



# ASSESMEN PEMBELAJARAN MATEMATIKA

**Anetha L. F. Tilaar**

# ASSESMEN PEMBELAJARAN MATEMATIKA



**UNIMA**

# **Assesmen Pembelajaran Matematika**

**Penulis**  
**Anetha L. F. Tilaar**

**Penerbit**  
**Percetakan Unima Press**  
**Alamat : Kampus Unima di Tondano 95618**  
**Telp : (0431) 321845, 321846, 3211847**  
**Fax : (0431) 321866**

ISBN 978-602-14611-3-6



## KATA PENGANTAR

Puji dan Syukur dinaikkan kepada Yang Maha Besar Tuhan, karena Berkat dan Penyertaan-Nya penulis dapat menyelesaikan penyusunan dan penulisan Buku Assesmen Pembelajaran Matematika.

Penulisan buku ini bertujuan untuk membantu para mahasiswa mengupdate pengetahuan dan wawasan di bidang Assesmen Pembelajaran Matematika sesuai dengan perkembangan ilmu pengetahuan dan teknologi yang sangat pesat di era globalisasi ini.

Dengan selesainya penulisan buku ini, penulis berterimakasih kepada:

1. Rektor Universitas Negeri Manado yang selalu memotivasi para dosen untuk meningkatkan pengetahuan.
2. Ketua LP2AI yang memberi masukan untuk penulisan buku ini.
3. Semua teman yang membantu dalam penyelesaian penulisan buku ini.

Akhirnya, penulis menyadari bahwa buku yang dihasilkan ini masih membutuhkan masukan untuk memperbaiki kualitas. Olehnya saran yang bertujuan untuk menyempurnakan buku ini, diterima dengan senang hati.

Tondano, Oktober 2018

*Anetha L. F. Tilaar*

## DAFTAR ISI

<b>Kata Pengantar .....</b>	<b>i</b>
<b>Daftar Isi .....</b>	<b>iii</b>
<b>Bab 1 Pendahuluan .....</b>	<b>1</b>
A. Peran Assesmen Dalam Pembelajaran .....	1
B. <i>Student-Centered Learning</i> (SCL) .....	2
C. Pembelajaran Melibatkan Assesmen.....	6
D. Assesmen Dalam SKM dan Hubungannya Dengan Pembentukan Kompetensi.....	7
<b>Bab II Assesmen Pembelajaran .....</b>	<b>11</b>
A. Pengertian Assesmen .....	11
B. Tujuan Pelaksanaan Assesmen .....	12
C. Manfaat Pelaksanaan Assesmen .....	14
D. Subjek, Objek/Sasaran .....	14
E. Fungsi Assesmen.....	15
F. Prinsip Assesmen .....	15
<b>Bab III Perbedaan Pengukuran, Penilaian dan Evaluasi .....</b>	<b>18</b>
A. Pendahuluan .....	18
B. Pengukuran.....	18
C. Penilaian .....	19
D. Evaluasi .....	20
E. Pengertian Pengukuran, Penilaian dan Evaluasi .....	21
F. Evaluasi dalam Pendidikan .....	22
G. Penilaian dalam Pendidikan .....	23
H. Pengukuran dalam Pendidikan .....	24
I. Perbedaan Evaluasi, Penilaian dan Pengukuran.....	25
<b>Bab IV Pelaksanaan Assesmen dan Evaluasi .....</b>	<b>27</b>
A. Proses Pelaksanaan.....	29
B. Tujuan .....	31
C. Prinsip - prinsip .....	32
D. Prosedur Pelaksanaan.....	32
E. Kegagalan Guru dalam melakukan Assesmen dan Evaluasi .....	33
F. Jenis – jenis Penilaian .....	39

<b>Bab V Teknik Evaluasi .....</b>	<b>42</b>
A. Teknik Tes.....	43
1. Bentuk soal benar salah.....	45
2. Bentuk soal pilihan ganda atau pilihan jamak.....	46
3. Bentuk soal menjodohkan .....	46
4. Bentuk soal jawaban singkat .....	47
B. Teknik Non tes .....	47
1. Pengamatan atau observasi .....	47
2. Wawancara.....	47
3. Angket.....	48
4. Skala.....	48
<b>Bab VI Instrumen Tes Objektif.....</b>	<b>51</b>
A. Pengertian Tes Objektif.....	51
B. Penggolongan Tes Objektif .....	53
C. Penggunaan Tes Objektif .....	61
<b>Bab VII Instrumen Tes Essay .....</b>	<b>63</b>
A. Pengertian Tes Essay/Uraian.....	63
B. Petunjuk Penyusunan Tes Essay/Uraian .....	67
C. Penggunaan Tes Uraian.....	67
D. Contoh Soal Tes Uraian .....	67
E. Kaidah Pengembangan Instrumen Tes Essay/ Uraian.....	68
1. Pengembangan Butir Tes Uraian .....	68
2. Pedoman Penilaian .....	68
<b>Bab VIII Asesmen Hasil Belajar .....</b>	<b>70</b>
A. Asesmen Formatif dan Sumatif.....	70
B. Asesmen Obyektif dan Subyektif.....	71
C. Asesmen Acuan Patokan dan Acuan Normatif .....	71
D. Asesmen Formal dan Informal .....	71
E. Asesmen Autentik .....	71
F. Asesmen Portofolio .....	72
<b>Bab IX Jenis-Jenis Asesmen .....</b>	<b>76</b>
A. Pendahuluan .....	76
B. Asesmen Formal dan Informal .....	78
C. Asesmen Berbasis Performa.....	82
D. Asesmen Portofolio .....	88
E. Asesmen Tes Obyektif .....	92
F. Asesmen Tes Subyektif .....	98

<b>Bab X Kisi-Kisi Tes.....</b>	<b>104</b>
A. Pengertian Kisi-kisi.....	104
B. Fungsi Kisi-kisi .....	104
C. Penulisan Kisi-kisi .....	105
D. Penentuan dan Penyebaran Soal .....	107
E. Format Penulisan Kisi-Kisi Soal.....	107
F. Syarat-syarat Kisi-kisi yang Baik .....	108
G. Jenis Perilaku yang Dapat Diukur.....	108
<b>Bab XI Validitas Instrumen Assesmen dan Evaluasi .....</b>	<b>117</b>
A. Pengertian Uji Validitas Instrumen.....	117
B. Pengertian Validitas Instrumen .....	118
C. Validitas Instrumen .....	119
D. Validitas dan Reliabilitas Penilaian .....	121
<b>Bab XII Reliabilitas Instrumen Assesmen dan Evaluasi.....</b>	<b>124</b>
A. Pengertian Reliabilitas .....	124
B. Reliabilitas Instrumen .....	124
C. Pengukuran Reliabilitas .....	125
D. Teknik untuk Menentukan Reliabilitas .....	126
1. Teknik Pengukuran Ulang .....	126
2. Teknik Belah Dua .....	126
3. Teknik Bentuk Paralel.....	126
<b>Bab XIII Skala Pengukuran.....</b>	<b>128</b>
A. Tipe Skala Pengukuran .....	128
1. Tipe Skala .....	128
2. Skala Nominal.....	129
3. Skala Ordinal.....	131
4. Skala Interval .....	132
5. Skala Rasio.....	134
B. Macam-macam Data .....	135
1. Data Nominal .....	135
2. Data Kuantum .....	135
C. Macam-macam Skala.....	136
1. Skala Likert.....	136
2. Skala Guttman.....	137
3. Semantic Deferential.....	137
4. Rating Scale .....	137
<b>Daftar Pustaka.....</b>	<b>141</b>

## **BAB I PENDAHULUAN**

### **A. Peran Assesmen Dalam Pembelajaran**

Pergeseran paradigma pendidikan dari sentralisasi ke desentralisasi yang dicirikan dengan kebijakan dan pelaksanaan pendidikan *bottom-up* telah mengubah praktik pendidikan nasional kita. Hal ini tercermin dalam pengelolaan lembaga maupun pengelolaan pembelajaran. Di tingkat kebijakan, ditetapkannya delapan Standar Pendidikan Nasional (SPN) menunjukkan adanya upaya untuk memfasilitasi peningkatan mutu pendidikan dari berbagai perspektif yang mungkin dilakukan untuk itu. Penetapan sekolah dengan kualitas tertentu sebagai Sekolah Kategori Mandiri (SKM) merupakan salah satu upaya untuk itu. PP No 19 tahun 2005, pasal 11 ayat (3) menyatakan bahwa beban belajar untuk SMA dan bentuk lain yang sederajat pada jalur pendidikan formal katagori mandiri dinyatakan dalam SKS. Ketentuan tersebut mengisyaratkan bahwa SKM harus menerapkan SKS. Penerapan sistem kredit semester didasarkan oleh kenyataan bahwa kecepatan belajar seseorang (siswa) tidak sama disebabkan oleh keunikan masing-masing dilihat dari potensi yang dimiliki dan minat. Dengan begitu, beban belajar masing-masing siswa perlu disesuaikan dengan potensi yang dimiliki tersebut.

Dalam SKM, Standar Kompetensi (SK) dan Kompetensi Dasar (KD) yang telah ada pada SI disusun menjadi satuan kredit semester (sks), menjadi 120 sks, yang terdistribusi dalam berbagai mata pelajaran yaitu :

- a) mata pelajaran wajib/pokok yang harus diambil oleh seluruh peserta didik;
- b) pilihan paket, sebagai dasar untuk mendukung bidang kemampuan yang akan dipilih di Perguruan Tinggi,
- c) pilihan bebas, sesuai dengan bakat dan minat peserta didik,
- d) kelompok MP Pilihan Paket, meliputi berbagai bidang kemampuan yang diperlukan peserta didik untuk melanjutkan ke pendidikan lebih lanjut, yang meliputi : Program akademik (teknik, Ilmu Kesehatan, Sains, Ekonomi, Ilmu Sosial, Bahasa, Hukum dan sebagainya, dan program profesional seperti politeknik. Beban belajar peserta didik dinyatakan dengan sks yaitu 16-27 sks per-semester, dimana kecepatan belajar normal rata-rata 20 sks per-semester.

Adanya sistem ini mengisyaratkan pentingnya kinerja pembelajaran yang optimal, sebab sangat penting guru berusaha menggali potensi siswa dengan cara-cara pembelajaran yang berterima sesuai dengan karakteristik siswa. Sangat mungkin sistem SKS ini tidak berhasil jika pembelajaran yang diselenggarakan oleh guru tidak bermutu.

### **B. *Student-Centered Learning (SCL)***

Sebagai Inti Pembelajaran SCL adalah refleksi dari ciri kehidupan global yang penuh dengan kompetisi dalam perubahan yang sangat cepat. Lulusan SMP-BI harus memiliki kemampuan dan strategi *problem solving* dan kemampuan berfikir kritis. Siswa harus berkembang kompetensinya yang dibangun dari pengetahuan tentang fakta, konsep, prosedur, dan metakognisi (Anderson dan Krathwohl, 2004; bandingkan dengan B. S. Bloom yang membagi kemampuan menjadi tiga domain yaitu domain kognitif, afektif, dan psikomotor). Sejauh ini, hal yang masih dirasakan kurang dalam proses pendidikan kita adalah kurangnya latihan *problem solving*. Belajar secara *problem solving* adalah *learning to learn*, yaitu kemampuan yang dicapai akan membantu siswa belajar selanjutnya. Untuk itu, yang harus dibangun adalah kompetensi. Pembelajaran yang *content-based* tidak tepat. Guru menggunakan konten/materi bukan sebagai sasaran pembelajaran, namun sebagai jalan membangun kompetensi.

Menurut sejarahnya, pembelajaran yang berpusat pada peserta didik (*Student-Centered Learning*, disingkat SCL) lahir pada awal abad ke-20, yaitu pada saat orang-orang mulai meyakini bahwa pendidikan harus memperhitungkan peserta didik sebagai unsur aktif dalam proses inkuiri, yaitu proses memecahkan masalah yang dihadapinya sendiri. Dibawah pengaruh perspektif pendidikan yang disebut *Progressive Education* (lahir di Amerika Serikat) yang meyakini bahwa pengalaman langsung adalah inti dari belajar, para pendukung *Progressive Education* menentang pembelajaran yang menganggap bahwa peserta didik sebagai kantong kosong yang baru berisi bila diisi oleh guru (teori Tabula Rasa). Bagi pendidikan progresif, peran guru adalah sebagai fasilitator dan pemandu dalam proses pemecahan masalah peserta didik.

John Dewey adalah pelopor pandangan progresif ini. Dia menegaskan bahwa kelas adalah laboratorium yang memotret kehidupan yang sebenarnya. Dia mengajak guru untuk menggunakan masalah riil sehari-hari untuk dipecahkan oleh siswa, sebagai bahan pembelajaran.

Dewey menekankan bahwa pembelajaran yang bermakna adalah pembelajaran yang memuat masalah-masalah nyata yang sedang dihadapi, tidak tentang hal-hal yang abstrak bagi siswa. Dewey dikenal dengan filosofi pendidikan *learning by doing*. Ciri-ciri pembelajaran progresif antara lain, ruang kelas yang diatur secara fleksibel, keleluasaan bagi peserta didik untuk bekerja kelompok maupun individual sesuai dengan kebutuhannya, peserta didik ikut berperan dalam menentukan aturan kelas, dan materi pembelajaran yang kaya dan variatif.

Selain pengaruh pendidikan progresif, juga ada pengaruh perspektif *open classroom* yang meyakini bahwa peserta didik memiliki motivasi intrinsik untuk belajar, dan dorongan dari dalam ini hanya bisa dipuaskan melalui kegiatan eksplorasi dan pemecahan masalah (*problem solving*). Pada akhir tahun 70an, dibawah pengaruh psikologi kognitif, berkembang perspektif konstruktivisme dalam pembelajaran.

Konstruktivisme berarti bahwa peserta didik membangun (*to construct*) pemahamannya tentang dunia. Berbicara mengenai konstruktivisme bukanlah berbicara tentang suatu teknik tertentu dalam pembelajaran, melainkan kita berfikir tentang proses perolehan pengetahuan dan asesmennya. Ada dua kata kunci dalam konstruktivisme, yaitu mahasiswa aktif (*active*) dan memperoleh makna (*meaning*) (Elliott, dkk, 2000); dimana pembelajaran konstruktivis tersebut digambarkan sebagai berikut:

“Peserta didik tidak semata-mata merekam atau mengingat materi yang dipelajari, melainkan mengkonstruksi suatu representasi mental yang unik tentang materi tersebut, tugas yang akan dipentaskan, memilih informasi yang dianggapnya relevan, dan memahami informasi tersebut berdasarkan pengetahuan yang ada padanya, dan kebutuhannya. Siswa menambahkan informasi yang diperlukannya tidak selalu dari materi yang disediakan guru. Ini merupakan suatu proses yang aktif karena siswa harus melakukan berbagai kegiatan kognitif, afektif, dan psikomotorik agar informasi tersebut bermakna bagi dirinya “(p. 15, terjemahan oleh penulis makalah).

Belakangan, berbagai interpretasi muncul tentang bagaimana konstruksi pengetahuan itu terwujud pada peserta didik; ada yang mengatakan bahwa peserta didik itu sendiri mampu membangunnya, tapi ada pula yang mengatakan bahwa konstruksi pengetahuan terjadi dalam interaksi sosial seperti teman sebaya, dan keluarga. Yang pertama diwakili oleh J. Piaget, yang mengatakan bahwa konstruksi makna terjadi

melalui proses asimilasi dan akomodasi. Asimilasi adalah akuisisi pengetahuan yang sesuai dengan yang telah ada sebelumnya; dan akomodasi adalah proses akuisisi terhadap hal-hal baru yang belum ada dalam skema (pengetahuan yang tersimpan dibenak) yang bersangkutan. Di lain pihak, Vygotsky mengatakan bahwa konstruksi pengetahuan terjadi melalui proses interaksi sosial dengan orang lain yang lebih mampu (dalam istilah Vygotsky: *skilled individuals*). Diyakini bahwa konstruksi makna akan terjadi jika proses akuisisi pengetahuan dilakukan dalam lingkungan sosial budaya yang sesuai.

Dibawah pengaruh perspektif konstruktivis, pembelajaran yang dianggap dapat menjawab tantangan pendidikan global sekarang ini (pendidikan yang bermakna, bukan pendidikan yang membebani hidup) adalah pembelajaran yang berpusat pada siswa.

Berdasarkan hakikat SCL tersebut di atas, maka dapat dilihat perbedaan antara SCL dengan pembelajaran yang berpusat pada guru dan berorientasi pencapaian materi (*Teacher-centered, content-oriented/TCCO*), sebagai berikut :

<i>Teacher Centered</i>	<i>Student-Centered Learning</i>
<b>1</b>	<b>2</b>
Pengetahuan ditransfer dari guru ke siswa	Siswa secara aktif mengembangkan pengetahuan dan keterampilan yang dipelajarinya
Siswa menerima pengetahuan secara pasif	Siswa secara aktif terlibat didalam mengelola pengetahuannya
Lebih menekankan pada penguasaan materi	Penguasaan materi dan juga mengembangkan karakter siswa ( <i>life-long learning</i> )
Biasanya memanfaatkan media tunggal	Multimedia
Fungsi guru sebagai pensuplai informasi utama dan evaluator	Guru sebagai fasilitator, evaluasi dilakukan bersama dengan siswa
Proses pembelajaran dan asesmen dilakukan secara terpisah	Terpadu dan berkesinambungan
Menekankan pada jawaban yang benar saja	Menekankan pada pengembangan pengetahuan. Kesalahan menunjukkan proses belajar dan dapat digunakan sebagai salahsatu sumber belajar
Cocok untuk pengembangan ilmu dalam satu disiplin saja	Untuk pengembangan ilmu interdisipliner

Iklm belajar lebih individual dan kompetitif	Iklm yang tercipta lebih bersifat kolaboratif, supaortif, dan kooperatif
Proses pembelajaran hanya terjadi pada siswa	Siswa dan guru belajar bersama dalam mengembangkan konsep dan keterampilan
Pengajaran mengambil porsi waktu terbanyak	Pengajaran dan berbagai kegiatan lain dalam proses belajar
Penekanan pada ketuntasan materi	Penekanan pada pencapaian target kompetensi
Penekanan pada cara pembelajaran yang dilakukan oleh guru	Penekanan pada bagaimana cara siswa belajar. Penekanan pada <i>problem-based learning</i> dan <i>skill competency</i>

Dari pembahasan di atas, lalu bagaimana peran guru dan siswa dalam SCL?

#### **Peran Guru**

- Bertindak sebagai fasilitator dalam proses pembelajaran.
- Mengkaji kompetensi mata pelajaran yang harus dikuasai siswa pada akhir pembelajaran.
- Merancang strategi dan lingkungan pembelajaran yang dapat menyediakan beragam pengalaman belajar.
- Membantu siswa mengakses informasi, menata, dan memprosesnya untuk dimanfaatkan dalam memecahkan permasalahan hidup sehari-hari.
- Mengidentifikasi dan menentukan pola penilaian hasil belajar siswa yang relevan dengan kompetensi yang akan diukur.

#### **Peran siswa**

- Mengkaji kompetensi yang disampaikan guru.
- Mengkaji strategi pembelajaran yang ditawarkan oleh guru.
- Membuat rencana pembelajaran untuk mata pembelajaran yang diikutinya.
- Belajar secara aktif dalam kelompok maupun individual (dengan cara mendengar, membaca, menulis, diskusi, pemecahan masalah; serta terlibat dalam kegiatan berfikir tingkat tinggi seperti analisis, sintesis, dan evaluasi).

### **C. Pembelajaran Melibatkan Asesmen**

Telah kita ketahui bersama bahwa dalam pembelajaran terdapat berbagai komponen, baik komponen statik maupun dinamik sebagai instrumental input. Interaksi antara guru dan siswa melibatkan kurikulum, materi, media dan sumber belajar, kegiatan pembelajaran, dan assesmen. Kegiatan pembelajaran inovatif membutuhkan dilakukannya assesmen yang sesuai. Assesmen yang baik adalah assesmen yang tepat untuk mengukur indikator pencapaian dan dengan cara apa pembelajaran dilakukan. Uraian di atas menjelaskan mengapa diperlukan assesmen baik terhadap proses maupun hasil belajar. Sebagai contoh, pembelajaran berbasis masalah memerlukan metode assesmen yang sesuai. Misalnya, cara penyelesaian masalah adalah proses belajar yang sangat penting untuk dipantau. Untuk itu diperlukan teknik asesmen seperti lembar observasi, ceklis kinerja, dan sejenisnya. Jika dipilih tes objektif sebagai metode asesmennya, misalnya teknik pilihan ganda, maka kualitas proses penyelesaian masalah tersebut sulit untuk dipantau secara objektif. Oleh karena itu, assesmen harus dipahami sebagai upaya mengefektifkan proses belajar mengajar dalam rangka mencapai tujuan pembelajaran itu sendiri.

Assesmen merupakan sebuah proses pengumpulan informasi yang terus menerus berlangsung untuk mengukur performansi murid dan proses pembelajaran. Assesmen perkembangan dan belajar anak memiliki nilai penting. Tidak hanya mengukur kemajuan anak-anak sebagai bentuk evaluasi program, asesmen juga berguna untuk mengidentifikasi kebutuhan-kebutuhan pengembangan staf dan perencanaan pembelajaran di masa yang akan datang

Assesmen yang tepat berguna untuk membantu anak-anak berkembang secara optimal, baik fisik, sosial, emosional, intelektual maupun spiritual. Assesmen yang tepat juga dapat digunakan untuk mendeteksi keterlambatan-keterlambatan perkembangan atau kebutuhan-kebutuhan khusus yang mungkin dimiliki anak-anak. Selain itu informasi yang akurat dari sebuah assesmen bermanfaat untuk peningkatan pembelajaran sehingga proses belajar anak-anak membaik dan sebagai informasi bagi para orangtua tentang kemajuan dan hal-hal terkait dengan belajar anak-anak mereka.

Assesmen yang tepat merupakan bagian penting dari program evaluasi dan perbaikan terus menerus kualitas program pendidikan yang sudah dirancang. Dalam program pendidikan yang berkualitas, pihak-

pihak terkait dengan pendidikan anak menggunakan informasi dari berbagai macam sumber untuk merencanakan dan membuat keputusan-keputusan tentang anak-anak secara individual.

Prosedur-prosedur dan instrumen-instrumen (alat) asesmen—seperti test, observasi, portofolio, penilaian guru, penilaian orangtua, dan lain sebagainya—dikatakan efektif ketika mereka memenuhi standar validitas (tepat dan akurat), reliabilitas (keajegan), dan kepekaan terhadap isu-isu kultural. Instrumen asesmen yang tepat memungkinkan jawaban-jawaban yang khas dari anak-anak menurut kelompok usia, jenis kelamin, latar belakang budaya, dan kondisi geografis. Anak-anak harus diukur secara individual oleh orang-orang yang mengetahui mereka dengan objektif dalam setting dan situasi-situasi yang mencerminkan penampilan mereka yang sesungguhnya. Semakin muda usia anak, maka akan semakin sulit untuk mendapatkan asesmen yang valid. Perkembangan anak-anak usia dini berlangsung sangat cepat dan sangat dipengaruhi oleh pengalaman. Performansi mereka dalam tes dipengaruhi oleh kondisi-kondisi emosional anak dan kondisi-kondisi saat asesmen dilakukan.

Satu hal yang pasti dalam melakukan asesmen adalah jangan pernah menggunakan satu instrumen asesmen untuk membuat keputusan-keputusan yang memiliki konsekuensi penting bagi anak-anak. Setiap asesmen seharusnya dipilih untuk memenuhi tujuan-tujuan yang spesifik. Asesmen seharusnya diselaraskan dengan kurikulum dan proses pembelajaran di kelas.

#### **D. Asesmen Dalam SKM dan Hubungannya Dengan Pembentukan Kompetensi**

Menurut Standar Proses Pendidikan, penilaian dilaksanakan dalam :

- (a) Bentuk tugas-tugas dan asesmen otentik lainnya (penilaian proses), ujian tengah semester (midsemester), ujian akhir semester,
- (b) Penilaian menggunakan acuan kriteria/patokan (PAP) dengan kategori A, B, C, dan D (dalam skala 4), (c) lulus minimum mencapai nilai C, dan (d) syarat lulus dari sekolah dengan IP minimum 2,0. Alternatif PAP adalah sebagai berikut :

Tingkat Penguasaan (%)	Nilai	Kategori
90 – 100	4	A
75 - 89	3	B
55 - 74	2	C
≤ 54	1	D

Selanjutnya, Permen No. 20 tahun 2007 tentang Standar Penilaian Pendidikan menyebutkan bahwa hasil belajar peserta didik pada jenjang pendidikan dasar dan menengah didasarkan pada prinsip-prinsip sebagai berikut :

- a. sah, berarti penilaian didasarkan pada data yang mencerminkan kemampuan yang diukur.
- b. objektif, berarti penilaian didasarkan pada prosedur dan kriteria yang jelas, tidak dipengaruhi subjektivitas penilai.
- c. adil, berarti penilaian tidak menguntungkan atau merugikan peserta didik karena berkebutuhan khusus serta perbedaan latar belakang agama, suku, budaya, adat istiadat, status sosial ekonomi, dan gender.
- d. terpadu, berarti penilaian oleh pendidik merupakan salah satu komponen yang tak terpisahkan dari kegiatan pembelajaran.
- e. terbuka, berarti prosedur penilaian, kriteria penilaian, dan dasar pengambilan keputusan dapat diketahui oleh pihak yang berkepentingan.
- f. menyeluruh dan berkesinambungan, berarti penilaian oleh pendidik mencakup semua aspek kompetensi dengan menggunakan berbagai teknik penilaian yang sesuai, untuk memantau perkembangan kemampuan peserta didik.
- g. sistematis, berarti penilaian dilakukan secara berencana dan bertahap dengan mengikuti langkah-langkah baku.
- h. beracuan kriteria, berarti penilaian didasarkan pada ukuran pencapaian kompetensi yang ditetapkan.
- i. akuntabel, berarti penilaian dapat dipertanggungjawabkan, baik dari segi teknik, prosedur, maupun hasilnya.

Dari poin a kutipan Standar Proses di atas, jelas bahwa dalam proses pembelajaran, asesmen dilakukan dengan menggunakan tugas-tugas dan asesmen otentik lainnya. Penggunaan tes-tes objektif dalam

proses dianggap tidak mampu menjawab tantangan pembelajaran yang mensyaratkan pembentukan kompetensi. Jenis tes objektif seperti tes pilihan ganda bila dikonstruksi secara baik, hanya mampu mengukur kemampuan kognitif siswa, sementara untuk pencapaian kompetensi, siswa harus membangun kemampuan kognitif, psikomotor, dan afektif yang terwujudkan dalam suatu unjuk kerja. Oleh karena itu, penggunaan asesmen otentik merupakan keharusan.

## Rangkuman

- Assesmen berperan mengelolah lembaga maupun mengelolah pembelajaran.
- SCL adalah refleksi dari ciri kehidupan global yang penuh dengan kompetisi dalam perubahan yang sangat cepat. SCL juga dikenal sebagai pembelajaran yang berpusat pada peserta didik (*Student-Centered Learning*, disingkat SCL) lahir pada awal abad ke-20.
- Assesmen merupakan sebuah proses pengumpulan informasi yang terus menerus berlangsung untuk mengukur performansi murid dan proses pembelajaran.

Menurut Standar Proses Pendidikan, penilaian dilaksanakan dalam:

1. Bentuk tugas-tugas dan asesmen otentik lainnya (penilaian proses), ujian tengah semester (midsemester), ujian akhir semester,
2. Penilaian menggunakan acuan kriteria/patokan (PAP) dengan kategori A, B, C, dan D (dalam skala 4), (c) lulus minimum mencapai nilai C, dan (d) syarat lulus dari sekolah dengan IP minimum 2,0. Alternatif PAP adalah sebagai berikut.

Tingkat Penguasaan (%)	Nilai	Kategori
90 – 100	4	A
75 - 89	3	B
55 - 74	2	C
≤ 54	1	D

## **BAB II**

### **ASSESMEN PEMBELAJARAN**

#### **A. Pengertian Assesmen**

Istilah evaluasi dan asesmen seringkali dipertukarkan, namun sebenarnya terdapat perbedaan yang esensial diantara keduanya. Assesmen dalam hal ini dinyatakan sebagai suatu cara yang tepat untuk mengungkap proses dan kemajuan belajar. Assesmen dapat memberikan umpan balik secara berkesinambungan tentang siswa untuk perbaikan pembelajaran. Sementara itu evaluasi dinyatakan sebagai pemberian nilai (*judgement*) terhadap hasil belajar berdasarkan data yang diperoleh melalui asesmen (Kumano,2001; Mehrens & Lehman,1989). Selain dari itu, terdapat pula beberapa istilah lainnya yaitu tes, testing, dan pengukuran yang juga seringkali dipertukarkan oleh guru.

Assesmen dalam pembelajaran adalah suatu proses atau upaya formal pengumpulan informasi yang berkaitan dengan variabel-variabel penting pembelajaran sebagai bahan dalam pengambilan keputusan oleh guru un-tuk memperbaiki proses dan hasil belajar siswa (Herman et al., 1992:95; Popham, 1995:3). Menurut beberapa pandangan ahli menyatakan bahwa assesmen merupakan suatu penilaian yang komprehensif dan melibatkan anggota tim untuk mengetahui kelemahan dan kekuatan yang mana hsil keputusannya dapat digunakan untuk layanan pendidikan yang dibutuhkan anak sebagai dasar untuk menyusun suatu rancangan pembelajaran Smith (2002); Proses sistematika dalam mengumpulkan data seseorang anak yang berfungsi untuk melihat kemampuan dan kesulitan yang dihadapi seseorang saat itu, sebagai bahan untuk menentukan apa yang sesungguhnya dibutuhkan. Berdasarkan informasi tersebut guru akan dapat menyusun program pembelajaran yang bersifat realitas sesuai dengan kenyataan objektif (James A. Mc. Lounghlin & Rena B. Lewis); Mengidentifikasi masalah dan menyeleksi target intervensi, memilih dan mendesain program treatmen, mengukur dampak treatmen yang diberikan secara terus menerus, serta mengevaluasi hasil-hasil umum dan ketepatan dari terapi (Bomstein dan Kazdin, 1985); Proses pengumpulan informasi untuk mendapatkan profil psikologis anak yang meliputi gejala dan intensitasnya, kendala-kendala yang dialami kelebihan dan kelemahannya, serta peran penting yang dibutuhkan anak (Lidz, 2003).

Pengertian assesmen dalam berbagai literatur asing tersebut di atas selaras dengan makna assesmen/penilaian yang digariskan dalam Buku Pedoman Penilaian pada kurikulum pendidikan dasar. Dalam buku tersebut tertulis bahwa, penilaian adalah suatu kegiatan yang dilakukan oleh guru untuk memberikan berbagai informasi secara berkesinambungan dan menyeluruh tentang proses dan hasil belajar yang telah dicapai (Depdikbud, 1994:3).

## **B. Tujuan Pelaksanaan Assesmen**

Tujuan utama penggunaan assesmen dalam pembelajaran (*classroom assessment*) adalah membantu guru dan siswa dalam mengambil keputusan profesional untuk memperbaiki pembelajaran. Menurut Popham (1995:4-13) assesmen bertujuan untuk antara lain untuk:

- (1) mendiagnosa kelebihan dan kelemahan siswa dalam belajar,
- (2) memonitor kemajuan siswa,
- (3) menentukan jenjang kemampuan siswa,
- (4) menentukan efektivitas pembelajaran,
- (5) mempengaruhi persepsi publik tentang efektivitas pembelajaran,
- (6) mengevaluasi kinerja guru kelas,
- (7) mengklarifikasi tujuan pembelajaran yang dirancang guru.

Beberapa ahli menyatakan bahwa tujuan pelaksanaan assesmen adalah: adalah untuk melihat kondisi anak saat tersebut dalam rangka menyusun suatu program pembelajaran yang tepat sehingga dapat melakukan layanan pembelajaran secara tepat (Lidz, 2003); Menurut Robb (2002): (i) Untuk menyaring dan mengidentifikasi anak, (ii) Untuk membuat keputusan tentang penempatan anak, (iii) Untuk merancang individualisasi pendidikan, (iv) Untuk memonitor kemajuan anak secara individu, dan (v) Untuk mengevaluasi keefektifan program; Menurut Sumardi dan Sunaryo (2006): (i) Memperoleh data yang relevan, objektif, akurat dan komprehensif tentang kondisi anak saat ini, (ii) Mengetahui profil anak secara utuh terutama permasalahan dan hambatan belajar yang dihadapi, potensi yang dimiliki, kebutuhan-kebutuhan khususnya, serta daya dukung lingkungan yang dibutuhkan anak, (iii) Menentukan layanan yang dibutuhkan dalam rangka memenuhi kebutuhan-kebutuhan khususnya dan memonitor kemampuannya (Sumardi & Sunaryo, 2006).

