiles are calculated from grouped data.

2. Distribusi Nilai Berdasarkan Standar Relatif

Standar relatif atau *norm-referenced*, kedudukan seorang selalu dibandingkan dengan kawan-kawannya dalam kelompok. Standar relatif ini tanpa menggunakan ketentuan apakah distribusi skor terletak dalam kurva juling positif atau juling negatif tetapi dalam *norm-referenced* selalu tergambar dalam kurva normal.

Asumsi kurva normal tersebut dibangun berdasarkan ketentuan :

- a. Apabila distribusi skor tergambar dalam kurva juling positif, yang kurang sempurna adalah soal-soal tesnya, yaitu terlalu sukar. Dengan demikian nilai siswa lalu direntangkan diantara nilai tertinggi dan nilai terendah, dengan sebagian besar terletak pada nilai sedang.
- b. Apabila distribusi skor tergambar dalam kurva juling negatif. Dalam ubahan menjadi nilai, disebar sedemikian rupa sehingga menjadi kurva normal, dengan nilai sedang adalah nilai yang paling banyak.



Keterangan =

___ = nilai berdasarkan standar mutlak

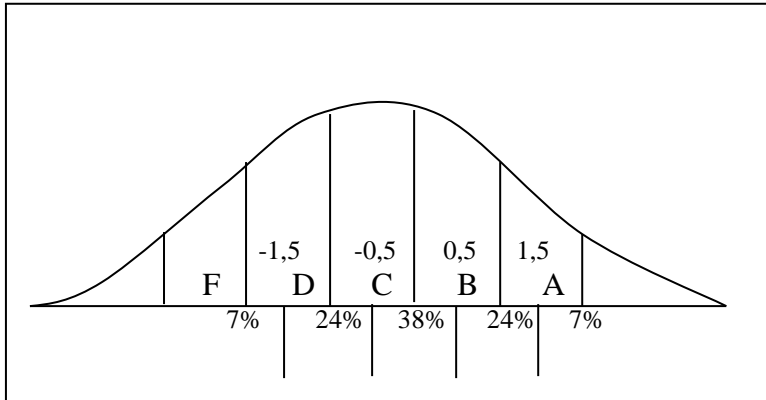
-- = nilai berdasarkan standar relatif

3. Standar Nilai

Nilai yang telah kita olah dalam bentuk distribusi frekuensi dapat dilakukan interpretasi dengan standar nilai yang telah ditetapkan. Ada beberapa pendapat mengenai standar nilai sebagaimana yang dikemukakan oleh para ahli yaitu :

a. Standar lima

Gronlund mengemukakan penyebaran nilai dengan angka, juga mengemukakan penyebaran nilai dengan huruf yang dapat digambarkan pada kurve normal sebagai berikut :



Dikutip dari : *Gronlund Improving Marking and Reporting in Classroom Instruction: hal 26.*

Pada gambar diatas digunakan symbol huruf F bukan E singkatan dari *Fail* (gagal). Selanjutnya rentangan persentase hanya berlaku bagi poplasi yang sangat heterogen. Apabila populasi telah terseleksi akibat kenaikan kelas atau pindah ke tingkat sekolah yang lebih tinggi, maka golongan F yang berada di ekor kiri akan berangsur-angsur berkurang, sehingga distribusinya dapat berbentuk menjadi :

- A 10 sampai 20 %
- B 20 sampai 30 %
- C 40 sampai 50 %
- D 10 sampai 20 %
- F 0 sampai 20 %

b. Standar sembilan

Skor-skor siswa direntangkan menjadi 9 nilai (disebut juga *Standar Nines* atau *Stanines*). Atas dasar inilah semua situasi skor siswa dapat direntangkan antara nilai 1-9 seperti pada tabel berikut ini :

Stanines		Interpretasi
(9)	4%	Tinggi (4%)
(8)	7%	Di atas rata-rata (19%)
(7)	12%	
(6)	17%	Rata-rata (54%)
(5)	20%	
(4)	17%	
(3)	12%	Dibawah rata-rata (19%)
(2)	7%	
(1)	4%	Rendah (4%)

BAB XII

MENILAI INTELEGENSI ATAU BAKAT SEKOLAH

A. Definisi Intelegensi

Setiap orang mempunyai intelegensi yang berbeda-beda. Perbedaan Intelegensi tersebut akan mampu mempengaruhi seseorang dalam menangkap respon apa yang dilihat, dipelajari dan dianalisis.

Bagi seorang guru mengetahui bakat sekolah (intelegensi) merupakan hal yang sangat penting. Pemahaman seorang guru terkait dengan karakteristik peserta didik sangat membantu dalam efektifitas pelayanan pembelajaran di kelas. Peserta didik yang sudah diketahui sejak dini tingkat intelegensinya akan mempermudah dalam memberikan arahan, dorongan dalam kegiatan pembelajaran sebagai wahana untuk menemukan jati dirinya sesuai dengan potensi yang diberikan oleh Allah. Swt baik potensi mental maupun fisik.

Sebelum kita membicarakan masalah penilaian inteligensi atau bakat sekolah (*scholastic aptitude*), maka terlebih dahulu perlu kita bicarakan apakah inteligensi itu ?

Ada beberapa definisi inteligensi yang telah dikemukakan oleh para ahli ilmu jiwa, tetapi antara definisi yang satu dengan definisi yang lain terdapat perbedaan-perbedaan.

William Stern mengatakan bahwa Inteligensi adalah suatu kapasitas yang bersifat umum (*general capacity*) daripada individu untuk mengadakan penyesuaian terhadap situasi yang baru atau suatu problem yang dihadapi.

George Studard menginterpretasikan inteligensi sebagai kecakapan untuk melaksanakan suatu aktivitas yang ditandai oleh :

- (1). kesukaran,
- (2). kekomplekan,
- (3). keabstrakan,
- (4). ekonomis,
- (5). penyesuaian ke arah tujuan,
- (6). mempunyai nilai sosial dan
- (7). bersifat asli.

Thorndike menitik beratkan inteligensi sebagai kesanggupan untuk mengadakan respon yang baik sesuai dengan fakta yang dihadapi. Terman mendefinisikan inteligensi sebagai kecakapan untuk berpikir secara abstrak. (Crow and Crow,. General Psychologi, halaman 115).

Dengan melihat dari berbagai macam pendapat para ahli, maka intelegensi dapat didefinisikan sebagai kecapan untuk merespon yang baik dengan cara berfikir dan beraktivitas berkaitan dengan fakta-fakta yang dihadapi dalam bentuk kesukaran, kekomplekan, keabsatrakan, ekonomis, penyesuaian ke arah tujuan, mempunyai nilai sosial dan bersifat asli.

B. Teori-Teori Intelegensi

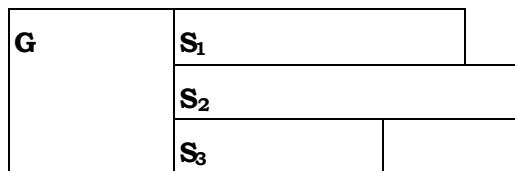
Ada beberapa teori intelegensi yang dikembangkan oleh para ahli. Dari sekian ahli yang dapat dikemukakan sebagai berikut :

(1).Teori dwi faktor dari Spearman

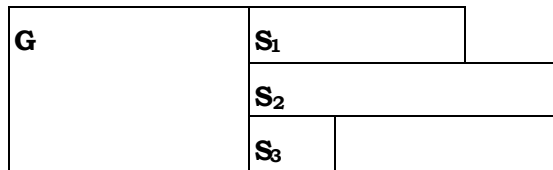
Teori ini dikembangkan oleh seorang ahli bangsa Inggris yang bernama : Charles Spearman dengan mempergunakan metode analisa faktor menemukan bahwa inteligensi terdiri dari dua faktor, yaitu faktor umum atau general faktor (disingkat faktor G) dan faktor khusus atau spesial faktor (disingkat sebagai faktor S). Faktor G itu berfungsi pada semua tingkah laku atau kecakapan sedangkan faktor S hanya berfungsi pada suatu tingkah laku atau kecakapan saja. (Witherington, Psych.ologi Pendidikan, p. 52).

Teori Spearman tentang dwi --faktor tersebut dapat dilukiskan sebagai berikut :

A



B



Dengan memperhatikan gambar diatas, maka dapat diambil kesimpulan bahwa :

1. A mempunyai faktor yang lebih kecil daripada B.
2. A dan B mempunyai kecakapan sama terlihat dari $G + S_1$ pada A sama dengan $G + S_1$ pada B.

