

BAB I

PENDAHULUAN

A. Latar Belakang

Sulit dibayangkan bagaimana proses pembelajaran akan berlangsung secara logis dan sistematis serta dapat memperoleh hasil yang maksimal apabila para guru atau instruktur yang akan melaksanakan tugas kegiatan mengajar atau memberikan suatu pelatihan tidak memiliki perencanaan mengajar atau pelatihan. Tanpa adanya perencanaan, kegiatan pembelajaran mungkin saja bisa dilaksanakan, akan tetapi karena tanpa adanya perencanaan yang akan berfungsi sebagai pedoman operasionalnya, maka pembelajaran akan banyak terjadi spontanitas (situasional) didasarkan pada apa yang diingat oleh guru/ instruktur pada saat terjadinya proses pembelajaran. Dengan kata lain, jika mengajar tanpa adanya perencanaan, guru akan menghadapi kesulitan untuk mengontrol dan mengendalikan pencapaian sasaran pembelajaran atau kompetensi yang harus dicapai, materi apa yang harus disampaikan yang sesuai dengan upaya pencapaian kompetensi, bagaimana proses pembelajaran harus dilakukan, sarana dan fasilitas pembelajaran apa yang harus disediakan, serta bagaimana kegiatan evaluasi harus dilakukan.

Oleh karena itu bagi guru dan para instruktur yang bertugas melaksanakan pembelajaran, perencanaan atau satuan pembelajaran sangat penting dan mutlak menjadi bagian yang tak terpisahkan dari tugas profesionalnya sebagai fasilitator pembelajaran, salah satunya bagaimana mengukur (evaluasi) tingkat keberhasilan penguasaan siswa menjadi jelas dan sistematis”.

Melalui tugas ini diharapkan mahasiswa dapat menyelaraskan antara berbagai konsep teori dengan kondisi objektif. Demikian pentingnya kegiatan perkuliaan ini, maka tugas rangkuman ini disusun guna melengkapi salah satu tugas mata kuliah Evaluasi Pembelajaran.

B. Tujuan Penulisan

Tujuan Penulisan tugas rangkuman ini adalah sebagai berikut:

1. Untuk lebih memantapkan dan menguasai pemahaman isi materi evaluasi pembelajaran
2. Untuk menambah pengetahuan mengenai evaluasi pembelajaran.
3. Untuk memahami proses pembelajaran.

C. Sistematika Penulisan

Sistematika penulisan tugas rangkuman adalah sebagai berikut :

KATA PENGANTAR

DAFTAR ISI

BAB I PENDAHULUAN

Mengemukakan tentang latar belakang, tujuan, dan sistematika penulisan laporan.

BAB II KONSEP DASAR EVALUASI

Pembahasan konsep dasar evaluasi pembelajaran yang berisi tentang definisi evaluasi pendidikan serta hubungan antara pengukuran, penilaian, dan evaluasi,

BAB III TUJUAN DAN KEGUNAAN EVALUASI

Pembahasan mengenai tujuan atau fungsi evaluasi dalam kaitannya dengan proses pembelajaran.

BAB IV SASARAN POKOK EVALUASI

Pembahasan mengenai subjek dan sasaran pokok evaluasi dalam hubungannya dengan tujuan daripada pembelajaran itu sendiri.

BAB V PROSEDUR EVALUASI

Pembahasan mengenai berbagai tahapan/langkah-langkah dan prosedur evaluasi pembelajaran.

BAB VI PENDEKATAN EVALUASI

Pembahasan mengenai pendekatan yang digunakan dalam prosedur evaluasi pembelajaran.

BAB VII PELAKSANAAN EVALUASI

Pembahasan mengenai penerapan konsep evaluasi yang direalisasikan dalam pelaksanaan evaluasi pembelajaran.

BAB VIII BENTUK DAN ALAT EVALUASI

Pembahasan mengenai berbagai bentuk dan alat evaluasi yang akan digunakan dalam proses pembelajaran.

BAB IX PENGOLAHAN SKOR TES HASIL BELAJAR

Pembahasan mengenai berbagai macam teknik pengolahan skor tes hasil belajar.

BAB X MENILAI KUALITAS ALAT EVALUASI

Pembahasan mengenai bagaimana menilai kualitas suatu alat evaluasi yang digunakan.

BAB XI PENUTUP

Pembahasan mengenai kesimpulan dan saran.

DAFTAR PUSTAKA

BAB II

KONSEP DASAR EVALUASI

A. Hubungan Antara Pengukuran, Penilaian, Dan Evaluasi

Dalam setiap melakukan pekerjaan evaluasi, terdapat beberapa istilah yang seringkali kita jumpai yaitu evaluasi, pengukuran, dan penilaian. Sementara orang lebih cenderung mengartikan ketiga kata tersebut sebagai suatu pengertian yang sama. Akan tetapi ada juga yang membedakan istilah tersebut. Untuk dapat mengadakan *penilaian*, kita mengadakan *pengukuran* terlebih dahulu, sedangkan *evaluasi* mencakup kedua langkah tersebut.

- *Mengukur* adalah membandingkan sesuatu dengan satu ukuran yang lebih cenderung bersifat *kuantitatif*.
- *Menilai* adalah mengambil suatu keputusan terhadap sesuatu dengan kriteria tertentu yang lebih cenderung bersifat *kualitatif*.
- *Kegiatan evaluasi* mencakup kedua langkah diatas, yakni mengukur dan menilai.

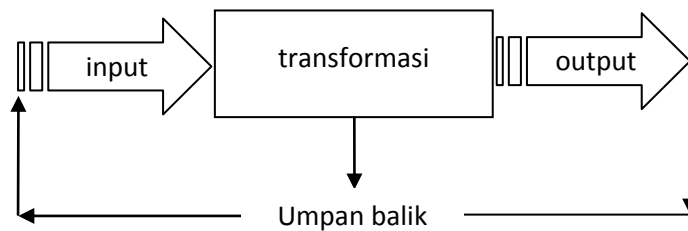
B. Definisi Evaluasi Pendidikan

- Ralph Tyler (Arikunto, S. 2003: 3) menyatakan bahwa “evaluasi merupakan sebuah proses pengumpulan data untuk menentukan sejauh mana keberhasilan tujuan pendidikan sudah tercapai”.
- Cronbach dan Stufflebeam (Arikunto, S. 2003: 3) menyatakan bahwa “evaluasi merupakan proses yang bukan sekadar mengukur sejauh mana tujuan tercapai, tetapi digunakan untuk membuat keputusan”.
- Wand dan Brown (Nurkencana, W. 1986: 1) menyatakan bahwa “evaluasi merupakan suatu tindakan atau suatu proses untuk menentukan nilai segala sesuatu dalam dunia pendidikan atau segala sesuatu yang ada hubungannya dengan dunia pendidikan”.

C. Pola Diagram Evaluasi Pendidikan Di Sekolah

Apabila sekolah diumpamakan sebagai tempat mengolah sesuatu dan calon siswa diumpamakan sebagai bahan mentah, maka lulusan dari sekolah itu dapat disamakan dengan hasil olahan yang sudah siap digunakan. Dalam istilah inovasi yang menggunakan teknologi, maka tempat pengolahan ini disebut *transformasi*.

Jika digambarkan dalam bentuk diagram akan terlihat sebagai berikut:



Keterangan :

- Input = bahan mentah (calon siswa) yang dimasukkan ke dalam transformasi.
- Output = bahan jadi (siswa lulusan sekolah) yang dihasilkan oleh transformasi.
- Transformasi = mesin yang bertugas mengubah bahan mentah menjadi bahan jadi. Terdiri dari berbagai unsur, diantaranya: siswa, guru dan personal lainnya, bahan pelajaran, metode mengajar dan sistem evaluasi, sarana penunjang, dan sistem administrasi.
- Umpan balik = segala informasi baik yang menyangkut output maupun transformasi. Intinya adalah untuk memperbaiki input maupun transformasi.

BAB III

TUJUAN DAN KEGUNAAN EVALUASI

A. Tujuan dan Fungsi Penilaian

1. Penilaian berfungsi selektif

Memiliki beberapa tujuan:

- a. Untuk memilih siswa yang dapat diterima di sekolah tertentu
- b. Untuk memilih siswa yang dapat naik ke kelas atau tingkat berikutnya
- c. Untuk memilih siswa yang seharusnya mendapat beasiswa
- d. Untuk memilih siswa yang sudah berhak meninggalkan sekolah dan sebagainya.

2. Penilaian berfungsi diagnosik

Diagnosa bertujuan adalah untuk mengetahui kebaikan dan sebab-musabab kelemahan yang terjadi, sehingga akan lebih mudah mencari solusinya.

3. Penilaian berfungsi sebagai penempatan

Pendekatan yang lebih bersifat melayani perbedaan kemampuan, adalah pengajaran secara kelompok. Tujuan sudah jelas yaitu untuk menentukan dengan pasti dikelompok mana seorang siswa harus ditempatkan, digunakan suatu penilaian.