Menurut Salvia dan Yesseldyke seperti dikutip Lerner (1988: 54)

Assesmen dilakukan untuk lima keperluan yaitu :

- Penyaringan (*screening*)
- Pengalih-tanganan (*referral*)
- Klasifikasi (*classification*)
- Perencanaan Pembelajaran (*instructional planning*)
- Pemantauan kemajuan belajar anak (*monitoring pupil progress*)

Berdasarkan hasil kajian dari teori-teori diatas penulis dapat menyimpulkan bahwa: Asesmen dilakukan untuk mengetahui keadaan anak pada saat tertentu baik potensi-potensinya maupun kelemahan-kelemahan yang dimiliki anak sebagai bahan untuk menyusun suatu program pembelajaran sehingga dapat melakukan layanan / intervensi secara tepat.

Setiap penggunaan assesmen alternatif bentuk apapun dicirikan oleh hal-hal berikut:

- (1) menuntut siswa untuk merancang, membuat, menghasilkan, menunjukkan atau melakukan sesuatu;
- (2) memberi peluang untuk terjadinya berpikir kompleks dan/atau memecahkan masalah;
- (3) menggunakan kegiatan-kegiatan yang bermakna secara instruksional;
- (4) menuntut penerapan yang autentik pada dunia nyata;
- (5) penekoran lebih didasarkan pada pertimbangan manusia yang terlatih daripada mengandalkan mesin.

Untuk memperoleh asesmen dengan standar tinggi, maka penggunaan asesmen harus relevan dengan standar atau kebutuhan hasil belajar siswa; adil bagi semua siswa; akurat dalam pengukuran; berguna; layak dan dapat dipercaya. (Herman,1997:198).

Ada beberapa tujuan penilaian dilakukan guru, antara lain untuk grading (membedakan kedudukan hasil kerja siswa dibandingkan dengan siswa lain dalam satu kelas), alat seleksi (memisahkan antara siswa yang masuk dalam kategori tertentu dan yang tidak, atau untuk menentukan seorang siswa dapat masuk atau tidak di sekolah tertentu), menguasai kompetensi (menentukan apakah seorang siswa telah menguasai kompetensi tertentu atau belum), bimbingan (mengevaluasi hasil belajar siswa dalam rangka membantu siswa memahami dirinya, membuat keputusan yang harus dilakukan siswa, atau untuk menetapkan penjurusan), alat prediksi (mendapatkan informasi yang digunakan untuk memprediksi kinerja siswa pada pendidikan berikutnya) dan alat

diagnosis (melihat kesulitan belajar atau dalam hal apa siswa memiliki prestasi untuk menentukan perlu remediasi atau pengayaan). Dalam kaitannya dengan pelaksanaan penilaian berbasis kelas, jenis penilaian diagnosis, bimbingan, dan pencapaian penguasaan kompetensi harus menjadi perhatian utama guru pada setiap kali mengajar. Guru dituntut mampu melaksanakan penilaian mulai dari awal sampai akhir proses belajar mengajar. Untuk menilai sejauhmana siswa telah menguasai beragam kompetensi, tentu saja berbagai jenis penilaian perlu diberikan se-suai dengan kompetensi yang akan dinilai, seperti unjuk kerja/kinerja (performance), penugasan (proyek), hasil karya (produk), kumpulan hasil kerja siswa (portofolio), dan penilaian tertulis (paper and pencil test).

### **C. Manfaat Pelaksanaan Assesmen**

Pelaksanaan assesmen bermanfaat untuk :

1. Mengetahui tingkat pencapaian kompetensi selama dan setelah proses pembelajaran berlangsung.
2. Memberikan umpan balik bagi peserta didik agar mengetahui kekuatan dan kelemahannya dalam proses pencapaian kompetensi.
3. Memantau proses pembelajaran berlangsung.
4. Memberikan umpan balik bagi peserta didik agar mengetahui kekuatan dan kelemahannya dalam proses pencapaian kompetensi.
5. Memantau kemajuan dan mendiagnosis kesulitan belajar yang dialami peserta didik sehingga dapat dilakukan pengayaan dan remedial.
6. Sebagai umpan balik bagi guru dalam memperbaiki metode, pendekatan, kegiatan, dan sumber belajar yang digunakan.
7. Memberikan pilihan alternatif penilaian kepada guru.
8. Memberikan informasi kepada orang tua dan komite sekolah tentang efektivitas pendidikan.

### **D. Subjek, Objek/Sasaran**

Subjek assesmen pembelajaran adalah pelaku-pelaku kegiatan assesmen dalam bidang pendidikan. Assesmen ditentukan oleh aturan pembagian tugas, misalnya untuk melakukan assesmen terhadap prestasi belajar, maka yang menjadi subjek adalah pendidik/guru; untuk kegiatan assesmen terhadap kepribadian dimana instrumen yang digunakan berupa instrumen yang sudah distandarisasi maka yang menjadi subjek adalah para ahli psikologi.

### **E. Fungsi Assesmen**

Dalam pedoman penilaian Kurikulum 1994 (Depdikbud, 1994:3) ditegaskan bahwa tujuan dan fungsi penilaian untuk memberikan umpan balik baik kepada guru, siswa, orangtua maupun lembaga pendidikan yang berkepentingan serta untuk menentukan nilai hasil belajar siswa. Bagi guru, hasil penilaian tidak hanya digunakan untuk memberikan pertanggung-jawaban secara obyektif kepada atasan ataupun sekedar bahan nilai raport. Namun penilaian dapat digunakan sebagai bahan dasar untuk melakukan introspeksi diri terhadap proses pembelajaran yang baru saja berlangsung. Bagi siswa, hasil penilaian dapat dijadikan alat untuk memotivasi siswa tersebut agar lebih giat dalam proses pembelajaran berikutnya. Selain itu, dari hasil penilaian siswa mendapatkan informasi tentang seberapa jauh tingkat penguasaan bahan pelajaran yang diberikan guru.

Bagi orangtua, dengan mengetahui hasil belajar siswa (anaknya) orangtua dapat turut berpartisipasi dan mengambil langkah yang tepat dalam memberikan bimbingan dan bantuan serta dorongan bagi putra-putrinya. Selain itu dengan informasi hasil penilaian yang benar, orangtua dapat secara akurat mengetahui kemampuan, kekurangan dan kedudukan siswa secara riil di kelasnya. Bagi pengelola program pendidikan, hasil penilaian merupakan masukan yang sangat berarti yang dapat digunakan untuk bahan kajian dalam membantu guru meningkatkan kompetensi profesionalnya, khususnya dalam bidang penilaian. Hasil penilaian yang komprehensif dapat juga digunakan untuk tujuan dan kebutuhan lain misalnya penentuan status siswa, pengelompokkan, seleksi, diagnosis dan bimbingan, serta menyempurnakan pengalaman pendidik, atau penelitian.

### **F. Prinsip Assesmen**

Hasil kegiatan assesmen dapat memberikan manfaat yang optimal jika dilakukan dengan mengacu pada prinsip-prinsip penilaian sebagaimana ditetapkan oleh pedoman formal penilaian dari pemerintah (Depdikbud, 1994:5), yakni dilaksanakan secara menyeluruh, berkesinambungan, berorientasi pada tujuan, obyektif, terbuka serta mempertimbangkan aspek kebermaknaan. Penelitian yang dilakukan secara menyeluruh artinya informasi yang dikumpulkan melalui proses penilaian menyangkut seluruh aspek kepribadian siswa. Penilaian dikatakan menyeluruh jika mampu mengungkap aspek produk dan proses

belajar anak, yakni menyangkut pengetahuan, sikap, dan keterampilan proses peserta didik.

Target hasil belajar yang diharapkan terjadi pada diri siswa setelah berlangsungnya proses pembelajaran tertuang dalam tujuan pembelajaran sejak tujuan umum pada Standar Kompetensi Mata Pelajaran hingga Kom-petensi Dasar, Hasil Belajar, dan Indikator dari setiap materi pokok pembelajaran. Oleh karena proses penilaian bertujuan untuk mengetahui sejauhmana tingkat ketercapaian tujuan pembelajaran, maka dalam melakukan penilaian harus selalu berorientasi pada tujuan; karena antara tujuan dan penilaian merupakan komponen sistem pembelajaran yang tidak dapat dipisahkan.

Prinsip penilaian selanjutnya adalah bersifat obyektif, artinya dalam melakukan penilaian terhadap hasil belajar siswa, guru berusaha untuk meminimalisasi faktor subyektivitas. Menurut Ign. Masidjo (1995 : 25) obyektivitas pelaksanaan penilaian dapat dicapai dengan menaati aturan-aturan yang telah ditetapkan. Penilaian yang didasarkan atas kriteria penilaian yang telah ditetapkan sebelumnya dapat mengurangi faktor subyektivitas dalam melakukan penilaian. Agar hasil penilaian dapat memberikan manfaat baik kepada guru, siswa, orang tua maupun pihak sekolah, maka penilaian hendaknya dilaku-kan secara terbuka. Maksudnya baik proses maupun hasil penilaian hendaknya diinformasikan kepada pihak-pihak terkait, sehingga hasil penilaian memiliki kebermaknaan bagi pihak-pihak yang memerlukan.

## **Rangkuman**

- Asesmen dalam pembelajaran adalah suatu proses atau upaya formal pengumpulan informasi yang berkaitan dengan variabel-variabel penting pembelajaran sebagai bahan dalam pengambilan keputusan oleh guru untuk memperbaiki proses dan hasil belajar siswa.
- Pelaksanaan assesmen bermanfaat untuk:
  1. Mengetahui tingkat pencapaian kompetensi selama dan setelah proses pembelajaran berlangsung.
  2. Memberikan umpan balik bagi peserta didik agar mengetahui kekuatan dan kelemahannya dalam proses pencapaian kompetensi.
  3. Memantau proses pembelajaran berlangsung.
  4. Memberikan umpan balik bagi peserta didik agar mengetahui kekuatan dan kelemahannya dalam proses pencapaian kompetensi.
  5. Memantau kemajuan dan mendiagnosis kesulitan belajar yang dialami peserta didik sehingga dapat dilakukan pengayaan dan remedial.
  6. Sebagai umpan balik bagi guru dalam memperbaiki metode, pendekatan, kegiatan, dan sumber belajar yang digunakan.
  7. Memberikan pilihan alternatif penilaian kepada guru.
  8. Memberikan informasi kepada orang tua dan komite sekolah tentang efektivitas pendidikan.

## **BAB III**

### **PERBEDAAN PENGUKURAN, PENILAIAN DAN EVALUASI**

#### **A. Pendahuluan**

Dalam kehidupan sehari-hari tanpa disadari sebenarnya kita sering membuat suatu kegiatan evaluasi dan selalu menggunakan prinsip mengukur dan menilai. Namun, banyak orang belum memahami secara tepat arti kata evaluasi, pengukuran, dan penilaian bahkan masih banyak orang yang lebih cenderung mengartikan ketiga kata tersebut dengan suatu pengertian yang sama.

Secara umum orang hanya mengidentikkan kegiatan evaluasi sama dengan menilai, karena aktifitas mengukur biasanya sudah termasuk didalamnya. Pengukuran, penilaian dan evaluasi merupakan kegiatan yang bersifat hierarki. Artinya ketiga kegiatan tersebut tidak dapat dipisahkan satu sama lain dan dalam pelaksanaannya harus dilaksanakan secara berurutan.

#### **B. Pengukuran**

Pengukuran dapat diartikan dengan kegiatan untuk mengukur sesuatu. Pada hakekatnya, kegiatan ini adalah membandingkan sesuatu dengan atau sesuatu yang lain (Anas Sudijono, 1996: 3) Jika kita mengukur suhu badan seseorang dengan termometer, atau mengukur jarak kota A dengan kota B, maka sesungguhnya yang sedang dilakukan adalah mengkuantifikasi keadaan seseorang atau tempat kedalam angka. Karenanya, dapat dipahami bahwa pengukuran itu bersifat kuantitatif.

Maksud dilaksanakan pengukuran sebagaimana dikemukakan Anas Sudijono (1996: 4) ada tiga macam yaitu : (1) pengukuran yang dilakukan bukan untuk menguji sesuatu seperti orang mengukur jarak dua buah kota, (2) pengukuran untuk menguji sesuatu seperti menguji daya tahan lampu pijar serta (3) pengukuran yang dilakukan untuk menilai. Pengukuran ini dilakukan dengan jalan menguji hal yang ingin dinilai seperti kemajuan belajar dan lain sebagainya.

Dalam dunia pendidikan, yang dimaksud pengukuran sebagaimana disampaikan Cangelosi (1995: 21) adalah proses pengumpulan data melalui pengamatan empiris. Proses pengumpulan ini dilakukan untuk menaksir apa yang telah diperoleh siswa setelah mengikuti pelajaran selama waktu tertentu. Proses ini dapat dilakukan dengan mengamati kinerja mereka, mendengarkan apa yang mereka katakan serta

mengumpulkan informasi yang sesuai dengan tujuan melalui apa yang telah dilakukan siswa.

Menurut Mardapi (2004: 14) pengukuran pada dasarnya adalah kegiatan penentuan angka terhadap suatu obyek secara sistematis. Karakteristik yang terdapat dalam obyek yang diukur ditransfer menjadi bentuk angka sehingga lebih mudah untuk dinilai. aspek-aspek yang terdapat dalam diri manusia seperti kognitif, afektif dan psikomotor dirubah menjadi angka. Karenanya, kesalahan dalam mengangkakan aspek-aspek ini harus sekecil mungkin. Kesalahan yang mungkin muncul dalam melakukan pengukuran khususnya dibidang ilmu-ilmu sosial dapat berasal dari alat ukur, cara mengukur dan obyek yang diukur.

Pengukuran dalam bidang pendidikan erat kaitannya dengan tes. Hal ini dikarenakan salah satu cara yang sering dipakai untuk mengukur hasil yang telah dicapai siswa adalah dengan tes. Selain dengan tes, terkadang juga dipergunakan nontes. Jika tes dapat memberikan informasi tentang karakteristik kognitif dan psikomotor, maka nontes dapat memberikan informasi tentang karakteristik afektif obyek.

### **C. Penilaian**

Penilaian merupakan bagian penting dan tak terpisahkan dalam sistem pendidikan saat ini. Peningkatan kualitas pendidikan dapat dilihat dari nilai-nilai yang diperoleh siswa. Tentu saja untuk itu diperlukan sistem penilaian yang baik dan tidak bias. Sistem penilaian yang baik akan mampu memberikan gambaran tentang kualitas pembelajaran sehingga pada gilirannya akan mampu membantu guru merencanakan strategi pembelajaran. Bagi siswa sendiri, sistem penilaian yang baik akan mampu memberikan motivasi untuk selalu meningkatkan kemampuannya.

Dalam sistem evaluasi hasil belajar, penilaian merupakan langkah lanjutan setelah dilakukan pengukuran. informasi yang diperoleh dari hasil pengukuran selanjutnya dideskripsikan dan ditafsirkan. Karenanya, menurut Djemari Mardapi (1999: 8) penilaian adalah kegiatan menafsirkan atau mendeskripsikan hasil pengukuran. Menurut Cangelosi (1995: 21) penilaian adalah keputusan tentang nilai. Oleh karena itu, langkah selanjutnya setelah melaksanakan pengukuran adalah penilaian. Penilaian dilakukan setelah siswa menjawab soal-soal yang terdapat pada tes. Hasil jawaban siswa tersebut ditafsirkan dalam bentuk nilai.

Menurut Djemari Mardapi (2004: 18) ada dua acuan yang dapat dipergunakan dalam melakukan penilaian yaitu acuan norma dan acuan kriteria. Dalam melakukan penilaian dibidang pendidikan, kedua acuan ini dapat dipergunakan. Acuan norma berasumsi bahwa kemampuan seseorang berbeda serta dapat digambarkan menurut kurva distribusi normal. Sedangkan acuan kriteria berasumsi bahwa apapun bisa dipelajari semua orang namun waktunya bisa berbeda.

Penggunaan acuan norma dilakukan untuk menyeleksi dan mengetahui dimana posisi seseorang terhadap kelompoknya. Misalnya jika seseorang mengikuti tes tertentu, maka hasil tes akan memberikan gambaran dimana posisinya jika dibandingkan dengan orang lain yang mengikuti tes tersebut. Adapun acuan kriteria dipergunakan untuk menentukan kelulusan seseorang dengan membandingkan hasil yang dicapai dengan kriteria yang telah ditetapkan terlebih dahulu. Acuan ini biasanya digunakan untuk menentukan kelulusan seseorang. Seseorang yang dikatakan telah lulus berarti bisa melakukan apa yang terdapat dalam kriteria yang telah ditetapkan dan sebaliknya. Acuan kriteria, ini biasanya dipergunakan untuk ujian-ujian praktek.

Dengan adanya acuan norma atau kriteria, hasil yang sama yang didapat dari pengukuran ataupun penilaian akan dapat diinterpretasikan berbeda sesuai dengan acuan yang digunakan. Misalnya, kecepatan kendaraan 40 km/jam akan memiliki interpretasi yang berbeda apabila kendaraan tersebut adalah sepeda dan mobil.

#### **D. Evaluasi**

Pengukuran, penilaian dan evaluasi merupakan kegiatan yang bersifat hierarki. Artinya ketiga kegiatan tersebut dalam kaitannya dengan proses belajar mengajar tidak dapat dipisahkan satu sama lain dan dalam pelaksanaannya harus dilaksanakan secara berurutan.

Evaluasi Menurut Suharsimi Arikunto (2004: 1) adalah kegiatan untuk mengumpulkan informasi tentang bekerjanya sesuatu, yang selanjutnya informasi tersebut digunakan untuk menentukan alternatif yang tepat dalam mengambil keputusan. Dalam bidang pendidikan, evaluasi sebagaimana dikatakan Gronlund (1990: 5) merupakan proses yang sistematis tentang mengumpulkan, menganalisis dan menafsirkan informasi untuk menentukan sejauhmana tujuan pembelajaran telah dicapai oleh siswa. Menurut Djemari Mardapi (2004: 19) evaluasi adalah

proses mengumpulkan informasi untuk mengetahui pencapaian belajar kelas atau kelompok.

Dari pendapat di atas, ada beberapa hal yang menjadi ciri khas dari evaluasi yaitu: (1) sebagai kegiatan yang sistematis, pelaksanaan evaluasi haruslah dilakukan secara berkesinambungan. Sebuah program pembelajaran seharusnya dievaluasi di setiap akhir program tersebut, (2) dalam pelaksanaan evaluasi dibutuhkan data dan informasi yang akurat untuk menunjang keputusan yang akan diambil. Asumsi-asumsi ataupun prasangka bukan merupakan landasan untuk mengambil keputusan dalam evaluasi, dan (3) kegiatan evaluasi dalam pendidikan tidak pernah terlepas dari tujuan-tujuan pembelajaran yang telah ditetapkan sebelumnya. Karena itulah pendekatan goal oriented merupakan pendekatan yang paling sesuai untuk evaluasi pembelajaran.

#### **E. Pengertian Pengukuran, Penilaian, dan Evaluasi**

Untuk memahami pengertian evaluasi, pengukuran dan penilaian kita dapat memahaminya lewat contoh berikut :

1. Apabila ada seseorang yang memberikan kepada kita 2 pensil yang berbeda ukuran, yang satu panjang dan yang satu lebih pendek dan kita diminta untuk memilihnya, maka otomatis kita akan cenderung memilih pensil yang panjang karena akan bisa lebih lama digunakan. Kecuali memang ada kriteria lain sehingga kita memilih sebaliknya.
2. Peristiwa menjual dan membeli di pasar. Kadang kala sebelum kita membeli durian di pasar, sering kali kita membandingkan terlebih dahulu durian yang ada sebelum membelinya. Biasanya kita akan mencium, melihat bentuknya, jenisnya ataupun tampak tangkai yang ada pada durian tersebut untuk mengetahui durian manakah yang baik dan layak dibeli.

Dari kedua contoh di atas maka dapat kita simpulkan bahwa kita selalu melakukan penilaian sebelum menentukan pilihan untuk memilih suatu objek/benda. Pada contoh pertama kita akan memilih pensil yang lebih panjang dari pada pensil yang pendek karena pensil yang lebih panjang dapat kita gunakan lebih lama. Sedangkan pada contoh yang kedua kita akan menentukan durian mana yang akan kita beli berdasarkan bau, bentuk, jenis, ataupun tampak tangkai dari durian yang dijual tersebut. Sehingga kita dapat memperkirakan mana durian yang manis.

Untuk mengadakan penilaian, kita harus melakukan pengukuran terlebih dahulu. Dalam contoh 1 di atas, jika kita mempunyai pengaris,

maka untuk menentukan pensil mana yang lebih panjang maka kita akan mengukur kedua pensil tersebut dengan menggunakan penggaris kemudian kita akan melakukan penilaian dengan membandingkan ukuran panjang dari masing-masing penggaris sehingga pada akhirnya kita dapat mengatakan bahwa “Yang ini panjang” dan “Yang ini pendek” lalu yang panjanglah yang kita ambil.

Dalam contoh yang ke 2, kita memilih durian yang terbaik lewat bau, tampak tangkai, maupun jenisnya. Hal itu juga diawali dengan proses pengukuran dimana kita membanding-bandingkan beberapa durian yang ada sekalipun tidak menggunakan alat ukur yang paten tetapi berdasarkan pengalaman. Barulah kita melakukan penilaian mana durian yang terbaik berdasarkan ukuran yang kita tetapkan yang akan dibeli.

Dari hal ini kita dapat mengetahui bahwa dalam proses penilaian kita menggunakan 3 ukuran, yakni ukuran baku (meter, kilogram, takaran, dan sebagainya), ukuran tidak baku (depa, jengkal, langkah, dan sebagainya) dan ukuran perkiraan yakni berdasarkan pengalaman. Langkah – langkah mengukur kemudian menilai sesuatu sebelum kita mengambilnya itulah yang dinamakan mengadakan evaluasi yakni mengukur dan menilai. Kita tidak dapat mengadakan evaluasi sebelum melakukan aktivitas mengukur dan menilai.

Berdasarkan contoh diatas dapat kita simpulkan pengertian pengukuran, penilaian, dan evaluasi sebagai berikut :

- Pengukuran adalah kegiatan membandingkan sesuatu dengan ukuran tertentu dan bersifat kuantitatif.
- Penilaian adalah kegiatan mengambil keputusan untuk menentukan sesuatu berdasarkan kriteria baik buruk dan bersifat kualitatif. Sedangkan
- Evaluasi adalah kegiatan yang meliputi pengukuran dan penilaian.

#### **F. Evaluasi dalam Pendidikan**

Secara harafiah evaluasi berasal dari bahasa Inggris *evaluation* yang berarti penilaian atau penaksiran (John M. Echols dan Hasan Shadily: 1983). Menurut Stufflebeam, dkk (1971) mendefinisikan evaluasi sebagai “*The process of delineating, obtaining, and providing useful information for judging decision alternatives*”. Artinya evaluasi merupakan proses menggambarkan, memperoleh, dan menyajikan informasi yang berguna untuk merumuskan suatu alternatif keputusan.

Evaluasi menurut Kumano (2001) merupakan penilaian terhadap data yang dikumpulkan melalui kegiatan asesmen. Sementara itu menurut Calongesi (1995) evaluasi adalah suatu keputusan tentang nilai berdasarkan hasil pengukuran. Sejalan dengan pengertian tersebut, Zainul dan Nasution (2001) menyatakan bahwa evaluasi dapat dinyatakan sebagai suatu proses pengambilan keputusan dengan menggunakan informasi yang diperoleh melalui pengukuran hasil belajar, baik yang menggunakan instrumen tes maupun non tes.

Secara garis besar dapat dikatakan bahwa evaluasi adalah pemberian nilai terhadap kualitas sesuatu. Selain dari itu, evaluasi juga dapat dipandang sebagai proses merencanakan, memperoleh, dan menyediakan informasi yang sangat diperlukan untuk membuat alternatif-alternatif keputusan. Dengan demikian, Evaluasi merupakan suatu proses yang sistematis untuk menentukan atau membuat keputusan sampai sejauhmana tujuan-tujuan pengajaran telah dicapai oleh siswa (Purwanto, 2002). Arikunto (2003) mengungkapkan bahwa evaluasi adalah serangkaian kegiatan yang ditujukan untuk mengukur keberhasilan program pendidikan. Tayibnaxis (2000) dalam hal ini lebih meninjau pengertian evaluasi program dalam konteks tujuan yaitu sebagai proses menilai sampai sejauhmana tujuan pendidikan dapat dicapai.

Berdasarkan tujuannya, terdapat pengertian evaluasi sumatif dan evaluasi formatif. Evaluasi formatif dinyatakan sebagai upaya untuk memperoleh feedback perbaikan program, sementara itu evaluasi sumatif merupakan upaya menilai manfaat program dan mengambil keputusan (Lehman, 1990).

#### **G. Penilaian Dalam Pendidikan**

Penilaian (assessment) adalah penerapan berbagai cara dan penggunaan beragam alat penilaian untuk memperoleh informasi tentang sejauh mana hasil belajar peserta didik atau ketercapaian kompetensi (rangkaiian kemampuan) peserta didik. Penilaian menjawab pertanyaan tentang sebaik apa hasil atau prestasi belajar seorang peserta didik. Hasil penilaian dapat berupa nilai kualitatif (pernyataan naratif dalam kata-kata) dan nilai kuantitatif (berupa angka). Pengukuran berhubungan dengan proses pencarian atau penentuan nilai kuantitatif tersebut.

Penilaian hasil belajar pada dasarnya adalah mempermasalahkan, bagaimana pengajar (guru) dapat mengetahui hasil pembelajaran yang telah dilakukan. Pengajar harus mengetahui sejauh mana pebelajar

(*learner*) telah mengerti bahan yang telah diajarkan atau sejauh mana tujuan/kompetensi dari kegiatan pembelajaran yang dikelola dapat dicapai. Tingkat pencapaian kompetensi atau tujuan instruksional dari kegiatan pembelajaran yang telah dilaksanakan itu dapat dinyatakan dengan nilai.

## H. Pengukuran dalam Pendidikan

Pengukuran adalah penentuan besaran, dimensi, atau kapasitas, biasanya terhadap suatu standar atau satuan pengukuran. Pengukuran tidak hanya terbatas pada kuantitas fisik, tetapi juga dapat diperluas untuk mengukur hampir semua benda yang bisa dibayangkan, seperti tingkat ketidakpastian, atau kepercayaan konsumen.

Pengukuran adalah proses pemberian angka-angka atau label kepada unit analisis untuk merepresentasikan atribut-atribut konsep. Proses ini seharusnya cukup dimengerti orang walau misalnya definisinya tidak dimengerti. Hal ini karena antara lain kita sering kali melakukan pengukuran.

Menurut Cangelosi (1995) yang dimaksud dengan pengukuran (*Measurement*) adalah suatu proses pengumpulan data melalui pengamatan empiris untuk mengumpulkan informasi yang relevan dengan tujuan yang telah ditentukan. Dalam hal ini guru menaksir prestasi siswa dengan membaca atau mengamati apa saja yang dilakukan siswa, mengamati kinerja mereka, mendengar apa yang mereka katakan, dan menggunakan indera mereka seperti melihat, mendengar, menyentuh, mencium, dan merasakan. Menurut Zainul dan Nasution (2001) pengukuran memiliki dua karakteristik utama yaitu: 1) penggunaan angka atau skala tertentu; 2) menurut suatu aturan atau formula tertentu.

*Measurement* (pengukuran) merupakan proses yang mendeskripsikan performance siswa dengan menggunakan suatu skala kuantitatif (*system angka*) sedemikian rupa sehingga sifat kualitatif dari performance siswa tersebut dinyatakan dengan angka-angka (Alwasilah et al.1996). Pernyataan tersebut diperkuat dengan pendapat yang menyatakan bahwa pengukuran merupakan pemberian angka terhadap suatu atribut atau karakter tertentu yang dimiliki oleh seseorang, atau suatu obyek tertentu yang mengacu pada aturan dan formulasi yang jelas. Aturan atau formulasi tersebut harus disepakati secara umum oleh para ahli (Zainul & Nasution, 2001). Dengan demikian, pengukuran dalam bidang pendidikan berarti mengukur atribut atau karakteristik peserta

didik tertentu. Dalam hal ini yang diukur bukan peserta didik tersebut, akan tetapi karakteristik atau atributnya. Senada dengan pendapat tersebut, Secara lebih ringkas, Arikunto dan Jabar (2004) menyatakan pengertian pengukuran (*measurement*) sebagai kegiatan membandingkan suatu hal dengan satuan ukuran tertentu sehingga sifatnya menjadi kuantitatif.

### **I. Perbedaan Evaluasi, Penilaian dan Pengukuran**

Berdasarkan pengertian di atas dapat kita simpulkan bahwa penilaian adalah suatu proses untuk mengambil keputusan dengan menggunakan informasi yang diperoleh melalui pengukuran hasil belajar baik yang menggunakan tes maupun nontes. Pengukuran adalah membandingkan hasil tes dengan standar yang ditetapkan. Pengukuran bersifat kuantitatif. Sedangkan menilai adalah kegiatan mengukur dan mengadakan estimasi terhadap hasil pengukuran atau membandingkan dan tidak sampai ke taraf pengambilan keputusan. Penilaian bersifat kualitatif.

Agar lebih jelas perbedaannya maka perlu dispesifikasi lagi untuk pengertian masing-masing :

- Evaluasi pembelajaran adalah suatu proses atau kegiatan untuk menentukan nilai, kriteria-*judgment* atau tindakan dalam pembelajaran.
- Penilaian dalam pembelajaran adalah suatu usaha untuk mendapatkan berbagai informasi secara berkala, berkesinambungan, dan menyeluruh tentang proses dan hasil dari pertumbuhan dan perkembangan yang telah dicapai oleh anak didik melalui program kegiatan belajar.
- Pengukuran atau *measurement* merupakan suatu proses atau kegiatan untuk menentukan kuantitas sesuatu yang bersifat numerik. Pengukuran lebih bersifat kuantitatif, bahkan merupakan instrumen untuk melakukan penilaian. Dalam dunia pendidikan, yang dimaksud pengukuran sebagaimana disampaikan Cangelosi (1995: 21) adalah proses pengumpulan data melalui pengamatan empiris.

### **Rangkuman**

- Pengukuran, penilaian dan evaluasi merupakan kegiatan yang bersifat hierarki. Artinya ketiga kegiatan tersebut dalam kaitannya dengan proses belajar mengajar tidak dapat dipisahkan satu sama lain dan dalam pelaksanaannya harus dilaksanakan secara berurutan.
- Evaluasi pembelajaran adalah suatu proses atau kegiatan untuk menentukan nilai, kriteria-*judgment* atau tindakan dalam pembelajaran.
- Penilaian dalam pembelajaran adalah suatu usaha untuk mendapatkan berbagai informasi secara berkala, berkesinambungan, dan menyeluruh tentang proses dan hasil dari pertumbuhan dan perkembangan yang telah dicapai oleh anak didik melalui program kegiatan belajar.
- Pengukuran atau measurement merupakan suatu proses atau kegiatan untuk menentukan kuantitas sesuatu yang bersifat numerik. Pengukuran lebih bersifat kuantitatif, bahkan merupakan instrumen untuk melakukan penilaian.

## **BAB IV**

### **PELAKSANAAN ASSESMEN DAN EVALUASI**

Terdapat urutan atau proses yang mendasari sebelum melakukan evaluasi, yakni :

1. Mengembangkan konsep dan mengadakan penelitian awal. Konsep perlu direncanakan secara matang sebelum diadakan eksekusi pesan dan perlu diadakan uji coba untuk mengecek kesesuaian antara draft yang dibuat dengan eksekusi pesannya.
2. Dengan uji coba yang dilakukan, pengevaluasi mencoba mencari tanggapan dari khalayak. Tanggapan dari khalayak ini penting untuk mengukur efektifitas pesan yang disampaikan.