3. Kapasitas A dalam kecakapan 2 lebih rendah daripada B.
4. Kapasitas A dalam kecakapan 3 lebih tinggi daripada B.

(2). Teori Burt

Pendirian Cyrill Burt sangat dekat dengan pendirian Spearman. Ia sependapat dengan Spearman, bahwa pada manusia terdapat faktor G yang mendasari semua tingkah lakunya, dan faktor S yang mendasari suatu tingkah laku tertentu. Tetapi disamping kedua faktor itu menurut Burt masih ada lagi faktor yang ketiga, yaitu faktor kelompok atau cluster faktor (disingkat sebagai faktor I). Faktor G adalah faktor yang berfungsi pada sejumlah tingkah laku. Ini berarti bahwa faktor G itu lebih luas dan faktor S tetapi lebih sempit dari faktor G.

Jadi menurut Burt tiap tingkah laku itu dimungkinkan oleh dua faktor yaitu faktor G faktor dan faktor S.

(3). Teori Thurstone

Thurstone sependapat dengan Burt bahwa ada faktor C1 yang berfungsi pada sejumlah tingkah laku, juga sependapat dengan Burt mengenai faktor S. Faktor S jumlahnya banyak sekali sebanyak tingkah laku khusus yang dilakukan oleh manusia bersangkutan, Akan tetapi mengenai faktor G Thurstone monolaknya. Ia berpendapat bahwa faktor G itu tidak ada. Jadi hanya ada dua faktor saja yaitu faktor C dan faktor S. Adapun faktor C menurut Thurstone ada 7 jenis yaitu

- a. Faktor ingatan atau *Memory* (M) yaitu kapasitas untuk mengingat.

- b. Faktor bahasa atau *Verbal* (V) yaitu kapasitas untuk menggunakan bahasa.
- c. Faktor bilangan atau *Numerical* (N) yaitu kapasitas untuk bekerja dengan bilangan.
- d. Faktor kelancaran berkata-kata atau *word fluency* (W), yaitu kapasitas untuk berbicara secara lancar.
- e. Faktor penalaran atau *Reasoning* (R) yaitu faktor yang mendasari berpikir logis.
- f. Faktor pengamatan atau *Perceptual* (P) yaitu kapasitas untuk mengamati dengan cepat dan tepat.
- g. Faktor ruang atau *Spatial* (S), yaitu kapasitas untuk mengadakan orientasi dalam ruang.

(4).Teori Guilford.

Guilford sependapat dengan Thurstone. yaitu bahwa yang pokok adalah faktor C. Bahkan pada hakekatnya hanya inilah faktor-faktor inteligensi itu. Apa yang sebenarnya disebut faktor G oleh Spearman pada hakekatnya adalah kesatuan (pooling) dari faktor-faktor C yang terpisah-pisah (*separate*). Menurut Guilford faktor C itu banyaknya tidak hanya 7. melainkan 40 macam. (Sumadi Suryabrata, Psikologi, Diagnostik, halaman 29 32).

C. SEJARAH PENGUKURAN INTELIGENSI

Sebelum abad ke 19 belum ada usaha-usaha untuk mengadakan pengukuran inteligensi. Sebab pada waktu itu pengajaran diberikan secara individual, sehingga tidak timbul masalah-masalah mengenai perbedaan-perbedaan individual dalam kecakapan. Setelah pengajaran itu semakin meluas dan mulai dipraktikkannya sistem klasikal maka mulailah

orang merasakan adanya masalah-masalah perbedaan individual antara murid-murid. Guru yang mengajar dalam satu kelas akan menghadapi sejumlah anak dengan kapasitas yang berbeda-beda. Ada anak yang mempunyai kapasitas yang cukup tinggi, ada anak yang mempunyai kapasitas yang sedang, ada anak yang mempunyai kapasitas yang rendah, bahkan ada anak yang mempunyai kapasitas yang sangat rendah sehingga tidak bisa mengikuti pelajaran di sekolah dasar biasa. Jadi harus dididik dalam sekolah luar biasa. Anak yang mempunyai kapasitas yang sangat rendah tersebut disebut anak yang lemah jiwa.

Untuk dapat membedakan, mana anak-anak yang normal yang dapat mengikuti pelajaran di sekolah dasar biasa, dan mana anak-anak yang lemah jiwa yang harus diberikan pelajaran dalam sekolah khusus, maka perlulah adanya suatu alat pengukur. Dengan demikian maka mulailah dirasakan adanya kebutuhan untuk mendapatkan alat yang dapat membedakan antara kedua golongan anak-anak tersebut.

Salah seorang tokoh yang sangat tertarik untuk menyelidiki masalah perbedaan individual ini adalah seorang ahli bangsa Perancis bernama Alfred Binet. Sejak tahun 1890 ia telah mengadakan percobaan-percobaan untuk menemukan alat yang dapat digunakan untuk mengukur perbedaan-perbedaan individual.

Pada tahun 1904 Menteri Pendidikan dan Pengajaran Perancis dihadapkan pada suatu masalah untuk membedakan antara anak-anak yang normal dengan anak-anak yang lemah jiwa. Untuk mengatasi masalah ini maka ia menugaskan Alfred Binet untuk menyusun alat yang dimaksudkan. Demikianlah Alfred Binet dengan dibantu oleh Theodore Simon bekerja keras selama satu tahun, sehingga pada tahun 1905 telah dapat diterbitkan tes yang pertama. Tes

inilah yang kemudian terkenal dengan sebutan tes Binet-Simon.

yang pertama. Tes ini terdiri dari sejumlah tugas/pertanyaan. Anak-anak yang dapat menjawab dengan betul sejumlah pertanyaan tertentu dapat digolongkan sebagai anak yang normal. Sedangkan apabila syarat ini tidak dipenuhi maka anak tersebut digolongkan anak lemah jiwa.

Tes yang diterbitkan pada tahun 1905 itu masih merupakan tes yang sederhana. Binet dan Simon terus berusaha untuk menyempurnakan hasil karyanya. Pada tahun 1908 dapat diterbitkan suatu tes yang merupakan pembaharuan daripada tes yang diterbitkan dalam tahun 1905. Tes-tes itu dibagi-bagi menjadi beberapa golongan menurut umur. Apabila seorang anak dapat menjawab suatu tes yang termasuk ke dalam golongan umur 5 tahun maka dikatakan bahwa umur psikhis anak tersebut adalah 5 tahun

Berdasarkan perbedaan antara umur kronologis anak yang bersangkutan dengan umur psikhisnya, maka anak-anak dibedakan menjadi tiga golongan yaitu :

- a. Golongan superior, yaitu apabila umur psikhis yang dicapai lebih tinggi dua tahun atau lebih dari umur kronologisnya.
- b. Golongan anak normal, yaitu apabila umur psikhis yang dicapai sama atau selisih satu tahun dengan umur kronologisnya.
- c. Golongan inferior, yaitu apabila umur psikhis yang dicapai dua tahun lebih rendah dari umur kronologisnya.

Tes yang diterbitkan oleh Binet Simon tahun 1908 tersebut banyak menarik perhatian para ahli, baik di Eropa maupun di Amerika. Mereka mencobakan tes Binet Simon di negara nya masing-masing, kemudian melaporkan hasil-hasil yang mereka capai. Dari laporan-laporan tersebut ternyata bahwa tes yang

diterbitkan pada tahun 1908 itu masih mengandung kelemahan-kelemahan. Oleh karena itu kembali Binet dan Simon bekerja keras untuk memperbaiki kelemahan-kelemahan tesnya. Pada tahun 1911 terbitlah revisinya yang terakhir. Dan setelah revisi yang terakhir Binet meninggal dunia pada tahun itu juga.

Dalam revisi yang dilakukan tahun 1911 ini umur psikhis seorang anak tidak hanya ditentukan oleh golongan umur tertinggi yang dapat dijawab secara benar oleh anak. Umur psikhis di mana suatu tes dapat dijawab dengan betul seluruhnya disebut umur psikhis dasar. Kemudian untuk setiap pertanyaan yang dapat dijawab dengan benar pada seri pertanyaan di atas umur psikhis dasar ini diberikan lagi nilai sebesar satu tahun per jumlah pertanyaan dalam seri tersebut. Misalnya satu seri pertanyaan terdiri dari 6 pertanyaan dan dapat dijawab dengan betul oleh anak sebanyak 2 pertanyaan; maka kepada anak ini ditambahkan lagi umur

psikhis sebanyak 2/6 tahun

Setelah penerbitan tahun 1911 itu, masih banyak lagi dilakukan revisi-revisi terutama sekali di Amerika Serikat. Beberapa revisi yang penting antara lain adalah sebagai berikut :

- a. Revisi Kuhlmann. Kuhlmann melakukan dua kali revisi, yaitu yang pertama tahun 1912, dan yang ke-dua tahun 1922.
- b. Revisi yang dilakukan oleh Lewis M. Terman dari Stanford University, yang selanjutnya terkenal dengan revisi Stanford, pada tahun 1916
- b. Revisi Terman dan Daud Merril pada tahun 1937.
- c. Revisi Herring pada tahun 1922.
- d. Revisi Norden pada tahun 1932.