4. Penilaian berfungsi sebagai pengukur keberhasilan

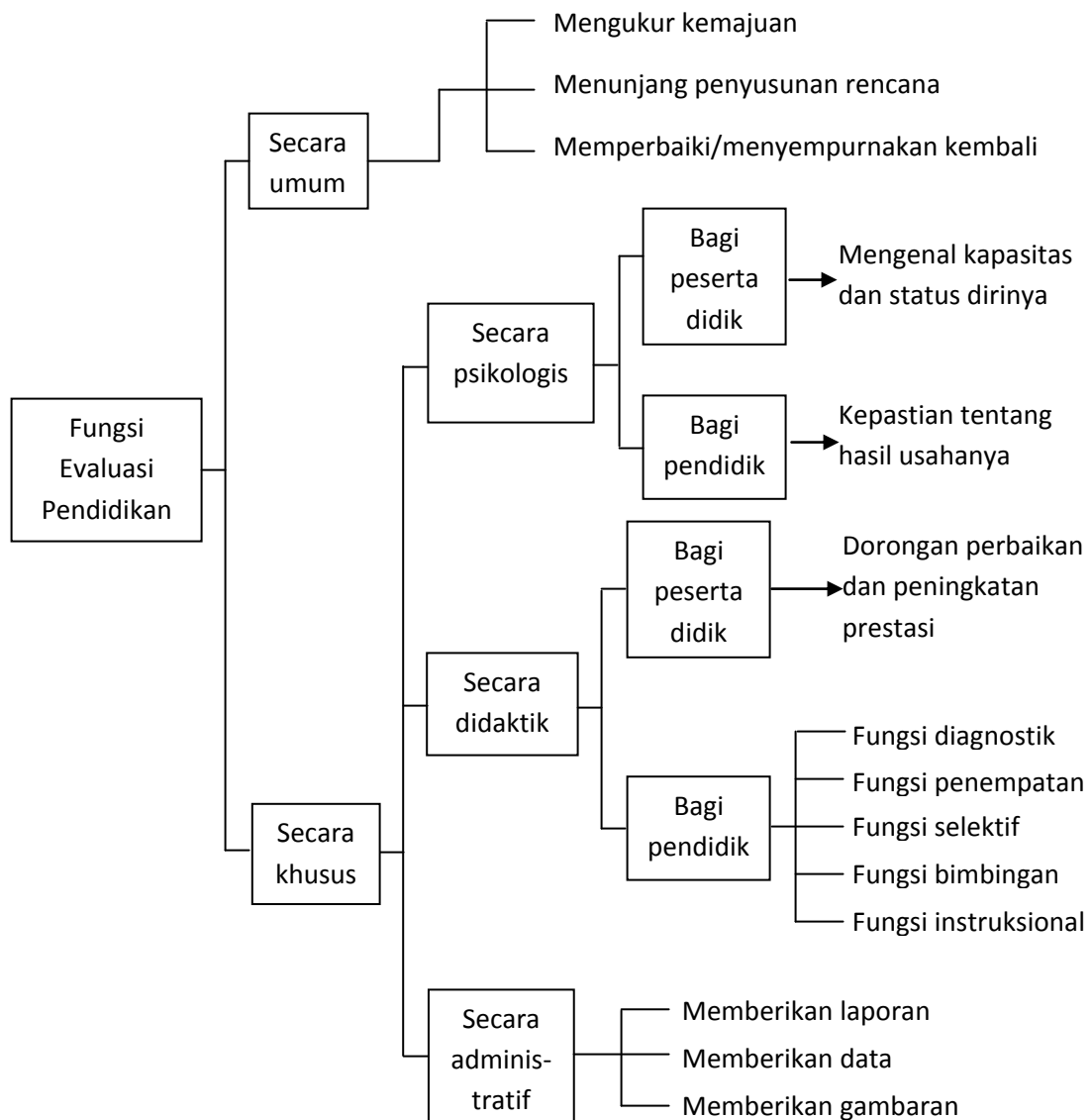
Tujuannya adalah untuk mengetahui sejauh mana suatu program berhasil diterapkan. Keberhasilan suatu program ditentukan oleh beberapa faktor, yaitu faktor guru, metode mengajar, kurikulum, sarana dan sistem administrasi.

B. Kegunaan Evaluasi Pendidikan

Diantara kegunaan yang dapat dipetik dari kegiatan evaluasi dalam bidang pendidikan adalah:

1. Terbukanya kemungkinan bagi evaluator guna memperoleh informasi tentang hasil-hasil yang telah dicapai dalam rangka pelaksanaan program pendidikan.
2. Terbukanya kemungkinan untuk dapat diketahuinya relevansi antara program pendidikan yang telah dirumuskan, dengan tujuan yang hendak dicapai.
3. Terbukanya kemungkinan untuk dapat dilakukannya usaha perbaikan, penyesuaian, dan penyempurnaan program pendidikan yang dipandang lebih berdaya guna dan berhasil guna, sehingga tujuan yang dicita-citakan akan dapat dicapai dengan hasil yang sebaik-baiknya.

C. Bagian Tentang Fungsi Evaluasi Pendidikan



BAB IV

SUBJEK DAN SASARAN POKOK EVALUASI

A. Subjek Evaluasi Pendidikan

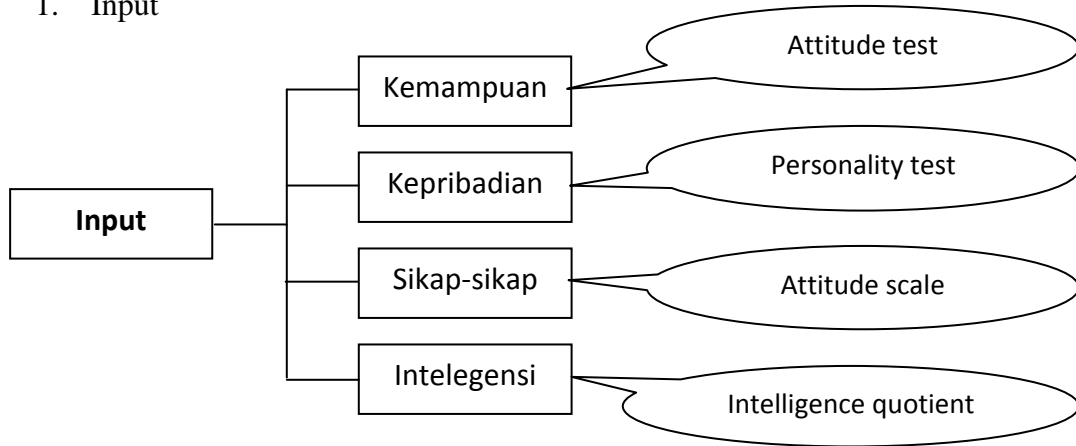
Subjek evaluasi adalah orang yang melakukan pekerjaan evaluasi, dalam hal ini evaluasi pendidikan. Siapa yang dapat disebut sebagai subjek evaluasi untuk setiap tes, ditentukan oleh suatu aturan pembagian tugas atau ketentuan yang berlaku. Suharsimi Arikunto (2003 : 20) mengategorikan pelaksana evaluasi sebagai subjek evaluasi. Terdapat pandangan lain yang disebut subjek evaluasi adalah siswa, yaitu orang yang dievaluasi. Dalam hal ini yang dipandang sebagai objek misalnya: prestasi matematika, kemampuan membaca, kecepatan lari, dan sebagainya. Pandangan lain lagi mengklasifikasikan siswa sebagai objek evaluasi dan guru sebagai subjeknya.

Dalam kegiatan evaluasi pendidikan dimana evaluasinya adalah kepribadian peserta didik, dimana pengukuran tentang kepribadian itu dilakukan dengan menggunakan instrumen berupa tes yang sifatnya baku (*standarized test*). Alat-alat evaluasi yang digunakan untuk mengukur kepribadian itu seseorang itu sifatnya rahasia, juga hasil-hasil pengukuran yang diperoleh dari tes kepribadian itu, hanya dapat diinterpretasi dan disimpulkan oleh para psikolog tersebut, tidak mungkin dapat dikerjakan oleh orang lain.

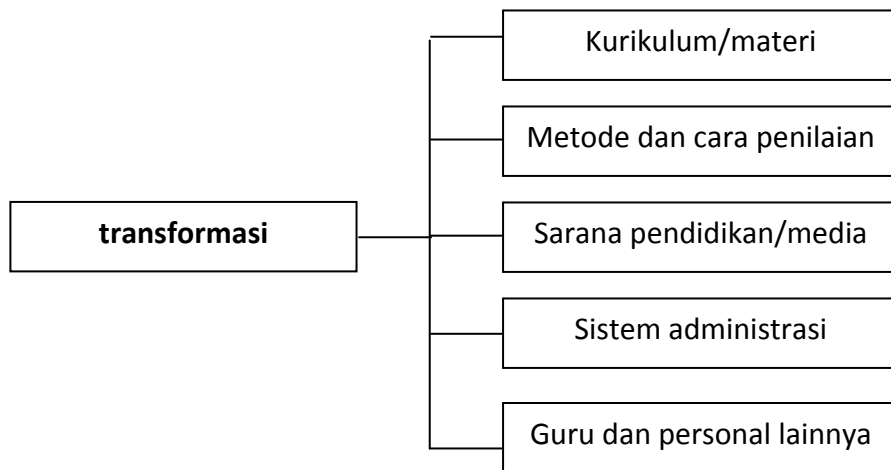
B. Objek/Sasaran Evaluasi Pendidikan

Objek/sasaran penilaian adalah segala sesuatu yang menjadi titik pusat pengamatan karena penilai menginginkan informasi tentang sesuatu tersebut. Dengan menggunakan diagram tentang transformasi maka sasaran penilaian untuk unsur-unsurnya meliputi: input, transformasi, dan output.