Secara logis dan sistematis langkah-langkah pelaksanaan evaluasi pelatihan sebagai berikut :

#### ***Langkah 1 : Persiapan Evaluasi***

Pada langkah ini terdapat tiga kegiatan pokok yang berkaitan dengan pelaksanaan evaluasi yaitu: menentukan tujuan atau maksud evaluasi, merumuskan informasi yang akan dicari atau memfokuskan evaluasi dan menentukan cara pengumpulan data.

Beberapa kriteria yang digunakan dalam merumuskan tujuan evaluasi adalah :

- (1) kejelasan,
- (2) keterukuran,
- (3) kegunaan dan kemanfaatan,
- (4) relevansi dan kesesuaian atau compatibility.

Jadi tujuan evaluasi harus jelas, terukur, berguna, relevan dan sesuai dengan kebutuhan pengembangan program diklat. Dalam merumuskan informasi atau memfokuskan evaluasi harus berdasarkan kepada tujuan evaluasi.

Terdapat beberapa metode dalam merumuskan pertanyaan evaluasi yaitu :

- (1) menganalisis objek,
- (2) menggunakan kerangka teoritis,
- (3) memanfaatkan keahlian dan pengalaman dari luar,
- (4) berinteraksi dengan audien kunci.

Menentukan cara pengumpulan data, misalnya survei atau yang lain, ditentukan pula pendekatan dalam pengumpulan data. Terdapat beberapa prosedur pengumpulan data dengan pendekatan kuantitatif, misalnya observasi, tes, survei atau lainnya.

### ***Langkah 2 : Mengembangkan Instrumen***

Setelah metode pengumpulan data ditentukan, selanjutnya ditentukan pula bentuk instrumen yang akan digunakan serta kepada siapa instrumen tersebut ditujukan (respondennya). Terdapat syarat-syarat yang harus dipenuhi oleh instrumen evaluasi sebagai berikut :

- (1) validitas adalah keabsahan instrumen dalam mengukur apa yang seharusnya diukur,
- (2) reliabilitas adalah ketetapan hasil yang diperoleh, misalnya bila melakukan pengukuran dengan orang yang sama dalam waktu yang berlainan atau orang yang lain dalam waktu yang sama,
- (3) objektivitas adalah upaya penerjemahan hasil pengukuran dalam bilangan atau pemberian skor tidak terpengaruh oleh siapa yang melakukan,
- (4) standarisasi untuk memastikan evaluator mempunyai persepsi yang sama dalam mengukur karena adanya petunjuk khusus pengisian data,
- (5) relevansi adalah kepatuhan untuk mengembangkan berbagai pertanyaan agar sesuai dengan maksud instrumen,
- (6) mudah di gunakan

### ***Langkah 3 : Mengumpulkan Data***

Dalam melakukan pengumpulan data ini dilakukan dengan berbeda-beda pada masing-masing level. Pada level reaksi data yang dikumpulkan berupa data kuantitatif dengan menggunakan metode survei. Kemudian pada level pembelajaran data yang dikumpulkan berupa data kuantitatif dengan menggunakan tes. Selanjutnya pada level perilaku, data yang dikumpulkan melalui observasi atau dapat juga dengan rencana aksi (*action plan*), yaitu rencana tahapan tindakan yang akan dilakukan oleh peserta pelatihan dalam mengimplementasikan hasil pelatihan yang telah diikuti. Dalam hal ini para peserta harus mempunyai suatu sasaran peningkatan kinerja/kompetensi yang bersangkutan dalam unit kerja masing-masing yang kemudian diukur dengan menggunakan patokan kinerja/kompetensi yang bersangkutan. Kemudian yang terakhir,

yaitu pada level hasil atau dampak, pada data yang dikumpulkan dapat melalui atasan, peserta pelatihan, bawahan atau rekan kerja (client).

#### ***Langkah 4 : Mengolah dan Menganalisis Data***

Setelah data yang diperlukan sudah terkumpul, maka langkah berikutnya adalah mengolah dan menganalisis data. Dalam menganalisis data dan menafsirkannya (menginterpretasikan) harus berdasarkan hasil data yang telah berhasil didapatkan. Kemudian menyajikannya dalam bentuk yang mudah dipahami dan komunikatif.

#### ***Langkah 5 : Menyusun Laporan***

Melaporkan merupakan langkah terakhir kegiatan evaluasi pelatihan. Laporan disusun dengan format yang telah disepakati oleh tim. Langkah terakhir evaluasi ini erat kaitannya dengan tujuan diadakannya evaluasi. Langkah-langkah tersebut dapat dengan digunakan untuk menjawab sejauh mana evaluasi pelatihan yang akan dilakukan dan bagaimana pelaksanaan proses pelatihan dari awal hingga akhir sehingga memberikan hasil untuk improvisasi pada pelatihan-pelatihan selanjutnya.

#### **A. Proses Pelaksanaan**

Dalam mengadakan sebuah proses evaluasi, terdapat beberapa hal yang akan dibahas yaitu apa yang menjadi bahan evaluasi, bagaimana proses evaluasi, kapan evaluasi diadakan, mengapa perlu diadakan evaluasi, dimana proses evaluasi diadakan, dan pihak yang mengadakan evaluasi. Hal yang perlu dilakukan evaluasi tersebut adalah narasumber yang ada, efektifitas penyebaran pesan, pemilihan media yang tepat dan pengambilan keputusan anggaran dalam mengadakan sejumlah promosi dan periklanan. Evaluasi tersebut perlu diadakan dengan tujuan untuk menghindari kesalahan perhitungan pembiayaan, memilih strategi terbaik dari berbagai alternatif strategis yang ada, meningkatkan efisiensi iklan secara general, dan melihat apakah tujuan sudah tercapai. Di sisi lain, perusahaan kadang-kadang enggan untuk mengadakan evaluasi karena biayanya yang mahal, terdapat masalah dengan penelitian, ketidaksetujuan akan apa yang hendak dievaluasi, merasa telah mencapai tujuan, dan banyak membuang waktu.

Secara garis besar, proses evaluasi terbagi menjadi di awal (*pretest*) dan diakhir (*posttest*). *Pretest* merupakan sebuah evaluasi yang diadakan untuk menguji konsep dan eksekusi yang direncanakan. Sedangkan, *posttest* merupakan evaluasi yang diadakan untuk melihat tercapainya tujuan dan dijadikan sebagai masukan untuk analisis situasi berikutnya.

Evaluasi dapat dilakukan di dalam atau diluar ruangan. Evaluasi yang diadakan di dalam ruangan pada umumnya menggunakan metode penelitian laboratorium dan sampel akan dijadikan sebagai kelompok percobaan. Kelemahannya, realisme dari metode ini kurang dapat diterapkan. Sementara, evaluasi yang diadakan di luar ruangan akan menggunakan metode penelitian lapangan dimana kelompok percobaan tetap dibiarkan menikmati kebebasan dari lingkungan sekitar. Realisme dari metode ini lebih dapat diterapkan dalam kehidupan sehari-hari.

Untuk mencapai evaluasi tersebut dengan baik, diperlukan sejumlah tahapan yang harus dilalui yakni menentukan permasalahan secara jelas, mengembangkan pendekatan permasalahan, memformulasikan desain penelitian, melakukan penelitian lapangan untuk mengumpulkan data, menganalisis data yang diperoleh, dan kemampuan menyampaikan hasil penelitian.

Evaluasi merupakan kegiatan pengumpulan kenyataan mengenai proses pembelajaran secara sistematis untuk menetapkan apakah terjadi perubahan terhadap peserta didik dan sejauh apakah perubahan tersebut mempengaruhi kehidupan peserta didik. (dikutip dari Bloom et.al 1971). Stufflebeam et.al 1971 mengatakan bahwa evaluasi adalah proses menggambarkan, memperoleh dan menyajikan informasi yang berguna untuk menilai alternatif keputusan.

Evaluasi sendiri memiliki beberapa prinsip dasar yaitu :

1. Evaluasi bertujuan membantu pemerintah dalam mencapai tujuan pembelajaran bagi masyarakat.
2. Evaluasi adalah seni, tidak ada evaluasi yang sempurna, meski dilakukan dengan metode yang berbeda.
3. Pelaku evaluasi atau evaluator tidak memberikan jawaban atas suatu pertanyaan tertentu. Evaluator tidak berwenang untuk memberikan rekomendasi terhadap keberlangsungan sebuah program. Evaluator hanya membantu memberikan alternatif.
4. Penelitian evaluasi adalah tanggung jawab tim bukan perorangan.
5. Evaluator tidak terikat pada satu sekolah demikian pula sebaliknya.
6. evaluasi adalah proses, jika diperlukan revisi maka lakukanlah revisi.

7. Evaluasi memerlukan data yang akurat dan cukup, hingga perlu pengalaman untuk pendalaman metode penggalian informasi.
8. Evaluasi akan mntap apabila dilakukan dengan instrumen dan teknik yang *aplicable*.
9. Evaluator hendaknya mampu membedakan yang dimaksud dengan evaluasi formatif, evaluasi sumatif dan evaluasi program.
10. Evaluasi memberikan gambaran deskriptif yang jelas mengenai hubungan sebab akibat, bukan terpaku pada angka soalan tes.

Dengan demikian dapat dimengerti bahwa sesungguhnya evaluasi adalah proses mengukur dan menilai terhadap suatu objek dengan menampilkan hubungan sebab akibat diantara faktor yang mempengaruhi objek tersebut.

## **B. Tujuan**

Tujuan evaluasi adalah untuk melihat dan mengetahui proses yang terjadi dalam proses pembelajaran. Proses pembelajaran memiliki 3 hal penting yaitu, input, transformasi dan output. Input adalah peserta didik yang telah dinilai kemampuannya dan siap menjalani proses pembelajaran.

Transformasi adalah segala unsur yang terkait dengan proses pembelajaran yaitu ; guru, media dan bahan belajar, metode pengajaran, sarana penunjang dan sistem administrasi. Sedangkan output adalah capaian yang dihasilkan dari proses pembelajaran.

Evaluasi pendidikan memiliki beberapa fungsi yaitu :

1. Fungsi Selektif.
2. Fungsi Diagnostik.
3. Fungsi Penempatan
4. Fungsi Keberhasilan.

Adapun maksud dilakukannya evaluasi adalah :

1. Perbaikan sistem.
2. Pertanggungjawaban kepada pemerintah dan masyarakat.
3. Penentuan tindak lanjut pengembangan.

### **C. Prinsip-prinsip**

1. Keterpaduan.
2. Evaluasi harus dilakukan dengan prinsip keterpaduan antara tujuan intrusional pengajaran, materi pembelajaran dan metode pengajaran.
3. Keterlibatan peserta didik  
Prinsip ini merupakan suatu hal yang mutlak, karena keterlibatan peserta didik dalam evaluasi bukan alternatif, tapi kebutuhan mutlak.
4. Koherensi.  
Evaluasi harus berkaitan dengan materi pengajaran yang telah dipelajari dan sesuai dengan ranah kemampuan peserta didik yang hendak diukur.
5. Pedagogis.  
Perlu adanya tool penilai dari aspek pedagogis untuk melihat perubahan sikap dan perilaku sehingga pada akhirnya hasil evaluasi mampu menjadi motivator bagi diri siswa.
6. Akuntabel.  
Hasil evaluasi haruslah menjadi alat akuntabilitas atau bahan pertanggungjawaban bagi pihak yang berkepentingan seperti orangtua siswa, sekolah, dan lainnya.

### **D. Prosedur Pelaksanaan**

Dalam melaksanakan evaluasi pendidikan hendaknya dilakukan secara sistematis dan terstruktur. Sebagaimana telah dikemukakan sebelumnya bahwa evaluasi pendidikan secara garis besar melibatkan 3 unsur yaitu input, proses dan output. Apabila prosedur yang dilakukan tidak bercermin pada 3 unsur tersebut maka dikhawatirkan hasil yang digambarkan oleh hasil evaluasi tidak mampu menggambarkan gambaran yang sesungguhnya terjadi dalam proses pembelajaran. Langkah-langkah dalam melaksanakan kegiatan evaluasi pendidikan secara umum adalah sebagai berikut:

- a. perencanaan (mengapa perlu evaluasi, apa saja yang hendak dievaluasi, tujuan evaluasi, teknik apa yang hendak dipakai, siapa yang hendak dievaluasi, kapan, dimana, penyusunan instrument, indikator, data apa saja yang hendak digali, dan sebagainya).
- b. pengumpulan data (tes, observasi, kuesioner, dan sebagainya sesuai dengan tujuan).

- c. verifikasi data (uji instrument, uji validitas, uji reliabilitas, dan sebagainya).
- d. pengolahan data (memaknai data yang terkumpul, kualitatif atau kuantitatif, apakah hendak di olah dengan statistik atau non statistik, apakah dengan parametrik atau non parametrik, apakah dengan manual atau dengan *software* (misal : SAS, SPSS).
- e. penafsiran data, (ditafsirkan melalui berbagai teknik uji, diakhiri dengan uji hipotesis ditolak atau diterima, jika ditolak mengapa? Jika diterima mengapa? Berapa taraf signifikannya?) interpretasikan data tersebut secara berkesinambungan dengan tujuan evaluasi sehingga akan tampak hubungan sebab akibat. Apabila hubungan sebab akibat tersebut muncul maka akan lahir alternatif yang ditimbulkan oleh evaluasi itu.

#### **E. Kegagalan Guru Dalam Melakukan Assesmen dan Evaluasi**

Kalau kita perhatikan dunia pendidikan, kita akan mengetahui bahwa setiap jenis atau bentuk pendidikan pada waktu-waktu tertentu selama satu periode pendidikan, selalu mengadakan evaluasi. Artinya pada waktu-waktu tertentu selama satu periode pendidikan, selalu mengadakan penilaian terhadap hasil yang telah dicapai, baik oleh pihak terdidik maupun oleh pendidik.

Demikian pula dalam satu kali proses pembelajaran, guru hendaknya menjadi seorang evaluator yang baik. Kegiatan ini dimaksudkan untuk mengetahui apakah tujuan yang telah dirumuskan itu tercapai atau belum, dan apakah materi pelajaran yang diajarkan sudah tepat. Semua pertanyaan tersebut akan dapat dijawab melalui kegiatan evaluasi atau penilaian.

Dengan menelaah pencapaian tujuan pengajaran, guru dapat mengetahui apakah proses belajar yang dilakukan cukup efektif memberikan hasil yang baik dan memuaskan atau sebaliknya. Jadi jelaslah bahwa guru hendaknya mampu dan terampil melaksanakan penilaian, karena dengan penilaian guru dapat mengetahui prestasi yang dicapai oleh siswa setelah ia melaksanakan proses belajar.

Dalam fungsinya sebagai penilai hasil belajar siswa, guru hendaknya terus menerus mengikuti hasil belajar yang telah dicapai oleh siswa dari waktu ke waktu. Informasi yang diperoleh melalui evaluasi ini merupakan umpan balik (feed back) terhadap proses belajar mengajar. Umpan balik ini akan dijadikan titik tolak untuk memperbaiki dan

meningkatkan proses belajar mengajar selanjutnya. Dengan demikian proses belajar mengajar akan terus dapat ditingkatkan untuk memperoleh hasil yang optimal.

Khusus untuk mata pelajaran matematika hampir semua guru telah melaksanakan evaluasi di akhir proses belajar mengajar di dalam kelas. Namun hasil yang diperoleh kadang-kadang kurang memuaskan. Kadang-kadang hasil yang dicapai dibawah standar atau di bawah rata-rata.

Pada mata pelajaran yang lainnya kadang dilaksanakan pada akhir pelajaran, dan ada juga pada saat proses belajar mengajar berlangsung. Kapan waktu pelaksanaan evaluasi tersebut tidak menjadi masalah bagi guru yang penting dalam satu kali pertemuan ia telah melaksanakan penilaian terhadap siswa di kelas.

Tetapi ada juga guru yang enggan melaksanakan evaluasi di akhir pelajaran, karena keterbatasan waktu, menurut mereka lebih baik menjelaskan semua materi pelajaran sampai tuntas untuk satu kali pertemuan, dan pada pertemuan berikutnya di awal pelajaran siswa diberi tugas atau soal-soal yang berhubungan dengan materi tersebut.

Ada juga guru yang berpendapat, bahwa penilaian di akhir pelajaran tidak mutlak dengan tes tertulis. Bisa juga dengan tes lisan atau tanya jawab. Kegiatan dirasakan lebih praktis bagi guru, karena guru tidak usah bersusah payah mengoreksi hasil evaluasi anak. Tetapi kegiatan ini mempunyai kelemahan yaitu anak yang suka gugup walaupun ia mengetahui jawaban dari soal tersebut, ia tidak bisa menjawab dengan tepat karena rasa gugupnya itu. Dan kelemahan lain tes lisan terlalu banyak memakan waktu dan guru harus punya banyak persediaan soal. Tetapi ada juga guru yang mewakilkan beberapa orang anak yang pandai, anak yang kurang dan beberapa orang anak yang sedang kemampuannya untuk menjawab beberapa pertanyaan atau soal yang berhubungan dengan materi pelajaran itu.

Cara mana yang akan digunakan oleh guru untuk evaluasi tidak usah dipermasalahkan, yang jelas setiap guru yang paham dengan tujuan dan manfaat dari evaluasi atau penilaian tersebut. Karena ada juga guru yang tidak mengiraukan tentang kegiatan ini, yang penting ia masuk kelas, mengajar, mau ia laksanakan evaluasi di akhir pelajaran atau tidak itu urusannya. Yang jelas pada akhir semester ia telah mencapai target kurikulum.

Akhir-akhir ini kalau kita teliti di lapangan, banyak guru yang mengalami kegagalan dalam melaksanakan evaluasi di akhir pelajaran. Hal ini tentu ada faktor penyebabnya dan apakah cara untuk mengatasinya.

Penulisan kritikan ini bertujuan untuk mengkritik kegagalan persekolah oleh guru dalam melakukan evaluasi di akhir pelajaran. Mencari faktor penyebabnya dan cara untuk mengatasinya.

Dalam kritikan ini pembatasan masalahnya adalah :

- Kondisi permasalahan evaluasi di akhir pelajaran dipersekolahan pada saat ini
- Telaah teori/pendapat ahli
- Kegagalan pelaksanaan evaluasi di akhir pelajaran
- Kesimpulan kritikan dan saran

Menurut Drs. Moh. Uzer Usman dalam bukunya (Menjadi Guru Profesional hal 11) menyatakan bahwa:

Tujuan penilaian adalah :

1. Untuk mengetahui keberhasilan pencapaian tujuan
2. Untuk mengetahui penguasaan siswa terhadap materi pelajaran
3. Untuk mengetahui ketepatan metode yang digunakan
4. Untuk mengetahui kedudukan siswa di dalam kelompok/kelas
5. Untuk mengklasifikasikan seorang siswa apakah termasuk dalam kelompok yang pandai, sedang, kurang atau cukup baik dibandingkan dengan teman-teman sekelasnya.

Dan menurut buku Mengukur Hasil Belajar (hal 72-74) yang disusun oleh Drs. Azhari Zakri menyatakan evaluasi bermanfaat bagi guru untuk :

1. Mengukur kompetensi atau kapabilitas siswa, apakah mereka telah merealisasikan tujuan yang telah ditentukan.
2. Menentukan tujuan mana yang belum direalisasikan sehingga dapat menentukan tindakan perbaikan yang cocok yang dapat diadakan
3. Memutuskan ranking siswa, dalam hal kesuksesan mereka mencapai tujuan yang telah disepakati.
4. Memberikan informasi kepada guru tentang cocok tidaknya strategi mengajar yang digunakan.
5. Merencanakan prosedur untuk memperbaiki rencana pengajaran dan menentukan apakah sumber belajar tambahan perlu digunakan.

6. Memberikan umpan balik kepada kita informasi bagi pengontrolan tentang sesuai tidaknya pengorganisasian belajar dan sumber belajar.
7. Mengetahui dimana letak hambatan pencapaian tujuan tersebut.

Atas dasar ini, faktor yang paling penting dalam evaluasi itu bukan pada pemberian angka. Melainkan sebagai dasar feed back (catu balik). Catu balik itu sendiri sangat penting dalam rangka revisi. Sebab proses belajar mengajar itu kontinyu, karenanya perlu selalu melakukan penyempurnaan dalam rangka mengoptimalkan pencapaian tujuan.

Bila evaluasi merupakan catu balik sebagai dasar memperbaiki sistem pengajaran, sesungguhnya pelaksanaan evaluasi harus bersifat kontinyu. Setiap kali dilaksanakan proses pengajaran, harus dievaluasi (formatif). Sebaliknya bila evaluasi hanya dilaksanakan di akhir suatu program (sumatif) catu balik tidak banyak berarti, sebab telah banyak proses terlampaui tanpa revisi.

Oleh karena itu, agar evaluasi memberi manfaat yang besar terhadap sistem pengajaran hendaknya dilaksanakan setiap kali proses belajar mengajar untuk suatu topik tertentu. Namun demikian evaluasi sumatif pun perlu dilaksanakan untuk pengembangan sistem yang lebih luas.

Dari tujuan dan manfaat evaluasi yang di atas, masih ada pendapat lain dari manfaat evaluasi seperti yang dikemukakan oleh Noehi Nasution dalam bukunya Materi Poko Psikologi Pendidikan hal 167, menjelaskan bahwa kegiatan penilaian tidak hanya untuk mengisi raport anak didik, tetapi juga untuk :

1. Menseleksi anak didik
2. Menjuruskan anak didi
3. Mengarahkan anak didik kepada kegiatan yang lebih sesuai denganpotensi yang dimilikinya
4. Membantu orang tua untuk menentukan hal yang paling baik untuk anaknya, untuk membina dan untuk mempersiapkan dirinya untuk masa depan yang lebih baik.

Tujuan dan manfaat evaluasi yang telah diikemukakan oleh para ahli di atas, yang penting dengan mengadakan evaluasi sebagai guru dapat mengetahui kelemahan-kelemahan atau kekurangannya dalam menyampaikan materi pelajaran. Sehingga ia dapat menata kembali atau menggunakan strategi baru dalam proses pembelajaran sehingga akan mendapatkan hasil yang lebih baik dari sebelumnya.

Di dalam telaah teori dan berdasarkan pendapat para ahli, telah mencantumkan tujuan serta manfaat evaluasi di akhir pelajaran. Selain menilai hasil belajar murid, evaluasi juga menilai hasil mengajar guru dengan kata lain, guru dapat menilai dirinya sendiri dimana kekurangan dan kelemahannya dalam mengajar, sehingga memperoleh hasil yang sesuai dengan apa yang diharapkan.

Jika dalam suatu kegiatan belajar, tujuan sudah diidentifikasi, biasanya dapat disusun suatu ters atau ujian yang akan digunakan untuk menentukan apakah tujuan tersebut dicapai atau tidak. Mager pernah mengatakan bahwa jika kita mempelajari dengan teliti semua tahap yang telah dibicarakan sampai saat ini, maka siswa sudah harus dapat melakukan apa yang telah direncanakan untuk mereka lakukan. Hasil dari penilaian dapat mendorong guru untuk memperbaiki keterampilan profesional mereka, dan juga membantu mereka mendapat pasilitas serta sumber belajar yang lebih baik.

Di dalam suatu tes belajar, sebagian besar nilai berdistribusi normal (yakni beberapa murid hasilnya baik, beberapa buruk, tetapi sebagian besar menunjukkan rata-rata). Dalam ter kriteria, sebagian tes berada di bagian atas. Hal ini lumrah, karena jika seorang guru memberikan tujuan yang berjumlah 10, misalnya, maka ia akan kecewa jika para siswa hanya merealisasikan 50% saja.

Tes dan ujian yang mengukur pencapaian tujuan, belum mendapat perhatian yang serius oleh guru dan instruktur, kecuali akhir-akhir ini. Program pendidikan dan latihan sebelum ini telah dianggap sudah berhasil tanpa perlu ada evaluasi. Sikap ini disebabkan oleh empat kesulitan utama yakni :

1. Tidak adanya kerangka konseptual yang sesuai bagi evaluasi.
2. Kurangnya ketepatan dalam perumusan tujuan dalam pendidikan
3. Kesulitan yang meliputi pengukuran pendidikan
4. Sifat program pendidikan itu sendiri.

Namun dengan adanya investasi besar-besaran dalam pendidikan, telah dirasakan kebutuhan akan suatu bentuk evaluasi.

Evaluasi dapat mengambil dua macam bentuk :

- 1) Ia dapat menilai cara mengajar seorang guru (dengan mengukur variabel-variabel seperti suatu kebiasaan-kebiasaan, humor, kepribadian, penggunaan papan tulis, teknik bertanya, aktivitas kelas, alat bantu audiovisual, strategi mengajar dan lain-lain.
- 2) Ia dapat menilai hasil belajar (yakni pencapaian tujuan belajar.

Selama ini guru mengadakan penilaian hanya untuk mencari angka atau nilai untuk anak didik. Apabila anak banyak memperoleh nilai dibawah 6 (enam), maka guru menganggap bahwa anak didiklah yang gagal dalam menyerap materi pelajaran atau materi pelajaran terlalu berat, sehingga sukar dipahami oleh anak. Kalau anak yang memperoleh nilai dibawah 6 mencapai 50% dari jumlah anak, hal ini sudah merupakan kegagalan guru dalam melaksanakan evaluasi di akhir pelajaran.

**Apa penyebab hal ini bisa terjadi ?**

1. Guru kurang menguasai materi pelajaran. Sehingga dalam menyampaikan materi pelajaran kepada anak kalimatnya sering terputus-putus ataupun berbelit-belit yang menyebabkan anak menjadi bingung dan sukar mencerna apa yang disampaikan oleh guru tersebut. Tentu saja di akhir pelajaran mereka kewalahan menjawab pertanyaan atau tidak mampu mengerjakan tugas yang diberikan. Dan akhirnya nilai yang diperoleh jauh dari apa yang diharapkan.
2. Guru kurang menguasai kelas, Guru yang kurang mampu menguasai kelas mendapat hambatan dalam menyampaikan materi pelajaran, hal ini dikarenakan suasana kelas yang tidak menunjang membuat anak yang betul-betul ingin belajar menjadi terganggu.
3. Guru enggan mempergunakan alat peraga dalam mengajar. Kebiasaan guru yang tidak mempergunakan alat peraga memaksa anak untuk berpikir verbal sehingga membuat anak sulit dalam memahami pelajaran dan otomatis dalam evaluasi di akhir pelajaran nilai anak menjadi jatuh.
4. Guru kurang mampu memotivasi anak dalam belajar sehingga dalam menyampaikan materi pelajaran, anak kurang menaruh perhatian terhadap materi yang disampaikan oleh guru, sehingga ilmu yang terkandung di dalam materi yang disampaikan itu berlalu begitu saja tanpa ada perhatian khusus dari anak didik.
5. Guru menyamaratkan kemampuan anak di dalam menyerap pelajaran. Setiap anak didik mempunyai kemampuan yang berbeda dalam menyerap materi pelajaran. Guru yang kurang tangkap tidak mengetahui bahwa ada anak didinya yang daya serapnya di bawah rata-rata mengalami kesulitan dalam belajar.

6. Guru kurang disiplin dalam mengatur waktu. Waktu yang tertulis dalam jadwal pelajaran, tidak sesuai dengan praktek pelaksanaannya,. Waktu untuk memulai pelajaran selalu telat, tetapi waktu istirahat dan jam pulang selalu tepat atau tidak pernah telat.
7. Guru enggan membuat persiapan mengajar atau setidaknya menyusun langkah-langkah dalam mengajar, yang disertai dengan ketentuan-ketentuan waktu untuk mengawali pelajaran, waktu untuk kegiatan proses dan ketentuan waktu untuk akhir pelajaran.
8. Guru tidak mempunyai kemajuan untuk menambah atau menimba ilmu misalnya membaca buku atau bertukar pikiran dengan rekan guru yang lebih senior dan profesional guna menambah wawasannya.
9. Dalam tes lisan di akhir pelajaran, guru kurang trampil mengajukan pertanyaan kepada murid, sehingga murid kurang memahami tentang apa yang dimaksud oleh guru.
10. Guru selalu mengutamakan pencapaian target kurikulum. Guru jarang memperhatikan atau menganalisa berapa persen daya serap anak terhadap materi pelajaran tersebut.

#### **F. Jenis-jenis Penilaian**

Dilihat dari fungsinya, penilaian terdiri atas beberapa macam yakni penilaian formatif, penilaian sumatif, penilaian diagnostik, penilaian selektif dan penilaian penempatan.

**Penilaian formatif** adalah penilaian yang dilaksanakan pada akhir program belajar mengajar untuk melihat tingkat keberhasilan proses belajar mengajar itu sendiri. Penilaian formatif berorientasi pada proses, yang akan memberikan informasi kepada guru apakah program atau proses belajar mengajar masih perlu diperbaiki.

**Penilaian sumatif** adalah penilaian yang dilaksanakan pada akhir unit program misalnya penilaian yang dilaksanakan pada akhir caturwulan, akhir semester atau akhir tahun. Tujuan penilaian ini adalah untuk mengetahui hasil yang dicapai oleh para siswa, yakni seberapa jauh siswa telah mencapai kompetensi yang ditetapkan dalam kurikulum. Penilaian ini berorientasi pada produk/hasil.

**Penilaian diagnostik** adalah penilaian yang bertujuan untuk mengetahui kelemahan-kelemahan siswa serta faktor-faktor penyebabnya. Pelaksanaan penilaian semacam ini biasanya bertujuan untuk keperluan

bimbingan belajar, pengajaran remedial, menemukan kasus-kasus dan lain-lain.

**Penilaian selektif** adalah penilaian yang dilaksanakan dalam rangka menyeleksi atau menyaring. Memilih siswa untuk mewakili sekolah dalam lomba-lomba tertentu termasuk jenis penilaian selektif. Untuk kepentingan yang lebih luas penilaian selektif misalnya seleksi penerimaan mahasiswa baru atau seleksi yang dilakukan dalam rekrutmen tenaga kerja.

**Penilaian penempatan** adalah penilaian yang bertujuan untuk mengetahui keterampilan prasyarat yang diperlukan bagi suatu program belajar dan penguasaan belajar seperti yang diprogramkan sebelum memulai kegiatan belajar untuk program itu. Dengan kata lain penilaian ini berorientasi pada kesiapan siswa untuk menghadapi program baru dan kecocokan program belajar dengan kemampuan yang telah dimiliki siswa.

## Rangkuman

- Evaluasi adalah proses menggambarkan, memperoleh dan menyajikan informasi yang berguna untuk menilai alternatif keputusan.
- langkah-langkah pelaksanaan evaluasi pelatihan sebagai berikut :
  1. Persiapan Evaluasi
  2. Mengembangkan Instrumen
  3. Mengumpulkan Data
  4. Mengolah dan Menganalisis Data
  5. Menyusun Laporan
- Tujuan evaluasi adalah untuk melihat dan mengetahui proses yang terjadi dalam proses pembelajaran.
- Secara garis besar, proses evaluasi terbagi menjadi di awal (*pretest*) dan diakhir (*posttest*). *Pretest* merupakan sebuah evaluasi yang diadakan untuk menguji konsep dan eksekusi yang direncanakan. Sedangkan, *posttest* merupakan evaluasi yang diadakan untuk melihat tercapainya tujuan dan dijadikan sebagai masukan untuk analisis situasi berikutnya.
- Jenis-jenis Penilaian :
  1. *Penilaian formatif*
  2. *Penilaian sumatif*
  3. *Penilaian diagnostik*
  4. *Penilaian selektif*
  5. *Penilaian penempatan*

## BAB V TEKNIK EVALUASI

Teknik evaluasi digolongkan menjadi 2 yaitu teknik tes dan teknik non Tes : Teknik non tes meliputi: skala bertingkat, kuesioner,daftar cocok, wawancara, pengamatan, riwayat hidup.

1. *Rating scale* atau skala bertingkat menggambarkan suatu nilai dalam bentuk angka. Angka-angka diberikan secara bertingkat dari angak terendah hingga angkat paling tinggi. Angka-angka tersebut kemudian dapat dipergunakan untuk melakukan perbandingan terhadap angka yang lain.
2. Kuesioner adalah daftar pertanyaan yang terbagi dalam beberapa kategori. Dari segi yang memberikan jawaban, kuesioner dibagi menjadi kuesioner langsung dan kuesioner tidak langsung.
  - Kuesioner langsung adalah kuesioner yang dijawab langsung oleh orang yang diminta jawabannya.
  - Kuesioner tidak langsung dijawab secara tidak langsung oleh orang yang dekat dan mengetahui si penjawab seperti contoh, apabila yang hendak dimintai jawaban adalah seseorang yang buta huruf maka dapat dibantu oleh anak, tetangga atau anggota keluarganya.