Dari penggunaan revisi Stanford dapat diketahui bahwa ukuran kecerdasan seseorang yang didasarkan atas perbedaan umur psikhis dengan umur kronologis ternyata tidak tepat. Sebab seorang anak yang berumur 4 tahun yang mencapai umur psikhis 5 tahun, ternyata setelah anak itu berumur 8 tahun ia mencapai umur psikhis 10 tahun. Jadi perbedaan antara umur kronologis dengan umur psikhis ternyata tidak tetap. Dan yang tetap ternyata adalah perbandingan antara umur kronologis dengan umur psikhis atau umur mental.

Berdasarkan pada kenyataan tersebut, William Stern mengemukakan, bahwa untuk menentukan tingkat kecerdasan seseorang hendaknya digunakan istilah "quotient Psychis," yang merupakan indek perbandingan antara umur psikhis dengan umur kronologis. Jadi apabila seorang anak berumur 4 tahun mencapai umur psikhis 5 tahun, maka quotient psikhisnya adalah : $5 : 4 = 1,25$.

Konsep William Stern mengenai ukuran kecerdasan seseorang dengan memperbandingkan umur psikhis dengan umur kronologis dapat diterima oleh Terman. Tetapi Terman tidak sependapat dengan penggunaan istilah quotient psychis. Terman mengemukakan istilah Inteligensi Quotion (IQ). Dan supaya mendapatkan bilangan bulat, maka hasil perbandingan antara umur psikhis dengan umur kronologis dikalikan lagi dengan 100. Dengan demikian maka anak yang berumur 4 tahun yang mencapai umur psikhis

5 tahun, Inteligensi Quotionnya adalah : $5/4 \times 100 = 125$

Bagi Terman ukuran ini merupakan petunjuk terhadap kecerdasan seorang anak, oleh karena ukuran ini pada hakekatnya menyatakan perbandingan antara perkembangan psykhis seorang anak normal yang sebaya. (Witherington, 1967, hal. 86).

Di samping berkembangnya tes Binet Simon dengan segala revisi-revisinya, kemudian berkembang pula jenis-jenis tes

inteligensi yang lain. Beberapa di antaranya adalah sebagai berikut:

- a. Tes Wechsler.
- b. Tes Army alpa dan army beta
- c. Tes menggambar orang.
- d. Tes labirin.
- e. Tes Progressive Matrices.

Dari uraian historis yang telah dikemukakan di depan, dapat disimpulkan bahwa usaha-usaha pengukuran inteligensi baru dimulai pada akhir abad kesembilan belas. Dan yang dapat dipandang sebagai perintis pengukuran inteligensi adalah Binet dan Simon. Setelah tes Binet Simon barulah muncul sejumlah tes inteligensi yang lain.

D. METODE PENGUKURAN INTELIGENSI

Seperti telah diuraikan dalam sejarah pengukuran inteligensi, metode yang dapat dipergunakan untuk mengadakan pengukuran inteligensi adalah metode tes, yang disebut tes inteligensi. Berdasarkan hasil-hasil tes inteligensi akan dapat diketahui tingkat inteligensi seseorang, Karena inteligensi mempunyai korelasi yang cukup tinggi dengan prestasi belajar yang dicapai di sekolah, maka tes intelligen

si ini dapat dipergunakan untuk meramalkan sukses tidaknya seorang anak dalam pelajaran-pelajaran di sekolah pada umumnya. Dengan demikian tes inteligensi juga dapat dipergunakan untuk mengadakan seleksi dalam penerimaan calon siswa. Oleh sebab itu tes inteligensi juga disebut tes bakat akademis (*Scholastic aptitude test*)

Dalam sejarah pengukuran inteligensi telah dikemukakan bahwa ada beberapa jenis atau pola tes inteligensi. Pada

uraian-uraian berikut akan dikemukakan hanya Tes Binet-Simon.

1. TES BINET - SIMON

Tes Binet Simon adalah merupakan tes inteligensi yang pertama. Telah pula dikemukakan bahwa tes Binet-Simon dibagi-bagi menurut tingkatan umur. Jadi ada tes untuk anak umur 3 tahun, ada tes untuk anak umur 4 tahun dan seterusnya. Item-item yang digunakan pada setiap tingkat umur antara lain adalah sebagai berikut:

- a. Untuk umur 3 tahun antara lain terdiri dari pengenalan bagian badan sendiri, misalnya : coba tunjukkan hidungmu !.
- b. Untuk umur 4 tahun antara lain terdiri dari pengenalan terhadap benda-benda sekitarnya, misalnya : Apakah nama benda ini ? (sambil menunjukkan sebuah sendok makan).
- c. Untuk umur 5 tahun antara lain terdiri dari menirukan suatu bentuk tertentu, misalnya : Coba kamu membuat gambar seperti ini! (sambil menunjukkan sebuah gambar segitiga).
- d. Untuk anak umur 6 tahun antara lain terdiri dari pengenalan terhadap waktu, misalnya : Pada waktu matahari baru terbit disebut
- e. Untuk anak umur 7 tahun antara lain terdiri dari pengenalan kiri-kanan, misalnya. Coba pegang telinga kirimu!
- f. Untuk anak umur 8 tahun antara lain terdiri dari kemampuan membedakan dua benda, misalnya : apakah perbedaan antara lalat dengan kupu-kupu ?.
- g. Untuk anak umur 9 tahun antara lain terdiri dari kemampuan untuk memahami situasi (*comprehension*),

misalnya : apa yang harus kamu lakukan kalau kamu ketinggalan kereta api ?.

- h. Untuk anak umur 10 tahun antara lain terdiri dari kemampuan untuk mengeritik kemustahilan, misalnya : Coba kamu pernatikan. Adakah yang janggal dalam gamter ini ? (sambil menunjukkan sebuah gambar kapal laut yang benderanya berkibar ke depan);
- i. Untuk anak umur 11 tahun antara lain terdiri dari kemampuan untuk membuat analogi, misalnya : rumput hijau, laut.
- j. Untuk anak umur 12 tahun antara lain terdiri dari kemampuan untuk merakit pecahan-pecahan benda, Misalnya : Cobalah kamu rakit kepingan-kepingan ini supaya menjadi sebuah benda yang utuh! (sambil memberikan kepingan-kepingan dari sebuah gambar muka).

Dalam memberikan tes terhadap seorang anak, biasanya dimulai dari umur tes yang sesuai dengan umur anak bersangkutan. Apabila anak masih membuat kesalahan dalam tingkatan tersebut, maka diberikan tes pada tingkat urnur di bawahriya, sampai anak itu betul untuk seluruh seri tes umur tertentu. Setelah itu selanjutkan pada seri-seri tes dalam umur yang lebih tinggi, sampai anak itu gagal menjawab semua pertanyaan dalam umur tes tertentu.

Seperti juga telah dibicarakan dalam sejarah pengukur-an inteligensi apabila anak dapat menjawab suatu seri tes dengan betul semua, maka anak itu diberikan skor umur mental sesuai dengan umur tes yang dapat dijawab dengan betul. Kemudian setiap satu pertanyaan yang dapat dijawab dengan betul pada seri di atasnya diberikan skor umur mental sebanyak 1 tahun dibagi dengan jumlah pertanyaan dalam seri tersebut.

Misalnya kita akan memberi tes kepada seorang anak yang berumur 8 tahun 4 bulan. Maka anak tersebut mulai kita

berikari tes umur 8 tahun. Misalkan saja dari 6 buah pertanyaan dia hanya dapat menjawab 4 buah pertanyaan. Oleh karena dalam tes umur 8 tahun ini ia masih membuat kesalahan, maka kita berikan seri tes umur 7 tahun. Misalnya dari 6 pertanyaan umur 7 tahun, dapat dijawab betul semua. Karena ia betul menjawab semua pertanyaan maka anak tersebut diberi skor umur mental 7 tahun. Sekarang kita lanjutkan memberikan tes pada umur di atasnya. Tes umur 8 tahun telah kita berikan tadi. Karena di antara 6 pertanyaan ia dapat menjawab 4 pertanyaan maka ia men-dapat tambahan umur mental 4/6 tahun.

Sekarang kita lanjutkan memberikan tes umur 9 tahun. Misalkan dari 6 pertanyaan ia dapat menjawab 3 pertanyaan. Maka ia mendapatkan tambahan skor umur mental sebanyak 3/6 tahun. Kita lanjutkan lagi pada tes umur 10 tahun. Misalkan dari 6 pertanyaan ia salah semua. Jadi ia tidak mendapatkan tambahan skor umur mental. Kita tidak perlu melanjutkan memberikan tes umur 11 tahun karena tes umur 10 tahun saja anak tersebut sudah salah semua, jelas bahwa dalam tes umur 11 tahun yang lebih sukar anak itu tidak akan bisa menjawab.