1. Input



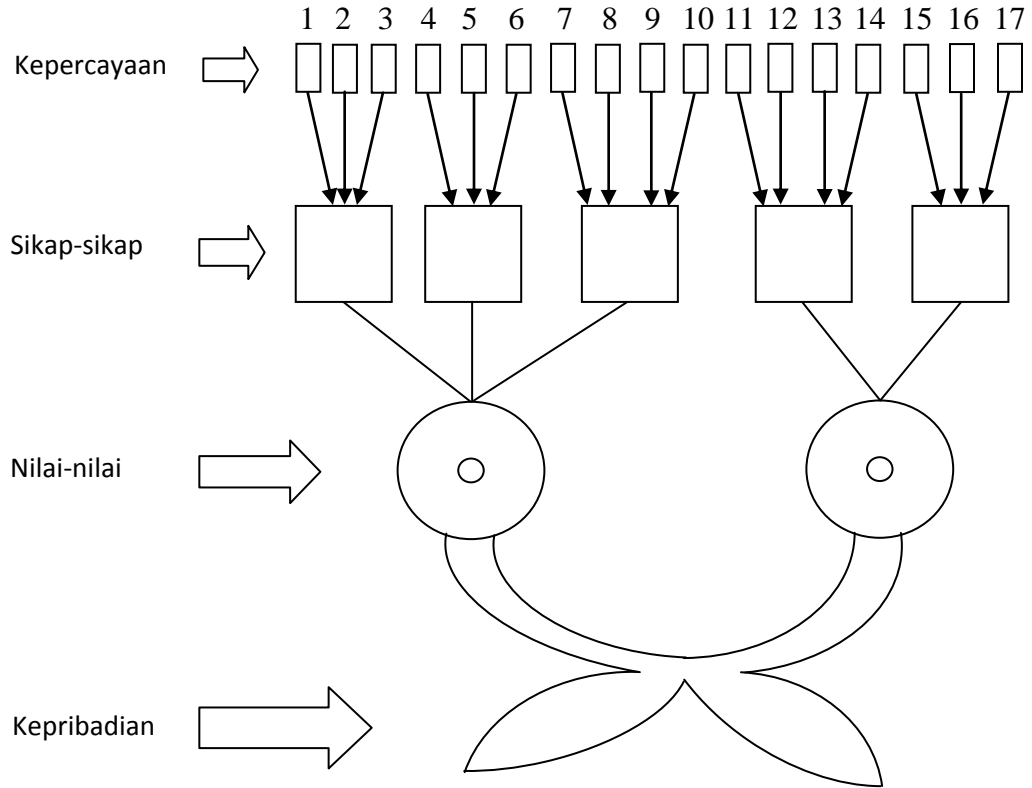
2. Transformasi



3. Output

Alat yang digunakan untuk mengukur pencapaian terhadap lulusan suatu sekolah disebut tes pencapaian atau *achievement test*. Kecenderungan yang ada sampai saat ini di sekolah adalah bahwa guru hanya menilai prestasi belajar aspek kognitif atau kecerdasan saja. Aspek psikomotorik, apalagi afektif, sangat langka dijamah oleh guru. Akibatnya dapat kita saksikan, yakni bahwa para lulusan hanya menguasai teori tetapi tidak terampil melakukan pekerjaan keterampilan, juga tidak mampu mengaplikasikan pengetahuan yang sudah

mereka kuasai. Lemahnya pembejaran dan evaluasi terhadap aspek afektif ini, jika kita mau instrospeksi, telah berakibat merosotnya akhlak para lulusan, yang selanjutnya berdampak luas pada merosotnya akhlak bangsa.







BAB V PROSEDUR EVALUASI PENDIDIKAN

Menurut Yulien Stanley (Nurkencana, W. 1986: 6-7), langkah-langkah evaluasi itu terdiri atas:

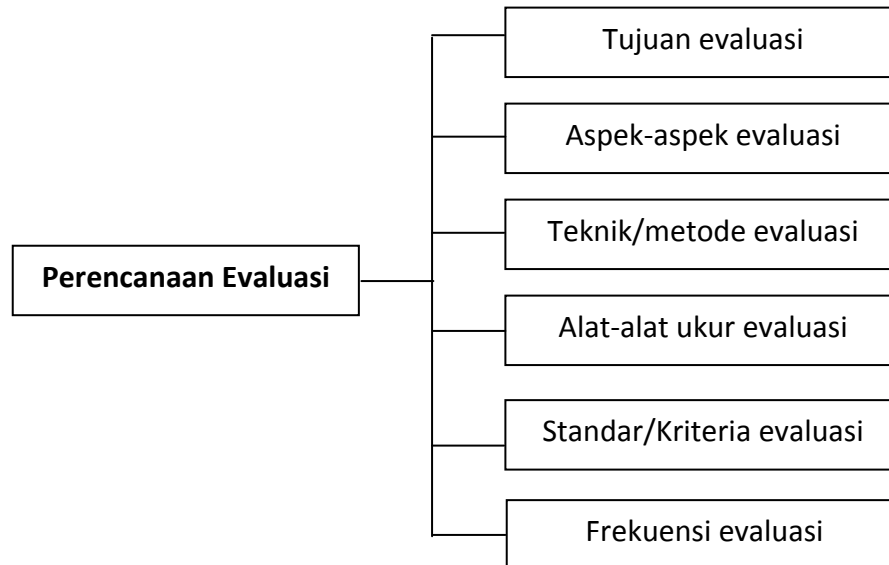
1. Menetapkan tujuan program
2. Memilih alat yang layak
3. Pelaksanaan pengukuran
4. Memberikan skor
5. Menganalisa dan menginterpretasikan skor
6. Membuat catatan yang baik, dan
7. Menggunakan hasil-hasil pengukuran

Menurut Mochtar Buchari (Nurkencana, W. 1986: 7) langkah-langkah pokok dalam evaluasi terdiri dari:

1. Perencanaan 
 - tujuan evaluasi
 - aspek-aspek penilaian
 - metode evaluasi
 - alat evaluasi
 - kriteria penilaian
 - frekuensi evaluasi
2. Pengumpulan data 
 - Pelaksanaan evaluasi
 - Pemeriksaan hasil evaluasi
 - Pemberian kode/skor
3. Verifikasi data (Pemeriksaan data-data yang akan diolah)
4. Analisa data 
 - Pengolahan secara statistic (statistical analisis)
 - Pengolahan bukan secara statistic (non statistical analisis)
5. Penafsiran data 
 - Memberikan interpretasi hasil-pengolahan data
 - Penggunaan hasil-hasil evaluasi

Pada umumnya Langkah-langkah pokok dalam evaluasi hasil belajar terdiri dari:
(Anas Sudijono, 2003 : 59-62)

1. Menyusun rencana evaluasi hasil belajar



2. Menghimpun data

Soal pertama yang kita hadapi dalam melakukan langkah ini ialah menentukan data apa saja yang kita butuhkan untuk melakukan tugas evaluasi yang kita hadapi dengan baik, Seperti telah disinggung di muka dan kemudian disinggung kembali dalam uraian kita tentang langkah perencanaan di atas, soal penentuan data yang harus dikumpulkan untuk keperluan suatu tugas evaluasi ini berhubungan erat dengan rumusan tentang tugas kita dalam suatu usaha pendidikan.

3. Melakukan verifikasi data

Data yang telah terkumpul harus disaring lebih dahulu sebelum diolah lebih lanjut. 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.

4. Mengolah dan menganalisis data

Langkah pengolahan data dilakukan untuk memberikan "makna" terhadap data yang ada pada kita. Jadi hal ini berarti bahwa tanpa kita dan diatur lebih dulu data itu sebenarnya tidak dapat menceritakan suatu apa pun kepada kita.

5. Memberikan interpretasi dan menarik kesimpulan

Dalam pekerjaan merumuskan tafsiran yang sedang kita bicarakan ini kesukaran yang sering dihadapi oleh para petugas evaluasi biasanya ialah kesukaran kata-kata yang tepat. Tetapi kesukaran yang lebih lazim lebih sering terletak dalam penyusunan kalimat yang tepat sehingga apa yang kemudian dinyatakan oleh kalimat itu tidak melampaui atau mengurangi batas kebenaran yang terdapat dalam data yang telah diolah tadi.

6. Tindak lanjut (*Follow Up*) hasil evaluasi

Menindaklanjuti dan melakukan perbaikan terhadap evaluasi yang telah dilakukan.

BAB VI

PENDEKATAN EVALUASI

Istilah pendekatan evaluasi ini diartikan sebagai beberapa pendapat tentang apa tugas evaluasi dan bagaimana dilakukan, dengan kata lain tujuan dan prosedur evaluasi. Berikut ini ada beberapa pendekatan evaluasi dan setiap pendekatan memberikan petunjuk bagaimana memperoleh informasi yang berguna dalam beberapa kondisi. Semua pendekatan paling tidak mempunyai tujuan yang sama yaitu bagaimana memperoleh informasi yang berarti atau tepat untuk klien atau pemakai. Namun masing-masing dalam usahanya berbeda penekanan pada aspek tertentu dalam tahap pengumpulan data, analisis, dan laporannya.

A. Beberapa Bentuk Pendekatan Evaluasi

Berikut ini beberapa pendekatan yang biasa digunakan dalam merancang suatu program evaluasi

1. Pendekatan Experimental

Pendekatan ini berasal dari kontrol eksperimen yang biasanya dilakukan dalam penelitian akademik. Tujuan evaluator yaitu untuk memperoleh kesimpulan yang bersifat umum tentang dampak suatu program tertentu yang mengontrol sebanyak-banyaknya faktor dan mengisolasi pengaruh program. Evaluator berusaha sekuat tenaga menggunakan metode saintifik sebanyak mungkin.

2. Pendekatan yang Berorientasi pada Tujuan (Goal Oriented Approach)

Kelebihan pendekatan yang berorientasi pada tujuan ini ialah terletak pada hubungan antara tujuan dan kegiatan dan penekanan pada elemen yang penting dalam program yang melibatkan individu pada elemen khusus bagi mereka. Namun keterbatasan pendekatan ini yaitu kemungkinan evaluasi ini melewati konsekuensi yang tak diharapkan terjadi.

3. Pendekatan yang Berfokus kepada Keputusan (The Decision Pocused Approach)

Keunggulan pendekatan ini ialah perhatiannya terhadap kebutuhan pembuat keputusan yang khusus dan pengaruh yang makin besar pada keputusan program yang relevan. Keterbatasan pendekatan ini yaitu banyak keputusan penting dibuat tidak pada waktu yang tepat, tapi dibuat pada waktu yang kurang tepat. Sering kali banyak keputusan tidak dibuat berdasarkan data, tapi tergantung pada impresi perorangan, politik, perasaan, kebutuhan pribadi, dan lain-lain.

4. Pendekatan yang Berorientasi kepada Pemakai (The User Oriented Approach).

Kelebihan pendekatan ini ialah perhatiannya terhadap individu yang berurusan dengan program dan perhatiannya terhadap informasi yang berguna untuk individu tersebut. Keterbatasan pendekatan ini, yaitu ketergantungan terhadap kelompok yang sama dan kelemahan ini bertambah besar pengaruhnya sehingga hal-hal lain diluar itu kurang mendapat perhatian.

5. Pendekatan yang Responsif (The Responsive Approach)

Kelebihan responsif ini ialah kepekaannya terhadap berbagai titik pandangan, dan kemampuannya mengakomodasi pendapat yang ambigis dan tidak fokus. Keterbatasan pendekatan responsif ialah membuat prioritas atau penyederhanaan informasi untuk pemegang keputusan dan kenyataan yang praktis tidak mungkin menampung semua sudut pandang dari berbagai kelompok.