Bila ditinjau dari segi cara menjawab maka kuesioner terbagi menjadi kuesioner tertutup dan kuesioner terbuka.

- Kuesioner tertutup adalah daftar pertanyaan yang memiliki dua atau lebih jawaban dan si penjawab hanya memberikan tanda silang (X) atau cek (√) pada awaban yang ia anggap sesuai.
  - Kuesioner terbuka adalah daftar pertanyaan dimana si penjawab diperkenankan memberikan jawaban dan pendapat nya secara terperinci sesuai dengan apa yang ia ketahui.
3. Daftar cocok adalah sebuah daftar yang berisikan pernyataan beserta dengan kolom pilihan jawaban. Si penjawab diminta untuk memberikan tanda silang (X) atau cek (√) pada awaban yang ia anggap sesuai.
  4. Wawancara, suatu cara yang dilakukan secara lisan yang berisikan pertanyaan-pertanyaan yang sesuai dengan tujuan informasi yang

hendak digali. wawancara dibagi dalam 2 kategori, yaitu pertama, wawancara bebas yaitu si penjawab (responden) diperkenankan untuk memberikan jawaban secara bebas sesuai dengan yang ia ketahui tanpa diberikan batasan oleh pewawancara. Kedua adalah wawancara terpimpin dimana pewawancara telah menyusun pertanyaan pertanyaan terlebih dahulu yang bertujuan untuk menggiring penjawab pada informasi-informasi yang diperlukan saja.

5. Pengamatan atau observasi, adalah suatu teknik yang dilakukan dengan mengamati dan mencatat secara sistematis apa yang tampak dan terlihat sebenarnya. Pengamatan atau observasi terdiri dari 3 macam yaitu:
  - (1) observasi partisipan yaitu pengamat terlibat dalam kegiatan kelompok yang diamati.
  - (2) Observasi sistematis, pengamat tidak terlibat dalam kelompok yang diamati. Pengamat telah membuat list faktor faktor yang telah diprediksi sebagai memberikan pengaruh terhadap sistem yang terdapat dalam objek pengamatan.
6. Riwayat hidup, evaluasi ini dilakukan dengan mengumpulkan data dan informasi mengenai objek evaluasi sepanjang riwayat hidup objek evaluasi tersebut.

#### **A. Teknik Tes**

Teknik tes meliputi tes lisan, tes tertulis dan tes perbuatan. Tes lisan dilakukan dalam bentuk pertanyaan lisan di kelas yang dilakukan pada saat pembelajaran di kelas berlangsung atau di akhir pembelajaran. Tes tertulis adalah tes yang dilakukan tertulis, baik pertanyaan maupun jawabannya. Sedangkan tes perbuatan atau tes unjuk kerja adalah tes yang dilaksanakan dengan jawaban menggunakan perbuatan atau tindakan.

Evaluasi dengan menggunakan teknik tes bertujuan untuk mengetahui:

- a. Tingkat kemampuan awal siswa.
- b. Hasil belajar siswa.
- c. Perkembangan prestasi siswa.
- d. Keberhasilan guru dalam melaksanakan kegiatan pembelajaran.

Tes lisan dilakukan melalui pertanyaan lisan untuk mengetahui daya serap siswa. Tujuan tes lisan ini terutama untuk menilai :

- a. Kemampuan memecahkan masalah.
- b. Proses berpikir terutama melihat hubungan sebab akibat.
- c. Kemampuan menggunakan bahasa lisan.
- d. Kemampuan mempertanggungjawabkan pendapat atau konsep yang dikemukakan.

Tes tertulis dapat berbentuk uraian (*essay*) atau soal bentuk obyektif (*objective tes*). Tes uraian merupakan alat penilaian hasil belajar yang paling tua. Secara umum tes uraian ini adalah pertanyaan yang menuntut siswa menjawab dalam bentuk menguraikan, menjelaskan, mendiskusikan, membandingkan, memberikan alasan, dan bentuk lain yang sejenis sesuai dengan tuntutan pertanyaan dengan menggunakan kata-kata dan bahasa sendiri.

**Cara-cara penyusunan tes esai yang dimaksud :**

- a. Guru hendaknya memfokuskan pertanyaan esai pada materi pembelajaran yang tidak dapat diungkap dengan bentuk tes lain misalnya tes objektif.
- b. Guru hendaknya memformulasikan item pertanyaan yang mengungkap perilaku spesifik yang diperoleh dari pengalaman hasil belajar.
- c. Item-item pertanyaan tes esai sebaiknya jelas dan tidak menimbulkan kebingungan sehingga siswa dapat menjawabnya dengan tidak ragu-ragu.
- d. Sertakan petunjuk waktu pengerjaan untuk setiap pertanyaan, agar para siswa dapat memperhitungkan kecepatan berpikir, menulis dan menuangkan ide sesuai dengan waktu yang disediakan.
- e. Ketika mengonstruksi sejumlah pertanyaan esai, para guru hendaknya menghindari penggunaan pertanyaan pilihan. Misalnya pilih empat soal dari lima pertanyaan yang tersedia.

Menurut Sukardi (2008) kelebihan dan kelemahan Tes Esai, dapat diuraikan sebagai berikut.

Kelebihannya:

- a. Mengukur proses mental siswa dalam menuangkan ide ke dalam jawaban item secara tepat
- b. Mengukur kemampuan siswa dalam menjawab melalui kata dan bahasa mereka sendiri.
- c. Mendorong siswa untuk mempelajari, menyusun, merangkai, dan menyatakan pemikiran siswa secara aktif.
- d. Mendorong siswa untuk berani mengemukakan pendapat serta menyusun dalam bentuk kalimat mereka sendiri.
- e. Mengetahui seberapa jauh siswa telah memahami dan mendalami suatu permasalahan atas dasar pengetahuan yang diajarkan di dalam kelas

Kelemahan:

- a. Dalam memeriksa jawaban pertanyaan tes esai, ada kecenderungan pengaruh subjektif yang selalu muncul dalam pribadi seorang guru.
- b. Pertanyaan esai yang disusun oleh seorang guru atau evaluator cenderung kurang bisa mencakup seluruh materi yang telah diberikan
- c. Bentuk pertanyaan yang memiliki arti ganda, sering membuat kesulitan pada siswa sehingga memunculkan unsur-unsur menerka dan menjawab dengan ragu-ragu.

Tes Objektif banyak digunakan dalam menilai hasil belajar. Hal ini disebabkan antara lain oleh luasnya bahan pelajaran yang dapat dicakup dalam tes dan mudahnya menilai jawaban yang diberikan.

### **1. Bentuk Soal Benar-Salah.**

Bentuk soal benar salah adalah bentuk tes yang soal-soalnya berupa pernyataan. Sebagian dari pernyataan itu merupakan pernyataan yang benar dan sebagian lagi merupakan pernyataan yang salah.

Kelebihan Benar Salah yaitu:

- a. Item tes betul salah memiliki karakteristik yang menguntungkan, yaitu mudah dan cepat dalam menilai.
- b. Untuk item betul salah yang dikonstruksi secara cermat, membawa implikasi kepada peserta didik, yaitu waktu mengerjakan soal lebih cepat diselesaikan.

- c. Seperti bentuk tes objektif lainnya, item tes benar salah hasil akhir penilaian dapat objektif

Kelemahan Benar Salah:

- a. Mengonstruksi item tes betul salah pada umumnya diperlukan waktu yang lebih lama jika dibandingkan dengan pembuatan tes esai
- b. Penggunaan pertanyaan alternatif lebih memungkinkan peserta didik mengira-ngira jawaban.

## **2. Bentuk Soal Pilihan Ganda atau Pilihan Jamak (*multiple choice*)**

Soal pilihan ganda adalah bentuk tes yang mempunyai satu jawaban yang benar atau paling tepat.

Kelebihan Soal Pilihan Ganda yaitu:

- a. Tes pilihan ganda memiliki karakteristik yang baik untuk suatu alat pengukur hasil belajar siswa
- b. Item tes pilihan ganda yang dikonstruksi dengan intensif dapat mencakup hampir seluruh bahan pembelajaran yang diberikan oleh guru di kelas.
- c. Item tes pilihan ganda adalah tepat untuk mengukur penguasaan informasi para siswa yang hendak dievaluasi.

Kelemahan Soal Pilihan Ganda yaitu;

- a. Mengonstruksi item tes betul salah pada umumnya diperlukan waktu yang lebih lama jika dibandingkan dengan pembuatan tes esai
- b. Penggunaan pertanyaan alternative lebih memungkinkan peserta didik mengira-ngira jawaban.

## **3. Bentuk soal menjodohkan (*matching*).**

Bentuk soal menjodohkan terdiri atas dua kelompok pernyataan yang paralel. Kedua kelompok pernyataan ini berada dalam satu kesatuan. Kelompok sebelah kiri merupakan bagian yang berisi soal-soal yang harus dicari jawabannya.

Kelebihan Bentuk Soal Menjodohkan:

- a. Penilaiannya dapat dilakukan dengan cepat dan objektif.
- b. Tepat digunakan untuk mengukur kemampuan bagaimana mengidentifikasi antara dua hal yang berhubungan.
- c. Dapat mengukur ruang lingkup pokok bahasan atau subpokok bahasan yang lebih luas.

Kelemahan Bentuk Soal Menjodohkan:

- a. Hanya dapat mengukur hal-hal yang didasarkan atas fakta dan hafalan
- b. Sukar untuk menentukan materi atau pokok bahasan yang mengukur hal-hal yang berhubungan.

#### **4. Bentuk Soal Jawaban Singkat (Isian).**

Bentuk soal jawaban singkat merupakan soal yang menghendaki jawaban dalam bentuk kata, bilangan, kalimat, atau simbol.

Kelebihan bentuk soal jawaban singkat;

- a. Menyusun soalnya relatif mudah
- b. Kecil kemungkinan siswa member jawaban dengan cara menebak
- c. Menuntut siswa untuk dapat menjawab dengan singkat dan tepat
- d. Hasil penilaiannya cukup objektif

Kelemahan bentuk soal jawaban singkat:

- a. Kurang dapat mengukur aspek pengetahuan yang lebih tinggi.
- b. Memerlukan waktu yang agak lama untuk menilainya sekalipun tidak selama bentuk uraian
- c. Menyulitkan pemeriksaan apabila jawaban siswa membingungkan pemeriksa.

### **B. Teknik Non Tes**

Teknik tes bukanlah satu-satunya teknik untuk melakukan evaluasi hasil belajar, sebab masih ada teknik lainnya yang dapat digunakan, yaitu teknik non tes. Dengan teknik non tes maka penilaian atau evaluasi hasil belajar peserta didik dilakukan dengan tanpa menguji peserta didik melainkan dilakukan melalui:

#### **1. Pengamatan atau observasi.**

Secara umum, pengertian observasi adalah cara menghimpun bahan-bahan keterangan yang dilakukan dengan mengadakan pengamatan dan pencatatan secara sistematis terhadap fenomena-fenomena yang sedang dijadikan sasaran pengamatan. Alat yang digunakan berupa lembar observasi yang disusun dalam bentuk check list atau skala penilaian.

#### **2. Wawancara.**

Secara umum yang dimaksud dengan wawancara adalah cara menghimpun bahan-bahan keterangan yang dilaksanakan dengan melakukan tanya jawab lisan secara sepihak. Alat yang

digunakan adalah pedoman wawancara yang mengacu pada tujuan yang telah ditetapkan.

### **3. Angket.**

Angket adalah wawancara yang dilakukan secara tertulis. Angket dapat digunakan sebagai alat penilaian hasil belajar. Angket dapat diberikan langsung kepada peserta didik, dapat pula diberikan kepada orang tua mereka.

### **4. Skala.**

Skala adalah alat untuk mengukur nilai, sikap, minat, perhatian, dan lain-lain yang disusun dalam bentuk pernyataan untuk dinilai oleh responden dan hasilnya dalam bentuk rentangan nilai sesuai dengan kriteria yang ditentukan.

### **Kriteria Tes yang Baik**

- Validitas (Ketepatan); Suatu alat pengukur dapat dikatakan alat pengukur yang valid apabila alat pengukur tersebut dapat mengukur apa yang hendak diukur secara tepat.
- Reliabilitas merujuk pada konsistensi skor yang dicapai oleh orang yang sama ketika diuji ulang dengan tes yang sama pada kesempatan yang berbeda, atau dengan seperangkat butir-butir ekuivalen yang berbeda, atau pada kondisi pengujian yang berbeda
- Objektivitas; Suatu tes dikatakan obyektif jika tes tersebut diajukan kepada beberapa penilai, tetapi memberikan skor yang sama, untuk disiapkan kunci jawaban (scoring key).
- Memiliki discrimination power (daya pembeda); Tes yang dikatakan baik apabila mampu membedakan anak yang pandai dan anak yang bodoh.
- Mencakup ruang lingkup (scope) yang sangat luas dan menyeluruh; Tes yang baik harus memiliki komprehensi veenes, ini akan menyisihkan siswa yang berspekulasi dalam menempuh tes.
- Praktis; mencakup :
  - Mudah dipakai/ diperiksa
  - Hemat biaya
  - Mudah diadministrasikan
  - Tidak menyulitkan guru dan sekolah.

Tingkat kesahihan dan keterandalan instrumen evaluasi ditentukan oleh :

- Keadaan objek yang dievaluasi. Misalnya sasaran didik sedang dalam keadaan sedih dan senang, maka kalau orang tersebut dites, ia akan memberikan respon yang berbeda terhadap suatu tes atau ujian yang sama.
- Alat pengukur. Alat yang kurang baik, misalnya penggaris yang melengkung, timbangan yang tidak pernah ditar dan sebagainya
- Situasi pengukuran. Situasi yang akrab akan menghasilkan data yang berbeda dengan situasi yang tegang. Misalnya pada waktu evaluasi dilakukan, kelas ditunggu oleh Menteri hasilnya akan lain apabila ditunggu oleh muridnya.
- Penyelenggaraan evaluasi. Evaluasi yang diselenggarakan secara tertib akan menghasilkan data yang berbeda dengan yang diselenggarakan secara kacau.

### **Rangkuman**

- Teknik evaluasi digolongkan menjadi 2 yaitu teknik tes dan teknik non Tes.
- Teknik tes meliputi tes lisan, tes tertulis dan tes perbuatan.
- Teknik non tes meliputi: skala bertingkat, kuesioner, daftar cocok, wawancara, pengamatan, riwayat hidup.
- Teknik tes lisan terdiri dari :
  1. *Bentuk Soal Benar-Salah.*
  2. *Bentuk Soal Pilihan Ganda atau Pilihan Jamak (multiple choice)*
  3. *Bentuk soal menjodohkan (matching).*
  4. *Bentuk Soal Jawaban Singkat (Isian).*
- Evaluasi dengan menggunakan teknik tes bertujuan untuk mengetahui:
  - a. Tingkat kemampuan awal siswa.
  - a. Hasil belajar siswa.
  - b. Perkembangan prestasi siswa.
  - c. Keberhasilan guru dalam melaksanakan kegiatan pembelajaran.

## **BAB VI INSTRUMEN TES OBJEKTIF**

### **A. Pengertian Tes Objektif**

Tes objektif adalah tes yang dalam pemeriksaannya dapat dilakukan secara objektif. Hal ini memang dilakukan untuk mengatasi kelemahan-kelemahan dari tes bentuk esai (uraian). Dalam penggunaan tes objektif ini jumlah soal yang diajukan jauh lebih banyak dari tes uraian. Kadang-kadang untuk tes yang berlangsung selama 60 menit dapat diberikan 30-40 buah soal. Tes objektif juga merupakan tes dimana keseluruhan informasi yang diperlukan untuk menjawab tes telah tersedia dan peserta harus memilih salah satu alternatif yang disediakan tersebut ((Estina Ekawati dan Sumaryanta, 2011). Dalam tes objektif siswa dituntut untuk memilih beberapa kemungkinan jawaban yang telah ter sedia dan/atau memberi jawaban singkat atau mengisis titik– titik ditempat yang tersedia (Sumadi Suryabrata,1987). Soal sudah disusun terstruktur dengan sempurna.

Tes objektif terdiri dari empat jenis yaitu :

- Tes benar salah ( *true false test* )
- Tes pilihan ganda ( *multiple choice test* )
- Tes menjodohkan ( *multiple test* )
- Tes isian atau tes melengkapi ( *completion test* ).

#### **a. Keuntungan atau kelebihan tes objektif**

1. Mengandung lebih banyak segi-segi yang positif, misalnya lebih representatif mewakili isi dan luas bahan, lebih objektif, dapat dihindari campur tangannya unsur-unsur subjektif baik dari segi siswa maupun segi guru yang memeriksa.
2. Lebih mudah dan lebih cepat memeriksanya karena dapat menggunakan kunci tes bahkan alat-alat hasil kemajuan teknologi.
3. Pemeriksaannya dapat diserahkan pada orang lain.
4. Dalam pemeriksaan tidak ada unsur subjektif yang mempengaruhi.
5. Mempunyai validitas yang tinggi.
6. Memiliki tingkat kepercayaan (reabilitas) yang tinggi yang susah dicapai tes essay.
7. Dapat meliputi aspek-aspek bahan pelajaran dan kecakapan yang cukup lengkap.
8. Petunjuknya mudah dimengerti dan pengajarannya lebih mudah.

9. Skoring lebih mudah dan lebih cepat dari pada tes essay.
10. Item-item tes objektif dapat dianalisa dengan item analisis untuk
11. Meningkatkan mutu tes-tes yang akan datang.
12. Tes objektif dapat digunakan lagi berulang-ulang kali selama masih valid dan tidak bocor.

**b. Kekurangan tes objektif**

1. Cara membuatnya memerlukan waktu, tenaga, pikiran dan ketekunan yang banyak
2. Tidak semua aspek pribadi anak dapat diukur dengan tes objektif. Tes objektif ini berhasil baik untuk mengukur ingatan atau pengetahuan saja. Sukar untuk mengukur berpikir, sikap dan keterampilan
3. Jawaban anak belum tentu menunjukkan hasil yang sebenarnya sebab anak-anak kemungkinan hanya akan mengira-ngira saja
4. Kurang ekonomis, sebab banyak membutuhkan kertas dan lain-lain.
5. Persiapan untuk menyusunnya jauh lebih sulit dari pada tes esai karena soalnya banyak dan harus teliti untuk menghindari kelemahan-kelemahan yang lain.
6. Soal-soalnya cenderung untuk mengungkapkan ingatan dan daya pengenalan kembali saja, dan sukar untuk mengukur proses mental yang tinggi.
7. Banyak kesempatan untuk main untung-untungan.
8. Kerja sama antar siswa pada mengerjakan soal tes lebih terbuka.

**c. Cara mengatasi kelemahan tes objektif**

- 1) Kesulitan menyusun tes objektif dapat diatasi dengan jalan banyak berlatih terus menerus hingga betul-betul mahir.
- 2) Menggunakan tabel spesifikasi untuk mengatasi kelemahan nomor satu dan dua.
- 3) Menggunakan norma (standar) penilaian yang memperhitungkan faktor tebakan (guessing) yang bersifat spekulatif.

## **B. Penggolongan Tes Objektif**

### **a. Tes Benar – Salah (*true-false*)**

#### **1. Pengertian**

Soal-soalnya berupa pernyataan-pernyataan (*statement*). *Statement* tersebut ada yang benar dan ada yang salah. Orang yang ditanya bertugas menandaimasing-masing pernyataan itu dengan melingkari huruf B jika pernyataan itu betul menurut pendapatnya dan melingkari huruf S jika pernyataannya salah. Bentuk tes benar-salah ada 2 macam (dilihat dari segi mengerjakan/menjawab soal), yakni:

- Dengan pembetulan (*with corection*) yaitu siswa diminta membetulkan jika ia memilih jawaban yang salah.
- Tanpa pembetulan (*without correction*) yaitu siswa hanya diminta melingkari huruf B atau S tanpa memberikan jawaban yang betul.

Soal tes ini berbentuk kalimat berita atau pertanyaan yang mengandung dua kemungkinan yaitu benar atau salah. Siswa diminta untuk menentukan pendapatnya mengenai pertanyaan-pertanyaan yang menjadi isi dari setiap soal. Bentuk tes ini bermacam-macam variasinya jika dilihat dari segi pola pengerjaannya, yaitu :

- 1) Tes Benar-Salah bentuk pernyataan  
Dalam bentuk ini soal terdiri dari pernyataan – pernyataan dan siswa diminta memilih kemungkinan betul atau salah saja.
- 2) Tes benar salah yang menuntut alasan  
Dalam bentuk ini selain seperti bentuk pertama juga menuntut supaya siswa memberi alasan apabila ia tak memilih kemungkinan salah (menyalahkan pernyataan soal).
- 3) Tes benar salah dengan membetulkan  
Dalam bentuk ini selain seperti bentuk pertama juga menuntut supaya siswa membetulkan pernyataan soal yang disalahkan (jika siswa memilih kemungkinan salah terhadap pernyataan/soal yang bersangkutan).
- 4) Tes benar salah berganda  
Pada tempat ini satu induk persoalan menghasilkan beberapa anak persoalan. Beberapa anak persoalan itu dirumuskan dalam pernyataan/soal yang mempunyai kemungkinan betul atau salah.

## 2. **Kebaikan Tes Benar-Salah**

- Mudah dan cepat dalam menilai
- Waktu mengerjakannya cepat
- Penilaiannya objektif
- Menyusun soalnya lebih mudah dibanding dengan tes pilihan berganda
- Mencakup bahwa yang luas dan tidak banyak memakan tempat karena biasanya pertanyaan-pertanyaannya singkat saja
- Dapat dipakai berulang-ulang.
- Dapat mencakup bahan yang luas dan tidak banyak memakan tempat karena biasanya pertanyaan-pertanyaannya singkat saja.
- Dapat digunakan berkali-kali.
- Petunjuk cara mengerjakannya mudah dimengerti.

## 3. **Kelemahan-kelemahannya**

- Lama menyusun soalnya dibanding dengan tes essay
- Kemungkinan mengira-ngira jawabannya benar
- Menyusun pernyataan (soal) supaya pernyataan itu benar atau hanya salah adalah sulit
- Kurang dapat membedakan murid yang pandai dari murid yang kurang pandai
- Reliabilitasnya rendah
- Bisa membingungkan siswa
- Hanya dapat mengungkapkan ingatan dan pengenalan kembali
- Banyak/hal persoalan yang tidak dapat diukur dinyatakan hanya dengan dua kemungkinan ( benar atau salah ) saja

## 4. **Cara menyusun soal**

- ✓ Syarat – syarat menyusun soal:
- ✓ Soal harus singkat, jelas dan bukan kalimat ,ajemuk
- ✓ Jumlah soal harus banyak dan disusun atas dasar tabel spesifikasi
- ✓ Satu soal harus berisi satu persoalan
- ✓ Tidak menggunakan kata-kata seperti: selalu, seringkali, pada umumnya biasanya, karena kata-kata itu memudahkan murid untuk menerka jawaban
- ✓ Setiap pernyataan/soal harus pasti salah atau betul (tidak mendua arti)
- ✓ Jumlah soal yang betul dan yang salah harus seimbang.

- ✓ Urutan soal (yang betul dan yang salah) seharusnya tidak mengikuti pola yang teratur
- ✓ Sebaiknya pernyataan tidak diambil langsung dari buku
- ✓ Tulislah huruf B–S pada permulaan nomor masing–masing item dengan maksud untuk mempermudah mengerjakan dan menilai (skoring).

### 5. Cara mengolah skor

Rumus untuk mencari skor akhir bentuk benar-salah ada 2 macam, yaitu:

#### Dengan denda

$$S = R - W$$

Keterangan:

S = skor yang diperoleh

R = right (jawaban yang benar)

W = wrong (jawaban yang salah)

*Contoh:*

Jumlah soal tes = 20

A menjawab betul 16 buah dan salah 4 buah. Maka skor untuk A adalah:

$$16 - 4 = 12$$

Dengan menggunakan rumus seperti ini maka ada kemungkinan seorang siswa memperoleh skor negatif.

#### Tanpa denda

$$S = R$$

Yang dihitung hanya yang betul (untuk nilai yang tidak dikerjakan nilainya 0).

## **b. Tes Pilihan Ganda (*multiple choice test*)**

### **1. Pengertian**

Tes ini pada pokoknya menghadapkan kepada siswa sejumlah alternative jawaban, umumnya antara 3 sampai 5 alternatif untuk setiap soal dan tugas siswa dalam memilih salah satu diantara alternatif tersebut berdasarkan sesuatu dasar pertimbangan tertentu. Sebelum memasuki uraian dengan contoh – contoh ada baiknya kalau diberikan definisi tentang beberapa pengertian yang akan selalu kembali dipergunakan dalam uraian – uraian selanjutnya. *Stem* adalah bagian pokok dari soal yang merumuskan isi soal. *Stem* ini bias jawaban yang menyertainya dinamakan *options* atau kalau diterjemahkan secara berbentuk pertanyaan, perintah maupun kalimat tidak sempurna. Alternatif–alternative langsung pilihan–pilihan. Alternatif yang benar dinamakan *key* atau kunci, sedangkan alternatif–alternatif lainnya yang bertujuan mempersulit proses pencapaian jawaban yang benar dinamakan *distractors* atau kalau secara langsung diterjemahkan pengganggu – pengganggu/pengecoh.

### **2. Jenis tes pilihan berganda**

- a) Tes pilihan ganda dengan menemukan satu – satunya jawaban yang benar
- b) Tes pilihan ganda dengan memilih jawaban yang terbaik
- c) Tes pilihan ganda dengan memilih lebih dari satu jawaban benar
- d) Tes pilihan ganda yang menggunakan gambar, bagan, peta dan lain-lain.

### **3. Keباikan tes pilihan ganda**

- Lebih fleksibel dan efektif
- Mencakup hampir seluruh bahan pelajaran
- Tepat untuk mengukur penguraian informasi, perbendaharaan kata–kata, pengertian–pengertian, aplikasi prinsip, rumus serta kemampuan untuk menginterpretasikan data
- Dapat juga untuk mengukur kemampuan siswa dalam hal membuat penafsiran, melakukan pemilihan, mendiskriminasikan, menentukan pendapat atas dasar alasan tertentu dan menarik kesimpulan
- Koreksi dan penilaiannya mudah
- Objektif
- Dapat dipakai berulang-ulang

#### **4. Kelemahan tes pilihan ganda**

- Sulit serta membutuhkan waktu yang lama dalam menyusun soal
- Tidak dapat dipakai untuk mengukur kecakapan siswa dalam mengorganisasikan bahan.

#### **5. Petunjuk umum pernyataan tes pilihan ganda**

1. Stem hendaknya secara spesifik mengembangkan persoalan yang dimaksud untuk dikemukakan kepada siswa sehingga begitu membaca, siswa mudah memperoleh gambaran kemana arah persoalan
2. Kunci jawaban harus tidak bisa diperdebatkan lagi
3. Berlawanan adengan ketentuan kedua diatas, alternati-alternatif yang dicantumkan sebagai distractors tidak boleh terlalu jelas kelihatan” salahnya ”atau” dan ”ketidak-masuk-akalannya ” sehingga testee bisa menjawab cukup dengan mengeliminasi alternati – alternatif yang tidak masuk akal
4. Soal – soal manapun alternatif tidak boleh diambil secara kata demi kata dari buku sehingga kemungkinan jawaban siswa menjawab benar bukan karena ia menguasai bahannya akan tetapi karena bunyi kalimatnya yang sangat familiar
5. Untuk masing–masing soal hendaknya disediakan antara 4 sampai 5 options, kurang dari itu menyebabkan soal menjadi terlalu mudah, akan tetapi lebih dari itu menyebabkan sulit untuk dibuat perbedaan – perbedaan yang jelas diantara alternatif yang satu dengan alternatif yang lain
6. Setiap soal harus berdiri sendiri dalam arti bahwa jawaban terhadap soal yang satu tidak boleh memberi bantuan dalam menjawab soal yang lain
7. Harus dihindarkan soal–soal yang menuntut terlalu banyak detail, lebih–lebih apabila detail–detail tersebut merupakan fakta–fakta yang tidak terlalu penting maknanya selain menuntut hafalan–hafalan mekanis
8. Sebagaimana juga tes objektif yang lain, stem maupun *options* harus dirumuskan secara sederhana dan *to the point*.

## 6. Cara mengolah skor

Untuk mengolah skor dalam tes bentuk pilihan ganda ini digunakan 2 macam rumus, yaitu:

Dengan denda, rumusnya adalah

Keterangan:

S = skor yang diperoleh (*Raw Score*)

R = jawaban yang betul

W = jawaban yang salah

O = banyaknya option

1 = bilangan tetap

*Contoh:*

Murid menjawab betul 17 dari 20 soal. Soal bentuk multiple choice ini dengan menggunakan option sebanyak empat buah.

$$\text{Skor} = 17 - \frac{3}{4 - 1} = 16$$

Tanpa denda, dengan rumus:

$$S = R$$

## c. Tes Menjodohkan

### 1. Pengertian

Tes ini sebenarnya merupakan bentuk khusus dari tes pilihan berganda. Isi yang membedakan keduanya adalah bahwa dalam bentuk menjodohkan tidak hanya ada satu alternatif jawaban. Jawabannya harus dituliskan dalam satu kemungkinan jawaban. Secara nyata dalam tes bentuk ini disediakan dua kelompok bahan, dan siswa harus mencari pasangan/jodoh-jodoh yang sesuai antara bahan yang terdapat pada kelompok bahan yang terdapat pada kelompok pertama dan pada kelompok kedua.

## 2. Kelebihan tes menjodohkan

- Baik untuk mengukur proses mental yang rendah (knowledge)
- Kemungkinan untuk mengukur proses mental yang tinggi tetap ada tetapi sulit sekali
- Objektif
- Mudah disusun
- Cocok untuk mengukur informasi–informasi yang berbentuk fakta dari suatu pengertian, hubungan antar pengertian atau konsep-konsep

## 3. Kelemahan tes menjodohkan

Kelemahan dari soal tes bentuk ini adalah sukar untuk mengukur proses mental yang tinggi dan siswa cenderung untuk membuat penafsiran-penafsiran.

## 4. Petunjuk praktis penyusunan soal :

1. Buatlah pengantar sejelas mungkin
2. Bentuklah tes terdiri dari sederetan (satu seri) pertanyaan/persoalan dan sederetan jawaban
3. Hal-hal yang disusun baik dalam pertanyaan maupun jawaban hendaknya homogeny
4. Jumlah jawaban buatlah lebih banyak daripada jumlah pertanyaan
5. Batasi tiap kelompok jangan lebih dari 10 pertanyaan, jika ingin lebih banyak buatlah beberapa kelompok
6. Semua pertanyaan dan jawaban hendaknya dibuat pada hanya satu halaman saja
7. Setiap satu pertanyaan usahakan hanya ada satu jawabanyang benar
8. Tiap pertanyaan diberi waktu pengerjaan maksimum  $\frac{1}{2}$  menit
9. Buatlah kunci jawaban dan pedoman penilaiannya.

Contoh :

Pilihlah pasangan premis di bagian kiri dengan bagian kanan sehingga terjadi suatu pengertian yang utuh dan logis

1.  $F(x)$  stasioner di  $x = a$
2.  $F(x)$  naik di  $x = a$
3.  $F(x)$  merupakan maksimum di stasioner  $x = a$

4.  $F(x)$  merupakan titik belok horizontal di stasioner  $x = a$

- a.  $F'(a) > 0$
- b.  $F'(a) = 0$
- c.  $F'(a) < 0$
- d.  $F''(a) > 0$
- e.  $F''(a) = 0$
- f.  $F''(a) < 0$

#### **d. Tes Isian**

##### **1. Pengertian**

Tes isian adalah tes tertulis yang menuntut siswa untuk mengisi pernyataan, ungkapan atau kalimat pendek sebagai jawaban terhadap kalimat yang tidak lengkap atau jawaban atau suatu pertanyaan atau jawaban atas asosiasi yang harus dilakukan.