Contoh diatas dapat menghitung umur mental anak sebagai berikut :

No	Kemampuan menjawab pertanyaan	Umur psikhis	
1	Betul semua pertanyaan umur 7 tahun	: 7 tahun
2	Betul 4 dari 6 pertanyaan umur 8 tahun	: 4/6 tahun
3	Betul 3 dari 6 pertanyaan umur 9 tahun	: 3/6 tahun
4	Salah semua pertanyaan umur 10 tahun	: 0 tahun
	Total		: 8 1/6 tahun

Jadi umur mental anak tersebut ialah $8 \frac{1}{6}$ tahun. Berdasarkan umur mental atau *mental age* (MA) dan umur kronologis atau *cronological age* (CA) anak tersebut dapat dihitung IQ dengan rumus :

$$IQ = \frac{MA}{CA} = \frac{8 \frac{1}{6}}{8 \frac{4}{12}} = 98$$

Untuk mengetahui status seorang anak yang mempunyai IQ tertentu, maka perlu kita ketahui kriteria yang dipergunakan untuk mengklasifikasi IQ. Adapun klasifikasi tersebut adalah sebagai berikut :

Interval Nilai IQ	Kategori
140 ke atas	Genius
130 ke 139	Sangat Suprior
120 ke 129	Suprior
110 ke 119	Di atas normal
90 ke 109	Normal
80 ke 89	Di bawah normal
Di bawah 70	Lemah jiwa

Dengan melihat kriteria tersebut, maka anak yang 98 adalah termasuk dalam kategori **ber IQ normal**.

Penggunaan hasil dari tes IQ yaitu :

1. Untuk mengetahui seorang anak sudah cukup matang untuk diterima di kelas 1 SD/Mi atau belum.
2. Sebagai prasyarat dalam pendaftaran siswa baru.
3. Sebagai salah satu informasi dalam mengadakan diagnosis kesulitan belajar.
4. Untuk pengelompokkan murid-murid berdasarkan IQ.
5. sebagai bahan prediksi peserta didik sukses tidaknya di perguruan tinggi.



MENILAI SIKAP

A. Pengertian Sikap

Sikap dapat didefinisikan sebagai suatu predisposisi atau kecenderungan untuk melakukan suatu respon dengan cara-cara tertentu terhadap dunia sekitarnya, baik berupa individu-individu maupun obyek-obyek disekitarnya. Sikap ini akan memberikan arah kepada perbuatan atau tindakan seseorang. Hal ini dikarenakan sikap seseorang sangat dipengaruhi oleh nilai-nilai yang membentuk dirinya, baik nilai religius, budaya dan adat, wawasan global dan lain-lain.

Seorang guru hendaknya mengetahui dengan baik sikap siswa, baik positif maupun negatif dengan harapan sikap positif akan dapat ditingkatkan kualitasnya dan sikap negatif akan dicari-cari problem solving melalui berbagai macam teknik yang ada.

B. Metode Pengukuran Sikap

Ada beberapa metode yang dapat dilakukan oleh seorang guru dalam melakukan penilaian sikap. Metode-metode tersebut akan diuraikan sebagai berikut :

1. Kuesioner

Metode kuesioner yang digunakan untuk mengukur sikap sosial ada bermacam-macam.

a. Sosil Distance Inventori

Pada tahun 1925 Bogardus dari Amerika Serikat pernah mengadakan pengukuran tentang sikap sosial anak-anak terhadap orang asing, yang diukur adalah jarak sosial anak dengan orang asing. Caranya yaitu memberikan pertanyaan-pertanyaan tertulis kepada anak-anak apakah ia setuju apabila orang asing tertentu terikat dalam hubungan kekeluargaan dengannya, dengan menjadi sahabat karibnya, tetangganya, sahabatnya, warga negaranya, turis atau orang asing harus pergi dari negerinya. Urutan-urutan tersebut disusun secara graduil dari jarak yang paling dekat sampai pada jarak yang paling jauh.

Contoh :

Saya menghendaki agar orang(orang asing tertentu) :

- a). Terjalin dalam satu hubungan kekeluargaan dengan saya.
- b). Menjadi sahabat karibnya saya.
- c). Menjadi tetangganya
- d). Menjadi kawan sekerjanya
- e). Menjadi warga negara saya
- f). Menjadi pengunjung negeri saya
- g). Harus keluar dari negeri saya

b. Remmers Attitude Scales

Remmers mengembangkan pola skala yang umum, untuk mengukur sikap tertentu (khusus) untuk level orang, group, lembaga-lembaga dll. Item-item digunakan untuk mengukur sikap terhadap lembaga-lembaga tertentu.

Contoh :

Di bawah ini adalah suatu daftar statement tentang (lembaga tertentu). Silahkan isi tanda check (√) di depan setiap statement yang menurut pendapat anda yang tepat hubungannya dengan lembaga tersebut.

1	Cukup sempurna dalam segala bidangnya.
5	Lembaga yang sangat baik dalam kehidupan modern
9	Mempunyai pengaruh yang besar untuk mencapai kehidupan yang benar (for right living)
12	Sangat berguna untuk mencetuskan idea-idea
16	Membantu individu untuk menggunakan waktu luang secara bijaksana

c. Likert

Pada skala ini subyek tidak disuruh memilih statement-statement yang disetujuinya saja. Tiap-tiap item pada statement positif dan negatif dibagi dalam lima point yaitu : sangat setuju, setuju, tidak tentu (undecided), tidak setuju, dan sangat tidak setuju. Untuk statement positif masing-masing diberi skor mulai dari 5 sampai 1, dan sebaliknya untuk statement negatif diberi skor mulai dari 1 sampai 5. Kemudian skor tersebut diberikan bobot item.

Kelebihan pada skala ini adalah memberikan informasi yang lebih banyak daripada Thurstone dan Remmers yang hanya memberikan dua kategori yaitu mengisi atau tidak mengisi (*checked or not checked*).

Dalam skoring skala sikap, kita tidak dapat mengatakan bahwa skor satu lebih baik dari skor yang lain, akan tetapi dapat dikatakan bahwa *sikap seseorang lebih berkenan dan tidak berkenan terhadap subyek daripada individu lain.*

Contoh :

Dibawah ini adalah beberapa pertanyaan tentang bagaimana kepuasan anda dalam pekerjaan. Silahkan memberi lingkaran pada alternatif yang merupakan jawaban atau yang paling mendekati jawaban anda :

1. Seberapa besar antusias anda terhadap jabatan anda sekarang ?
 - a. Saya benci pada pekerjaan saya sekarang
 - b. Saya sangat antusias
 - c. Saya tidak menyukainya
 - d. Saya menyukainya
 - e. Sedang-sedang saja
2. Seberapa lama biasanya anda merasa puas dalam pekerjaan anda ?
 - a. Sepatu waktu
 - b. Kadang-kadang (Occasionally)
 - c. Sepanjang waktu
 - d. Jarang (Seldom)
 - e. Sebagian besar

Option pada skala Likert tidak disusun secara berurutan, melainkan dicampur adukkan (*in scrambled order*). Hal ini untuk menghindari adanya jawaban yang mempunyai kecenderungan untuk memilih posisi yang sama.

d. Study Values.

Studi ini dipelopori oleh Allport-Vernon-Lindezey dengan menyusun bersama dalam bentuk kuesioner untuk mengukur sikap seseorang terhadap nilai-nilai sosial

tertentu. Pada awalnya untuk mengukur sikap para mahasiswa dan pengajar terhadap nilai-nilai sosial tertentu. Klasifikasi nilai-nilai sosial yang dipergunakan didasarkan atas kategori watak dari Sprenger yang terkait nilai politik, sosial, ekonomi, keagamaan, keindahan atau kebenaran. Dalam penyusunan item dipergunakan 30 item dengan dua alternatif, dan 15 item dengan empat alternatif.

Contoh :

Tujuan pokok daripada research ilmiah adalah lebih menekankan kepada penemuan kebenaran yang murni daripada aplikasinya yang praktis :

- (a) Ya.....
- (b) Tidak.....

Dalam item-item ini, individu yang menyetujui (a) dan menolak (b), supaya menulis angka 3 pada (a) dan angka 0 pada (b). Sebaliknya individu menyetujui (b) dan menolak (a), supaya menulis angka 3 pada (b) dan mengisi angka 0 pada (a). Apabila lebih menyenangi (a) daripada (b), supaya menulis angka 2 pada (a) dan mengisi angka 1 pada (b). Sebaliknya apabila lebih menyenangi (b) daripada (a) , supaya menulis angka 2 pada (b) dan angka 1 pada (a).

Item-item dengan empat alternatif :

Contoh :

Pemerintah yang baik adalah pemerintah yang mengutamakan :

- (a). Pertolongan kepada orang-orang miskin, sakit, dan orang-orang lanjut usia.
- (b). Memajukan perdagangan dan industri

(c). Memperkenalkan lebih banyak prinsip-prinsip etika kepada para polisi dan diplomat-diplomatnya

(d). Menekankan pada prestise dan respek di antara bangsa-bangsa.