6. Goal Free Evaluation

Alasan mengemukakan evaluasi *goal free evaluation* (Evaluasi Bebas Tujuan), dapat diuraikan secara singkat sebagai berikut: *pertama*, tujuan, pendidikan tidak dapat dikatakan sebagai pemberian, seperti tujuan lain, ia harus dievaluasi. Lebih jauh lagi, tujuan biasanya atau umumnya hanya formalitas, dan jarang menunjukkan tujuan yang sebenarnya. Dampak negatif

yang tak pernah termasuk dalam evaluasi yang berorientasi pada tujuan atau *goal based evaluation*. Berikut ini merupakan ciri-ciri evaluasi bebas tujuan.

- a) Evaluator sengaja menghindari untuk mengetahui tujuan program.
- b) Tujuan yang telah dirumuskan terlebih dahulu tidak dibenarkan menyempitkan fokus evaluasi.
- c) Evaluasi bebas tujuan berfokus pada hasil yang sebenarnya, bukan pada hasil yang direncanakan.
- d) Hubungan evaluator dan manajer atau dengan karyawan proyek dibuat seminimal mungkin.
- e) Evaluasi menambah kemungkinan ditemukannya dampak yang tak diramalkan.

BAB VII

PELAKSANAAN EVALUASI

Dalam prakteknya, pelaksanaan evaluasi pembelajaran 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 harus dilaksanakan oleh *testee*, dan cara penilaiannya dilakukan terhadap proses penyelesaian tugas dan hasil akhir yang dicapai.

A. Teknik Pelaksanaan Tes Tertulis

Dalam proses pelaksanaan tes tertulis ini terdapat beberapa hal yang harus diperhatikan:

1. Tempat/ruangan dimana *testee* mengerjakan soal-soal, seyogyanya dialokasikan tempat yang tenang, yang jauh dari keramaian, kebisingan, suara hiruk-pikuk dan lalu-lalangnya orang.
2. Ruangan tes harus cukup longgar, tidak berdesak-desakkan, tempat duduk diatur dengan jarak tertentu yang memungkinkan tercegahnya kerjasama yang tidak sehat di antara *testee*.
3. Ruangan tes sebaiknya memiliki sistem pencahayaan dan pertukaran udara yang baik.
4. Ruangan yang akan digunakan tes, sudah harus dipersiapkan sarana dan prasarana pendukung yang dibutuhkan.
5. Agar peserta tes dapat mengerjakan soal secara bersamaan, maka perlu dilakukan pengaturan lembar soal-soal. Lembaran soal-soal tersebut

diletakkan terbalik, sehingga tidak memungkinkan bagi *testee* untuk membaca dan mengerjakan soal lebih awal daripada teman-temannya.

6. Selama 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 juga jangan selalu duduk di kursi sehingga dapat membuka peluang bagi *testee* yang tidak jujur untuk bertindak curang.
7. Sebelum berlangsungnya tes, sebaiknya pengawas menginformasikan aturan main yang sudah ditentukan.
8. Sebagai bukti mengikuti tes, harus dipersiapkan daftar hadir yang harus ditandatangani oleh seluruh peserta tes. Dengan catatan pelaksanaannya tidak boleh mengganggu ketenangan jalannya tes.
9. Jika waktu yang ditentukan telah habis, hendaknya pengawas mengintruksikan kepada *testee* untuk segera menghentikan pekerjaannya dan segera mengumpulkan lembaran pekerjaannya tersebut, serta jumlahnya diperiksa dan disesuaikan dengan jumlah *testee* yang tertera dalam daftar hadir.
10. Pengawas sebaiknya mengisi berita acara pelaksanaan tes secara lengkap, yang mencakup daftar hadir *testee*, identitas *testee*, serta hal-hal yang terjadi selama kegiatan tes berlangsung.

B. Teknik Pelaksanaan Tes Lisan

Dalam proses pelaksanaan tes tertulis ini terdapat beberapa hal yang harus diperhatikan:

1. Sebelum tes lisan dilaksanakan, seyogyanya tester sudah melakukan inventarisasi berbagai jenis soal yang akan diajukan kepada *testee* dalam tes lisan tersebut, sehingga diharapkan memiliki validitas yang tinggi.
2. Setiap butir soal yang telah ditetapkan untuk dalam tes lisan itu, juga harus disiapkan sekaligus pedoman atau ancar-ancar jawaban betulnya.

3. Jangan sekali-kali menentukan skor atau nilai hasil tes lisan setelah seluruh *testee* menjalanites lisan.
4. Tes hasil belajar yang dilaksanakan secara lisan hendaknya jangan sampai menyimpang atau berubah arah dari evaluasi menjadi diskusi.
5. Dalam rangka menegakkan obyektivias dan prinsip keadilan, dalam tes yang dilaksanakan secara lisan itu, tester hendaknya jangan memberikan peluang untuk menaruh simpati terhadap *testee*.
6. Tes lisan harus berlangsung secara wajar, jangan sampai menimbulkan rasa takut, gugup, atau panik di kalangan *testee*.
7. Sekalipun acapkali sulit untuk dapat diwujudkan, namun sebaiknya tester mempunyai pedoman atau ancar-ancar yang pasti, berapa lama soal atau pertanyaan-pertanyaan dalam tes lisan tersebut.
8. Pertanyaan-pertanyaan yang diajukan dalam tes lisan hendaknya bervariasi.
9. Sebisa mungkin tes dilakukan secara individual.

C. Teknik Pelaksanaan Tes Perbuatan

1. Tester harus mengamati dengan teliti cara yang ditempuh oleh *testee* dalam menyelesaikan tugas yang telah ditentukan.
2. Agar dapat dicapai kadar obyektivitas setinggi mungkin, hendaknya tester jangan berbicara atau berbuat sesuatu yang dapat mempengaruhi *testee* yang sedang mengerjakan tugas tersebut.
3. Dalam mengamati *testee* yang sedang mengerjakan tugas itu, hendaknya tester telah menyiapkan instrumen berupa lembar penilaian yang di dalamnya telah ditentukan hal-hal apa sajakah yang harus diamati dan diberikan penilaian.

BAB VII

BENTUK DAN ALAT EVALUASI

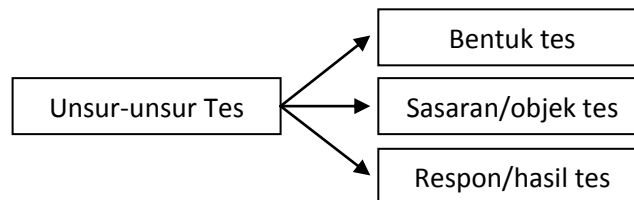
Terdapat dua metode yang dipergunakan dalam penilaian proses evaluasi, yaitu metode test dan metode non tes.

A. Tes

Definisi tes menurut Ejaan Yang Disempurnakan dalam bahasa Indonesia, yaitu alat atau prosedur yang digunakan untuk mengetahui atau mengukur sesuatu dalam suasana, dengan cara dan aturan-aturan yang sudah ditentukan. Dalam hal ini Wayan Nurkencana (1986 :25) mendefinisikan tes sebagai suatu cara untuk mengadakan penilaian yang berbentuk suatu tugas atau serangkaian tugas yang harus dikerjakan oleh anak atau sekelompok anak sehingga menghasilkan suatu nilai tentang tingkah laku atau prestasi anak tersebut, yang dapat dibandingkan dengan nilai standar yang ditetapkan.

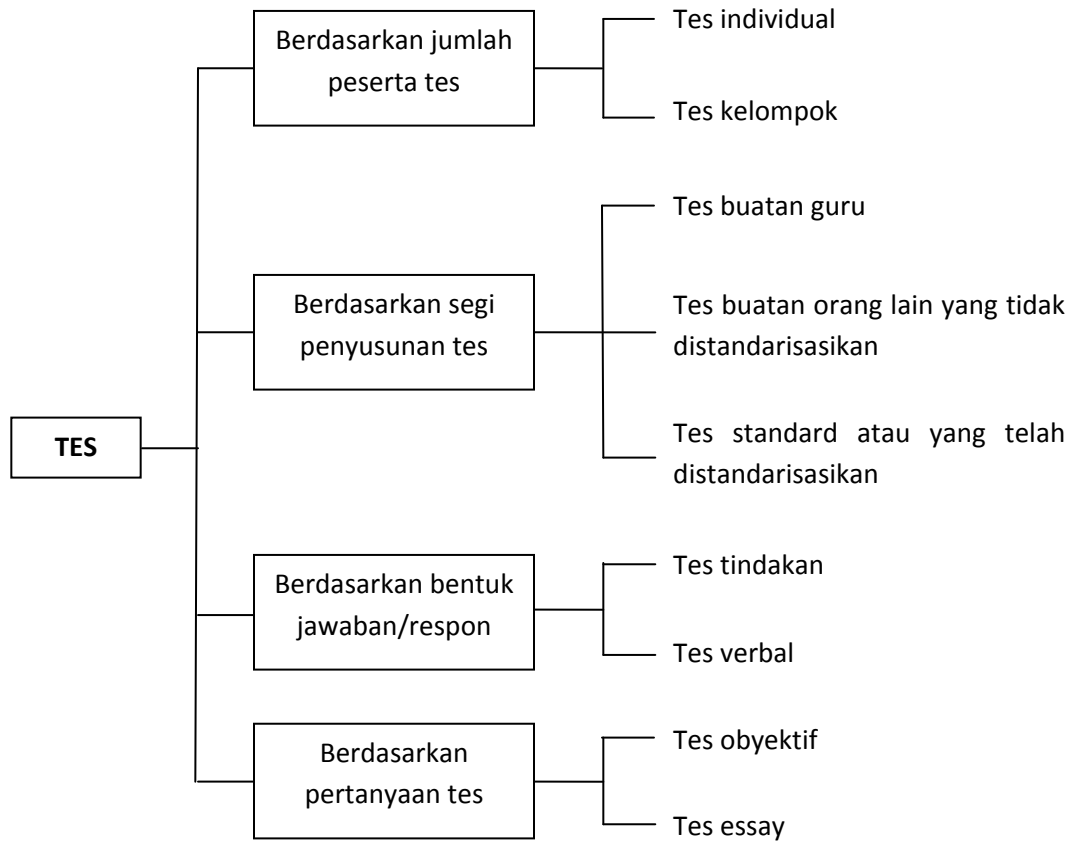
Berdasarkan definisi diatas, bila dinalisa lebih jauh terdapat beberapa unsur yang terkandung didalamnya, antara lain:

1. Tes berbentuk tugas yang terdiri dari pertanyaan-pertanyaan atau perintah-perintah.
2. Tes diberikan kepada seorang anak atau sekelompok anak untuk dikerjakan
3. Respon anak atau sekelompok anak tersebut dinilai



Tes hasil belajar dapat dibedakan atas beberapa jenis. Pembagian jenis-jenis tes ini dapat ditinjau dari beberapa sudut pandang

Klasifikasi Jenis-Jenis Tes

**B. Non Tes**

Ada beberapa teknis non tes yaitu:

1. Skala bertingkat (*rating scale*)

Skala menggambarkan suatu nilai yang berbentuk angka terhadap suatu hasil pertimbangan. Sebagai contoh adalah skor yang diberikan oleh seorang guru di sekolah untuk menggambarkan tingkat prestasi belajar siswa. Siswa yang mendapat skor 9, digambarkan ditempat yang lebih kanan dalam skala, dibandingkan dengan siswa lain yang mendapat skor 5.