Sesuai dengan bentuknya, terdapat tiga jenis tes isian yaitu :

- Bentuk pertanyaan dengan satu jawaban  
Contoh : Berapa  $x$ , jika  $2x + 5 = 11$  ?
- Bentuk kalimat tidak lengkap  
Dimana siswa tinggal mengisi satu jawaban yang dibutuhkan.  
Contoh :  $127 \times 4 = \dots$
- Bentuk asosiasi  
Persoalan diajukan dalam bentuk pertanyaan dan kemudian diikuti (digabungkan) dengan kalimat–kalimat tidak lengkap dan siswa diminta untuk mengisi /melengkapi kalimat tersebut  
Contoh : Tulislah tempat angka-angka berikut:  
1. 2, 3, 5, . . .  
2. 2, ..., 5, ..., 7, . . .  
3. 5, ..., 15, ..., ..., 30

##### **2. Kebaikan tes isian**

- Mudah dalam penyusunannya, terutama untuk mengukur ingatan/pengetahuan
- Sedikit kesempatan untuk menduga–menduga jawaban
- Cocok untuk siswa kelas/tingkat rendah

### 3. Kelemahan tes isian

- Sukar untuk mengukur proses mental yang tinggi
- Sulit menyusun soal yang hanya satu jawaban, lebih-lebih untuk proses mental yang tinggi
- Sulit penilaiannya jikaterdapat bermacam-macam jawaban yang benar

### 4. Petunjuk praktis penulisan soal

1. Dalam membuat pertanyaan jangan terlalu banyak kata yang dihilangkan. Hendaknya pengertian yang penting saja, tetapi maksud dari kalimat tetap jelas dan mudah dipahami
2. Jawaban yang diinginkan hendaknya benar-benar dibatasi
3. Titik-titik (tempat siswa menulis jawaban) sebaiknya diletakkan pada ujung pernyataan
4. Jika masalah/persoalannya memerlukan jawaban yang berupa angka, nyatakanlah satuan-satuan tertentu dalam perhitungan itu
5. Berilah waktu maksimal 1 ½ menit untuk setiap nomor soal
6. Jangan mengambil alih soal langsung dari buku teks

Contoh :

- 1) Syarat apa yang harus dipenuhi agar persamaan kuadrat  $ax^2 + bx + c = 0$  mempunyai dua akar yang kembar?

### C. Penggunaan Tes Objektif

#### Tes objektif digunakan apabila:

- a. Kelompok yang akan dites banyak dan tes akan dilakukan berkali-kali.
- b. Skor yang diperoleh diperkirakan akan dapat dipercaya(mempunyai reabilitas tinggi).
- c. Guru lebih mampu menyusun tes objektif dari pada tes bentuk uraian.
- d. Hanya mempunyai waktu sedikit untuk koreksi dibandingkan dibandingkan waktu yang digunakan untuk menyusun tes. Pada umumnya, guru seyogyanya menggunakan dua macam bentuk tes ini dalam perbandingan 3 : 1, yaitu 3 bagian untuk tes objektif, dan 1 bagian untuk tes uraian.

### **Rangkuman**

- Tes objektif adalah tes yang dalam pemeriksaannya dapat dilakukan secara objektif.
- Tes objektif terdiri dari empat jenis yaitu :
  1. Tes benar salah ( *true false test* )
  2. Tes pilihan ganda ( *multiple choice test* )
  3. Tes menjodohkan ( *multipletest* )
  4. Tes isian atau tes melengkapi ( *completion test* ).
- Bentuk tes benar-salah ada 2 macam (dilihat dari segi mengerjakan/menjawab soal), yakni:
  1. Dengan pembetulan ( *with corection* ) yaitu siswa diminta membetulkan jika ia memilih jawaban yang salah.
  2. Tanpa pembetulan ( *without correction* ) yaitu siswa hanya diminta melingkari huruf B atau S tanpa memberikan jawaban yang betul.
- Jenis tes pilihan berganda
  1. Tes pilihan ganda dengan menemukan satu – satunya jawaban yang benar
  2. Tes pilihan ganda dengan memilih jawaban yang terbaik
  3. Tes pilihan ganda dengan memilih lebih dari satu jawaban benar
  4. Tes pilihan ganda yang menggunakan gambar, bagan, peta dan lain-lain.

## **BAB VII INSTRUMEN TES ESSAY**

### **A. Pengertian Tes Essay/Uraian**

Tes ini pada umumnya berbentuk essay (uraian). Tes bentuk essay adalah sejenis tes kemajuan belajar yang memerlukan jawaban yang bersifat pembahasan atau uraian kata-kata. Ciri-ciri pertanyaannya adalah dengan kata-kata seperti : uraikan, jelaskan, mengapa, bagaimana, bandingkan, simpulkan, dsb.

Soal-soal uraian ini menuntut kemampuan siswa untuk dapat mengorganisir, menginterpretasi, menghubungkan pengertian-pengertian yang telah dimiliki. Dengan singkat dapat dikatakan bahwa tes ini menuntut siswa untuk dapat mengingat-ingat dan mengenal kembali, dan terutama harus mempunyai daya kreatifitas tinggi. Soal uraian biasanya jumlahnya tidak banyak, hanya sekitar 5-10 buah soal dalam waktu kira-kira 90 sampai dengan 120 menit.

Pengertian tes uraian merupakan butir soal yang mengandung pertanyaan atau tugas yang jawaban atau pengerjaan soal tersebut harus dilakukan dengan cara mengekspresikan pikiran peserta tes secara naratif. Ciri khas tes uraian ialah jawaban terhadap soal tersebut tidak disediakan oleh orang yang mengkontruksi butir soal, tetapi dipasok oleh peserta tes. Peserta tes bebas untuk menjawab pertanyaan yang diajukan. Setiap peserta tes dapat memilih, menghubungkan, dan atau menyampaikan gagasan dengan menggunakan kata-katanya sendiri. Jadi perbedaan utama tes objektif dan uraian ialah siapa yang menyediakan jawaban atau alternatif jawaban sudah disediakan oleh pembuat soal. Dengan pengertian diatas maka pemberian skor terhadap soal uraian tidak mungkin dilakukan secara objektif. Setiap bentuk butir soal memiliki kelebihan dan kekurangan.

Adapun kelebihan soal uraian adalah :

1. Tes uraian dapat dengan baik mengukur hasil belajar yang kompleks. Hasil belajar yang kompleks artinya hasil belajar yang tidak sederhana. Hasil belajar yang kompleks tidak hanya membedakan yang benar dari yang salah, tetapi juga dapat mengekspresikan pemikiran peserta tes serta pemilihan kata yang dapat memberi arti yang spesifik pada suatu pemahaman tertentu. Apabila yang diukur adalah kemampuan hasil belajar yang sederhana, yaitu memilih suatu

yang lebih benar atau yang lebih tepat, maka sebaiknya menggunakan tes objektif.

2. Tes bentuk uraian terutama menekankan kepada pengukuran kemampuan dan kemampuan mengintegrasikan berbagai buah pikiran dan sumber informasi kedalam suatu pola berpikir tertentu, yang disertai dengan keterampilan pemecahan masalah. Integrasi buah pikiran itu membutuhkan dukungan kemampuan untuk mengekspresikannya. Tanpa dukungan kemampuan mengekspresikan buah pikiran secara teratur dan taat asas, maka kemampuan tidak terlihat secara utuh. Bahkan kemampuan itu secara sederhana sudah akan dapat kelihatan dengan jelas dalam pemilihan kata, penyusunan kalimat, penggunaan tanda baca, penyusunan paragraf dan susunan rangkain paragraf dalam suatu keutuhan pikiran.
3. Bentuk tes uraian lebih meningkatkan motivasi peserta didik untuk melahirkan kepribadiannya dan watak sendiri, sesuai dengan sifat tes uraian yang menuntut kemampuan siswa untuk mengekspresikan jawaban dalam kata-kata sendiri. Untuk dapat mengekspresikan pemahaman dan penguasaan bahan dalam jawaban tes, maka bentuk tes uraian menuntut penguasaan bahan secara utuh. Penguasaan bahan yang tanggung atau parsial dapat dideteksi dengan mudah. Karena itu untuk menjawab tes uraian dengan baik peserta tes akan berusaha menguasai bahan yang diperkirakan akan diujikan dalam tes secara tuntas. Seorang peserta tes yang mengerjakan tes uraian dengan penguasaan bahan parsial akan tidak mampu menjawab soal dengan benar atau akan berusaha dengan cara membual.
4. Kelebihan lain tes uraian ialah memudahkan guru untuk menyusun butir soal. Kemudahan ini terutama disebabkan oleh dua hal, yaitu pertama, jumlah butir soal tidak perlu banyak dan kedua, guru tidak selalu harus memasok jawaban atau kemungkinan jawaban yang benar sehingga akan sangat menghemat waktu konstruksi soal. Tetapi hal ini tidak berarti butir soal uraian dapat dikonstruksikan secara asal-asalan. Kaidah penyusunan tes uraian tidaklah lebih sederhana dari kaidah penyusunan tes objektif.
5. Tes uraian sangat menekankan kemampuan menulis. Hal ini merupakan kebaikan sekaligus kelemahannya. Dalam arti yang positif tes uraian akan sangat mendorong siswa dan guru untuk belajar dan mengajar, serta menyatakan pikiran secara tertulis. Dengan demikian diharapkan kemampuan para peserta didik dalam menyatakan pikiran

secara tertulis akan meningkat. Tetapi dilihat dari segi lain, penekanan yang berlebihan terhadap penggunaan tes uraian yang sangat menekankan kepada kemampuan menyatakan pikiran dalam bentuk tulisan yang dapat menjadikan tes sebagai alat ukur yang tidak adil dan tidak reliable. Bagi siswa yang tidak mempunyai kemampuan menulis, akan menjadi beban.

Di samping memiliki kelebihan, tes butir uraian terdapat pula kelemahan-kelemahannya, yaitu :

1. Reliabilitasnya rendah artinya skor yang dicapai oleh peserta tes tidak konsisten bila tes yang sama atau tes yang paralel yang diuji ulang beberapa kali. Menurut Robert L. Ebel A. Frisbie (1986 : 129) terdapat tiga hal yang menyebabkan tes uraian realibilitasnya rendah yaitu pertama keterbatasan sampel bahan yang tercakup dalam soal tes. Kedua, batas-batas tugas yang harus dikerjakan oleh peserta tes sangat longgar, walaupun telah diusahakan untuk menentukan batasan-batasan yang cukup ketat. Ketiga, subjektifitas penskoran yang dilakukan oleh pemeriksa tes.
2. Untuk menyelesaikan tes uraian guru dan siswa membutuhkan waktu yang banyak.
3. Jawaban peserta tes kadang-kadang disertai bualan-bualan.
4. Kemampuan menyatakan pikiran secara tertulis menjadi hal yang paling membedakan prestasi belajar siswa.

Setelah Anda memahami kelemahan dan kelebihan bentuk tes uraian. Anda harus mempertimbangkan bagaimana tes uraian digunakan.

Sebaiknya tes uraian digunakan apabila :

1. Jumlah siswa atau peserta tes terbatas.
2. Waktu yang dipunyai guru untuk mempersiapkan soal sangat terbatas.
3. Tujuan instruksional yang ingin dicapai adalah kemampuan mengekspresikan pikiran dalam bentuk tertulis, menguji kemampuan dengan baik, atau penggunaan kemampuan penggunaan bahasa secara tertib.
4. Guru ingin memperoleh informasi yang tidak tertulis secara langsung di dalam soal ujian tetapi dapat disimpulkan sari tulisan peserta tes, seperti : sikap, nilai, atau pendapat. Soal uraian dapat digunakan untuk

memperoleh informasi langsung tersebut, tetapi harus digunakan dengan sangat hati-hati oleh guru.

5. Guru ingin memperoleh hasil pengalaman belajar siswanya.

Bentuk tes uraian dapat diklasifikasi ke dalam dua tipe yaitu tes uraian bebas (*extended response*) dan tes uraian terbatas (*restricted response*). Perbedaan kedua tipe tes uraian ini adalah atas dasar besarnya kebebasan yang diberikan kepada peserta tes untuk mengorganisasikan, menulis dan menyatakan pikiran dan gagasannya. Selanjutnya akan dijelaskan kedua tipe tes uraian tersebut. Sebagaimana telah dikemukakan, perbedaan utama antara tes uraian bebas dan uraian terbatas tergantung kepada kebebasan memberikan jawaban. Jawaban yang diberikan oleh peserta tes dalam tes uraian bebas hampir-hampir tidak ada pembatasan. Peserta tes memiliki kebebasan yang luas sekali untuk mengorganisasikan dan mengekspresikan pikiran dan gagasannya dalam menjawab soal tersebut. Jadi jawaban siswa bersifat terbuka, fleksibel, dan tidak terstruktur. Contoh uraian bebas :

Uraikanlah perlawanan-perlawanan yang dilakukan oleh bangsa Indonesia terhadap penjajah Belanda pada masa sebelum abad ke-20. Dalam uraian ini anda hendaknya menyertai dengan latar belakang dan contoh para pemimpinnya. Uraian anda hendaknya tidak melebihi 2 halaman folio.

Untuk menjawab contoh soal tersebut dibutuhkan kemampuan belajar siswa yang kompleks. Dalam menjawab soal tersebut siswa diberikan kebebasan untuk menjawab. Jawaban yang diberikan oleh siswa mulai pengetahuan fakta sampai mengevaluasi terhadap fakta-fakta yang diketahuinya, kemudian mengorganisasikan dalam pikiran dan bahasanya sendiri kedalam bentuk yang logis dan argumentatif dalam bentuk narasi. Kemampuan membuat narasi dengan kata-katanya sendiri merupakan kemampuan dalam kategori jenjang yang tinggi. Dengan demikian kompleksitas jawaban pada soal uraian bebas terletak pada tercakupnya berbagai jenjang kemampuan. Pembatasan jawaban hanya terletak pada banyaknya uraian yang harus dibuat untuk mempertimbangkan waktu yang digunakan dalam tes.

## **B. Petunjuk Penyusunan Tes Essay/Uraian**

Dalam penyusunan tes subjektif, maka harus diperhatikan beberapa hal berikut.

1. Hendaknya soal- soal tes dapat meliputi ide-ide pokok dari bahan yang di teskan, dan kalau mungkin disusun soal yang sifatnya komprehensif.
2. Hendaknya soal tidak mengambil kalimat-kalimat yang disalin langsung dari buku atau catatan.
3. Pada waktu menyusun, soal-soal itu sudah dilengkapi dengan kunci jawaban serta pedoman penilaiannya.
4. Hendaknya diusahakan agar pertanyaannya bervariasi antara "jelaskan", "bagaimana", "mengapa", "seberapa jauh", agar dapat diketahui lebih jauh penguasaan siswa terhadap bahan.
5. Hendaknya rumusan soal dibuat sedemikian rupa sehingga mudah dipahami oleh terdoba.
6. Hendaknya ditegaskan model jawaban apa yang dikehendaki oleh penyusun tes. Untuk ini pertanyaan tidak boleh terlalu umum, tapi harus spesifik.

## **C. Penggunaan Tes Uraian**

Tes bentuk uraian digunakan apabila :

1. Kelompok yang akan dites kecil, dan tes itu tidak akan dilakukan berulang-ulang.
2. Tester(guru) ingin menggunakan berbagai cara untuk mengetahui kemampuan siswa dalam bentuk tertulis.
3. Guru ingin mengetahui lebih banyak tentang sikap-sikap siswa dari pada hasil yang telah dicapai.
4. Memiliki waktu yang cukup untuk menyusun tes.

## **D. Contoh Soal Tes Uraian**

Selesaikan Sistem Persamaan Linear berikut dengan cara eliminasi :

$$2x + 3y = 12$$

$$3x + 5y = 15$$

## **E. Kaidah Pengembangan Instrumen Tes Essay/Uraian**

### **1. Pengembangan Butir Tes Uraian**

- Sebelum menuliskan butir pertanyaan ditentukan dulu tingkat proses berfikir yang harus digunakan siswa untuk menjawab pertanyaan tersebut.
- Tes uraian hendaknya mengungkapkan permasalahan yang menyangkut proses berfikir dan pemahaman karena kedua aspek ini lebih sesuai untuk pertanyaan obyektif.
- Pertanyaan biasa menggunakan kata : perbandingan, pertentangan, berikan alasan, jelaskan bagaimana, apakah yang akan terjadi jika, berikan kritik, bedakan, dan lain-lain.
- Buatlah pertanyaan uraian yang jelas rumusannya sehingga tidak meragukan siswa.
- Usahakan agar pertanyaan uraian mengungkapkan pendapat bukan untuk menyampaikan fakta.

### **2. Pedoman Penilaian**

Butir soal bentuk obyektif dapat diperiksa dengan mudah, cepat dan hasil penilaiannya obyektif, untuk mendapatkan penilaian yang lebih obyektif perlu diperhatikan:

- Apakah jawaban yang paling baik untuk satu butir pertanyaan uraian.
- Butir-butir apa saja yang harus terdapat dalam jawaban pertanyaan uraian.
- Apakah ada butir yang lebih penting diantara butir-butir jawaban yang diharapkan.

## **Rangkuman**

- Tes Essay/Uraian adalah sejenis tes kemajuan belajar yang memerlukan jawaban yang bersifat pembahasan atau uraian kata-kata. Ciri-ciri pertanyaannya adalah dengan kata-kata seperti; uraikan, jelaskan, mengapa, bagaimana, bandingkan, simpulkan, dan sebagainya.
- Pengertian tes uraian merupakan butiran soal yang mengandung pertanyaan atau tugas yang jawaban atau pengerjaan soal tersebut harus dilakukan dengan cara mengekspresikan pikiran peserta tes secara naratif.
- Ciri khas tes uraian ialah jawaban terhadap soal tersebut tidak disediakan oleh orang yang mengkonstruksi butir soal, tetapi dipasok oleh peserta tes.
- Tes bentuk uraian digunakan apabila :
  - Kelompok yang akan dites kecil, dan tes itu tidak akan dilakukan berulang-ulang.
  - Tester(guru) ingin menggunakan berbagai cara untuk mengetahui kemampuan siswa dalam bentuk tertulis.
  - Guru ingin mengetahui lebih banyak tentang sikap-sikap siswa dari pada hasil yang telah dicapai.
  - Memiliki waktu yang cukup untuk menyusun tes.

## **BAB VIII**

### **ASSESMEN HASIL BELAJAR**

Assesmen merupakan proses mendokumentasikan pengetahuan, keterampilan, sikap dan keyakinan peserta didik untuk memperoleh informasi tentang apa yang diketahui, dilakukan dan dikerjakan peserta didik. Metode dan alat assesmen meliputi:

1. Observasi
2. Proyek
3. Laporan Tertulis
4. Review Kinerja
5. Assesmen Mandiri oleh Peserta Didik
6. Skala Penilaian
7. Tes Tertulis
8. Assesmen Portofolio

Evaluasi memiliki kesamaan dengan assesmen, dan kadang-kadang kedua istilah itu digunakan secara bergantian. Isi evaluasi dipandang lebih luas dibandingkan dengan assesmen. Assesmen dipandang sebagai proses pengukuran terhadap suatu karakteristik tertentu, seperti deskripsi tujuan. Sedangkan evaluasi dipandang sebagai proses pengukuran terhadap suatu karakteristik dan penentuan nilai atau harga suatu obyek.

Kegiatan assesmen pertama kali muncul di China pada tahun 206 SM ketika dinasti Han memperkenalkan ujian untuk membantu proses seleksi pegawai kerajaan.

Jenis-jenis assesmen, yaitu:

#### **A. Assesmen Formatif dan Sumatif**

- |                     |   |
|---------------------|---|
| Assesmen sumatif →  | Dilaksanakan di akhir pembelajaran dan digunakan untuk membuat keputusan tentang kenaikan kelas peserta didik.  |
| Assesmen formatif → | Dilaksanakan selama proses pembelajaran berlangsung. Wujudnya berupa pemberian balikan atas pekerjaan peserta didik dan tidak dijadikan dasar penentuan kenaikan kelas. |

Dalam konteks belajar assesmen sumatif dan normatif disebut dengan assesmen belajar.

## **B. Assesmen Obyektif dan Subyektif**

Assesmen obyektif → Bentuk pertanyaan yang memiliki satu jawaban benar.

Assesmen subyektif → Bentuk pertanyaan yang memiliki lebih dari satu jawaban benar.

## **C. Assesmen Acuan Patokan dan Acuan Normatif**

Assesmen acuan patokan → Assesmen yang digunakan untuk mengukur kemampuan peserta didik berdasarkan kriteria yang telah ditetapkan sebelumnya menggunakan tes acuan patokan.

Assesmen acuan normatif → Assesmen yang menggunakan tes acuan normatif dan tidak digunakan untuk mengukur kemampuan peserta didik berdasarkan kriteria yang telah ditetapkan sebelumnya. Assesmen ini dikenal dengan penentuan rangking berdasarkan kurve normal.

## **D. Assesmen Formal dan Informal**

Assesmen formal → Diwujudkan dalam bentuk dokumen tertulis dan diberikan skor dalam bentuk angka atau penentuan rangking berdasarkan kinerja peserta didik.

Assesmen informal → Dilakukan dengan cara yang lebih terbuka seperti observasi, inventori, diskusi yang tidak dimaksudkan untuk menentukan rangking.

## **E. Assesmen Autentik (Assesmen Kinerja)**

Assesmen berbasis kinerja merupakan bentuk ujian di mana peserta didik menjawab suatu pertanyaan atau membuat produk dan mendemonstrasikan ketrampilan atau menampilkan kemampuan/ pengetahuan. Wujudnya antara lain :



- Tugas membuat proyek secara individual atau kelompok : eksperimen ilmiah
- Contoh tulisan atau karangan : portofolio
- Memecahkan masalah terbuka : simulasi komputer
- Pertanyaan yang membutuhkan konstruksi jawaban : wawancara atau presentasi.

**Tahap-tahap Assesmen Kinerja adalah :**

- Mengidentifikasi hasil pembelajaran.
- Mengembangkan tugas-tugas untuk menemukan tujuan pembelajaran.
- Mengidentifikasi hasil belajar tambahan yang di dukung oleh tugas.
- Merumuskan kriteria dan tingkat kinerja untuk mengevaluasi kinerja peserta didik.

**F. Assesmen Portofolio**

Assesmen portofolio merupakan bentuk evaluasi kinerja yang paling populer. Biasanya berbentuk file atau folder yang berisi koleksi karya peserta didik.

Tahap-tahap assesmen portofolio adalah :

**a. Perencanaan dan Pengorganisasian**

- Mengembangkan perencanaan portofolio yang bersifat fleksibel.
- Merencanakan waktu secukupnya agar peserta didik mempersiapkan dan mendiskusikan aspek-aspek portofolio.
- Dimulai dengan satu aspek belajar dan hasil belajar peserta didik, kemudian semakin meningkat sejalan dengan apa yang dipelajari peserta didik.
- Memilih aspek yang dimasukkan di dalam portofolio yang mampu menunjukkan kemajuan peserta didik atau penguasaan tujuan pembelajaran.
- Memilih setidaknya dua aspek, yakni indikator yang diperlukan atau aspek-aspek inti dan sampel pekerjaan yang dipilih.
- Menempatkan daftar tujuan di depan masing-masing portofolio. Bersamaan dengan indikator yang dipersyaratkan dan tempat mencatat aspek-aspek pilihan.

## **b. Implementasi**

- Melekatkan perkembangan aspek-aspek portofolio di dalam kegiatan kelas yang sedang berlangsung.
- Memberikan tanggung jawab kepada peserta didik untuk mempersiapkan, memilih, menilai dan menyimpan portofolionya sendiri.
- Membagi aspek-aspek portofolio yang telah dipilih.
- Mencatat komentar pendidik dan peserta didik dengan segera terhadap portofolio tersebut.

## **c. Hasil**

- Menganalisis aspek-aspek portofolio untuk memahami pengetahuan dan keterampilan peserta didik.
- Menggunakan informasi portofolio itu untuk mendokumentasi kegiatan-kegiatan belajar peserta didik, untuk disampaikan kepada orang tua dan memperbaiki pembelajaran di kelas.

## **d. Prinsip-prinsip Assesmen**

1. Tujuan utama assesmen adalah memperbaiki belajar peserta didik.
2. Assesmen bertujuan untuk mendukung belajar peserta didik.
3. Obyektif bagi semua peserta didik
4. Berkolaborasi secara profesional dengan sekelompok pendidik lain.
5. Melibatkan partisipasi komite sekolah dalam pengembangan assesmen.
6. Menjelaskan keteraturan dan kejelasan komunikasi mengenai peserta didik kepada peserta didik, keluarga dan masyarakat.
7. Meninjau dan memperbaiki assesmen.

## **e. Pentingnya Assesmen Perkembangan dan Belajar Siswa**

Assesmen merupakan sebuah proses pengumpulan informasi yang terus menerus berlangsung untuk mengukur performansi murid dan proses pembelajaran. Assesmen perkembangan dan belajar anak memiliki nilai penting. Tidak hanya mengukur kemajuan anak-anak sebagai bentuk evaluasi program, asesmen juga berguna untuk mengidentifikasi

kebutuhan-kebutuhan pengembangan staf dan perencanaan pembelajaran di masa yang akan datang.

Assesmen yang tepat berguna untuk membantu anak-anak berkembang secara optimal, baik fisik, sosial, emosional, intelektual maupun spiritual. Assesmen yang tepat juga dapat digunakan untuk mendeteksi keterlambatan-keterlambatan perkembangan atau kebutuhan-kebutuhan khusus yang mungkin dimiliki anak-anak. Selain itu informasi yang akurat dari sebuah assesmen bermanfaat untuk peningkatan pembelajaran sehingga proses belajar anak-anak membaik dan sebagai informasi bagi para orangtua tentang kemajuan dan hal-hal terkait dengan belajar anak-anak mereka.

Assesmen yang tepat merupakan bagian penting dari program evaluasi dan perbaikan terus menerus kualitas program pendidikan yang sudah dirancang. Dalam program pendidikan yang berkualitas, pihak-pihak terkait dengan pendidikan anak menggunakan informasi dari berbagai macam sumber untuk merencanakan dan membuat keputusan-keputusan tentang anak-anak secara individual.

Prosedur-prosedur dan instrumen-instrumen (alat) assesmen—seperti test, observasi, portofolio, penilaian guru, penilaian orangtua, dan lain sebagainya—dikatakan efektif ketika mereka memenuhi standar validitas (tepat dan akurat), reliabilitas (keajegan), dan kepekaan terhadap isu-isu kultural. Instrumen assesmen yang tepat memungkinkan jawaban-jawaban yang khas dari anak-anak menurut kelompok usia, jenis kelamin, latar belakang budaya, dan kondisi geografis. Anak-anak harus diukur secara individual oleh orang-orang yang mengetahui mereka dengan objektif dalam setting dan situasi-situasi yang mencerminkan penampilan mereka yang sesungguhnya. Semakin muda usia anak, maka akan semakin sulit untuk mendapatkan assesmen yang valid. Perkembangan anak-anak usia dini berlangsung sangat cepat dan sangat dipengaruhi oleh pengalaman. Performansi mereka dalam tes dipengaruhi oleh kondisi-kondisi emosional anak dan kondisi-kondisi saat asesmen dilakukan.

Satu hal yang pasti dalam melakukan asesmen adalah jangan pernah menggunakan satu instrumen asesmen untuk membuat keputusan-keputusan yang memiliki konsekuensi penting bagi anak-anak. Setiap asesmen seharusnya dipilih untuk memenuhi tujuan-tujuan yang spesifik. Assesmen seharusnya diselaraskan dengan kurikulum dan proses pembelajaran di kelas.

## **Rangkuman**

- Assesmen merupakan proses mendokumentasikan pengetahuan, keterampilan, sikap dan keyakinan peserta didik untuk memperoleh informasi tentang apa yang diketahui, dilakukan dan dikerjakan peserta didik.
- Metode dan alat assesmen meliputi : Observasi, Proyek, Laporan Tertulis , Review Kinerja, Assesmen Mandiri oleh Peserta Didik, Skala Penilaian, Tes Tertulis, Assesmen Portofolio.
- Jenis-jenis assesmen : Assesmen Obyektif, Assesmen Subyektif, Assesmen Acuan Patokan , Assesmen Acuan Normatif, Assesmen Formal, Assesmen Informal, Assesmen Autentik (Assesmen Kinerja) dan Assesmen Portofolio.
- Prinsip-prinsip Assesmen
  - Memperbaiki belajar peserta didik.
  - Mendukung belajar peserta didik.
  - Obyektif bagi semua peserta didik
  - Berkolaborasi secara profesional dengan sekelompok pendidik lain.
  - Melibatkan partisipasi komite sekolah dalam pengembangan asesmen.
  - Menjelaskan keteraturan dan kejelasan komunikasi mengenai peserta didik kepada peserta didik, keluarga dan masyarakat.
  - Meninjau dan memperbaiki asesmen.

## **BAB IX**

### **JENIS-JENIS ASSESMEN**

#### **A. Pendahuluan**

Assesmen merupakan kegiatan harian guru, yang ia laksanakan setiap hari, hari demi hari selama proses pembelajaran berlangsung. Tidak ada bentuk tanggungjawab guru lainnya yang lebih penting dibandingkan dengan melakukan assesmen pada murid-muridnya. Guru harus dapat mengkomunikasikan performa akademik dan performa sosial siswa serta kemajuan atau pertumbuhannya kepada berbagai pihak yang terkait meliputi siswa, orang tua siswa, sekolah dan administrator pendidikan, serta masyarakat umum.

Assesmen dan keputusan-keputusan evaluatif yang dihasilkannya haruslah akurat sehingga mampu mencegah pemahaman dan komunikasi yang keliru oleh pihak-pihak terkait. Oleh karenanya diperlukan berbagai jenis assesmen yang secara bersama-sama akan menghasilkan informasi evaluatif yang lengkap dan akurat.

Metode assesmen formal direncanakan lebih bagus dalam pengadministrasiannya. Metode ini kurang spontanitasnya dan biasanya dilaksanakan pada akhir proses pembelajaran. Para siswa menyadari atau mengetahui tentang penggunaan metode assesmen formal ini. Contoh metode ini diantaranya adalah tes meliputi beberapa bab, ujian final, PR terstruktur dan sebagainya. Metode assesmen informal dilaksanakan lebih spontan dan kurang kentara/terlihat. Biasanya terjadi selama proses pembelajaran. Contoh metode ini seperti: observasi dan pertanyaan-pertanyaan yang dilakukan oleh guru selama proses pembelajaran. Metode assesmen kuantitatif menghasilkan skor berupa angka atau numeris. Secara umum bentuknya meliputi tes yang disusun oleh guru, tes standar, cek list, dan skala rating. Metode assesmen kualitatif menghasilkan deskripsi secara verbal tentang karakteristik tertentu. Biasanya berbentuk observasi oleh guru, rekaman anecdotal, pertanyaan-pertanyaan informal.

Evaluasi formatif berkaitan dengan pembuatan keputusan yang terjadi selama proses pembelajaran yang bertujuan untuk menyelaraskan proses pembelajaran. Metode ini lebih banyak mengevaluasi pengajaran ketimbang mengevaluasi kerja siswa. Sangat mungkin didasarkan pada metode formal maupun informal. Evaluasi sumatif dilaksanakan pada akhir proses pembelajaran. Contohnya seperti tes pada akhir bab, tes pada

akhir satu unit pelajaran, tes akhir semester dan sebagainya. Evaluasi sumatif merupakan bentuk yang sering digunakan untuk keperluan mengambil keputusan secara administratif ( misalnya kenaikan kelas). Evaluasi sumatif semata-mata hanya didasarkan pada metode asesmen formatif.

Metode asesmen baku diadministrasikan, diskor, dan diinterpretasikan dalam model yang identik untuk semua ujian. Asesmen baku bertujuan untuk memberi peluang pada pendidik untuk membandingkan siswa dari sekolah yang berbeda, negara yang berbeda dan seterusnya. Contoh asesmen baku adalah UAN, UNAS, SPMB dan sebagainya.