Dalam item-item ini individu-individu disuruh mengadakan pilihan secara ordinal dengan memberi angka 1 pada alternatif yang paling disukai, angka 2 pada alternatif yang kurang disukai, angka 3 pada alternatif tidak disukai, dan angka 4 pada alternatif yang paling disukai.

e. Kuesioner Pasangan Berbanding

Metode ini banyak dipakai di Indonesia dilakukan pengukuran yang berhubungan dengan nilai-nilai sosial. Pada kuesioner ini dikemukakan dua alternatif yang sifatnya bertentangan. Anak-anak disuruh memilih salah satu daripada alternatif yang disediakan.

Ada beberapa nilai sosial yang bertentangan antara lain : kepentingan pribadi, kepentingan keluarga, kepentingan sahabat karib, kepentingan sekolah dan kepentingan masyarakat. Pada tiap-tiap item yang diajukan dikemukakan, dipertentangkan antara dua kepentingan yang ada.

Contoh :

Pribadi dengan keluarga

Joko pada suatu malam minggu mau nonton dengan pacarnya, dan mereka sudah membeli karcis. Ketika mereka mau masuk adik Jokok datang menyusul, menggambarkan bahwa tantenya baru opname di rumah sakit, dan Joko diharapkan datang menjenguk. Apakah yang harus dilakukan oleh Joko?

- a. Membatalkan nonton dan segera menjenguk tantenya.
- b. Nonton saja dulu, sebab masih ada kesempatan lain untuk menjenguk tantenya.

Setelah memilih semua kepentingan/nilai-nilai sosial yang ada dipertentangkandengan mempergunakan sejumlah item yang cukup representatif, maka akan dapat diketahui nilai-nilai mana saja yang lebih diutamakan oleh seorang anak.

f. Kuesioner Masalah-Masalah yang Searah

Yang dikembangkan oleh Simanhadi Widyaprakasa. Kuesioner ini mirip dengan kuesioner di atas tetapi tidak mempergunakan pasangan-pasangan berbanding, tetapi mengemukakan masalah-masalah yang sifatnya searah. Kuesioner ini digunakan untuk mengukur bermacam-macam aspek sikap sosial. Mis. Sikap kerjasama, keuletan, sikap tanggungjawab, dan sikap percaya pada diri sendiri.

Contoh :

Item untuk mengukur sikap kerjasama

Andaikata anda disuruh bekerjasama dengan kawan yang tidak anda sukai, sikap manakah yang anda ambil ?

- a. Tidak mau toh pekerjaan tidak akan beres
- b. Mau asal menguntungkan diri sendiri
- c. Mau asal ia suka kerjasama
- d. Mau secara terpaksa

Metode kuesioner memang hanya akan mampu mengukur sikap seseorang pada taraf verbal atau normatif saja. Namun

paling tidak sikap normatif tersebut merupakan orientasi pemilihan subyek terhadap suatu obyek, sehingga dapat memberikan arah pada sikap sesungguhnya. Untuk mengatasi kelemahan tersebut dapat dilakukan pengecekan dengan metode yang lain.

2. Interview

Interview dilakukan dengan cara berhadapan langsung dengan responden bersifat pribadi, lebih-lebih dilakukan dalam keadaan informal. Sehingga dapat dilakukan berbagai macam hubungan yang lebih bebas, dan mengetahui sikap-sikap tertentu terhadap masalah yang dihadapi melalui berbagai macam informasi yang berasal anak, dan orang tua.

Ada beberapa keuntungan daripada interview yaitu :

1. sering diperoleh respon yang lebih berarti dibandingkan dengan pertanyaan tertulis.
2. kita tidak hanya memperoleh data tentang sikap subyek terhadap sesuatu, tetapi juga mengetahui alasan-alasan mengapa subyek berbuat demikian.

Kelemahan-kelemahannya yaitu :

1. Jawaban-jawaban yang diberikan oleh responden dapat dipengaruhi oleh situasi yang ditimbulkan oleh interviewer.
2. Interview membutuhkan waktu yang jauh lebih banyak bila dibandingkan dengan kuesioner.

3. Observasi

Pengamatan secara langsung terhadap sikap yang diambil anak-anak dalam situasi tertentu yang konkrit akan dapat memberi gambaran yang lebih tepat daripada mendengarkan atau membaca jawaban anak-anak tentang apa yang

dilakukannya oleh anak-anak dalam situasi tersebut. Metode ini sangat baik sekali untuk melengkapi metode kuesioner dan interviu terkait data yang diperoleh, sehingga dapat digunakan untuk mengecek apakah sikap yang diambil dalam taraf normatif betul-betul dilaksanakan dalam tindakan yang riil.

Observasi terhadap sikap dapat dibedakan menjadi dua macam yaitu :

1. Observasi tak berstruktur, yaitu dimana dalam observasi tersebut kita mencatat semua tingkah laku anak yang nampak selama observasi berlangsung. Hasil-hasilnya dicatat dalam Anecdotal-record. Walaupun ada beberapa kelemahan pada cara ini yaitu penafsiran terhadap hasil-hasil observasi sering bersifat subyektif.
2. Observasi berstruktur, yaitu dimana variabel tingkah laku yang akan diobservasi telah disediakan dalam suatu daftar, dapat berupa check list, dan rating scale.

Contoh :

ANECDOTAL RECORD

Nama : Kelas :

 Umur : Sekolah :

Hari/Tanggal	Peristiwa Yang Terjadi	Sikap Anak	Interpretasi

BAB XIV

MENILAI KEPRIBADIAN

A. Pengertian Sikap

Apa yang telah dibahas pada bab sebelumnya sebenarnya merupakan bagian dari kepribadian.

Kepribadian berkaitan dengan penilaian kepribadian akan dikemukakan beberapa ahli yaitu :

1. Menurut Allport mengatakan bahwa " Personality is the dynamic organization within the individual of those psycho-physical systems that determine his characteristic behavior and thought." (Calvin s. Hall, 1970 : 263). Yang artinya kepribadian menyangkut keseluruhan aspek daripada seseorang, baik fisik maupun psikhis, baik yang dibawa sejak lahir maupun yang diperoleh dari pengalaman.
2. Menurut Crow anda Crow mengemukakan bahwa "Components of personality there are, physical characteristics, abilities, temperament, disposition, and character." (Crow and Crow, 1963:230). Yang artinya komponen kepribadian tersebut terdiri dari cirri-ciri fisik, kemampuan, perangai, disposisi dan karakter.
 - a. Ciri-ciri fisik menyangkut tentang bentuk muka, bentuk anggota tubuh konstitusi tubuh.
 - b. Kemampuan menyangkut masalah kemampuan dalam suatu jabatan atau pekerjaan tertentu, kemampuan dalam menerima pendidikan, dan kemampuan dalam aktivitas sosial.
 - c. Temprament adalah kecenderungan emosional yang telah dibawa sejak lahir. Tempramen relatif permanen.

Tempramen merupakan corak dasar emosional seseorang yang mewarnai setiap responnya.

- d. Disposisi adalah respon emosional yang bersifat temporer. Disposisi ini dipengaruhi oleh pengalaman masa lampau, sebagai hasil daripada reaksi emosional yang dilakukannya terhadap lingkungannya.
 - e. Karakter adalah aspek tingkah laku seseorang yang berhubungan dengan norma-norma etik.
3. Thorndike mengatakan bahwa aspek-aspek kepribadian yaitu : "Three two general areas of measurement, thar are : (1) measurement of ability and (2) measurement of personality. Measurement of ability that mean to measure what the individual can do under test condition and motivation to do best. Measure of personality that mean to measure what the individual will do under the natural sircumstances of life . (Thorndike, 1961:317). Menurut Thorndike aspek kemampuan terpisah dengan aspek kepribadian. Aspek kepribadian terdiri dari :
- a. temperamen berkaitan dengan suasana hati dan tingkat kepekaan.
 - b. Karakter merupakan aspek kepribadian yang berhubungan dengan nilai-nilai sosial. Mis. Kejujuran, baik hati, kerjasama dan kerajinan
 - c. Penyesuaian berkaitan dengan seberapa jauh individu itu sanggup untuk "berdamai" dengan dirinya sendiri dan dengan dunia disekitarnya.
 - d. Minat berkaitan dengan kecenderungan untuk mencari dan berpartisipasi dengan kegiatan tertentu.
 - e. Sikap berkaitan dengan penerimaan atau penolakan terhadap individu/kelompok lain, idea-idea tertentu, atau lembaga tertentu.

Walaupun beberapa ahli membagi kepribadian dalam berbagai macam aspek, namun dalam realitas aspek-aspek tersebut merupakan bagian dari satu tubuh yang utuh. Pembagian tersebut sebenarnya bertujuan untuk memudahkan mempelajari aspek-aspek pribadi yang ada pada manusia.