Biasanya angka-angka yang digunakan diterangkan pada skala dengan jarak yang sama. Meletakkannya secara bertingkat dari yang rendah ke yang tinggi. Dengan demikian maka skala ini disebut skala bertingkat.

2. Kuesioner (*questionnaire*)

Kuesioner juga sering dikenal sebagai angket. Terdapat beberapa macam kuesioner diantaranya:

1) ditinjau dari segi siapa yang menjawab maka:

- kuesioner langsung
- kuesioner tidak langsung

2) ditinjau dari segi cara menjawab maka:

- kuesioner tertutup
- kuesioner terbuka

3. Daftar cocok (*check-list*)

Deretan pernyataan dimana responden yang dievaluasi tinggal membubuhkan tanda cocok (*check list*) ditempat yang sudah disediakan.

4. Wawancara (*interview*)

Suatu cara yang digunakan untuk mendapatkan jawaban dari responden dengan jalan Tanya jawab sepihak. Dikatakan sepihak karena dalam wawancara ini responden tidak diberi kesempatan sama sekali untuk mengajukan pertanyaan. Terdapat dua cara:

- a) interview bebas, dan
- b) interview terpimpin

5. Pengamatan (*observation*)

Suatu teknik yang dilakukan dengan cara mengadakan pengamatan secara teliti serta pencatatan secara sistematis. Ada tiga macam observasi diantaranya:

- a) observasi partisipan.
- b) observasi sistematis.
- c) observasi eksperimental.

6. Riwayat hidup

Riwayat hidup adalah gambaran tentang keadaan seseorang selama dalam masa kehidupannya, dengan mempelajari riwayat hidup maka subjek evaluasi akan dapat menarik suatu kesimpulan tentang kepribadian, kebiasaan dan sikap dari objek yang diteliti.

BAB VIII

PENGOLAHAN SKOR TES HASIL BELAJAR

A. Pendekatan Acuan Penilaian

Woodworth (Suherman, E. 2003: 200) mengemukakan bahwa ada dua jenis pedoman yang bisa digunakan untuk menentukan nilai (mengubah skor menjadi nilai) sebagai hasil evaluasi, yaitu:

1. Dengan cara membandingkan skor yang diperoleh seorang individu dengan suatu standar yang sifatnya mutlak (PAP).
2. Dengan cara membandingkan skor yang diperoleh seorang siswa dengan yang diperoleh siswa lainnya dalam kelompok tes tersebut (PAN).

Cara pertama disebut dengan Penilaian Acuan Patokan (PAP), terjemahan dari *Criterion Referenced Test (CRT)* atau *Criterion Referenced Evaluation (CRE)*, sedangkan cara kedua disebut Penilaian Acuan Normatif (PAN), terjemahan dari *Normative Referenced Test (NRT)* atau *Criterion Referenced Evaluation (NRE)*.

Perbandingan PAP dan PAN

Keterangan	Pendekatan Penilaian	
	PAP	PAN
Orientasi	tingkat penguasaan siswa terhadap seluruh materi yang diteskan.	kedudukan siswa dalam kelompok
Dasar pengolahan skor	didasarkan pada SMI (Skor Maksimal Ideal), dimana skor aktual yang diperoleh siswa dikonversikan pada SMI.	didasarkan atas skor aktual dan tidak memperhatikan lagi SMI
Keunggulan	kualitas hasil belajar	kedudukan relatif siswa

	dapat terkontrol.	dalam kelompoknya dapat diketahui, sesuai dengan sifat dari nilai tersebut yang tidak mutlak
Kelemahan	kurang memperhatikan bahwa pada hakekatnya setiap penilaian itu bersifat relatif	tingkat penguasaan siswa terhadap materi tes tidak dapat diketahui, sehingga kualitas hasil belajar siswa tidak dapat terkontrol
kombinasi dari sistem PAP dan sistem PAN.	<ol style="list-style-type: none"> 1. Rerata \bar{x} dan s dari hasil perhitungan sistem PAP dan sistem PAN, dan 2. Batas lulus (<i>passing grade</i>) untuk menjaga kualitas lulusan, kemudian dilakukan perhitungan dengan sistem PAN. 	

B. Teknik Pengolahan Skor

1. Mengolah Skor Mentah Menjadi Nilai Huruf

- a. Mengolah skor mentah menjadi nilai huruf dengan menggunakan *mean* dan deviasi rata-rata (DR).

Penggunaan *mean* dan Deviasi Rata-rata (DR) dalam penjabaran skor mentah menjadi nilai huruf, terutama diperlukan jika data atau skor mentah yang akan dijabarkan itu jumlahnya kecil (kurang dari 30), sehingga untuk menghitung *Mean* dan DR dari data tersebut kita tidak perlu menyusun tabel distribusi frekuensi. Berikut ini sebuah contoh:

Misalkan seorang Dosen memperoleh skor mentah dari hasil tes yang telah diberikan kepada 20 orang mahasiswanya, sebagai berikut:

73, 69, 68, 68, 67, 67, 65, 64, 62, 61, 60, 59, 59, 57, 54, 53, 51, 48, 47, 40

Skor mentah tersebut akan dijabarkan menjadi nilai huruf, A, B, C, D, TL, dengan menggunakan *mean* dan DR serta sekaligus penentuan nilai hurufnya, kita membuat tabel seperti berikut:

	Skor Mentah (X)	Deviasi skor dari Mean (X - M)	Penentuan Nilai Huruf
	73	13	73
	69	9	69
	68	8	68
	68	8	68
	67	7	67
	67	7	67
	65	5	65
	64	4	64
	62	2	62
	61	1	61
	60	0	60
	59	1	59
	59	1	59
	57	3	57
	54	6	54
	53	7	53
	51	9	51
	48	12	48
	47	13	47
	40	20	40 → TL
	1192 (ΣX)	136 ($\Sigma [X - M]$)	

Langkah-langkah penyusunan table:

- 1) Masukkan skor mentah kedalam kolom 1 menurut urutan besarnya (rangking), kemudian jumlahkan.

- 2) Hitunglah *Mean* dengan membagi jumlah skor mentah itu dengan N (banyaknya mahasiswa yang dites).

$$\begin{aligned} \text{Rumus Mean} &= \frac{\sum X}{N} \\ &= \frac{1.192}{20} = 59,6 \text{ dibulatkan} = 60 \end{aligned}$$

- 3) Isilah kolom dua dengan selisih (deviasi) tiap-tiap skor dari *Mean*, kemudian jumlahkan pula (dalam menjumlahkan, tanda (-) pada kolom dan dihilangkan).
- 4) Hitunglah Deviasi Rata-ratanya (DR) dengan membagi $\sum (X - M)$ dengan $N = 136 : 20 = 6,8$.

Dengan demikian sekarang kita telah mendapatkan dari skor mentah tersebut

$$\text{Mean (M)} = 60$$

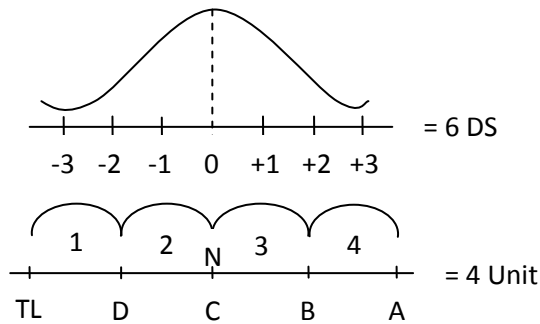
$$\text{DR} = 6,8$$

Penjabaran nilai huruf:

- 1) Pertama-tama kita tentukan besarnya Skala Unit Deviasi (SUD). Misalnya dalam penjabaran ini kita pergunakan seluruh jarak $\text{Range} - 3DS \text{ s/d} + 3 DS = 6 DS$. Karena nilai huruf yang berarti = 4 unit, maka dalam hal ini kita tentukan besarnya $\text{SUD} = \frac{6DS}{4} = 1,5 DS$

Oleh karena dalam penjabaran ini kita menggunakan DR sebagai pengganti DS, jadi $\text{SUD} = 1,5 \text{ DR} = 1,5 \times 6,8 = 10,25$; dibulatkan = 10.

- 2) Titik-tengah C terletak pada *Mean* = 60, karena C merupakan nilai tengah pada skala penilaian A – B – C – D – TL. Untuk lebih jelasnya, perhatikan gambar berikut:



Oleh karena itu $SUD = 6 DS : 4 = 1,5 DS$

Titik tengah C terletak pada Mean = 60.

- 3) Langkah selanjutnya kita menentukan batas-bawah (*lower limit*) dan batas-atas (*upper limit*) dari masing-masing nilai huruf.

Dari langkah b di atas sudah diketahui bahwa:

— Titik tengah $C = M = 60$. Dengan demikian maka:

— Batas-bawah $C = M - 0,5 SUD$
 $= 60 - 0,5 \times 10 = 55$

— Batas-atas $C = M + 0,5 SUD$
 $= 60 + 0,5 \times 10 = 65$

— Batas bawah $D = M - 1,5 SUD$
 $= 60 - 1,5 \times 10 = 45$

— Skor dibawah 45 = TL

— Batas-atas $B = M + 1,5 SUD$
 $= 60 + 1,5 \times 10 = 75$

— Skor di atas 75 = A.