Metode asesmen tak baku biasanya dalam bentuk buatan guru untuk digunakan di kelas. Asesmen tak baku memiliki tujuan untuk menentukan luasan subyek materi yang sedang diajarkan dan dipelajari. Metode asesmen rujukan norma menunjukkan dimana letak atau posisi performa seorang siswa dalam kaitannya dengan siswa-siswa lainnya. Asesmen baku atau tes baku biasanya adalah asesmen rujukan norma. Metode asesmen rujukan norma menghasilkan informasi kuantitatif. Dalam metode ini performa siswa dibandingkan dengan kelompok normal (*norm group*). Metode asesmen rujukan kriteria membandingkan performa siswa dengan kriteria atau objektif yang telah ditetapkan. Metode ini menghasilkan informasi kuantitatif, kualitatif atau keduanya. Metode asesmen rujukan kriteria biasanya juga dikenal dengan ketuntasan (*mastery*), atau *objectives-referenced*, atau ujian kompetensi (*competency tests*).

Metode asesmen tes merupakan prosedur *pencil-and-paper tests* dan kuis. Dalam metode ini hanya ada satu respon/jawaban yang benar untuk setiap butir pertanyaan/pernyataan. Metode ini dinilai merupakan metode yang mudah dan efisien untuk menilai/meng-ases banyak siswa secara simultan. Lebih mengarahkan/menganjurkan pada memorisasi fakta-fakta.

Metode asesmen alternatif merupakan metode asesmen yang lebih cocok untuk proses pembelajaran yang “melakukan” (*hands-on*), pembelajaran eksperimental (pengalaman belajar). Metode ini meliputi asesmen autentik (mencakup aplikasi riil dari ketrampilan, yang melampaui konteks pembelajaran).

Pada metode asesmen obyektif, kata “obyektif” disini lebih merujuk pada metode penskoran. Pada metode ini hanya mengandung

satu jawaban/respon yang benar untuk setiap butir pertanyaan/ Pernyataan. Contohnya adalah : pilihan ganda, benar-salah, mencocokkan dan sebagainya. Metode ini juga dikenal sebagai *structured-response*, atau *selected-response*, atau *teacher-supplied items*. Metode asesmen subyektif adalah metode asesmen yang teknik penskorannya melibatkan keputusan subyektif guru. Ada beberapa jawaban/respon yang benar atau satu jawaban/respon yang benar dapat dicapai dengan berbagai cara. Contohnya adalah : butir jawab singkat dan butir essay. Asesmen subyektif juga dikenal sebagai *open-ended*, *constructed-response*, *supply-type items*.

## **B. Asesmen Formal Dan Informal**

Biasanya asesmen informal meliputi bentuk-bentuk seperti observasi oleh guru, pertanyaan guru selama proses pembelajaran, dan refleksi siswa. Kebanyakan asesmen yang digunakan di kelas adalah asesmen informal. Asesmen ini terjadi selama proses pembelajaran berlangsung, dan dilakukan secara berkelanjutan.

### **1. Observasi oleh guru**

Observasi oleh guru merupakan kegiatan memperhatikan/melihat dan/atau mendengarkan siswa seperti : bagaimana ia melakukan kegiatan belajar, atau bagaimana ia membuat hasil produk belajar dan sebagainya. Observasi oleh guru bertujuan untuk merekam dan mendeskripsikan perilaku siswa seperti apa adanya. Dapat memberikan informasi yang mengarah pada :

- 1) Kualitas performa siswa,
- 2) Proses dan prosedur siswa dalam melengkapi atau mengerjakan tugas belajarnya,
- 3) Proses dan prosedur yang guru gunakan dalam melaksanakan pembelajaran.

#### **Karakteristiknya antara lain :**

1. Banyak kejadian yang diobservasi secara bersamaan atau berhasil dalam waktu yang cepat,
2. Observasi kelas tentu saja fokus pada satu kejadian,
3. Guru harus bergantung pada observasi siswa,
4. Sangat banyak kejadian yang ada di kelas yang terlewatkan tidak dicatat,

5. Observasi cepat dilupakan atau terkontaminasi tatkala dipanggil ulang
6. Observasi tentu saja perlu inferensi.

**Petunjuk penggunaannya :**

- 1) Ketahuilah apa yang akan diobservasi,
- 2) Ketahuilah keterbatasan tempat dan berapa banyak yang sedang diobservasi,
- 3) Akrablah dengan apa yang diobservasi,
- 4) Hindari perluasan inferensi, perhatikan kejadian yang substansial,
- 5) Sadari bahwa observasi bisa menjadi overestimate achievement,
- 6) Observasi dokumen yang harus dipanggil ulang di waktu kemudian

**Penyimpanan Rekaman:**

**1. Rekaman anekdotal:**

Narasi pendek yang mendeskripsikan konteks dan perilaku, digunakan untuk dokumen perilaku sebagai referensi di waktu kemudian oleh guru atau yang lainnya, boleh jadi meliputi interpretasi, bias harus dihindari dalam beberapa interpretasi, hanya rekaman observasi yang sangat penting dan tidak dapat ditemui melalui metode assesmen kelas yang lebih formal.

**a) Checklist :**

Daftar perilaku atau *outcomes*, guru yang mengindikasikan dengan mudah untuk diobservasi, dibatasi untuk situasi yang menghadirkan atau tak menghadirkan suatu kondisi untuk ditentukan, lebih terstruktur ketimbang rekaman anecdotal. Hanya ada dua pilihan yang mungkin untuk observasi yaitu : diobservasi atau tidak diobservasi.

**b) Rating Scales :**

Merupakan penyederhanaan dari checklists; balikan yang jauh lebih spesifik (sepanjang sebuah kontinum). Dapat mengidikasikan frekuensi atau derajat tentang bagaimana siswa mempertunjukkan sebuah karakteristik. Dapat digunakan secara formatif maupun sumatif. Biasanya juga dirujuk sebagai rubrik (*holistic dan analytic*).

**c) Pertanyaan guru :**

Pertanyaan guru : informal, tak terencana (*unplanned*), oral spontan secara ikhri yang diajukan oleh guru kepada siswa, berguna sebagai sebuah monitoring terhadap pemahaman siswa sepanjang proses pembelajaran. Dapat diajukan *lower-questions* dan *higher-questions*, juga dapat digunakan sebagai satu bentuk refleksi diri (*self-reflection*) dari siswa.

**Karakteristik pertanyaan guru antara lain :**

- 1) Dapat menjadi menonjolkan diri,
- 2) Harus diinterpretasikan oleh guru lain, harus menjadi jelas,
- 3) Dapat diarahkan pada individu-individu, pada kelompok-kelompok kecil, atau pada seluruh kelas,
- 4) Pertanyaan-pertanyaan detil ditanyakan dan bagian responnya dengan cepat dilupakan.

**Petunjuk penggunaan :**

- 1) Kembangkan pertanyaan dari objektif pembelajaran,
- 2) Berikan sebuah masalah yang jelas untuk siswa,
- 3) Berikan waktu yang cukup bagi siswa untuk merespon,
- 4) Hindari kebingungan siswa,
- 5) Tunjukkan perhatian ketika menanggapi respon siswa.

Rekaman penyimpanan; masih sedikit teknik yang ada untuk keperluan ini, guru dapat menyesuaikan bentuk variasi checklist untuk mendokumentasikan partisipasi dan respon siswa.

**d) Refleksi siswa :**

Refleksi siswa merupakan narasi ikhtisar atau laporan diri yang ditulis oleh siswa tentang subyek materi yang dipelajarinya. *Brief narratives or self-reports written by siswa concerning the subject matter being studied*. Dapat berupa jurnal belajar atau buku catatan belajar. Dilengkapi secara periodik oleh siswa unit demi unit. Dapat juga berupa ringkasan materi, pertanyaan-pertanyaan yang diajukan sepanjang atau selama kelas berlangsung, karakteristik dari sebuah proyek dsb.

Kesempatan diberikan kepada guru dan siswa untuk mendiskusikan komentar-komentar dan pertanyaan-pertanyaan. Yang digunakan

secara luas dengan berbagai variasi adalah bentuk *one-minute paper*. Para siswa diberikan kesempatan beberapa menit dari acara di kelas untuk merespon (dalam bentuk tulisan dan tanpa nama) satu atau dua pertanyaan yang diberikan secara spesifik oleh guru.

1. Hal terpenting apakah yang kamu pelajari di kelas sepanjang hari ini ?
2. Pertanyaan apakah yang dapat kamu ajukan setelah kamu mengikuti pembelajaran di kelas hari ini ?
3. Saya akan mencari tahu lebih jauh tentang .....
4. Saya masih tidak percaya tentang ...
5. Respon siswa dikumpulkan dan disintesis oleh guru. Guru memulai kelas berikutnya dengan komentar hasil sintesis yang telah dilakukannya. Guru menekankan pada tanggungjawab siswa untuk didengar dan diproses.

#### e) **Validitas and Reliabilitas**

##### **Asesmen informal**

Tentu saja perlu bersabar untuk mereduksi validitas dan reliabilitas yang diakibatkan oleh subyektivitas alamiah. Validitas dan Reliabilitas assesmen informal dapat ditingkatkan dengan cara mencegah munculnya masalah-masalah yang umum.

Validitas, dapat dipengaruhi oleh pengambilan keputusan atau pengantisipasi perilaku siswa secara merugikan. Ketidakcocokan indikator-indikator dari karakteristik siswa terkadang dapat dipilih. Reliabilitas, dipengaruhi oleh kurang-mencukupinya pengambilan sample perilaku siswa. Gambaran inferensi dalam satu setting mungkin tidak diperluas pada setting yang lainnya.

##### **Keuntungan dan Keterbatasan Asesmen Informal**

Keuntungannya, adalah efisien dan dapat diadaptasikan, dapat dibangun ke dalam arus pelajaran (bukan interupsi), dan dapat digunakan untuk memonitor kegiatan belajar dan pembelajaran. Keterbatasannya, adalah observasi dibatasi hanya untuk perilaku yang terjadi secara alamiah. Guru mengamati hanya bagian-bagian dari perilaku siswa. Beberapa observasi informal berlangsung tetapi tidak didokumentasikan dan oleh karenanya mudah dilupakan.

### **C. Assesmen Berbasis Performa**

Assesmen berbasis performa (juga dikenal sebagai assesmen performa) mempersyaratkan siswa untuk menerapkan pengetahuan dan ketrampilannya. Assesmen performa dapat digunakan secara formatif atau secara sumatif. Assesmen ini dapat menjadi *labor- and time-intensive*, dan juga cenderung menjadi berbeda sama sekali.

#### **a. Karakteristik Assesmen Berbasis Performa**

Assesmen performa menghadirkan siswa dengan *hands-on tasks*, atau assesmen berbasis performa lainnya yang mengharuskan siswa melengkapi secara individual atau dalam kelompok kecil; pekerjaan siswa dievaluasi menggunakan kriteria yang telah dibangun sebelumnya. Assesmen performa terdiri dari dua komponen yaitu: 1) sebuah *performance task* (aktivitas atau actual prompt) dan 2) sebuah rubrik penskoran (*scoring rubric*) atau panduan penskoran yang terdiri dari kriteria performa yang telah ditetapkan/dibangun sebelumnya. Assesmen performa memungkinkan observasi langsung terhadap kemampuan dan ketrampilan siswa (sangat berbeda dengan *pencil-and-paper tests*). Assesmen performa harus dikaitkan dengan objektiv pembelajaran. Disamping itu assesmen performa cenderung lebih abstrak ketimbang bentuk-bentuk asesmen tradisional, artinya assesmen performa lebih "*real world*". Assesmen performa yang didasarkan pada "*real world*" dikenal dengan *authentic assessment*. Assesmen performa, karena dilakukan oleh siswa sendiri, merupakan aktivitas pembelajaran yang penuh makna. Konsep assesmen performa bukan merupakan konsep baru; digunakan untuk bertahun-tahun di bidang yang lain.

#### **Beberapa persyaratan dasar dari assesmen performa :**

- 1) Kemampuan dan perilaku spesifik (krusial) harus diobservasi.
- 2) Sangat cocok untuk mengukur ketrampilan dan kemampuan siswa yang kompleks yang tak dapat diukur dengan menggunakan *pencil-and-paper tests*.
- 3) Tugas-tugas harus fokus pada proses-proses yang dapat diajarkan oleh guru atau dipelajari oleh siswa.
- 4) Dapat dipakai untuk memutuskan kecocokan perilaku atau pemahaman.

- 5) Dapat dipakai untuk memutuskan kecocokan perilaku atau pemahaman yang memberikan informasi tentang kekuatan (*strengths*) dan kelemahan (*weaknesses*).
- 6) Mempersyaratkan adanya produk perilaku yang bernilai untuk kebaikannya.
- 7) Tugas-tugas harus mendorong atau memotivasi refleksi siswa.

#### **b. Assesmen Proses vs Assesmen Produk**

Assesmen proses target khususnya adalah prosedur yang digunakan oleh siswa untuk menjawab suatu permasalahan atau problem. Assesmen Produk target utamanya adalah hasil-hasil belajar siswa yang berupa produk yang dapat diukur (*results in tangible outcome*). Para guru biasanya lebih tertarik pada salah satunya (assesmen proses atau assesmen produk), walaupun tugas-tugas itu mungkin saja mempersyaratkan keduanya (assesmen proses dan assesmen produk). Keputusan tentang fokus dari tugas-tugas itu harus dibuat oleh guru.

#### **c. Pengembangan Tugas-tugas Assesmen Performa**

Ada empat gambaran esensial yang perlu dipertahankan dalam pengembangan tugas-tugas dalam assesmen performa :

- 1) mempunyai satu tujuan yang jernih yang menspesifikasikan keputusan hasil yang akan dibuat dari assesmen :
  - a) merupakan tahap-tahap yang krusial,
  - b) apakah hasil assesmen akan digunakan untuk tujuan formatif atau sumatif ?
  - c) pada apakah tugas itu akan difokuskan, pada proses, produk atau keduanya ?
- 2) Mengidentifikasi aspek-aspek produk atau performa siswa yang akan diputusi:
  - a. Kriteria produk atau performa siswa yang akan diasses harus secara spesifik dapat diamati,
  - b. Difokuskan pada proses, produk atau keduanya,
  - c. Harus dapat dinyatakan dengan jernih,
  - d. Kriteria harus dibatasi pada sejumlah hal yang dapat dikelola dan dipikirkan dengan jelas.
- 3) Hasilnya dinyatakan dalam satu atau lebih skor yang mendeskripsikan performa siswa.

- 4) Memberikan sebuah seting yang cocok untuk kegiatan melengkapi tugas-tugas itu.

Pendesainan tugas-tugas Performa : Prosedur tahap demi tahap.

Tahap 1 :  
Menentukan tujuan assesmen.

Tahap 2 :  
Menspesifikasikan ketrampilan dan *outcomes* merujuk pada *taxonomic level*.

Tahap 3 :  
Menspesifikasikan kriteria performa yang akan digunakan untuk *menjudge* pekerjaan siswa, dan mengidentifikasi indikator-indikator tampak dari criteria itu.

Tahap 4 :  
Mengembangkan konteks yang *authentic* dan *meaningful* untuk tugas-tugas itu.

Tahap 5 :  
Mengembangkan instrumen penskoran (*rubrik scoring*).

Tahap 6 :  
Membangkitkan atau memilih contoh-contoh respon siswa

Tahap 7 :  
merevisi tugas, sesuai dengan keperluan.

**c. Metode Penskoran Assesmen berbasis performa**

Biasanya, tidak ada jawaban benar atau salah yang sederhana; mereka harus diasses beberapa saat semacam kontinu. Fokus pada derajat kualitas, pemahaman, kecakapan dsb. Sasarannya untuk mencoba mereduksi potensi subyektivitas dalam penskoran. Terbuka mengenai instrumen penskoran kepada siswa, memberikan bimbingan kepada siswa dengan merujuk pada hakekat *open-ended*-nya tugas-tugas performa.

#### d. Skema Metode Penskoran Assesmen Performa

**Checklists**, merupakan daftar perilaku atau keterampilan (*skills*) siswa, yang mengindikasikan setiap perilaku atau keterampilan siswa yang telah diobservasi. Sangat baik (paling baik) ketika digunakan secara formatif, untuk memberikan indikasi kekuatan dan kelemahan secara cepat. Rating scales : memperkenankan guru untuk mengindikasikan frekuensi atau derajat bagaimana ketrampilan atau perilaku dipertunjukkan oleh siswa.

**Rubrics**, merupakan *rating scales* yang secara khusus digunakan untuk penskoran hasil assesmen performa. Rubrics merupakan panduan penskoran yang terdiri dari kriteria performa spesifik yang telah ditetapkan/dibangun sebelumnya, digunakan untuk mengevaluasi pekerjaan siswa dalam assesmen performa.

Ada dua tipe rubrik yaitu :

- i. **Holistic rubric** : skor ditentukan untuk seluruh proses atau produk.
  - a. Digunakan ketika kesalahan dalam proses dapat ditoleransi
  - b. Digunakan dengan tugas-tugas dimana tidak ada respon benar yang definitif.Sangat cepat untuk menskor, tetapi kurang dapat memberikan balikan.

#### Contoh Holistic Rubrics Skore Deskripsi

- 5 : mendemonstrasikan secara lengkap pemahaman terhadap masalah.  
Semua persyaratan tugas ada dalam respons/produk siswa.
- 4 : mendemonstrasikan sebagian besar pemahaman terhadap masalah.  
Semua persyaratan tugas ada dalam respons/produk siswa.
- 3 : mendemonstrasikan sebagian pemahaman terhadap masalah.  
Sebagian besar persyaratan tugas ada dalam respons/produk siswa.
- 2 : mendemonstrasikan sebagian kecil pemahaman terhadap masalah.  
Beberapa persyaratan tugas ada dalam respons/produk siswa.
- 1 : tidak mendemonstrasikan pemahaman dari masalah.  
Beberapa persyaratan tugas ada dalam respons/produk siswa.
1. : tak ada respon atau tugas-tugas tak dikerjakan.

- ii. *Analytic rubric*: rubrik penskoran dimana komponen individual dari produk atau performa siswa diskor secara terpisah.
  - a. Digunakan ketika lebih fokus pada tipe respon yang dipersyaratkan.
  - b. Biasanya menghasilkan beberapa skor, yang bisa saja dijumlahkan untuk memperoleh skor total.
  - c. Lebih lambat dalam proses penskoran namun lebih detil dalam memberikan balikan.

#### Contoh Rubrik Analitik

Permulaan pengembangan Tuntas dengan baik sempurna skor :

- Kriteria 1 : Deskripsi mencerminkan level performa pemula.  
 Deskripsi mencerminkan kemajuan menuju level performa tuntas.  
 Deskripsi mencerminkan prestasi level performa tuntas.  
 Deskripsi mencerminkan performa level tingkat tinggi.
- Kriteria 2 : Deskripsi mencerminkan level performa pemula.  
 Deskripsi mencerminkan kemajuan menuju level performa tuntas.  
 Deskripsi mencerminkan prestasi level performa tuntas.  
 Deskripsi mencerminkan performa level tingkat tinggi.
- Kriteria 3 : Deskripsi mencerminkan level performa pemula.  
 Deskripsi mencerminkan kemajuan menuju level performa tuntas.  
 Deskripsi mencerminkan prestasi level performa tuntas.  
 Deskripsi mencerminkan performa level tingkat tinggi.
- Kriteria 4 : Deskripsi mencerminkan level performa pemula.  
 Deskripsi mencerminkan kemajuan menuju level performa tuntas.  
 Deskripsi mencerminkan prestasi level performa tuntas.  
 Deskripsi mencerminkan performa level tingkat tinggi.

Total Skor \_\_\_\_\_

Satu tipe rubrik tidak selalu lebih baik dari tipe yang lainnya, sangat bergantung pada tujuan. Biasanya tujuan yang berbeda akan berimplikasi pada pemilihan tipe rubrik yang berbeda. Guru harus memutuskan bagaimana format yang diperlukan dan cocok dengan tugas-tugas yang

ada. Boleh jadi rubrik berupa tingkat-tingkat kemahiran dan sebagainya, boleh jadi bersifat kualitatif, kuantitatif ataupun keduanya.

Prosedur mengembangkan Rubrik Penskoran tahap demi tahap :

- Tahap 1 : mempertanyakan ulang objektif pembelajaran untuk menentukan tugas-tugas yang cocok.
- Tahap 2 : menentukan atribut-atribut spesifik yang tampak yang ingin kamu lihat (seperti halnya yang tidak ingin kamu lihat).
- Tahap 3 : renungkan karakteristik yang mendeskripsikan masing-masing atribut.
- Tahap 4a : (untuk rubrik holistik) tuliskan deskripsi naratif secara rinci untuk pekerjaan yang sangat sempurna dan sangat tidak sempurna, mempersatukan/mempertautkan tiap atribut ke dalam deskripsi.
- Tahap 4b : (untuk rubrik analitik) tuliskan deskripsi naratif secara rinci untuk pekerjaan yang sangat sempurna dan sangat tidak sempurna untuk tiap atribut individual.
- Tahap 5a : lengkapi rubrik dengan pendeskripsian level lainnya pada kontinum yang merentang dari sangat sempurna sampai dengan sangat tidak sempurna untuk atribut kolektif.
- Tahap 5b : lengkapi rubrik dengan pendeskripsian level lainnya pada kontinum yang merentang dari sangat sempurna sampai dengan sangat tidak sempurna untuk setiap atribut.
- Tahap 6 : kumpulkan contoh-contoh pekerjaan siswa yang memberikan contoh setiap level.
- Tahap 7 : revisi rubrik sebagaimana diperlukan.

#### **e. Validitas dan Reliabilitas Assesmen Berbasis Performa**

Validitas, dapat ditingkatkan dengan cara berbagi dengan siswa tentang kriteria yang akan digunakan untuk *menjudge* pekerjaannya. Harus dipastikan bahwa performa bukan merupakan tugas spesifik. Siswa harus lulus/melampaui ketrampilan prasyarat yang sesuai untuk mendemonstrasikan ketrampilan yang kompleks yang mungkin dipersyaratkan oleh tugas. Tugas-tugas harus *fair* untuk semua siswa. Reliabilitas, untuk meningkatkannya rubrik harus didesain sedemikian

rupa sehingga dapat mereduksi subjektivitas. Guru harus mencegah bias personalnya dalam penskoran tugas.

**f. Keuntungan dan Keterbatasan Assesmen Berbasis Performa**

Keuntungannya antara lain:

- a. Dapat mengases kemampuan siswa untuk melakukan (“*to do*”)
- b. Dapat mengases keterampilan siswa yang tidak dapat diases dengan metode tradisional,
- c. Dapat mengases proses berpikir seperti produk,
- d. Dapat digunakan untuk meningkatkan proses pembelajaran praktis.

Keterbatasannya antara lain :

- a) Batasan umumnya merupakan keseluruhan waktu yang rumit,
- b) Tidak efisien ketika digunakan untuk mengases keterampilan tingkat rendah,
- c) Akibat dari subyektivitas, reliabilitasnya cenderung rendah,
- d) Siswa berkemampuan rendah cenderung menunjukkan frustrasi.

**D. Assesmen Portofolio**

*Portofolio assessments* mempersyaratkan adanya akumulasi bukti-bukti yang diperoleh selama beberapa/sepanjang waktu. Beberapa *subject areas* atau *setting* kelas khusus memandang assesmen portofolio mempunyai fokus yang berbeda.

**1) Karakteristik Portofolio Assesments**

*Portofolio assessments* merupakan bentuk assesmen yang penuh dengan tujuan, mengorganisasikan kumpulan pekerjaan siswa yang dapat digunakan untuk mendeskripsikan usaha, kemajuan atau pencapaian/prestasi siswa. Ide assesmen portofolio diperoleh dari portofolio artis atau penulis yang kemudian diadopsi di lingkungan akademik. Assesmen ini memberikan makna kepada siswa untuk menunjukkan apa yang dapat mereka kerjakan dengan nyata. Assesmen portofolio didasarkan pada kenyataan yang terjadi sesungguhnya atau *the real world*. Produk dan performa siswa dikompilasikan dalam sebuah upaya untuk menunjukkan kemahiran atau kemajuan siswa sepanjang waktu.

Beberapa sifat turunan dari assesmen portofolio dari folder contoh pekerjaan antara lain adalah:

- 1) Siswa harus diikuti dalam proses,
- 2) Kriteria pemilihan pekerjaan harus diidentifikasi,
- 3) Kriteria pengambilan keputusan nilai pekerjaan harus dikembangkan,
- 4) Bukti refleksi siswa harus ada.

Guru memberikan petunjuk dan pekerjaan secara kolaboratif dengan siswa untuk mengidentifikasi *goal* dari sampel portofolio dan pekerjaan untuk disertakan. Guru dan siswa melanjutkan saling akses ini untuk menentukan isi (lebih banyak berpusat pada siswa ketimbang bentuk assesmen kelas lainnya). Dapat diadaptasikan untuk menemukan pembelajaran individual yang diperlukan (setiap portofolio diperlakukan secara individual).

Refleksi siswa memberikan kesempatan pada siswa untuk berpikir tentang bagaimana berpikir (beberapa siswa tidak biasa dengan ide ini). Apakah ada satu metode yang sangat baik untuk menunjukkan metode dan cara berpikir yang digunakan oleh siswa, sebaik produk akhir (portofolio)? Berbeda dengan bentuk assesmen lainnya, portofolio cenderung fokus pada pencapaian siswa, bukan pada kesalahan. Portofolio menekankan kekuatan yang ada pada siswa. Portofolio dapat menghabiskan waktu yang banyak (proses *mereview* yaitu menguji portofolio secara kontinu sepanjang pertemuan dengan siswa secara individual).

Tiga ciri utama dari berbagai bentuk portofolio antara lain :

- Merupakan koleksi pekerjaan siswa yang mempunyai tujuan khusus,
- Siswa harus dilibatkan aktif dalam proses pemilihan pekerjaan untuk dimasukkan dalam portofolio,
- Siswa harus secara aktif melakukan refleksi diri.

## 2) Penggunaan Portofolio

### 2.1 Format ideal

Didesain untuk dua tujuan utama, yaitu :

- 1) Portofolio harus merupakan model pertanggungjawaban secara personal dalam merefleksikan pekerjaan orang/siswa itu sendiri,
- 2) Portofolio harus menjadi dokumen perkembangan dan pertumbuhan siswa selama rentang waktu tertentu.

Tiga komponen utama portofolio meliputi :

- 1) Dokumentasi perkembangan pekerjaan utama/pokok,
- 2) Bukti pekerjaan siswa selama rentang waktu tertentu,
- 3) Refleksi siswa.

Keterkaitan antara karakteristik utama dengan komponen format ideal ditunjukkan oleh bagan berikut :

*Karakteristik Utama Komponen Format Ideal*

- a) merupakan koleksi pekerjaan siswa yang mempunyai tujuan khusus
- b) kriteria pemilihan pekerjaan harus diidentifikasi
- c) siswa harus secara aktif melakukan refleksi diri

Dokumentasi portofolio bertujuan untuk memberikan rekaman kemajuan siswa secara aktual. Dokumentasi portofolio merupakan sebuah akumulasi bukti-bukti (yakni : usaha, kemajuan, dan prestasi) yang focus pada pendekatan holistik.

## 2.2 Showcase Portofolio

Bertujuan untuk menyoroti dan menayangkan keahlian-keahlian. Terdiri dari pekerjaan favorit dan terbaik dari siswa. *Sample*/contoh harus mendemonstrasikan level tertinggi dari pencapaian. Refleksi siswa merupakan kunci dari portofolio jenis ini.

## 2.3 Variasi Portofolio lainnya

Portofolio Kelas, merupakan ringkasan dokumen yang digunakan untuk memberi ilustrasi kemahiran-kemahiran dalam kelas. Tidak termasuk refleksi siswa. Digunakan untuk laporan kepada orang tua siswa dan pengelola sekolah.

Portofolio Evaluasi, digunakan secara eksklusif untuk laporan kepada masyarakat umum dan kepada agensi pemerintah (merupakan sebuah hasil dari proyek hadiah (*grant project, curriculum project, dsb*)).

## 3) Pengkreasian Portofolio

Isu-isu dalam pengkreasian portofolio, guru harus membuat jernih tujuan dari portofolia dibuat (dapat menyatakan jenis portofolio mana yang digunakan). Kriteria yang digunakan untuk mengasses *sample* pekerjaan harus dispesifikasikan. *Sample* pekerjaan harus sejalan/paralel dengan objektiv pembelajaran dan proses pembelajaran yang terjadi.

Kerangka waktu untuk *mereview* harus dispesifikasikan di awal. Pertimbangkan untuk memberi bantuan pada refleksi siswa. Penskoran portofolio harus dilakukan secara terbuka dengan siswa.

Pengembangan Portofolio : prosedur *Step-by-Step*

Step 1 : Tentukan tujuan yang akan dilayani dengan portofolio.

Step 2 : Identifikasi isi, konsep dan keterampilan yang akan diakses.

Step 3 : Identifikasikan pendekatan untuk mengorganisasikan.

Step 4 : Rencanakan satu proses *review* yang efisien dengan mengembangkan rangka waktunya.

Step 5 : Tentukan kapan dan bagaimana siswa dilibatkan.

Step 6 : Kembangkan rubrik penskoran (*scoring rubrics*)

Step 7 : Susunlah *multiple reviewers* untuk meningkatkan reliabilitas portofolio.

Step 8 : Rencanakan sebuah konferensi final (*final conference*).

#### 4) Keputusan Isi

Dalam banyak sekali situasi, siswa yang paling bertanggungjawab untuk memilih isi dari portofolio (tentu saja berkolaborasi dengan guru). Bagaimanapun, guru harus membuat jernih criteria spesifik untuk ditemukan dalam *sample* pekerjaan. Tidak terbatas pada apa yang dimasukkan dalam portofolio : harus diselaraskan antara *goal*, objektif pembelajaran, dan produk siswa sebagai dasar dari assesmen.

#### 5) Validitas dan Reliabilitas Assesmen Portofolio

Validitas, berhubungan penting dengan :

- a. Kemampuan khusus siswa dinyatakan oleh sampel yang ditentukan lebih lanjut,
- b. Pilihan pekerjaan harus merepresentasikan keseluruhan domain,
- c. Rubrik penskoran harus cocok dengan pekerjaan yang ada dalam portofolio.

Reliabilitas, merupakan konsistensi yang diinterpretasikan sebagai persetujuan antar penilai (*interater agreement*). *Rater* atau penilai yang jumlahnya lebih dari tiga (ganjil) untuk menilai satu *sample* dapat memfasilitasi konsistensi ini. Selama reliabilitas cenderung rendah, portofolio harus tidak harus digunakan sebagai semata-mata penilaian terhadap performa siswa.

## 6) Kelebihan dan Keterbatasan Assesmen Portofolio

Kelebihan :

- Secara substansial melibatkan siswa dalam proses,
- Dapat menunjukkan pertumbuhan siswa selama rentang waktu tertentu,
- Mendidik komunikasi antara guru dan siswa,
- Mempersyaratkan suatu peran mandiri untuk siswa dan guru,

## E. Assesmen Tes Obyektif

Merupakan teknik assesmen tradisional (tes obyektif dan subyektif) telah digunakan selama bertahun-tahun yang lalu. Ada kecenderungan seorang pendidik percaya bahwa perkembangan itu sederhana dan terus maju ke depan. Pengembangan tes itu mudah, tetapi mengembangkan tes yang baik memerlukan pengetahuan dan ketrampilan.

### 1. Karakteristik Umum

Tes obyektif memiliki satu respon yang dianggap benar, tanpa menghiraukan siapa yang menskor respon-respon itu, akan ditemukan skor-skor yang identik. Obyektif berimplikasi pada bahwa keputusan subyektif tidak mempengaruhi skor individual siswa. Juga dikenal sebagai "*selected-response items*" dan "*structured-response items*". Meliputi bentuk pilihan ganda, dan menjodohkan. Biasanya digunakan untuk mengasses ketrampilan berpikir tingkat rendah seperti pengetahuan, pemahaman dan aplikasi (ketrampilan berpikir tingkat tinggi lebih sulit untuk ditulis). Relatif mudah untuk diselenggarakan, diskor dan dianalisis. Penulisan butir dengan kualitas tinggi mempersyaratkan waktu secara substansial. Walaupun subyektivitas dapat dihindari dari proses penskoran, derajat subyektivitas secara substantif ada dalam penentuan muatan isi yang dikandung oleh butir-butir tersebut. Tebak-tebakan dalam merespon merupakan sebuah kenyataan yang hampir pasti terjadi.