B. Metode Pengukuran Kepribadian

Ada beberapa ahli yang telah mengembangkan dalam mengukur kepribadian. Menurut Thorndike membagi menjadi empat cara yaitu :

1. Mendengarkan atau membaca apa yang dikatakan atau ditulis oleh orang yang bersangkutan tentang dirinya. Metode yang dapat digunakan adalah self deskriptive, auto biografi, iterviu langsung, angket langsung dan inventori.
2. Mendengarkan atau membaca apa yang dikatakan atau ditulis oleh orang lain tentang individu yang bersangkutan. Metode yang dapat digunakan adalah biografi, iterviu tak langsung, angket tak langsung cumulative record dan case study.
3. Melihat at apa yang secara nyata dilakukan oleh individu bersangkutan dalam situasi yang wajar. Metode yang dapat digunakan yaitu metode observasi
4. Melihat bagaimana reaksi individu yang bersangkutan terhadap dunia imaginasi.

Ada beberapa metode yang perlu dijelaskan dalam bab ini yaitu :

1. Metode Inventori

Inventori merupakan suatu daftar statement tentang keadaan, sifat, kegiatan tertentu yang harus dipilih oleh subyek yang akan diukur mana statement yang cocok dengan dirinya.

Ada beberapa metode inventori yang telah dikembangkan oleh para ahli yaitu :

a. Gilford Zimmerman Temperament Survey.

Gilford mengklasifikasikan temperament atas 10 aspek yaitu :

- (1). Aktivitas umum, yaitu tentang energinitas, vitalitas, kecepatan dalam bergerak, antusias, dan produktivitas.
- (2). Pengendalian diri yaitu tentang keseriusan dalam menghadapi sesuatu, tidak mudah terganggu, dan self kontrol.
- (3). Kepemimpinan yaitu tentang kebiasaan memimpin, mengambil, inisiatif dalam pembicaraan, mempengaruhi orang lain, kecenderungan untuk mempertahankan harga diri, dan self defensif.
- (4). Sosialibilitas yaitu tentang banyaknya punya teman, kecenderungan mencari kenakalan, kecenderungan mencari kontak-kontak sosial, senang bercakap-cakap dan tidak pemalu.
- (5). Kestabilan emosi yaitu tentang ketenangan, optimisme, merasa dirinya sehat, tidak merasa cemas, tidak merasa kesunyian, dan tidak suka melamun.
- (6). Obyektifitas, yaitu tentang kebebasan dari rasa egois, tidak self sentred, tidak self sentred, tidak curiga dan cenderung untuk terbuka.
- (7). Persahabatan yaitu tentang menghormati orang lain, tidak menghendaki dominasi, bebas dari rasa permusuhan, dan bebas dari rasa dendam.
- (8). Perenungan, yaitu tentang banyaknya pertimbangan, dan merenungkan kembali tingkah lakunya sendiri, maupun tingkah laku orang lain.

- (9). Hubungan pribadi yaitu tentang toleransi terhadap orang lain dan kepercayaan terhadap orang lain.
- (10). Kejahatan yaitu tentang tidak mudah takut dan ekspresi emosional yang bebas.

Berdasarkan dari 10 aspek tersebut, maka dapat disusun menjadi beberapa statement. Statement yang *pencerminan* adalah item positif, sedangkan *kebalikannya* adalah item negatif.

Contoh :

1. Apakah anda memulai pekerjaan dengan antusias ?
2. Apakah anda mengerjakan sesuatu pekerjaan dengan bersenda gurau?
3. Apakah anda akan mempertahankan pendapat anda, apabila kebanyakan teman-teman anda menyalahkannya?
4. Apakah anda lebih senang bekerja sendiri dalam tempat yang terpisah dengan teman lain?
5. Apakah anda cepat dapat melupakan kesalahan yang telah anda lakukan?
6. Apakah anda sering merasa bahwa anda dimusuhi oleh orang lain?
7. Apakah anda mudah memaafkan kesalahan orang lain?
8. Apakah anda sering merenungkan kembali tentang apa yang telah anda lakukan?
9. Apakah anda yakin bahwa semua teman anda akan berbuat baik kepada anda ?
10. Apakah anda merasa gentar kalau mendapat ancaman dari seseorang?

Individu yang hendak dinilai diminta mengisi tanda check (√) pada item yang cocok dengan dirinya.

Skoring dalam pengukuran dengan metode inventori adalah untuk item yang positif, setiap tanda check diberi skor 1, dan yang tidak diberi skor 0. sebaliknya untuk item negatif yang diberi tanda check diberi skor 0 dan yang diberi tanda check diberi skor 1.

b. Minnesota Multiple Inventory

Inventori ini dikembangkan untuk mengukur tentang kelainan tingkah laku. Aspek penyimpangan yang diidentifikasi yaitu:

- (1). Hipokondria yaitu tentang kekuatan akan kesehatannya, dan sering mengeluh tentang sakit yang dirasakannya tanpa sebab yang jelas.
- (2). Depresi yaitu tentang rasa ketidakmampuan atau rasa takut menghadapi masa depan.
- (3). Histeria yaitu tentang adanya gejala jasmaniah yang ditimbulkan oleh problem psikhis yang tak terpecahkan. Seperti kejang, sakit perut, dll.
- (4). Psikopatik yaitu tentang ketidak acuan terhadap respon sosial.
- (5). Paranoid yaitu tentang rasa curiga dan perasaan yang over sensitif.
- (6). Psihastini yaitu kecenderungan yang kompulsip untuk mencetuskan sesuatu idea atau untuk berbuat sesuatu.
- (7). Schizophreni yaitu tentang ketidaktentuan dalam pikiran atau perbuatan
- (8). Hipomania yaitu tentang kecenderungan untuk menjadi over aktif baik bodi maupun mental, dengan

kecenderungan untuk meloncat-loncat dari satu hal ke hal yang lain.

Contoh :

1. Biasanya saya tidak cepat menjadi payah.
2. saya sering terbangun karena gangguan suara yang ribut.
3. Saya tidak mau berbicara dengan orang lain sebelum orang tersebut mendahuluinya.
4. Saya tidak pernah mencemaskan tentang apa yang dikatakan orang lain terhadap diri saya.

Sistem skoring dalam pengukuran dengan metode inventori ini adalah untuk item yang positif, setiap tanda cek diberi skor 1, dan yang tidak diberi skor 0. sebaliknya untuk item negatif yang diberi tanda cek diberi skor 0 dan yang diberi tanda cek diberi skor 1.

c. Edward Personal Preference Schedule

Mencoba menilai kekuatan daripada berbagai aspek kepribadian. Ada beberapa aspek yang harus dilihat yaitu :

1. Usaha untuk mencapai prestasi atau peraturan
2. Kepatuhan terhadap perintah atau peraturan
3. Keteraturan didalam bekerja
4. Usaha untuk menonjolkan diri
5. Usaha untuk berdiri sendiri (otonom)
6. Kecenderungan untuk berafiliasi dengan orang lain.
7. Kecenderungan untuk campur tangan terhadap urusan orang lain.

8. Kecenderungan untuk mendapatkan bantuan dari orang lain.
9. Kemauan untuk menguasai orang lain
10. Kemauan untuk mengalah
11. Kemauan untuk menyenangkan orang lain
12. Kemauan untuk mengadakan perubahan
13. Ketahanan dalam mengatasi rintangan-rintangan
14. Hetroseksualitas
15. Kecenderungan untuk agresif.

Berdasarkan dari aspek-aspek itu disusunlah item. Masing – masing item terdiri pasangan-pasangan statement dimana masing-masing statement mencerminkan salah satu aspek.

Contoh :

1. a. Saya akan menolong teman saya apabila ia dalam kesusahan
- b. Saya ingin bekerja untuk mendapatkan hasil yang sebaik-baiknya

Item no. 1.a. mencerminkan kecenderungan untuk campur tangan terhadap urusan orang lain, dan item no. 1b mencerminkan adanya usaha untuk mencapai prestasi.

Skoring diberikan pada item yang dipilih dengan skor 1.

d. Daftar Cek Masalah

Inventori ini berisi tentang masalah-masalah yang dihadapi siswa SMA. Yang diklasifikai menjadi 12 masalah :

1. Masalah kesehatan

2. Masalah Kehidupan /ekonomi keluarga.
3. Masalah rekreasi/hobi
4. Masalah hubungan sosial dan organisasi
5. Masalah moral/agama
6. Masalah situasi keluarga
7. Masalah masa depan pendidikan/jabatan
8. Masalah penyesuaian diri terhadap sekolah
9. Masalah pelajaran di sekolah
10. Masalah kebiasaan belajar
11. Masalah dalam menghadapi ulangan
12. Masalah percintaan.

Contoh :

1. saya sering merasa pusing-pusing
2. Saya sering kekurangan beaya untuk membeli alat-alat sekolah.
3. saya merasa kurang mendapat kesempatan untuk berkreasi.