- 4) Berdasarkan hasil perhitungan pada langkah c tadi di atas, kita mentransfer skor-skor mentah tersebut kedalam nilai huruf, seperti terlihat dalam kolom tiga pada table yang telah kita buat terdahulu:

a) Skor 76 ke atas $= A \rightarrow =$ tidak ada

b) Skor 66 – 75 $= B \rightarrow = 6$ orang

c) Skor 55 – 65 $= C \rightarrow = 8$ orang

d) Skor 45 – 54 = D → = 5 orang

e) Skor 44 – ke bawah = TL → = 1 orang.

Cara penjabaran seperti tersebut diatas ternyata bahwa hasilnya lebih baik. Kita melihat bahwa yang tidak lulus (TL) hanya satu orang, meskipun yang memperoleh A ternyata tidak ada. Hal ini dimungkinkan karena dalam penjabaran tersebut kita menggunakan seluruh *Range* antara – 3 DS sampai dengan + 3 DS, sehingga titik tengah $C = M$ dan batas-lulus (batas-bawah D) terletak jauh dibawah *Mean* ($M - 1,5$ SUD atau – 2,25 DS).

- b. Mengolah Skor Mentah Menjadi Nilai Huruf Dengan Batasan Lulus = *Mean*

Misalkan seorang dosen memperoleh skor dari hasil ujian semester dari 50 orang mahasiswanya sebagai berikut:

97, 93, 92, 90, 87, 86, 86, 83, 81, 80,
80, 78, 76, 76, 75, 74, 73, 72, 72, 71,
69, 67, 67, 67, 64, 63, 63, 62, 62, 60,
58, 57, 57, 56, 56, 54, 52, 50, 47, 45,
43, 39, 36, 36, 32, 29, 27, 26, 20, 16.

Skor mentah ini akan kita olah menjadi nilai huruf A, B, C, D, dan TL. Untuk mencari *Mean* dan DS kita susun skor mentah tersebut ke dalam tabel frekuensi. Kita cari dulu *range* untuk menentukan besarnya interval dan kelas interval.

$$\text{Range} = 97 - 16 = 81$$

$$\text{Kelas interval} = R/i + 1 = 81/10 + 1 = 9$$

Jadi, dengan menentukan besarnya interval = 10, kita peroleh kelas interval = 9

Kelas	Interval	f	d	fd	fd ²
1	96 – 105	1	+ 4	+ 4	16
2	86 – 95	6	+ 3	+ 18	54
3	76 – 85	7	+ 2	+ 14	28
4	66 – 75	10	+ 1	+ 10	10
5	56 – 65	11	0	0	0
6	46 – 55	4	- 1	- 4	4
7	36 – 45	5	- 2	- 10	20
8	26 – 35	3	- 3	- 9	27
9	16 – 25	3	- 4	- 12	48
Jumlah		50 (N)		+ 11 (Σfd)	207 (Σfd^2)

$$M = 60,5 + 10 \left(\frac{+11}{50} \right) = 60,5 + \frac{110}{50}$$

$$= 60,5 + 2,2 = 62,7 \text{ dibulatkan } 63$$

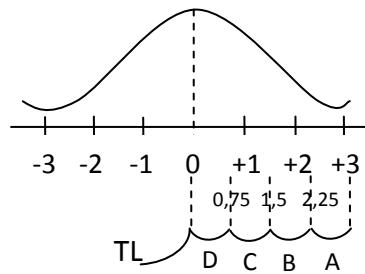
$$DS = 10 \sqrt{\frac{207}{50} - \left(\frac{+11}{50} \right)^2}$$

$$= 10 \sqrt{4,14 - 0,48} = 10 \sqrt{3,66}$$

$$= 10 \times 1,9 = 19$$

Selanjutnya, jika kita akan mengubah skor mentah yang diperoleh dosen itu menjadi nilai huruf A, B, C, D, dan TL dengan batas lulus = *mean*. Caranya adalah sebagai berikut:

Telah ditentukan bahwa batas lulus = *mean* = 63. Jadi, skor mentah dari 63 keatas kita bagi menjadi nilai huruf A, B, C, D, dan skor dibawah 63 dinyatakan Tidak Lulus (TL). Perhatikan gambar berikut :



Bahwa :

$$SUD = 0,75; DS = 0,75 \times 19 = 14,25$$

Dengan demikian, selanjutnya kita dapat menghitung dengan mudah batas atas dan batas bawah dari masing-masing nilai huruf itu sebagai berikut:

- Batas bawah D atau batas lulus = mean = 63
- Skor di bawah 63 = TL
- Batas atas D = $M + 1 SUD = M + 0,75 DS =$
 $= 63 + 14,25 = 77$ (dibulatkan)
- Batas atas C = $M + 2 SUD = M + 1,5 DS =$
 $= 63 + 28,5 = 92$ (dibulatkan)
- Batas atas B = $M + 3 SUD = M + 2,25 DS =$
 $= 63 + 42,75 = 106$ (dibulatkan)
- Skor di atas 106 = A

Dengan perhitungan tersebut maka hasil kelulusan dari 50 mahasiswa itu adalah sebagai berikut:

Yang tidak lulus (TL), skor di bawah 63	= 23 orang
Yang mendapat nilai D, skor 63 – 77	= 15 orang
Yang mendapat nilai C, skor 78 – 92	= 10 orang
Yang mendapat nilai B, skor 93 – 106	= 2 orang
Yang mendapat nilai A, skor di atas 106	= tidak ada

Jika skor mentah yang diperoleh 50 orang mahasiswa itu kita jabarkan menjadi nilai 1 – 10 dengan menggunakan *mean* dan DS aktual

dengan batas-lulus $M + 0,25 DS = 63 + 4,75 = 68$ (dibulatkan), maka yang dapat dinyatakan lulus = 21 orang, dan yang tidak lulus = 29 orang.

- c. Mengolah skor mentah menjadi nilai huruf dengan menggunakan Mean Ideal dan DS Ideal

Berikut ini akan penulis uraikan bagaimana mengolah skor mentah menjadi nilai huruf A, B, C, D, TL, dengan menggunakan *Mean* Ideal dan DS Ideal. Dibandingkan dengan cara seperti yang telah diuraikan pada pasal B di muka, cara berikut lebih mudah karena tidak perlu menyusun tabel distribusi frekuensi. Jika skor maksimum ideal dari tes yang diberikan kepada 50 orang mahasiswa tersebut = 120, maka *Mean* Ideal $\frac{1}{2} \times$ skor max. Ideal = $\frac{1}{2} \times 120 = 60$. Dan DS Ideal dari tes tersebut = $\frac{1}{3} \times 60 = 20$.

Dengan cara penjabaran yang diuraikan dalam pasal B, yakni dengan ketentuan titik-tengah $C = \text{Mean}$, dan $SUD = 1,5 DS$, maka kita peroleh perhitungan sebagai berikut:

$$\text{Mean} = 60. DS = 20. SUD = 1,5 DS = 1,5 \times 20 = 30.$$

$$\text{—Titik tengah } C = \text{Mean} = 60$$

$$\text{—Batas-bawah } C = M - 0,5 SUD = 60 - 15 = 45$$

$$\text{—Batas-atas } C = M + 0,5 SUD = 60 + 15 = 75$$

$$\text{—Batas-bawah } D = M - 1,5 SUD = 60 - 45 = 15$$

$$\text{—Batas-atas } B = M + 1,5 SUD = 60 + 45 = 105$$

$$\text{—Skor di bawah } 15 = \text{TL, dan}$$

$$\text{—Skor di atas } 105 = \text{A.}$$

Untuk lebih jelasnya lihat gambar di bawah ini:

Jika skor mentah yang diperoleh 50 orang mahasiswa itu kita transfer ke dalam nilai huruf sesuai dengan perhitungan di atas, kita peroleh hasil seperti berikut:

Yang tidak lulus (TL), skor di bawah 15 = tidak ada.

Yang mendapat nilai D, skor 15 – 44 = 10 orang.

Yang mendapat nilai C, skor 45 – 75 = 26 orang.

Yang mendapat nilai B, skor 76 – 105 = 14 orang.

Yang mendapat nilai A, skor di atas 105 = tidak ada.

Bandingkan hasil tersebut di atas dengan hasil perhitungan yang diperoleh pada pasal B di muka.

2. Mengolah skor mentah menjadi nilai 1 – 10

Untuk mengolah skor mentah menjadi nilai 1 – 10, kita perlu mencari mean (angka rata-rata) dan DS. Untuk itu skor mentah tersebut kita susun ke dalam tabel distribusi frekuensi. Langkah-langkah menyusun tabel frekuensi adalah sebagai berikut:

a. Kita tentukan dulu banyaknya kelas interval dengan jalan:

- 1) Mencari *range* (R), dengan mengurangi skor maksimum dengan skor minimum (*range* = selisih antara skor maksimum dan skor minimum).
- 2) Bagian *range* kedalam interval-interval yang sama sedemikian rupa sehingga jumlah kelas interval antara 6 – 15 atau 11 – 19.

Rumus untuk mencari kelas interval:

$$\frac{R}{i} + 1$$

- 3) Cara lain untuk mencari atau menentukan besarnya kelas interval dapat juga menggunakan rumus *Sturges* sebagai berikut:

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

K = banyaknya kelas yang dihendaki atau dicari.

1 = merupakan bilangan tetap.

N = banyaknya skor (jumlah siswa yang dites).

- b. Mengisi kolom dua (kolom interval) didalam tabel yang telah tersedia, mulailah dari skor minimum berturut-turut dari interval yang telah ditemukan dan sejumlah kelas yang ditentukan pada langkah pertama.