Panduan umum dalam menulis butir tes obyektif:

1. Tes obyektif harus mencakup isi dan keterampilan yang penting.
2. Kosa kata dan tingkat pembacaan dari tiap butir harus menjadi kemungkinan yang mendasar,
3. Setiap butir obyektif harus dinyatakan dalam pernyataan yang tidak multi tafsir, dan kerancuan kata maupun struktur kalimat harus dihindari,

4. Butir obyektif tidak harus terdiri dari pernyataan-pernyataan kata demikata atau frase harus dihilangkan dari teks,
5. Kunci untuk menjawab yang benar tidak diberikan.

**Petunjuk pemformatan tes obyektif :**

1. Variasi bentuk-bentuk butir yang tampil dalam tes kelas,
2. Kelompokkan butir-butir yang sejenis dalam format bersama kemudian masing-masing tipe tampilkan dalam sebuah bagian yang terpisah,
3. Setiap bagian harus didahului oleh petunjuk yang jernih,
4. Dalam tiap bagian, mengandung butir yang mudah sampai dengan sulit,
5. Walaupun semua butir tidak akan tampil pada setiap tes, butir-butir itu harus disusun dalam susunan benar-salah, menjodohkan, jawab singkat, pilihan ganda, dan esai,
6. Berikan ruang yang cukup bagi siswa untuk merespon tiap-tiap butir,
7. Hindari pemenggalan sebuah butir dalam 2 halaman.

**Petunjuk umum untuk penulisan tes obyektif :**

- 1) Mulai pengembangan tes dengan tabel spesifikasi :
  - bagan yang menunjukkan tujuan pembelajaran, isi dan taksonomi Bloom;
  - tentukan kategori utama dari isi, dalam baris-baris;
  - nyatakan 6 tingkat domain kognitif Bloom dalam kolom-kolom
  - tuliskan dalam sel-sel jumlah butir atau persentase butir yang dikembangkan untuk isi dan tingkatannya.
- 2) Berikan jaminan (pastikan) bahwa tes mencakup *sample-sample* isi secara akurat dan representatif.
- 3) Pilihan penggunaan tes obyektif yang disarankan :
  - a) Pertimbangkan pertanyaan dasar: apakah butir cocok/sejodoh dengan pembelajaran yang diterima siswa?;
  - b) Buku ajar harus hati-hati diuji kesesuaiannya
- 4) Penskoran relatif mudah: gunakan kunci jawaban untuk sampai pada jumlah total butir yang dijawab dengan benar oleh siswa.

## **Tipe-tipe Tes obyektif :**

### **2.1 Pilihan Ganda**

Format dasar yang terdiri dari sebuah pernyataan/pertanyaan dan pilihan jawaban (hanya satu pilihan yang benar, dan yang lainnya disebut dengan *distractors*). Pernyataan dapat berupa pertanyaan atau melengkapi pernyataan. Ada tiga sampai dengan lima pilihan. Sangat berguna untuk mengassess ingatan fakta dan aplikasi pengetahuan. Dapat juga digunakan untuk mengassess ketrampilan berpikir tingkat lebih tinggi (lebih sulit untuk dituliskan).

Panduan umum pengembangan tes pilihan ganda :

1. Pernyataan harus jernih menyatakan masalah yang menjadi perhatian siswa
2. Semua pilihan jawaban harus parallel dalam jenis isinya
3. Pilihan harus menghindari pengulangan kata
4. Kata benda atau kata sifat yang secara substansial merubah makna pernyataan atau pilihan harus ditekankan
5. Semua distraktor dalam seperangkat respon harus masuk akal
6. Tata bahasa dalam setiap pilihan harus konsisten dengan pernyataan/pertanyaan
7. Butir harus menghindari penyertaan ungkapan “semua ... di atas” atau kebalikannya atau pilihan jawaban yang ekuivalen
8. Butir harus tetap independen satu dengan lainnya
9. Hindari petunjuk kunci jawaban seperti satu ungkapan kunci, antara pertanyaan/pernyataan dengan pilihan yang benar.

*Keuntungannya antara lain :*

- a. Dapat untuk menguji sample domain isi secara efisien dan komprehensif,
- b. Dapat digunakan dalam semua subjek area secara semu,
- c. Dapat diskor dengan relatif cepat,
- d. Penskoran merupakan sebuah proses obyektif,
- e. Dapat memberikan informasi diagnostik.

*Keterbatasannya antara lain :*

- a. Mudah terpengaruh oleh tebak-tebakan,
- b. Penyusunannya memakan waktu.

*Ragam dari pilihan ganda :*

- a) variasi jawaban benar dan jawaban terbaik,
- b) kombinasi butir pilihan ganda dan esai jawaban singkat,

- c) butir pilihan ganda metakognitif (“jelaskan jawabanmu.....”), memberi kesempatan kepada siswa untuk menjelaskan dan merefleksikan pemikirannya.

## 2.2. Menjodohkan

Format dasarnya terdiri dari dua daftar yaitu daftar pernyataan/pertanyaan (stimuli) dan daftar jawaban. Harus terjadi korespondensi satu-satu antara daftar stimuli dan daftar jawaban. Dianggap sebagai satu kasus khusus dari butir pilihan ganda. Dapat mengasah ketrampilan pemahaman dan pengetahuan. Secara khusus sangat berguna dalam pengukuran pemahaman konsep atau bentuk-bentuk yang saling berhubungan.

Panduan umum pengembangan butir tes menjodohkan :

1. Daftar stimuli maupun jawaban harus homogen
2. Petunjuk penjodohan harus dibuat jernih
3. Hindari penjodohan sempurna dengan penempatan daftar jawaban lebih banyak ketimbang daftar stimuli
4. Gunakan daftar stimuli dan jawaban yang relatif singkat
5. Tempatkan frase yang lebih panjang dalam daftar stimuli dan yang lebih pendek pada daftar jawaban
6. Susunlah daftar tersebut dalam urutan yang logis.

*Keuntungannya antara lain :*

- a) memberi peluang untuk mengasah konsep, ide dan fakta yang saling terkait dengan lebih efisien,
- b) relatif lebih mudah disusun/dikembangkan,
- c) pada dasarnya merupakan kombinasi butir pilihan ganda dengan penggunaan seperangkat jawaban yang sama,
- d) penskorannya relatif mudah.

*Keterbatasannya antara lain :*

- a) mempersyaratkan ide atau konsep yang berkaitan dalam jumlah yang besar,
- b) sangat sulit didesain untuk ketrampilan tingkat tinggi.

## 2.3 Pilihan Alternatif

Secara esensial merupakan kasus khusus dari butir pilihan ganda di mana pilihan dibatasi hanya dua pilihan. Jenis paling populer adalah butir

benar-salah (*true-false item*). Ragamnya meliputi : benar-tidak benar, ya-tidak, dan fakta-opini. Dapat efektif jika ditulis dengan hati-hati (meskipun berita negatif). Cenderung *overestimate* pada prestasi siswa, sementara siswa memiliki peluang 50% untuk asal tebak terhadap jawaban yang benar.

Panduan umum pengembangannya :

1. hindari penggunaan ungkapan absolut dan ketentuan khusus lainnya.
2. hindari pengujian pengetahuan trivial/kurang berguna.
3. butir-butir dinyatakan dalam kalimat positif; jika pernyataan negatif harus digunakan, tandailah ungkapan negatif dengan garis bawah, tebal, atau cetak miring.
4. kira-kira separuh dari seluruh butir harus memiliki kunci jawaban benar dan separuhnya lagi mempunyai kunci jawaban salah.
5. pernyataan benar dan pernyataan salah harus memiliki sama panjang.
6. butir benar-salah haruslah benar sepenuhnya atau salah sepenuhnya.

*Keuntungannya antara lain adalah :*

- a) relatif cepat untuk membuatnya, menjawabnya, dan menskornya;
- b) dapat diskor secara efisien dan obyektif.

*Keterbatasannya antara lain adalah :*

- a) memberi peluang sangat besar untuk menebak-nebak,
- b) butir-butirnya secara dominan sesuai untuk mengases ketrampilan tingkat rendah.

Ragamnya meliputi :

- a. butir ya-tidak,
- b. benar-salah terkoreksi (mempersyaratkan siswa menulis ulang beberapa pernyataan yang salah),
- c. *embedded alternate-choice items*, yang menyatakan sederet butir pilihan alternatif di dalam satu *paragraph*,
- d. *multiple true-false items*, merupakan hasil persilangan pilihan ganda dengan pilihan alternatif(pernyataan-pernyataan benar-

salah ganda, masing-masing menggunakan pernyataan/pertanyaan yang sama).

## ii. Analisis Butir

Analisis butir, merupakan analisis secara statistik karakteristik setiap butir yang tampil dalam sebuah tes dengan tujuan membuat keputusan tentang butir yang dipertahankan dan butir yang harus dibuang.

Butir-butir harus dievaluasi :

- a) ketika butir-butir itu sedang dirancang (menggunakan tabel spesifikasi, panduan dsb),
- b) mengikuti penyelenggaraan tes dan penskoran.

Ada empat ukuran dasar statistik, yaitu :

- a) Taraf kesukaran butir : proporsi siswa yang menjawab butir dengan benar,
- b) Daya pembeda : perbedaan antara proporsi dari jawaban benar siswa dalam kelompok tinggi dan kelompok rendah,
- c) Analisis distraktor : menguji pola respon/jawaban untuk pilihan salah,
- d) Reliabilitas : konsistensi menyeluruh untuk seluruh butir.

## iii. Taraf Kesukaran

Taraf kesukaran/kesulitan, disimbolkan dengan  $p$ , secara sederhana merupakan pembagian antara jumlah siswa yang menjawab benar dalam suatu butir dengan jumlah siswa yang merespon/menjawab butir itu. Nilainya dapat merentang antara 0,00 (sulit) s/d 1,00 (mudah). Pertimbangan untuk merevisi butir jikalau  $p < 0,80$ . Keputusan yang baik juga harus juga digunakan dalam hubungannya dengan analisis statistik. Guru dapat mengharapkan seluruh siswa menjawab dengan benar beberapa butir; karenanya  $p = 1,00$  mengindikasikan bahwa seluruh siswa telah memahami konsep dengan sangat sempurna.

**iv. Derajat Pembeda**

Derajat Pembeda, disimbolkan dengan D, bertujuan untuk menunjukkan bagaimana butir dapat membedakan siswa yang berkemampuan rendah dan siswa yang berkemampuan tinggi (pada keseluruhan tes). Jika sebuah butir berfungsi sebagai pembeda yang baik, maka sebagian besar siswa dalam kelompok kemampuan tinggi akan menjawab butir itu dengan benar dan sebagian besar siswa dalam kelompok kemampuan rendah menjawab butir itu dengan salah. Biasanya nilainya merentang dari +0,10 sampai dengan + 0,60.

**v. Analisis Distraktor**

Secara informal menguji pola jawaban lintas semua pilihan pilihan.

**vi. Reliabilitas**

Reliabilitas, dihitung dengan KR-21 untuk memperoleh nilai koefisien reliabilitas. Nilainya merentang dari 0,00 sampai dengan 1,00. Rentang nilai yang diinginkan untuk tes di kelas adalah antara 0,70 sampai dengan 1,00.

**Validitas dan Reliabilitas Tes Obyektif**

Validitas, harus dapat menjawab pertanyaan-pertanyaan berikut :

- a) Apakah saya sedang mengukur apa yang hendak diukur ?
- b) Seberapakah tingkat kepercayaan yang harus saya miliki untuk keputusan yang akan saya buat berdasarkan pada ukuran-ukuran tersebut ?

Kecenderungan utamanya adalah validitas isi. Reliabilitas, dibangun melalui penggunaan analisis statistik, khususnya KR-21, yaitu untuk memperoleh nilai koefisien reliabilitas.

**F. Asesmen Tes Subyektif**

Butir tes subyektif merupakan alat yang sangat cocok untuk mengases pembelajaran siswa. Lagi pula, ada satu kecenderungan bagi pendidik untuk mempercayai bahwa perkembangannya sederhana dan maju kedepan, tetapi penulisan butir dengan kualitas tinggi memerlukan beberapa praktik/latihan.

**1. Karakteristik Butir Tes Subyektif**

Butir tes subyektif biasanya tidak memiliki satu jawaban/respon yang benar. Subyektif berimplikasi bahwa keputusan subyektif penskor

merupakan satu bagian integral dari proses penskoran. Juga dikenal dengan “*free response*” atau *constructed-response*” atau “*supply-type*” items. Meliputi butir jawab pendek dan esai. Mempersyaratkan siswa agar memproduksi apa yang mereka tahu, dibandingkan dengan hanya pengakuan jawaban terbaik dari satu set pilihan. Relatif mudah untuk dibuat. Siswa memerlukan waktu untuk menjawab/memberikan respon. Dapat memakan waktu dan melelahkan secara subyektif untuk menskor, bergantung pada formatnya.

Beberapa karakteristik umum dari butir tes subyektif antara lain adalah sebagai berikut :

1. Butir tes subyektif harus harus meng-cover ketrampilan dan isi penting.
2. Tata bahasa dan level pembacaan dari setiap butir harus diupayakan pada tingkat dasar.
3. Setiap butir tes subyektif harus dinyatakan dalam kalimat yang tidak multitafsir atau rancu.
4. Tidak memberikan petunjuk tentang kunci jawaban.
5. Variasikan jenis butir yang ditampilkan pada tes di kelas.
6. Kelompokkan butir dengan format yang sama sehingga masing-masing tipe tampil dalam kelompok yang saling terpisah.
7. Setiap bagian harus didahului dengan petunjuk yang jelas
8. Dalam setiap bagian, butir disusun mulai yang mudah menuju yang sulit.
9. Walau semua jenis butir tidak akan tampil pada setiap tes, test, mereka harus disusun dalam susunan seperti berikut: benar salah, menjodohkan, jawab singkat, pilihan ganda dan butir subyektif.
10. Berikan ruang yang cukup bagi siswa untuk memberi respon pada setiap butir.
11. Hindari pemenggalan butir ke dalam dua halaman.

## **2. Butir Jawaban Singkat**

Tipe ini mempersyaratkan siswa untuk memberikan satu kata, frase pendek, angka, atau respon/jawaban ringan lainnya. Diformat sebagaimana pertanyaan-pertanyaan lainnya (dikenal sebagai butir jawaban pendek) atau kalimat yang belum lengkap (dikenal sebagai butir melengkapi atau butir mengisi tempat kosong/titik-titik). Sebenarnya beberapa butir melengkapi dapat ditulis sebagai butir jawab singkat atau sebaliknya. Tipe ini dapat mengasses ketrampilan atau isi serupa seperti

tipe pilihan ganda, tetapi siswa harus mengingat atau membuat jawaban dia (pada tipe pilihan ganda hanya mengidentifikasi dari daftar pilihan yang ada). Paling efisien untuk mengasah ketrampilan berpikir tingkat rendah. Relatif mudah dibuat. Relatif mudah diskor (tidak semudah butir obyektif, sementara jawaban harus ditulis tangan).

**Ada beberapa isu penting terkait dengan penskoran :**

- 1) ejaan dan struktur kalimat,
- 2) siswa harus diberitahu jika akan diases dengan tes,
- 3) siswa bisa saja memberi jawaban benar tetapi tidak sama dengan yang guru maksudkan,
- 4) guru harus menentukan jika jawaban ini merupakan hasil dari kurangnya kemampuan menulis atukah kurang belajar (miskonsepsi).

Ragamnya bisa berbentuk pemberian satu daftar jawaban-jawaban yang mungkin : tidak lebih luas dari butir subyektif, hanya merupakan variasi dari butir menjodohkan.

Panduan umum untuk membuat butir jawab singkat :

1. *Short-answer items* harus disusun kata-katanya secara khusus dan jernih.
2. Dalam butir melengkapi, letakkan tempat kosong di dekat atau di akhir pernyataan.
3. Hindari pengkopian pernyataan kata demi kata.
4. Tuliskan hanya kata-kata yang penting saja
5. Gunakan hanya satu atau dua tempat kosong (untuk butir melengkapi)
6. Standarkan panjang dari tempat kosong.

Keuntungan dan keterbatasan butir jawab singkat. Keuntungan, relatif mudah dibuat, lebih mudah daripada pilihan ganda karena tidak perlu membuat pilihan. Kemungkinan untuk asal tebak dapat direduksi. Keterbatasan, utamanya hanya digunakan untuk ketrampilan berpikir tingkat rendah. Penskoran dapat menjadi sesuatu yang sulit dikarenakan salah eja dan kurang baiknya struktur kalimat.

### 3. Essay

Merupakan pertanyaan atau perintah yang mempersyaratkan siswa untuk menulis *paragraph* atau mengembangkan tema sebagai jawaban. Meliputi ragam ketrampilan berpikir yang lebih luas (siswa harus mengingat, memilih, mengorganisasikan, dan menerapkan). Jawaban bisa merentang dari sedikit/beberapa kalimat (*restricted-response items*) sampai dengan beberapa halaman (*extended-response items*); rentangan dalam suatu kontinum.

Sangat cocok untuk mengasses ketrampilan berpikir yang kompleks (seperti analisis sintesis dan evaluasi). Kadangkala digunakan secara eksklusif atau sebagai bagian dari assesmen performa yang lebih luas. Tidak cocok digunakan untuk mengasses ketrampilan berpikir tingkat rendah. Penskorannya merupakan sesuatu kegiatan yang kompleks dan biasanya tidak *reliable* (akibat pengaruh dari karakteristik jawaban). Penskoran dapat memakan banyak waktu. Penskoran harus mirip dengan penskoran pada assesmen performa, menggunakan rubrik holistik atau rubrik analitik.

Penskoran dapat dibuat lebih obyektif dengan :

1. Menetapkan terlebih dulu waktu untuk sebuah jawaban benar.
2. Penggunaan satu checklist atau rubrik yang dijelaskan dengan hati-hati.
3. Penentuan apakah ejaan, tata bahasa dsb akan diskor.
4. Penskoran jawaban siswa dengan tanpa nama (disembunyikan).
5. Penskoran untuk seluruh jawaban pada suatu butir tertentu dahulu, baru penskoran untuk butir yang lain, dan seterusnya.

Panduan umum pengembangannya :

1. Essay harus terdiri dari aplikasi pengetahuan yang esensial pada situasi yang baru.
2. Essay harus dinyatakan dalam kalimat yang jernih difokuskan pada tugas siswa.
3. Tentukan panjang yang diinginkan, batasan waktu, dan kriteria penilaian, untuk siswa.
4. Kembangkan satu model jawaban.

Keuntungan dan keterbatasan butir essay. Keuntungannya antara lain : dapat digunakan untuk memancing ragam jawaban yang luas, siswa diijinkan membuat jawabannya sendiri, peluang munculnya jawaban asal tebak dapat direduksi secara substansial. Keterbatasannya antara lain :

penipuan (*bluffing*) dapat menjadi sebuah problem, sangat memboroskan waktu untuk menskor, cakupan isinya sangat terbatas. Validitas dan Reliabilitas butir tes essay. Validitas, Harus dapat menjawab pertanyaan-pertanyaan berikut : Apakah saya sedang mengukur apa yang harusnya saya ukur ?, seberapaakah tingkat keyakinan yang harus sya miliki dalam keputusan yang akan saya buat berdasarkan atas ukuran itu?

Reliabilitas, tidak dapat ditentukan melalui analisis butir. Keputusan akan selalu dibuat (reliabilitas rendah) tetapi reliabilitas dapat ditingkatkan dengan mengikuti panduan pembuatan butir esai.

**Rangkuman**

Asesmen dan keputusan-keputusan evaluatif yang dihasilkannya haruslah akurat sehingga mampu mencegah kekeliruan pemahaman dan komunikasi oleh pihak-pihak terkait. Oleh karenanya diperlukan berbagai jenis asesmen yang secara bersama-sama akan menghasilkan informasi evaluatif yang lengkap dan akurat. Beberapa tipe dan kategori asesmen diantaranya asesmen formal dan informal, asesmen kuantitatif dan kualitatif, evaluasi formatif dan sumatif, asesmen baku dan tak baku, asesmen rujukan norma dan asesmen rujukan criteria, asesmen tes dan asesmen alternatif, asesmen obyektif dan subyektif. Pemahaman guru terhadap berbagai jenis asesmen ini memberikan peluang baginya untuk merancang berbagai instrumen asesmen dan membangun system asesmen yang menjamin diperolehnya hasil asesmen berupa informasi evaluatif yang lengkap, konsisten dan akurat. Semua keputusan lanjutan yang diambil berdasarkan informasi ini akan lebih pas dan cocok dengan keperluan.

## **BAB X**

### **KISI-KISI TES**

#### **A. Pengertian Kisi-kisi**

Kisi-kisi merupakan suatu format atau matriks yang memuat kriteria tentang soal-soal yang diperlukan atau yang hendak disusun. Kisi-kisi juga dapat diartikan *test blue-print* atau *table of specification* merupakan deskripsi kompetensi dan materi yang akan diujikan. Wujudnya adalah sebuah tabel yang memuat tentang perperincian materi dan tingkah laku beserta imbang/proporsi yang dikehendaki oleh penilai. Tiap kotak diisi dengan bilangan yang menunjukkan jumlah soal (Suhasimi, 2007:185). Tujuan penyusunan kisi-kisi adalah untuk menentukan ruang lingkup dan sebagai petunjuk dalam menulis soal.

#### **B. Fungsi Kisi-kisi**

- 1) Panduan/pedoman dalam penulisan soal yang hendak disusun. Pedoman penulisan soal merupakan aspek terpenting ketika guru hendak memberikan soal kepada siswa, pedoman tersebut akan menjadi acuan bagi guru dalam penulisan soal sehingga akan memudahkan dalam pembuatan soal.
- 2) Penulis soal akan menghasilkan soal-soal yang sesuai dengan tujuan tes. Tes merupakan bahan evaluasi guru terhadap keberhasilan peserta didik dalam pembelajaran yang disampaikan, guru dalam mengevaluasi peserta didik akan memberikan soal tes evaluasi yang bermacam-macam sesuai dengan tujuan pencapaian evaluasi terhadap pembelajaran tertentu. Dalam pembuatan soal yang menggunakan kisi-kisi, penulis akan menghasilkan soal-soal yang sesuai dengan tujuan tes.
- 3) Penulis soal yang berbeda akan menghasilkan perangkat soal yang relatif sama, dari segi tingkat kedalamannya segi cakupan materi yang ditanyakan.  
Penulisan kisi-kisi berfungsi untuk menselaraskan perangkat soal, sehingga hal ini juga akan mempermudah dalam proses evaluasi.

Kisi-kisi yang baik harus memenuhi persyaratan berikut ini :

- a. Kisi-kisi harus dapat mewakili isi silabus/kurikulum atau materi yang telah diajarkan secara tepat dan proporsional.

- b. Komponen-komponennya diuraikan secara jelas dan mudah dipahami.
- c. Materi yang hendak ditanyakan dapat dibuatkan soalnya.

### **C. Penulisan Kisi-kisi**

Penulisan kisi-kisi soal adalah kerangka dasar yang dipergunakan untuk penyusunan soal dalam evaluasi proses pendidikan dan pembelajaran. Dengan kisi-kisi soal ini, maka seorang guru dengan mudah dapat menyusun soal-soal evaluasi. Kisi-kisi soal inilah yang memberikan batasan guru dalam menyusun soal evaluasi.

Dengan kisi-kisi penulisan soal maka tidak akan terjadi penyimpangan tujuan dan sasaran dari penulisan soal untuk evaluasi penulisan soal. Guru hanya mengikuti arah dan isi yang diharapkan dalam kisi-kisi penulisan soal yang dimaksudkan.

Dalam penulisan kisi-kisi soal, guru harus memperhatikan hal-hal berikut :

1. Nama Sekolah

Nama sekolah ini menunjukkan tempat penyelenggaraan pendidikan dan pembelajaran yang akan dievaluasi proses pembelajarannya. Ini merupakan identitas sekolah.

2. Satuan Pendidikan

Satuan pendidikan menunjukkan tingkatan pendidikan yang menyelenggarakan proses pendidikan dan akan dievaluasi. Satuan pendidikan ini misalnya SD, SMP, SMA/SMK.

3. Mata Pelajaran

Mata pelajaran yang dimaksud dalam hal ini adalah mata pelajaran yang akan dibuatkan kisi-kisi soal dan dievaluasi hasil belajar anak-anak, misalnya matematika.

4. Kelas/Semester

Kelas/semester menunjukkan tingkatan yang akan dievaluasi, dengan mencantumkan kelas atau semester ini, maka kita semakin tahu batasan materi yang akan kita jadikan soal evaluasi proses.

5. Kurikulum Acuan

Seperti yang kita ketahui model kurikulum di negeri ini selalu berganti, akhirnya ada tumpang tindih antara kurikulum yang digunakan dan kurikulum baru. Untuk hal tersebut maka kita informasikan kurikulum yang digunakan dalam penyusunan kisi-kisi penulisan soal. Misalnya, KTSP.

6. Alokasi Waktu  
Alokasi waktu ini ditulis sebagai penyediaan waktu untuk penyelesaian soal. Dengan alokasi ini, maka kita dapat memperkirakan kesulitan soal. Dan jumlah soal yang harus dibuat guru agar anak-anak tidak kehabisan waktu saat mengerjakan soal.
7. Jumlah Soal  
Jumlah soal menunjukkan berapa banyak soal yang harus dibuat dan dikerjakan anak-anak sesuai dengan jatah alokasi waktu yang sudah dikerjakan untuk ujian bersangkutan. Dalam hal ini guru sudah memperkirakan penggunaan waktu untuk masing-masing soal.
8. Penulis/Guru Mata Pelajaran  
Ini menunjukkan identitas guru mata pelajaran atau penulis kisi-kisi soal. Hal ini sangat penting untuk mengetahui tingkat kelayakan seseorang dalam penulisan kisi-kisi dan soalnya.
9. Standar Kompetensi  
Standar kompetensi menunjukkan kondisi standar yang akan dicapai oleh peserta didik setelah mengikuti proses pendidikan dan pembelajaran. Dengan standar kompetensi ini maka guru dan anak didik dapat mempersiapkan segala yang harus dilakukan.
10. Kompetensi Dasar  
Kompetensi dasar menunjukkan hal yang seharusnya dimiliki oleh anak didik setelah mengikuti proses pendidikan dan pembelajaran. Dalam penulisan kisi-kisi soal aspek ini kita munculkan untuk mengevaluasi tingkat pencapaiannya.
11. Materi Pelajaran  
Ini menunjukkan semua materi yang diberikan untuk proses pendidikan dan pembelajaran. Dalam penulisan kisi-kisi soal, aspek ini merupakan batasan isi dari materi pelajaran yang kita jadikan soal.
12. Indikator Soal  
Indikator soal menunjukkan perkiraan kondisi yang diambil dalam soal ujian. Indikasi yang bagaimana dari materi pelajaran yang diterapkan disekolah.
13. Bentuk soal  
Bentuk soal yang dimaksudkan adalah subjektif tes atau objektif tes. Untuk memudahkan kita dalam menyusun soal, maka kita harus menentukan bentuk tes dalam setiap materi pelajaran yang kita ujikan dalam proses evaluasi.

14. Nomor Soal

Nomor soal menunjukkan urutan soal untuk materi atau soal yang guru buat. Dal hal ini, setiap standar kompetensi dan kompetensi dasar, penulisan nomor soal dikisi-kisi penulisan soal tidak selalu berurutan. Guru dapat menulis secara acak. Misalnya, standar kompetensi A dan kompetensi dasar A1 dapat saja diletakkan pada nomor 3 dan seterusnya sehingga tidak selalu standar kompetensi pertama dan kompetensir dasar pertama harus diurutkan di nomor satu.

**D. Penentuan dan Penyebaran Soal**

Sebelum menyusun kisi-kisi dan butir soal perlu ditentukan jumlah soal setiap kompetensi dasar dan penyebaran soalnya. Untuk lebih jelasnya, perhatikan contoh penilaian akhir semester berikut ini.

Contoh penyebaran butir soal untuk penilaian akhir semester ganjil

No	Kompetensi Dasar	Materi	Jumlah soal tes tulis		Jumlah soal Praktik
			PG	Uraian	
1	1.1 .....	.....	6	--	--
2	1.2 .....	.....	3	1	--
3	1.3 .....	.....	4	--	1
4	2.1 .....	.....	5	1	--
5	2.2 .....	.....	8	1	--
6	3.1 .....	.....	6	--	1
7	3.2 .....	.....	--	2	--
8	3.3 .....	.....	8	--	--
Jumlah soal			40	5	2

**E. Format Penulisan Kisi-Kisi Soal**

**KISI-KISI PENULISAN SOAL**

- a. Jenis sekolah : Jumlah soal :
- b. Mata pelajaran : Bentuk soal/tes :
- c. Kurikulum : Penyusun :
- d. Alokasi waktu :

No.	Standar Kompetensi	Kompetensi Dasar	Kls/smt	Materi Pokok	Indikator soal	Nomor Soal

Keterangan:

Isi pada kolom 2, 3, 4, dan 5 adalah harus sesuai dengan pernyataan yang ada di dalam silabus/kurikulum. Penulis kisi-kisi tidak diperkenankan mengarang sendiri, kecuali pada kolom 6.

#### **F. Syarat-Syarat Kisi-Kisi Yang Baik**

- 1) Mewakili isi kurikulum yang akan diujikan
- 2) Komponen-komponennya rinci, jelas dan mudah dipahami
- 3) Soal-soalnya dapat dibuat sesuai dengan indikator dan bentuk yang ditetapkan.

#### **G. Jenis Perilaku Yang Dapat Diukur**

Ranah kognitif yang dikembangkan Benjamin S. Bloom adalah :

- 1) Ingatan di antaranya seperti: menyebutkan, menentukan, menunjukkan, mengingat kembali, mendefinisikan;
- 2) Pemahaman di antaranya seperti: membedakan, mengubah, memberi contoh, memperkirakan, mengambil kesimpulan;
- 3) Penerapan di antaranya seperti: menggunakan, menerapkan;
- 4) Analisis di antaranya seperti: membandingkan, mengklasifikasikan, mengkategorikan, menganalisis;
- 5) Sintesis antaranya seperti: menghubungkan, mengembangkan, mengorganisasikan, menyusun;
- 6) Evaluasi di antaranya seperti: menafsirkan, menilai, memutuskan.

#### **a. Jenis perilaku yang dikembangkan Robert M. Gagne :**

- 1) Kemampuan intelektual : diskriminasi, indentifikasi/konsep yang nyata, klasifikasi, demonstrasi, generalisasi/menghasilkan sesuatu;

- 2) Kemampuan intelektual: diskriminasi, identifikasi/konsep yang nyata, klasifikasi, demonstrasi, generalisasi/menghasilkan sesuatu;
- 3) Strategi kognitif: menghasilkan suatu pemecahan;
- 4) Informasi verbal: menyatakan sesuatu secara oral;
- 5) Keterampilan motorist melaksanakan/menjalankan sesuatu;
- 6) Sikap: kemampuan untuk memilih sesuatu.

**b. Keterampilan Berpikir Yang Dikembangkan Linn dan Gronlund:**

- a) Membandingkan - Apa persamaan dan perbedaan antara ... dan... - Bandingkan dua cara berikut tentang ....
- b) Hubungan sebab-akibat - Apa penyebab utama ... - Apa akibat
- c) Memberi alasan (*justifying*) - Manakah pilihan berikut yang kamu pilih, mengapa? - Jelaskan mengapa kamu setuju/tidak setuju dengan pernyataan tentang ....
- d) Meringkas – Tuliskan pernyataan penting yang termasuk ... - Ringkaslah dengan tepat isi...
- e) Menyimpulkan – susunlah beberapa kesimpulan yang berasal dari data... Tulislah sebuah pernyataan yang dapat menjelaskan peristiwa berikut...
- f) Berpendapat (*inferring*) – berdasarkan..., apa yang akan terjadi bila.. apa reaksi A terhadap...
- g) Mengelompokkan – kelompokkan hal berikut berdasarkan...- apakah hal berikut memiliki...
- h) Menciptakan – tuliskan beberapa cara sesuai dengan ide anda tentang... – lengkapilah cerita... tentang apa yang akan terjadi bila...
- i) Menerapkan – selesaikan hal berikut dengan menggunakan kaidah – tuliskan dengan menggunakan pedoman...
- j) Analisis – manakah penulisan yang salah pada paragraf... – daftar dan beri alasan singkat tentang ciri utama...
- k) Sintesis – tuliskan satu rencana untuk pembuktian... – tuliskan sebuah laporan...
- l) Evaluasi – apakah kelebihan dan kelemahan.... – berdasarkan kriteria... tuliskanlah evaluasi tentang...