Dalam pelaksanaan inventori ini seorang siswa diminta untuk mengisi tanda cek didepan statement yang cocok dengan problim yang dihadapi. Untuk statement yang tidak cocok dengan problem yang dihadapi di kosongkan saja.

Contoh tabel

Tabel Problem Check List Siswa Kelas

No	Masalah/Nama	Kesehatan						
		1	2	3	4	5	f	%
1	Marijan						1	10
2	Paijah						2	20
3	Abdul Ghofur						1	10

4	Abdul Rohman						0	0
5	Syihab						1	10
6	Najma						0	0
7	Najwa						0	0
8	Adhim						3	30
9	Rohim						1	10
10	Amir						1	10
							10	100%

Untuk mencari nilai prosentase yaitu dengan rumus :

$$P = f/N \times 100\%$$

Sehingga dapat dicari No. Resp 8. $3/10 \times 100\% = 30\%$

Siswa yang mempunyai banyak masalah apabila dibandingkan teman-temannya adalah no. Resp 8.

e. Daftar Cek Kebiasaan Belajar (Study Habit)

Inventori ini diperuntukan bagi siswa berkaitan dengan kebiasaan belajar siswa. Bentuk cek list sama dengan cek list yang lain, tetapi dalam susunan itemnya berbeda dengan item yang lain. Ada item positif, negatif dan netral.

Contoh :

1. Saya belajar setiap hari sesuai dengan kesempatan yang ada.
2. Saya belajar kalau hanya ada ulangan
3. Saya belajar setiap hari sesuai dengan daftar yang telah saya susun dll.

Kemudian mentabulasi dalam bentuk check list dibawah ini :

Tabel Problem Check List Siswa Kelas

		Item Positif				Item Netral				Item Negatif				Selisih	
				f	%			f	%			f	%	F	%
1	Marijan														
2	Paijah														
3	Ghofur														
4	Rohman														
5	Syihab														
6	Najma														
7	Najwa														
8	Adhim														
9	Rohim														
10	Amir														

f. Teknik "Siapa Saya" (Who am I)

Tenik ini terdiri dari statement tentang sifat-sifat kepribadian yang positif dan sebagian sifat kepribadian yang negatif. Didepan statemen ada 3 kolom pilihan. Contoh :

Cocok sekali dg diri saya (CSS)	Agak Cocok dg diri saya (ACS)	Tidak Cocok dg diri saya (TCS)	Pernyataan-Pernyataan
--	--------------------------------------	---------------------------------------	------------------------------

			1. saya adalah orang yang mudah bergaul
			2. saya sering cemas kalau menghadapi sesuatu yang baru
			3. Saya dapat bekerja sama dengan orang lain

Skoring untuk statement positif CSS diberi skor 1, ACS diberi skor 0 dan TCS diberi skor -1. sebaliknya untuk statement negatif CSS= -1, ACS = 0, TCS = 1.

Tabel SKOR "SIAPA SAYA"

No	Nama/Item	1	2	3	4	5	TOTSKOR	%
1	Marijan	1	0	-1	0	-1		
2	Pajah							
3	Abdul Ghofur							
4	Abdul Rohman							
5	Syihab							
6	Najma							
7	Najwa							
8	Adhim							
9	Rohim							
10	Amir							

Untuk penafsirannya skor yang diperoleh siswa dengan kriteria sebagai berikut :

1. 60% s/d 100%
2. 20% s/d 59%
3. -20% s/d 19%
4. -60% s/d -21%
5. -100% s/d 61%

g. Metode Pertanyaan Yang Berlebihan

Metode over statement digunakan untuk menilai kejujuran seseorang. Contoh :

1. Karena saya dididik secara baik-baik, maka saya tidak pernah bertengkar dengan saudara-saudaranya.
2. Selama hidup saya tidak pernah tersinggung atas segala kata-kata yang ditujukan kepada saya.

Dalam pelaksanaannya siswa diminta untuk mengisi tanda cek pada setiap statemen yang cocok dengan dirinya.

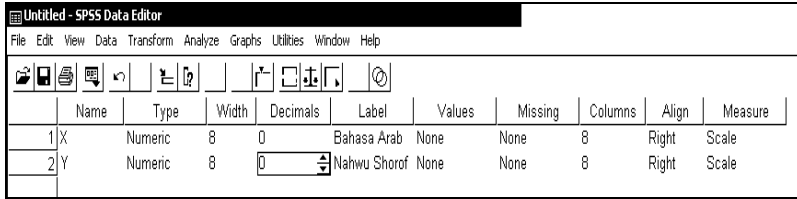
Contoh :

Dilakukan pengujian untuk menemukan hubungan antara hasil ujian pelaksanaan ujian mata pelajaran bahasa arab dan nahwu shorof. Setelah dilakukan ujian diketemukan angka sebagai berikut :

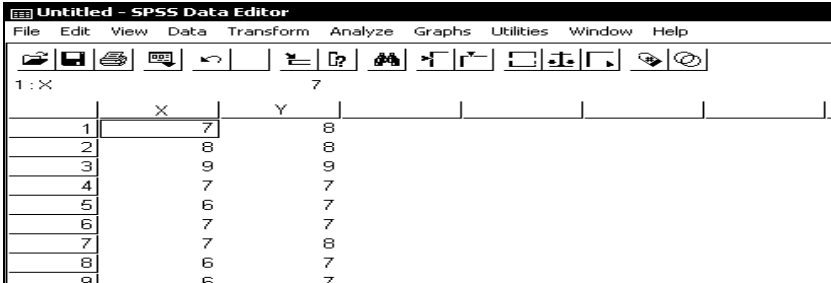
No.	Bahasa Arab	Nahwu Shorof
1	7	8
2	8	8
3	9	9
4	7	7
5	6	7
6	7	7
7	7	8
8	6	7
9	6	7
10	7	8

Adapun langkah-langkah dalam mencari nilai hubungan yaitu :

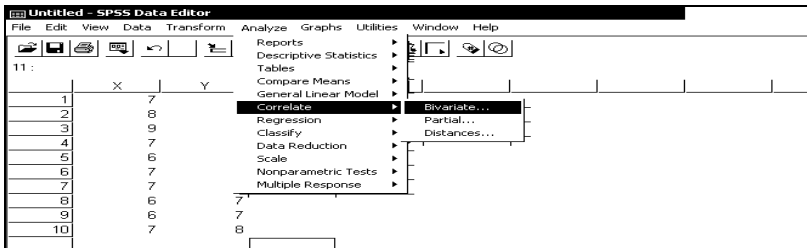
1. Bukalah program SPSS fro Windows dan definisikan variabel



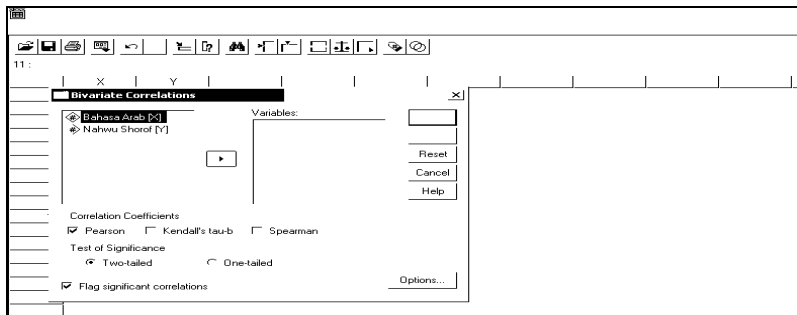
2. Masukkan data kedalam variabel



3. Lakukan analisis dengan memilih correlate dan bivariate



4. Masukkan variabel ke dalam kotak variabel



5. Hasil olah data SPSS

Correlations

		Bahasa Arab	Nahwu Shorof
Bahasa Arab	Pearson Correlation	1	,843**
	Sig. (2-tailed)	.	,002
	N	10	10
Nahwu Shorof	Pearson Correlation	,843**	1
	Sig. (2-tailed)	,002	.
	N	10	10

** . Correlation is significant at the 0.01 level (2-tailed).

Pada tabel tersebut diatas terlihat bahwa hasil korelasinya diketemukan 0,843 dengan angka signifikansi 0,002. Adapun kriteria penerimaan adalah:

1. Jika angka signifikan $\leq 0,05$ maka dapat dikatakan $H_a =$ diterima dan $H_o =$ ditolak.
2. Jika angka signifikan $> 0,05$ maka dapat dikatakan $H_a =$ ditolak dan $H_o =$ diterima.

Dengan melihat kriteria diatas, maka dapat disimpulkan nilai koefisien korelasi 0,843 dengan angka signifikan 0,002 diatas, termasuk dalam kategori 1. Dengan demikian dapat disimpulkan bahwa ada korelasi antara nilai bahasa arab dan nahwu shorof.