- c. Membuat *tally* pada kolom 3 (menabulasikan tipa-tiap skor kedalam kelasnya).
- d. Mengisikan angka (jumlah) *tally* kedalam kolom 4 (lajur frekuensi = f).
- e. Menentukan deviasi pada lajur d dengan menetapkan letak *mean dugaan* (M') dengan angka nol pada kelas tertentu. Untuk menduga letak nol tersebut dapat kita pilih kelas yang mengandung frekuensi yang paling tinggi. Selanjutnya kita letakan angka-angka deviasi itu dari nol keatas dan kebawah. Angka-angka diatas nol kita beri tanda + (*plus*) dan angka-angka dibawah nol kita beri tanda - (*minus*).
- f. Mengisi lajur fd dengan mengalikan angka-angka pada lajur f dan d, kemudian hasilnya dijumlahkan pada bagian bawah dari tabel (= fd). Sampai dengan kolom 6 ini (lajur fd) kita telah dapat menghitung besarnya *mean* yang sebenarnya dari tabel tersebut. Akan tetapi, karena kita masih memerlukan mencari DS (*Deviasi Standard*), kita perlu menambah satu kolom lagi untuk mencari fd^2 .
- g. Mengisi lajur fd^2 , kemudian dijumlahkan pula pada bagian bawah dari tabel sehingga kita peroleh Σfd^2 yang diperlukan dalam rumus untuk mencari DS.

Contoh:

Umpamakan seorang guru memperoleh skor mentah dari hasil ulangan sejarah di kelas III SMK yang berjumlah 50 orang siswa sebagai berikut:

16 64 87 36 65 42 43 45 47 51
 77 55 68 42 40 47 42 46 45 50
 20 57 28 7 44 51 40 39 39 57
 28 39 21 48 46 37 41 43 49 71
 29 44 34 50 45 35 44 52 56 45

Untuk mengolah skor mentah di atas menjadi nilai 1 – 10, kita perlu mencari *mean* dan DS. Untuk itu skor mentah tersebut kita susun ke dalam tabel distribusi frekuensi.

$$\text{Skor maksimum} = 87$$

$$\text{Skor minimum} = 7$$

$$\text{Range} = 87 - 7 = 80$$

Banyaknya kelas interval: +

$$\frac{R}{i} + 1 = \frac{80}{8} + 1 = 11$$

Jadi, interval (i) = 8, kelas interval = 11

Kelas	Interval	f	d	fd	Fd ²
1	87 - 94	1	+ 6	6	36
2	79 - 86	0	+ 5	0	0
3	71 - 78	2	+ 4	8	32
4	63 - 70	3	+ 3	9	27
5	55 - 62	4	+ 2	8	16
6	47 - 54	11	+ 1	11	11
7	39 - 46	18	0	0	0
8	31 - 38	4	- 1	- 4	4
9	23 - 30	3	- 2	- 6	12
10	15 - 22	3	- 4	- 9	27
11	7 - 14	1	- 4	- 4	16
Jumlah		N = 50		+ 19 ($\sum fd$)	181 ($\sum fd^2$)

Dengan melihat pada tabel distribusi frekuensi maka:

$$M = 42,5 + 8 \left(\frac{+ 19}{50} \right) = 42,5 + 3,04 = 45,54$$

Dari tabel tersebut juga kita sekarang mencari DS:

$$DS = 8 \sqrt{\frac{181}{50} - \left(\frac{+ 19}{50} \right)^2}$$

$$\begin{aligned}
 &= 8\sqrt{3,62 - 0,1444} \\
 &= 8\sqrt{3,5756} \\
 &= 8 \times 1,89 = 15,12 \text{ dibulatkan} = 15
 \end{aligned}$$

3. Mengolah skor mentah menjadi skor standar Z

Yang dimaksud dengan skor Z adalah skor yang penjabarannya didasarkan atas unit deviasi standar dari *mean*. Dalam hal ini *mean* dinyatakan = 0 (Nol). Oleh karena itu, dengan penjabaran skor-skor mentah menjadi skor standar Z itu kita dapat melihat bagaimana kedudukan skor-skor tersebut dibandingkan dengan rata-rata skor kelompoknya, apakah ia terletak di atas rata-rata kelompok (*mean*) atau dibawahnya. Pengolahan skor mentah menjadi skor Z ini seringkali dirasakan perlunya karena dengan hanya melihat skor mentah saja kita belum dapat memberikan tafsiran yang baik dan tepat. Atau dengan kata lain, dengan hanya mengetahui skor mentah saja dapat menimbulkan tafsiran yang salah mengenai kecakapan seseorang.

Contoh:

Misalkan kita melihat hasil tes dari dua orang anak bernama Umar dan Basir.

Hasil tes Umar sebagai berikut:

Bahasa Indonesia	= 65
Matematika	= 55
IPS	= 70

Hasil tes Basir sebagai berikut:

Bahasa Indonesia	= 70
Matematika	= 60
IPS	= 60

Pertanyaan yang timbul ialah: siapa diantara kedua anak tersebut yang sebenarnya lebih baik prestasinya? Umar atau Basir?. Dengan menggunakan skor standar (skor Z) kita dapat mengetahui siapa yang sebenarnya lebih

baik atau lebih tinggi prestasinya. Dalam hal ini perlu diingat bahwa *mean* dan DS untuk tiap mata pelajaran yang dicapai oleh kedua anak itu sama.

Skor Z Umar:

Bahasa Indonesia:

$$\frac{65 - 60}{4.0} = \frac{+5}{4} = +1,25$$

Matematika:

$$\frac{55 - 45}{4.0} = \frac{+10}{4} = +2,5$$

IPS:

$$\frac{70 - 75}{5.0} = \frac{-5}{5} = -1,0$$

$$\text{Jumlah} = +2,75$$

Skor Z Basir:

$$\frac{70 - 60}{4.0} = \frac{+10}{4} = +2,5$$

$$\frac{60 - 45}{4.0} = \frac{+15}{4} = +3,75$$

$$\frac{60 - 75}{5.0} = \frac{-15}{5} = -3,0$$

$$\text{Jumlah} = +3,25$$

Dengan melihat hasil tersebut di atas, dapat diambil kesimpulan bahwa prestasi Basir ternyata lebih baik daripada Umar.

4. Mengolah skor mentah menjadi skor standar T

Dengan bersumber pada skor Z seperti telah dibicarakan dimuka, banyak pula dikembangkan skor-skor standar lainnya yang dikenal orang sebagai *angka skala*. Jenis skor standar yang meupakan angka skala yang telah banyak dikenal orang antara lain ialah skor T. Yang dimaksud dengan skor T ialah angka skala yang menggunakan dasar *mean* = 50 dan jarak tiap deviasi standar (DS) = 10, didalam *range* - 3 DS sampai dengan + 3 DS, T tersebar dari 20 s.d. 80, tanpa bilangan-bilangan *minus*.

Penjabaran skor mentah kedalam skor T ini pun sering kali kita perlukan untuk mengetahui bagaimana kedudukan seorang anak yang memperoleh skor tersebut dibandingkan dengan kelompoknya didalam suatu hasil tes. Selain itu dengan penjabaran kedalam skor T ini, hasil-hasil tes (skor mentah) yang diperoleh dari beberapa mata pelajaran yang memiliki *mean* dan *DS* yang berbeda-beda dapat di ubah menjadi skor-skor standar dengan satu skala unit deviasi. Dengan demikian, suatu panitia tujuan sekolah, misalnya dapat menentukan “Batas lulus” dari berbagai mata pelajaran dengan kedudukan nilai skor yang sama setelah setiap skor dari mata pelajaran-mata pelajaran tersebut dijabarkan kedalam skor T.

Rumusnya:

$$\text{skor } T = \left(\frac{X - M}{DS} \right) 10 + 50$$

Contoh:

Jika skor-skor yang diperoleh Umar tadi kita jabarkan ke dalam skor T, maka akan kita peroleh data sebagai berikut:

Bahasa Indonesia :

$$= \left(\frac{65 - 60}{4} \right) \times 10 + 50 = (+1,25) \times 10 + 50 = 62,5$$

Matematika:

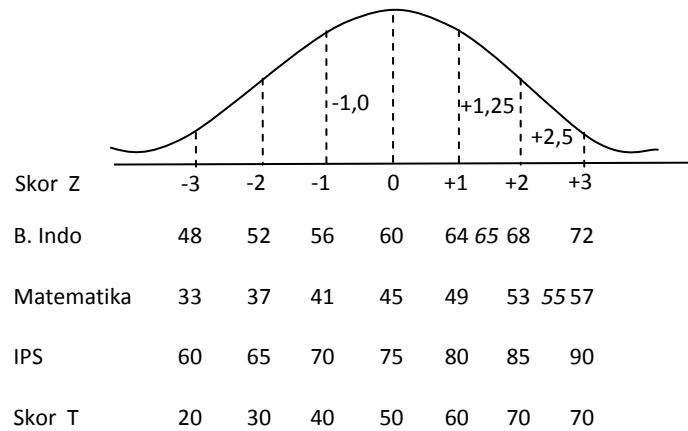
$$= \left(\frac{55 - 45}{4} \right) \times 10 + 50 = (+2,5) \times 10 + 50 = 75,0$$

IPS:

$$= \left(\frac{70 - 75}{5} \right) \times 10 + 50 = (-1,0) \times 10 + 50 = 40,0$$

Dengan melihat hasil penjabaran ke dalam skor T tersebut di atas, secara cepat kita dapat mengatakan bahwa Umar memiliki prestasi yang cukup baik dalam matematika dibandingkan dengan teman-teman sekelompoknya, dan kurang baik prestasinya dalam IPS. Ingat bahwa

dengan menjabarkan ke dalam skor T itu, kita telah menyamakan besarnya *Mean* dari ketiga mata pelajaran tersebut, yaitu *Mean* = 50.



Jika skor-skor yang diperoleh Umar tadi kita letakkan dalam skala skor T yang disejajarkan dengan skor Z, akan kita lihat seperti terlukis pada gambar di atas. Gambar di atas menjelaskan bagaimana kedudukan skor-skor yang diperoleh Umar dibandingkan dengan rata-rata kelompoknya. Perhatikan skor-skor yang dicetak miring pada gambar, dan skor Z-nya yang terletak di atas dasar kurva.

BAB IX

MENILAI KUALITAS ALAT EVALUASI

A. Teknik Analisis Tes

Teknik analisis tes dilakukan untuk mengetahui kelayakan perangkat tes dalam pengambilan data. Analisis yang dilakukan meliputi uji validitas, tingkat kesukaran, daya pembeda dan reliabilitas instrumen.