DAFTAR PUSTAKA

- Amudi Pasaribu, Dr. *Pengantar Statistik*, Medan : Imballo, 1965.
- Anas Sudijono, Drs. *Pengantar Evaluasi Pendidik*, PT. Raja Grafindo Persada, Jakarta, 1990.
- Anas Sudijono, Drs. *Pengantar Statistik Pendidik*, PT. Raja Grafindo Persada, Jakarta, 1980
- Anton Dajan, Drs. *Pengantar Metode Statistik Deskriptif*, Jakarta LP3ES, 1973.
- Bloom, B. S. ed. et al. *Taxonomy of Educational Objectives: Handbook 1, Cognitive Domain*. New York: David McKay, 1956.
- Cronbach, Lee, J. *Essentials of Psychological Testing*, Harper & Row, New York, 1960.
- Crow, Lester D. Ph.D. and Alice Crow Ph.,D. *General Psychology*. Littlefield, Adam & Co. 1963.
- Daryanto, Drs. *Evaluasi Pendidikan*, Rineka Cipta, 2000.
- Djarwanto, Ps. Drs, *Statistika Non Parametrik*, Universitas Sebelas Maret, Surakarta, 1999.
- Krathwohl, D. R. ed. et al., *Taxonomy of Educational Objectives: Handbook II, Affective Domain*. New York: David McKay, 1964.
- M. Chabib Thoba. *Teknik Evaluasi Pendidikan*, Rosyda Karya, Bandung, 2000.

- Masrukhin, M.Pd. *Statistik Deskriptif Berbasis Komputer*, Media Ilmu Kudus, 2007
- Masrukhin, M.Pd. *Statistik Inferensial Aplikasi Program SPSS*, Media Ilmu Kudus, 2007
- Moh. Nazir, *Metode Penelitian*, Ghalia, Jakarta, 1999.
- Muchtar Buchari, *Technik-Technik Evaluasi dalam Pendidikan*, Bapemsi, Bandung, 1967.
- Mulyasa, Dr. *Kurikulum Tingkat Satuan Pendidikan*. Rosyda Karya, Bandung, 2006.
- Nana Sudjana, Dr. *Dasar-Dasar Proses Belajar Mengajar*. Sinar Baru Algensindo Bandung, 2000.
- Nana Syaodih Sukmadinata. 2005. *Landasan Psikologi Proses Pendidikan*. Bandung: P.T. Remaja Rosdakarya. W. Gulo. *Strategi Belajar Mengajar*. Jakarta : Grasind, 2005.
- Ngalim Purwanto, MP. *Prinsip-Prinsip Teknik Evaluasi Pembelajaran*, Rosyda Karya, Bandung, 2002.
- Remmers H.H. and Gade N.L. *A Practical Introduction to Measurement and Evaluation*, Harper & Row, New York, 1960.
- Retno Sriningsih Satmoko, *Statistika Inferensial*, IKIP Semarang Press, 1993
- Singgih Santoso, *SPSS Statistik Multivariat*, Elex Media Komputindo, 2002
- Slameto, Drs. *Evaluasi Pendidikan*. Bumi Aksara, Jakarta, 2000.
- Soemadi Soeryabrata. *Psycho Diagnostik*. Rakepress Yogyakarta 1973
- Stanley Elan, *Performance Based Teacher Education*. Washington DC : American Association of Colleges for Teacher Education, One Dupont Circle, 1971.

- Stanley, Julien C., *Measurement in Today's Schools*. Printice-Hall, Inc., 1964.
- Sudjana, *Metoda Statistika*, Tarsito Bandung, 1996.
- Sugiyono, Dr. dan Eri Wibowo, S.Pd. *Statistika Penelitian dan Aplikasinya Dengan SPSS 10.00 For Windows*, CV. Alfabeta Bandung, 2001.
- Sugiyono, Dr., *Statistika Untuk Penelitian*, CV. Alfa Beta Bandung, 2002.
- Suharsimi Arikunto, Prof., Dr., *Dasar-Dasar Evaluasi Pendidikan*, Jakarta : PT. Bina Aksara, 1999
- Suharsimi Arikunto, Prof., Dr., *Manajemen Penelitian*, Jakarta : PT. Renika Cipta, 1983.
- Suharsimi Arikunto, Prof., Dr., *Prosedur Penelitian : Suatu Pendekatan Praktis*, Jakarta : PT. Bina Aksara, 1983.
- Sutrisno Hadi, Prof., Drs., MA., *Statistik, Jilid I*, Andi Offset : Yogyakarta, 2000.
- Thorndike, Robert L, and Elisabeth Hogen, *Measurement and Evaluation in Psychology and Education*. John Willey & Sons, Inc. New York, 1961.
- Vardiansyah, *Dani. Filsafat Ilmu Komunikasi: Suatu Pengantar*, Indeks, Jakarta 2008.
- Wandt, Edwin and Gerald W. Brown. *Essential of Educational Evaluation*. Holt Rinehart and Winston, New York, 1957.
- Wijaya IR, *Statistika Non Parametrik (Aplikasi Program SPSS)*, Alfa Beta Bandung, 2000.
- Witherington, *Psycologi Pendidikan*, (Terjamahan). Mochtar Buchori, M.Ed. Bapemsi, Bandung, 1967.
- Woodwart, Robert S. And Donald G. Marquis, *Psychology*, Holt Rinehart and Winston, Inc, New York, 1961.

Zaenal Arifin, *Evaluasi Pembelajaran Program Peningkatan Kualifikasi Guru Madrasah dan Guru PAI pada Sekolah*, Dirjen Pendidikan Islam Depag RI, 2009.

Zakiah Daradjat, Prof. Dr. *Metodik Khusus Pengajaran Agama Islam*. Bumi Aksara Jakarta, 1995.

Lampiran

Tabel 1

Tabel r Product Moment

dk=N	5%	1%	dk =N	5%	1%
3	0.997	0.999	38	0.32	0.413
4	0.95	0.99	39	0.316	0.408
5	0.878	0.959	40	0.312	0.403
6	0.811	0.917	41	0.308	0.398
7	0.754	0.874	42	0.304	0.393

8	0.707	0.834	43	0.301	0.389
9	0.666	0.798	44	0.297	0.384
10	0.632	0.765	45	0.294	0.38
11	0.602	0.735	46	0.291	0.376
12	0.576	0.708	47	0.288	0.372
13	0.553	0.684	48	0.284	0.368
14	0.532	0.661	49	0.281	0.364
15	0.514	0.641	50	0.279	0.361
16	0.497	0.623	55	0.266	0.345
17	0.482	0.606	60	0.254	0.33
18	0.468	0.59	65	0.244	0.317
19	0.456	0.575	70	0.235	0.306
20	0.444	0.561	75	0.227	0.296
21	0.433	0.549	80	0.22	0.286
22	0.423	0.537	85	0.213	0.278
23	0.413	0.526	90	0.207	0.27
24	0.404	0.515	95	0.202	0.263
25	0.396	0.505	100	0.195	0.256
26	0.388	0.496	125	0.176	0.23
27	0.381	0.487	150	0.159	0.21
28	0.374	0.478	175	0.148	0.194
29	0.367	0.47	200	0.138	0.181
30	0.361	0.463	300	0.113	0.148
31	0.355	0.456	400	0.098	0.128
32	0.349	0.449	500	0.088	0.115
33	0.344	0.442	600	0.08	0.105
34	0.339	0.436	700	0.074	0.097
35	0.334	0.43	800	0.07	0.091
36	0.329	0.424	900	0.065	0.086
37	0.325	0.418	1000	0.062	0.081

Tabel 3. Tabel Chi Square (χ^2)

Dk	Taraf Signifikansi	
	5%	1%
1	3.841	6.635
2	5.991	9.21
3	7.815	11.345
4	9.488	13.227
5	11.07	15.086
6	12.592	16.812
7	14.067	18.475
8	15.507	20.09
9	16.919	21.666
10	18.307	23.209
11	19.675	24.275
12	21.026	26.217
13	22.362	27.688
14	23.685	29.141
15	24.996	30.578
16	26.296	32
17	27.587	33.409
18	28.869	34.805
19	30.144	36.191
20	31.41	37.566
21	32.617	38.932
22	33.924	40.289
23	35.172	41.638
24	36.145	42.98
25	37.652	44.314
26	38.885	45.642
27	40.113	46.963

28	41.337	48.278
29	42.557	49.588
30	43.778	50.892

Tabel 4. Tabel Rho (Spearman)

Dk	TarafSignifikansi	
	5%	1%
5	1	
6	0.886	1
7	0.786	0.929
8	0.738	0.881
9	0.683	0.833
10	0.648	0.794
12	0.91	0.777
14	0.544	0.715
16	0.506	0.665
18	0.475	0.625
20	0.45	0.591
22	0.428	0.562
24	0.409	0.537

26	0.392	0.515
28	0.377	0.496
30	0.364	0.478