1. Validitas

Uji validitas dilakukan untuk mengetahui ukuran kesahihan butir soal yang digunakan sebagai instrumen untuk mengambil data pada saat penelitian. Uji ini sangat penting agar diperoleh data yang valid pada saat penelitian. Menurut Arikunto (2006:168) validitas tes adalah suatu ukuran yang menunjukkan tingkat-tingkat kevalidan atau kesahihan suatu instrumen. Untuk mengetahui validitas item dari tes, digunakan teknik kolerasi "*Pearson's Product Moment*". Adapun perumusannya sebagai berikut:

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

(Arikunto, S. 2006 : 170)

dengan : r_{xy} = koefisien kolerasi antara variabel x dan y

x = skor siswa pada butir item yang diuji validitasnya

y = skor total yang diperoleh siswa

Untuk menginterpretasikan koefisien korelasi yang telah diperoleh digunakan tabel nilai *r product moment*. Untuk menginterpretasikan tingkat validitasnya, maka koefisien kolerasinya dikategorikan pada kriteria seperti dalam tabel berikut ini.

Tabel 9.1
Interpretasi Validitas

Koefisien Korelasi	Kriteria validitas
$0,80 < r_{11} \leq 1,00$	sangat tinggi
$0,60 < r_{11} \leq 0,80$	tinggi
$0,40 < r_{11} \leq 0,60$	cukup
$0,20 < r_{11} \leq 0,40$	rendah
$0,00 < r_{11} \leq 0,20$	sangat rendah

2. Reliabilitas

Uji validitas dilakukan untuk mengetahui tingkat keajegan suatu perangkat tes yang digunakan sebagai instrumen pada suatu penelitian. Suatu perangkat tes yang baik merupakan perangkat yang menghasilkan skor yang tidak berubah-ubah atau ajeg. Menurut Syambasri (Dahlan, 1990:) “reliabilitas adalah tingkat keajegan (konsistensi) suatu tes, yakni sejauh mana suatu tes dapat dipercaya untuk menghasilkan skor yang konsisten (tidak berubah-ubah)”. Dalam beberapa penelitian, untuk menentukan reliabilitas tes uraian digunakan rumus alpha sebagai berikut :

$$r_{11} = \left(\frac{n}{n-1} \right) \left(1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right)$$

(Arikunto, S. 2003 : 109)

dengan : r_{11} = koefisien reliabilitas perangkat tes

$\sum \sigma_i^2$ = jumlah varians skor tiap-tiap item

σ_t^2 = varians total

n = jumlah siswa

Rumus varians yang digunakan yaitu :

$$\sigma_b^2 = \frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N} \text{ (Varians skor tiap butir soal)}$$

$$\sigma_t^2 = \frac{\sum Y^2 - \frac{(\sum Y)^2}{N}}{N} \text{ (Varians total)}$$

(Arikunto, S. 2003:110)

Adapun tolak ukur untuk menginterpretasikan derajat reliabilitas instrumen yang diperoleh dapat dilihat pada tabel berikut ini.

Tabel 9.2

Interpretasi Reliabilitas

r_{11}	Interpretasi
$0,80 < r_{11} \leq 1,00$	Sangat tinggi
$0,60 < r_{11} \leq 0,80$	Tinggi
$0,40 < r_{11} \leq 0,60$	Sedang
$0,20 < r_{11} \leq 0,40$	Rendah
$0,00 < r_{11} \leq 0,02$	Sangat rendah

3. Daya Pembeda

Uji daya pembeda dilakukan untuk mengetahui kemampuan suatu soal untuk membedakan antara kemampuan siswa pada kelompok atas (siswa berkemampuan tinggi) dan kemampuan siswa pada kelompok bawah (siswa berkemampuan rendah). Suharsimi Arikunto (1991 : 213) mengemukakan bahwa daya pembeda adalah kemampuan suatu soal untuk membedakan antara siswa yang pandai (berkemampuan tinggi) dan siswa yang bodoh (berkemampuan rendah). Untuk menghitung daya pembeda tiap item soal, menggunakan rumus:

$$DP = \frac{S_A - S_B}{I_A} \times 100\% \quad \text{(Karno To, 1996 : 15)}$$

dengan : DP = indek daya pembeda item satu butir soal tertentu
 S_A = jumlah skor kelompok atas pada butir soal yang diolah
 S_B = jumlah skor kelompok bawah pada butir soal yang diolah

I_A = jumlah skor ideal salah satu kelompok atas atau bawah

Nilai daya pembeda (DP) yang diperoleh, kemudian diinterpretasikan pada kategori berikut ini .

Tabel 9.3
Interpretasi Daya Pembeda

Koefisien Korelasi	Kriteria validitas
negatif - 9 %	sangat buruk
10 % - 19 %	buruk
20 % - 29 %	Agak baik
30 % - 49 %	baik
50 % - keatas	sangat baik

4. Tingkat Kesukaran

Uji tingkat kesukaran dilakukan untuk mengetahui tingkat kesukaran suatu soal, uji ini penting agar dalam suatu perangkat soal tidak didominasi oleh soal mudah atau sukar saja. Suharsimi (1991:210) menyatakan bahwa bilangan yang menunjukkan sukar dan mudahnya suatu soal disebut indeks kesukaran (*difficulty index*). Selanjutnya Karno To (1999) menjelaskan untuk menghitung taraf kemudahan dipergunakan rumus :

$$TK = \frac{S_A + S_B}{I_A + I_B} \times 100\% \quad (\text{Karno To, 1996 : 16})$$

dengan : S_t = jumlah skor kelompok atas

S_t = jumlah skor kelompok bawah

I_A = jumlah skor ideal kelompok atas

I_B = jumlah skor ideal kelompok bawah

Untuk menginterpretasikan tingkat kesukaran tiap item soal tiap tahap dilakukan dengan interpretasi pada tabel berikut

Tabel 9.4
Interpretasi Tingkat Kesukaran

Koefisien Korelasi	Kriteria validitas
0 % - 15 %	sangat buruk
16 % - 30 %	buruk
31 % - 70 %	cukup
71 % - 85 %	baik
86 % - 100 %	sangat baik

BAB X

PENUTUP

A. Kesimpulan

Peran sekolah dan guru-guru yang pokok adalah menyediakan dan memberikan fasilitas untuk memudahkan dan melancarkan cara belajar siswa. Guru harus dapat membangkitkan kegiatan-kegiatan yang membantu siswa untuk meningkatkan cara dan hasil belajarnya. Namun, di samping itu kadang-kadang guru merasa bahwa evaluasi itu merupakan sesuatu yang bertentangan dengan pengajaran. Hal ini timbul karena sering kali terlihat bahwa adanya kegiatan evaluasi justru merisaukan dan menurunkan gairah belajar pada siswa. Jadi seolah-olah kegiatan evaluasi bertentangan dengan kegiatan pengajaran. Pendapat demikian itu pada hakekatnya tidak benar. Memang, evaluasi yang dilakukan secara tidak benar dapat mematikan semangat siswa dalam belajar. Sebaliknya, evaluasi yang dilakukan dengan baik dan benar, seharusnya dapat meningkatkan mutu dan hasil belajar, karena kegiatan evaluasi itu membantu guru untuk memperbaiki cara mengajar dan membantu siswa dalam meningkatkan cara belajarnya. Bahkan dapat dikatakan bahwa evaluasi tidak dapat dilepaskan dari pengajaran.

Berdasarkan pemaparan diatas, dapat disimpulkan bahwa evaluasi merupakan bagian yang tak terpisahkan dari proses pengajaran. Kegiatan evaluasi merupakan suatu cara yang paling efektif untuk mengetahui sejauh mana keberhasilan penguasaan siswa terhadap materi yang diajarkan, sekaligus mengukur keberhasilan guru dalam mengajar. Oleh karena itu, sudah selayaknya bagi seorang guru/pendidik untuk terus meningkatkan kinerjanya dalam melayani dan memfasilitasi para peserta didik dalam setiap proses pembelajaran.

B. Saran

Evaluasi Pembelajaran merupakan kelompok Mata Kuliah Keahlian Profesi (MKKP) pada program studi Pendidikan Teknik Mesin/S1, khususnya dalam hal ini ditujukan bagi mahasiswa konsentrasi Otomotif. Dengan metode pendekatan pembelajaran bersifat ekspositori, penugasan, tanya jawab, dan drill, merupakan metode yang sangat tepat bagi tercapainya proses pembelajaran efektif, sehingga tujuan pembelajaran pun akan tepat sasaran. Satu hal yang mungkin bisa dijadikan bahan evaluasi perkuliahan ke depan, yaitu masalah alokasi waktu perkuliahan yang dirasakan sangat minim sekali untuk berdiskusi. Oleh karena itu, untuk perkuliahan selanjutnya diusahakan untuk lebih memaksimalkan waktu yang ada. Pada umumnya penulis sangat mendukung sekali terhadap metode pembelajaran yang diterapkan, karena dengan melalui metode ini mahasiswa akan lebih terlatih terhadap peka terhadap penggalian ilmu secara mandiri. Oleh karena itu, untuk proses pembelajaran kedepannya mudah-mudahan terdapat hubungan komunikasi yang baik antara pengajar (dosen) dan peserta didik (mahasiswa) demi tercapainya tujuan pembelajaran bersama.

DAFTAR PUSTAKA

- Arikunto, Suharsimi. (2006). *Prosedur Penelitian Suatu Pendekatan Praktik*. Jakarta : Rineka Cipta
- Arikunto, Suharsimi. (2003). *Dasar-dasar Evaluasi Pendidikan*. Jakarta : Bumi Aksara.
- Dahlan. (1990). *Model-model mengajar*. Bandung: Dipenogoro
- Karno To. (1996). *Mengenal Analisis Tes (Pengenalan Ke Program Komputer ANATES)*. Jurusan Psikologi Pendidikan dan Bimbingan: FIP IKIP Bandung
- Nurkencana & Sumartana, W (1983). *Evaluasi Pendidikan*. Surabaya : Usaha Nasional.
- Purwanto, Ngalim (1994). *Prinsip-prinsip dan Teknik Evaluasi Pengajaran*. Bandung : PT Remaja Rosda Karya
- Sudijono, Anas (2003). *Pengantar Evaluasi Pendidikan*. Jakarta : PT Raja Grafindo Persada.
- Suherman, Erman (2003). *Evaluasi Pembelajaran Matematika*. Bandung : JICA FPMIPA UPI.