**c. Syarat Menyusun Kisi-kisi yang Baik**

- 1) Kisi-kisi harus dapat mewakili isi silabus/kurikulum atau materi yang telah diajarkan secara tepat dan proporsional.
- 2) Komponen-komponennya diuraikan secara jelas dan mudah dipahami.
- 3) Materi yang hendak ditanyakan dapat dibuatkan soalnya.

**d. Perumusan Indikator Soal**

- 1) Menggunakan kata kerja operasional (perilaku khusus) yang tepat,
- 2) Menggunakan satu kata kerja operasional untuk soal objektif, dan satu atau lebih kata kerja operasional untuk soal uraian/tes perbuatan,
- 3) Dapat dibuatkan soal atau pengecohnya (untuk soal pilihan ganda).

**e. Penulisan indikator yang lengkap**

A = *audience* (peserta didik) ,

B = *behaviour* (perilaku yang harus ditampilkan),

C = *condition* (kondisi yang diberikan), D = *degree* (tingkatan yang diharapkan).

**f. Model Penulisan Indikator**

- 1) Menempatkan kondisinya di awal kalimat Contoh model pertama untuk soal menyimak pada mata pelajaran Bahasa Indonesia. Indikator: Diperdengarkan sebuah pernyataan pendek dengan topik "belajar mandiri", peserta didik dapat menentukan dengan tepat pernyataan yang sama artinya.  
Soal : (Soal dibacakan atau diperdengarkan hanya satu kali, kemudian peserta didik memilih dengan tepat satu pernyataan yang sama artinya. Soalnya adalah: "Hari harus masuk kelas pukul 7.00., tetapi dia datang pukul 8.00 pagi hari.") Lembar tes hanya berisi pilihan seperti berikut: a. Hari masuk kelas tepat waktu pagi ini. b. Hari masuk kelas terlambat dua jam pagi ini c. Hari masuk Kelas terlambat siang hari ini, d. Hari masuk Kelas terlambat satu jam hari ini; Kunci: d
- 2) Menempatkan peserta didik dan perilaku yang harus ditampilkan di awal kalimat Contoh model kedua Indikator: Peserta didik dapat menentukan dengan tepat penulisan tanda baca pada

nilai uang. Soal : Penulisan nilai uang yang benar adalah .... a. Rp 125,- b. RP 125,00 c. Rp125 d. Rp125. Kunci: b

**g. Langkah-langkah Penyusunan Butir Soal**

- 1) Menentukan tujuan tes,
- 2) Menentukan kompetensi yang akan diujikan,
- 3) Menentukan materi yang diujikan,
- 4) Menetapkan penyebaran butir soal berdasarkan kompetensi, materi dan bentuk penilaiannya (tes tertulis : bentuk pilihan ganda, uraian, dan tes praktik),
- 5) Menyusun kisi-kisinya,
- 6) Menulis butir soal,
- 7) Memvalidasi butir soal atau menelaah secara kualitatif,
- 8) Merakit soal menjadi perangkat tes,
- 9) Menyusun pedoman penskorannya,
- 10) Uji coba butir soal,
- 11) Analisis butir soal secara kuantitatif dari data empirik hasil uji coba,
- 12) Perbaiki soal berdasarkan hasil analisis.

**h. Keunggulan dan Kelemahan Kisi-kisi Tes Pilihan Ganda**

Keunggulan soal bentuk pilihan ganda di antaranya adalah dapat mengukur kemampuan/perilaku secara objektif, sedangkan untuk soal uraian di antaranya adalah dapat mengukur kemampuan mengorganisasikan gagasan dan menyatakan jawabannya menurut kata-kata atau kalimat sendiri. Kelemahan soal bentuk pilihan ganda di antaranya adalah sulit menyusun pengecohnya, sedangkan untuk soal uraian di antaranya adalah sulit menyusun pedoman penskorannya.

Penulisan soal bentuk uraian objektif suatu soal atau pertanyaan yang menuntut sehimpunan jawaban dengan pengertian/konsep tertentu, sehingga penskorannya dapat dilakukan secara objektif. Artinya perilaku yang diukur dapat diskor secara dikotomis (benar - salah atau 1 – 0 uraian non-objektif suatu soal yang menuntut sehimpunan jawaban dengan pengertian/konsep menurut pendapat masing-masing peserta didik, sehingga penskorannya sukar untuk dilakukan secara objektif. Untuk mengurangi tingkat kesubjektifan dalam pemberian skor ini, maka dalam menentukan perilaku yang diukur dibuatkan skala.

**i. Kaidah penulisan soal pilihan ganda**

**Materi**

- a. Soal harus sesuai dengan indikator. Artinya soal harus menanyakan perilaku dan materi yang hendak diukur sesuai dengan rumusan indikator dalam kisi-kisi.
- b. Pengecoh harus berfungsi
- c. Setiap soal harus mempunyai satu jawaban yang benar. Artinya, satu soal hanya mempunyai satu kunci jawaban.

**Konstruksi**

- a. Pokok soal harus dirumuskan secara jelas dan tegas. Setiap butir soal hanya mengandung satu persoalan/gagasan.
- b. Rumusan pokok soal dan pilihan jawaban harus merupakan pernyataan yang diperlukan saja.
- c. Pokok soal jangan memberi petunjuk ke arah jawaban yang benar.
- d. Pokok soal jangan mengandung pernyataan yang bersifat negatif ganda.
- e. Pilihan jawaban harus homogen dan logis ditinjau dari segi materi.
- f. Panjang rumusan pilihan jawaban harus relatif sama.
- g. Pilihan jawaban jangan mengandung pernyataan "semua pilihan jawaban di atas salah" atau "semua pilihan jawaban di atas benar".
- h. Pilihan jawaban yang berbentuk angka atau waktu harus disusun berdasarkan urutan besar kecilnya nilai angka atau kronologis.
- i. Gambar, grafik, tabel, diagram, wacana, dan sejenisnya yang terdapat pada soal harus jelas dan berfungsi.
- j. Rumusan pokok soal tidak menggunakan ungkapan atau kata yang bermakna tidak pasti seperti: sebaiknya, umumnya, kadang-kadang.
- k. Butir soal jangan bergantung pada jawaban soal sebelumnya. Ketergantungan pada soal sebelumnya menyebabkan peserta didik yang tidak dapat menjawab benar soal pertama tidak akan dapat menjawab benar soal berikutnya.

**Bahasa/budaya**

- a. Setiap soal harus menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia. Kaidah bahasa Indonesia dalam penulisan soal di antaranya meliputi:  $\theta$  pemakaian kalimat: (1) unsur subyek, (2)

unsur predikat, (3) anak kalimat;  $\theta$  pemakaian kata: (1) pilihan kata, (2) penulisan kata, dan  $\theta$  pemakaian ejaan: (1) penulisan huruf, (2) penggunaan tanda baca.

- b. Bahasa yang digunakan harus komunikatif, sehingga pernyataannya mudah dimengerti warga belajar/peserta didik.
- c. Pilihan jawaban jangan yang mengulang kata/frase yang bukan merupakan satu kesatuan pengertian. Letakkan kata/frase pada pokok soal.

### **Tabel Spesifikasi**

Tabel spesifikasi juga dikenal dengan istilah kisi-kisi soal atau bentuk *blue print* adalah sebuah tabel analisis yang didalamnya dimuat rincian materi tes dan tingkah laku beserta proporsi yang dikehendaki oleh tester, dimana pada tiap petak (sel) dari tabel tersebut diisi dengan angka-angka yang menunjukkan banyaknya butir soal yang akan dikeluarkan dalam tes hasil belajar bentuk objektif.

Tabel spesifikasi adalah tabel yang berguna untuk menjaga agar tes yang kita susun tidak menyimpang dari bahan atau materi dan aspek kejiwaan yang dicakup dalam tes. Dalam membuat tabel spesifikasi langkah pertama yang harus diambil adalah mendaftar pokok-pokok materi yang akan diteskan kemudian memberikan imbang bobot untuk masing-masing pokok materi.

Contoh:

**Tabel Spesifikasi**

<b>Aspek yang Diungkap</b> <b>Pokok Materi</b>	<b>Ingatan</b>	<b>Pemahaman</b>	<b>Aplikasi</b>	<b>Jumlah</b>
Bagian I	.....	.....	.....	.....
Bagian II	.....	.....	.....	.....
Bagian n (terakhir)	.....	.....	.....	.....
Jumlah	.....	.....	.....	.....

Tabel spesifikasi mempunyai kolom dan baris, sehingga tampak hubungannya antara materi dengan aspek yang tergambar dengan TIK. Sebenarnya penyusun tes bukan hanya mengingat hubungan antara dua hal tersebut tetapi empat hal, yaitu hubungan antara materi, TIK, kegiatan belajar dan evaluasi.

#### **Langkah-langkah Pembuatan Tabel Spesifikasi**

Sebenarnya ada beberapa macam tabel spesifikasi. Macam tabel ini ditentukan oleh bidang studi dan homogenitas materi yang akan diteskan. Langkah-langkah dalam membuat tabel spesifikasi yaitu :

- Langkah pertama yang harus dilakukan adalah mendaftarkan pokok-pokok materi yang akan diteskan kemudian memberi imbalan bobot untuk masing-masing pokok materi.
- Langkah kedua yaitu memasukkan pokok-pokok materi ke dalam tabel dan merubah indeks menjadi persentase.
- Langkah ketiga yaitu merinci banyaknya butir soal untuk tiap-tiap pokok materi, dan angka ini dituliskan pada kolom paling kanan.

Caranya adalah dengan membagi jumlah butir soal menjadi 4 bagian berdasarkan imbang bobot yang tertera sebagai persentase.

- Untuk langkah selanjutnya, tergantung homogenitas dan heterogenitas (keragaman) materi yang diteskan.

Contoh : **Tabel Spesifikasi Untuk Soal Evaluasi**

<b>Aspek yang Diungkap Pokok Materi</b>	<b>Ingatan</b>	<b>Pemahaman</b>	<b>Aplikasi</b>	<b>Jumlah</b>
Pengertian Evaluasi (14%)				7
Fungsi Evaluasi (21%)				10
Macam-macam cara evaluasi (36%)				18
Persyaratan Evaluasi (29%)				15
<b>Jumlah</b>				<b>50 butir soal</b>

### **Rangkuman**

- Kisi-kisi merupakan suatu format atau matriks yang memuat kriteria tentang soal-soal yang diperlukan atau yang hendak disusun. Kisi-kisi juga dapat diartikan *test blue-print* atau *table of specification* merupakan deskripsi kompetensi dan materi yang akan diujikan.
- Tujuan penyusunan kisi-kisi adalah untuk menentukan ruang lingkup dan sebagai petunjuk dalam menulis soal.
- Kisi-kisi yang baik harus memenuhi persyaratan berikut ini :
  1. Kisi-kisi harus dapat mewakili isi silabus/kurikulum atau materi yang telah diajarkan secara tepat dan proporsional.
  2. Komponen-komponennya diuraikan secara jelas dan mudah dipahami.
  3. Materi yang hendak ditanyakan dapat dibuatkan soalnya.

## BAB XI VALIDITAS INSTRUMEN ASSESMENT DAN EVALUASI

### A. Pengertian Uji Validitas Instrumen

Adapun Pengertian atau defenisi validitas instrumen menurut para ahli adalah merupakan arti seberapa besar ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Dengan kata lain suatu tes atau instrumen dapat dikatakan mempunyai validitas tinggi apabila alat tersebut menjalankan fungsi ukurnya atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran.

Menurut Ancok (1993) validitas instrumen merupakan suatu indeks yang menunjukkan sejauh mana suatu alat ukur betul-betul mengukur apa yang perlu diukur. Untuk itu dilakukan analisis item dengan metode korelasi *product moment pearson*. Uji validitas dengan metode ini dilakukan dengan cara mengkorelasikan skor jawaban yang diperoleh pada masing-masing item dengan skor total dari keseluruhan item. Hasil korelasi tersebut harus signifikan berdasarkan ukuran statistik tertentu. Koefisien korelasi yang tinggi menunjukkan kesesuaian antara fungsi item dengan fungsi ukur secara keseluruhan atau dengan kata lain instrumen tersebut valid. Instrumen tersebut dikatakan valid jika dapat mengukur apa yang seharusnya diukur atau mengukur apa yang diinginkan dengan tepat (Supranto,1997). Dalam pengujian validitas, instrumen diuji dengan menghitung koefisien korelasi antara skor item dan skor totalnya dalam taraf signifikansi 95% atau  $\alpha = 0,05$ . Instrumen dikatakan valid mempunyai nilai signifikansi korelasi  $\geq$  dari 95% atau  $\alpha = 0,05$ . (Santoso, 2002). Uji validitas dilakukan dengan menggunakan koefisien korelasi *product moment* kriteria pengujian yang digunakan pada instrumen yang dikatakan valid jika nilai  $r \geq 0,30$  (*cut of point*) Sugiyono, (2001).

Data yang valid dan instrumen penelitian yang valid merupakan hal yang berbeda. Uji validitas instrumen berarti yang di uji validitas nya adalah instrumen penelitiannya. Instrumen penelitian merupakan alat ukur pada penelitian, jadi instrumen yang valid berarti alat ukur penelitian yang digunakan tsb sudah valid (dapat mengukur dengan benar dan semestinya). Misalnya : untuk mengukur waktu maka alat ukurnya harus benar dan dapat menjalankan fungsi ukurnya dengan tepat. Kalau alat ukurnya sudah menjalankan fungsi ukurnya dengan tepat, maka alat ukurnya sudah valid (contoh alat ukur waktu = jam dinding).

Uji validitas instrumen merupakan salah satu uji yang disyaratkan untuk melihat kualitas dari instrumen penelitian. Pengujian validitas instrumen dimaksudkan untuk mengetahui sejauh mana ketepatan dan kecermatan instrumen penelitian dalam melakukan fungsi ukurnya.

## **B. Pengertian Validitas Instrumen**

Validitas instrumen didefinisikan sebagai tingkat ketepatan dan kecermatan suatu alat ukur (instrumen penelitian) dalam melakukan fungsi ukurnya. Catatan : dalam menentukan validitas, haruslah disesuaikan dengan tujuan yang ingin dicapai, sebab suatu alat ukur biasanya hanya merupakan ukuran yang valid untuk satu tujuan yang spesifik. Mengenai ini, ada beberapa macam validitas instrumen untuk tujuan spesifik tersebut.

### **Mengapa uji validitas instrumen diperlukan ?**

Instrumen yang valid merupakan syarat diperolehnya hasil penelitian (data) yang valid. Walaupun belum tentu juga hasil (data) penelitian akan otomatis valid setelah instrumen dinyatakan valid. Instrumen penelitian ada yang sudah baku (standar) sehingga validitasnya sudah diakui. Namun banyak instrumen yang digunakan pada penelitian belum teruji validitasnya, oleh karena itulah uji validitas perlu dilakukan.

Persoalan alat ukur yang digunakan mahasiswa ketika melakukan pencarian data penelitian sering dihadapkan pada persoalan akurasi, konsisten dan stabilitas sehingga hasil pengukuran yang diperoleh bisa mengukur dengan akurat sesuatu yang sedang diukur. Instrumen ini memang harus memiliki akurasi ketika digunakan. Konsisten dan stabil dalam arti tidak mengalami perubahan dari waktu pengukuran satu ke pengukuran yang lain. Fenomena ini merupakan titik awal dari aksi penelitian mahasiswa yang mau tidak mau harus dihadapi ketika mahasiswa akan menyelesaikan tugas-tugas kuliah dalam jenjang pendidikan Strata 1. Mereka akan berhadapan dengan cara bagaimana membuat alat ukur, atau instrumen itu memiliki validitas dan reliabilitas agar bisa digunakan dalam memperoleh data. Karena data yang kurang memiliki validitas dan reliabilitas, akan menghasilkan kesimpulan yang kurang lazim.

Data yang kurang memiliki validitas dan reliabilitas, akan menghasilkan kesimpulan yang bias, kurang sesuai dengan yang seharusnya, dan bahkan bisa saja bertentangan dengan kelaziman. Untuk membuat alat ukur instrumen itu, diperlukan kajian teori, pendapat para ahli serta pengalaman-pengalaman yang kadangkala diperlukan bila definisi operasional variabelnya tidak kita temukan dalam teori. Alat ukur atau instrumen yang akan disusun itu tentu saja harus memiliki validitas dan reliabilitas, agar data yang diperoleh dari alat ukur itu bisa reliabel, valid dan disebut dengan validitas dan reliabilitas alat ukur atau validitas dan reliabilitas instrumen.

### C. Validitas Instrumen

Fenomena kedua setelah mahasiswa menguji reliabilitas alat ukur, pekerjaan berikutnya adalah bagaimana memperoleh instrumen yang memiliki validitas, agar data yang diperoleh dari penyebaran instrumen itu benar-benar valid. Dikatakan memiliki validitas adalah bila instrumen atau alat ukur yang dibuat bisa dengan tepat mengukur objek yang akan diukur. Misalnya saja untuk mengukur, panjang dan lebar lapangan bola volley agar sama dengan luas standar lapangan internasional, maka sebaiknya kita menggunakan meteran. Karena meteran adalah merupakan alat ukur yang valid dan sudah memiliki validitas. Selain meteran, alat ukur untuk mengukur panjang dan lebar benda tidak lazim digunakan. Dalam teori ada bermacam-macam validitas.

Menurut Ebel, dalam Moh. Nazir, 2003, Validitas dibagi menjadi *concurrent validity* (validitas konsuren), *construct validity* (validitas konstruk), *face validity* (validitas rupa), *factorial validity* (validitas faktorial), *empirical validity* (validitas empiris), *intrinsic validity* (validitas intrinsik), dan *predictive validity* (validitas prediksi).

Sementara itu, Anastasi, 1973 dan Nunnally, 1979, dalam Masri 2006, validitas alat pengumpul data dapat digolongkan dalam beberapa jenis, yakni validitas konstruk (*construct validity*), validitas isi (*content validity*), validitas prediktif (*predictive validity*), validitas eksternal (*external validity*), dan validitas rupa (*face validity*). Selain itu ada beberapa jenis validitas lain yang tidak dibicarakan, tetapi relatif penting dilakukan bagi penelitian di Indonesia yang memiliki jumlah penduduk yang besar dan memiliki aneka ragam budaya. Validitas tersebut adalah validitas budaya (*cross-cultural validity*).

Sedangkan menurut Suharsimi, 2003, berdasarkan cara pengujiannya, terdapat dua validitas, yakni validitas eksternal dan validitas internal, selain itu validitas dikelompokkan menjadi beberapa kriteria, yakni, validitas konstruk, validitas isi, dan validitas prediksi. Pengertian *concurrent validity* adalah validitas yang berkaitan dengan hubungan (korelasi) antara skor dalam item instrumen dengan kinerja, atau objek penelitian yang lain.

*Construct validity* atau validitas konstruk, adalah bila kita mendefinisikan validitas sebagai kualitas psikologi apa yang diukur oleh sebuah pengujian, serta menilainya dengan memperlihatkan bahwa konstruk tertentu yang bisa diterangkan, dapat menyebabkan penampilan baik buruknya (*performance*) dalam pengujian.

*Face validity* atau validitas rupa yang berhubungan apa yang kelihatan dalam mengukur sesuatu, tetapi bukan terhadap apa yang seharusnya akan diukur.

*Factorial validity* dari sebuah alat ukur adalah korelasi antara alat ukur dengan faktor-faktor yang bersamaan dalam suatu kelompok atau ukuran-ukuran perilaku lainnya. Validitas ini biasanya diperoleh dengan menggunakan teknik analisis faktor.

*Empirical validity* adalah validitas empiris yang berkaitan dengan hubungan antara skor dengan suatu kriteria, dimana kriteria itu adalah merupakan ukuran yang bebas dan langsung berhubungan dengan apa yang ingin diramalkan oleh pengukuran.

*Intrinsic validity* adalah validitas yang berkaitan dengan penggunaan teknik uji coba untuk memperoleh fakta kuantitatif dan objektif untuk. Teknik uji coba itu yang dilakukan untuk mendukung bahwa instrumen yang digunakan sebagai alat ukur adalah benar-benar mengukur apa yang seharusnya diukur.

*Predictive validity* adalah validitas perkiraan yang berkenaan dengan hubungan antara skor suatu alat ukur dengan kinerja atau seseorang di masa mendatang berdasarkan pengukuran awal. Validitas prediksi adalah validitas instrumen yang diharapkan bisa memiliki hubungan dengan hasil yang diharapkan dari instrumen yang dibuat. Misalnya instrumen yang ditujukan terhadap mahasiswa baru. Bila jawaban responden (mahasiswa baru) memiliki hubungan dengan prestasi belajar mahasiswa ketika mengikuti kuliah mulai dari semester awal sampai semester akhir, berarti instrumen itu memiliki validitas prediksi yang tinggi. Sebaliknya jika instrumen yang dibuat dan ditujukan

terhadap mahasiswa baru itu tidak memiliki hubungan dengan prestasi belajar mahasiswa mulai dari semester awal hingga semester akhir, berarti instrumen itu memiliki validitas prediksi yang rendah.

*Content validity* adalah validitas yang berkaitan dengan baik atau buruknya sampel yang diambil dari populasi.

*Curricular validity* adalah validitas yang ditentukan oleh bagaimana cara peneliti memilih isi dari pengukuran dan menilai seberapa jauh pengukuran yang dilakukan itu adalah merupakan alat ukur yang benar-benar mengukur aspek-aspek sesuai dengan tujuan instruksional. Masri, 2006, validitas terbagi dalam beberapa kriteria antara lain, validitas konstruk, validitas isi, dan validitas prediksi.

Validitas isi adalah validitas instrumen yang memiliki kandungan isi butir-butir item pertanyaan yang dibuat sesuai dengan topik penelitian dan bisa menggali jawaban responden sesuai dengan permasalahan yang sudah dirumuskan oleh peneliti.

Validitas konstruk adalah validitas yang berkenaan dengan kualitas dalam aspek psikologis tentang apa yang diukur oleh suatu pengukuran serta terdapat evaluasi bahwa suatu konstruk tertentu itu bisa menyebabkan kinerja dan hasil yang baik dalam pengukuran. Validitas prediksi adalah validitas yang berkenaan dengan hubungan antara skor suatu alat ukur yang sudah dilakukan pada periode awal terhadap seseorang dengan kinerja seseorang pada waktu yang akan datang. Validitas ini biasanya dilakukan ketika ada rekrutmen siswa atau pegawai lembaga tertentu.

#### **D. Validitas dan Reliabilitas Penilaian**

Sebelum aksi penilaian dilakukan, penilai perlu membedakan kriteria tentang validitas dan reliabilitas hasil penelitian dengan validitas dan reliabilitas instrumen. Hasil penilaian yang valid dan reliabel dengan instrumen yang valid dan reliabel, Sugiono, 2005, merupakan hasil penelitian yang memiliki kesamaan antara data yang terkumpul dengan data yang sesungguhnya terjadi pada objek yang diteliti. Misalnya saja warna dalam objek yang berwarna merah, akan tetapi data yang terkumpul memberikan data berwarna lain, maka hasil penelitiannya tidak valid. Sedangkan hasil penelitian yang reliabel, diperoleh bila terdapat kesamaan data yang terkumpul dalam kurun waktu yang berbeda, ataupun pengukuran yang dilakukan secara berulang. Misalnya warna dalam obyek beberapa waktu lalu berwarna merah, maka pada saat

ini dan besok tetap berwarna merah. Jika kita memperoleh data tentang jumlah mahasiswa Fakultas Ekonomi dua hari yang lalu adalah 50 mahasiswa, maka jumlah mahasiswa pada hari ini dan besok adalah sebanyak 50 mahasiswa, demikian seterusnya.

**Rangkuman**

- Uji Validitas Instrumen menurut para ahli merupakan arti seberapa besar ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Dengan kata lain suatu tes atau instrumen dapat dikatakan mempunyai validitas tinggi apabila alat tersebut menjalankan fungsi ukurnya atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran.
- Uji validitas instrumen merupakan salah satu uji yang disyaratkan untuk melihat kualitas dari instrumen penelitian. Pengujian validitas instrumen dimaksudkan untuk mengetahui sejauh mana ketepatan dan kecermatan instrumen penelitian dalam melakukan fungsi ukurnya.
- Validitas instrumen didefinisikan sebagai tingkat ketepatan dan kecermatan suatu alat ukur (instrumen penelitian) dalam melakukan fungsi ukurnya.
- Dalam menentukan validitas, haruslah disesuaikan dengan tujuan yang ingin dicapai, sebab suatu alat ukur biasanya hanya merupakan ukuran yang valid untuk satu tujuan yang spesifik. Mengenai ini, ada beberapa macam validitas instrumen untuk tujuan spesifik tersebut.

## **BAB XII**

### **RELIABILITAS INSTRUMEN ASSEMEN DAN EVALUASI**

#### **A. Pengertian Reliabilitas**

Reliabilitas adalah indeks yang menunjukkan sejauh mana suatu alat pengukur dipercaya atau dapat diandalkan. Bila suatu alat ukur dipakai dua kali untuk mengukur gejala yang sama dan hasil pengukuran yang diperoleh relatif konsisten, maka alat ukur tersebut reliabel. Atau dengan kata lain, reliabilitas menunjukkan konsistensi suatu alat ukur di dalam mengukur gejala yang sama.

#### **B. Reliabilitas Instrumen**

Pengertian Reliabilitas, Sugiono 2005, adalah serangkaian pengukuran atau serangkaian alat ukur yang memiliki konsistensi bila pengukuran yang dilakukan dengan alat ukur itu dilakukan secara berulang. Kondisi itu ditengarai dengan konsistensi hasil dari penggunaan alat ukur yang sama yang dilakukan secara berulang dan memberikan hasil yang relatif sama dan tidak melanggar kelaziman.

Untuk pengukuran subjektif, penilaian yang dilakukan oleh minimal dua orang bisa memberikan hasil yang relatif sama (reliabilitas antar penilai). Pengertian Reliabilitas tidak sama dengan pengertian validitas. Artinya pengukuran yang memiliki reliabilitas dapat mengukur secara konsisten, tapi belum tentu mengukur apa yang seharusnya diukur.

Kata reliabilitas dalam bahasa Indonesia yang digunakan saat ini, sebenarnya diambil dari kata *reliability* dalam bahasa Inggris dan berasal dari kata *reliable* yang artinya dapat dipercaya, keajegan, konsisten, keandalan, kestabilan. Suatu tes dapat dikatakan reliabel jika tes tersebut menunjukkan hasil yang dapat dipercaya dan tidak bertentangan.

Menurut Sugiono (2005) Reliabilitas adalah serangkaian pengukuran atau serangkaian alat ukur yang memiliki konsistensi bila pengukuran yang dilakukan dengan alat ukur itu dilakukan secara berulang. Reliabilitas tes adalah tingkat keajegan (konsistensi) suatu tes, yakni sejauh mana suatu tes dapat dipercaya untuk menghasilkan skor yang ajeg, relatif tidak berubah walaupun diteskan pada situasi yang berbeda-beda. Sedangkan Sukadji (2000) mengatakan bahwa reliabilitas suatu tes adalah seberapa besar derajat tes mengukur secara konsisten sasaran yang diukur. Reliabilitas dinyatakan dalam bentuk angka, biasanya sebagai koefisien. Koefisien tinggi berarti reliabilitas tinggi.

### C. Pengukuran Reliabilitas

Sifat reliabilitas dari sebuah instrumen berhubungan dengan sejauh mana kemampuan alat ukur itu memberikan hasil yang konsisten dari satu even percobaan ke even percobaan lainnya. Jika konsistensi pengukuran itu tidak kita peroleh dalam setiap pengukuran, dapat dibayangkan bila pengukuran yang dilakukan dengan instrumen itu memberikan hasil yang berbeda dari pengukuran satu ke pengukuran berikutnya. Saat ini kita memperoleh hasil pengukuran berat badan seseorang adalah 70 kg. Beberapa saat kemudian, meskipun dengan alat ukur yang sama kita memperoleh hasil 73 kg. Demikian seterusnya, hasilnya tidak pernah konsisten. Data yang kita peroleh tidak pernah konsisten dari waktu ke waktu. Pertanyaan yang akan muncul dari benak kita adalah hasil pengukuran mana yang kita gunakan ?

Dalam kajian teoritis, reliabilitas adalah sejauh mana pengukuran dari suatu uji coba yang dilakukan tetap memiliki hasil yang sama meskipun dilakukan secara berulang-ulang terhadap subjek dan dalam kondisi yang sama. Instrumen alat ukur dianggap bisa diandalkan apabila memberikan hasil yang konsisten untuk pengukuran yang sama dan tidak bisa diandalkan bila pengukuran yang dilakukan secara berulang-ulang itu memberikan hasil yang relatif tidak sama. Pengujian reliabilitas instrumen untuk memperoleh hasil yang reliabel bisa dilakukan dengan berbagai metode statistik.

Contoh lain adalah misalnya saja dalam sebuah kesempatan kita ingin mengukur panjang dan lebar tiga (3) buah lapangan bola *volley*. Alat yang digunakan dalam pengukuran itu adalah meteran dan jangkauan langkah. Setelah dilakukan pengukuran, bisa dipastikan bahwa pengukuran yang dilakukan dengan menggunakan meteran memperoleh hasil panjang dan lebar yang relatif sama terhadap ketiga lapangan bola *volley* itu. Sedangkan pengukuran yang dilakukan dengan menggunakan jangkauan langkah terhadap ketiga lapangan bola *volley* itu, menghasilkan satuan ukur, yakni panjang dan lebar yang berbeda.

#### **D. Teknik untuk Menentukan Reliabilitas**

Ada beberapa teknik untuk mengukur reliabilitas, antara lain :

##### **1. Teknik Pengukuran Ulang**

Teknik ini dilakukan dengan cara mengadakan pengukuran ulang kepada responden, kita meminta responden yang sama agar menjawab semua pertanyaan dalam alat pengukur sebanyak dua kali. Selang waktu antara pengukuran pertama dan ke dua menurut Masri Singarimbun antara 15 s/d 30 hari, apa bila selang waktunya terlalu dekat dikhawatirkan responden masih ingat jawaban yang diberikan pada waktu yang pertama.

Hasil pengukuran pertama dan kedua kemudian dikorelasikan dengan teknik korelasi "*product moment*", kemudian dianalisa seperti dalam teknik validitas.

##### **2. Teknik Belah Dua**

Teknik belah dua, yaitu dengan membagi instrumen menjadi dua bagian misal ganjil genap.

##### **3. Teknik Bentuk Paralel**

Teknik bentuk paralel, yaitu dilakukan dengan menggunakan dua alat ukur yang mengukur aspek yang sama.

## **Rangkuman**

- Reliabilitas adalah indeks yang menunjukkan sejauh mana suatu alat pengukur dipercaya atau dapat diandalkan.
- Bila suatu alat ukur dipakai dua kali untuk mengukur gejala yang sama dan hasil pengukuran yang diperoleh relatif konsisten, maka alat ukur tersebut reliabel. Atau dengan kata lain, reliabilitas menunjukkan konsistensi suatu alat ukur di dalam mengukur gejala yang sama.
- Pengertian Reliabilitas (Sugiono 2005), adalah serangkaian pengukuran atau serangkaian alat ukur yang memiliki konsistensi bila pengukuran yang dilakukan dengan alat ukur itu dilakukan secara berulang.
- Pengertian Reliabilitas tidak sama dengan pengertian validitas. Artinya pengukuran yang memiliki reliabilitas dapat mengukur secara konsisten, tapi belum tentu mengukur apa yang seharusnya diukur.
- Dalam kajian teoritis, reliabilitas adalah sejauh mana pengukuran dari suatu uji coba yang dilakukan tetap memiliki hasil yang sama meskipun dilakukan secara berulang-ulang terhadap subjek dan dalam kondisi yang sama.

## **BAB XIII**

### **SKALA PENGUKURAN**

Sebelum kita membahas skala pengukuran, ada baiknya kita mengetahui definisi pengukuran. Pengukuran merupakan penggunaan aturan untuk menetapkan bilangan pada obyek atau peristiwa. Dengan kata lain, pengukuran memberikan nilai-nilai variabel dengan notasi bilangan. Setelah kita mengetahui apa itu pengukuran maka sekarang dapat dijelaskan skala pengukuran. Skala pengukuran merupakan aturan-aturan yang diperlukan untuk mengkuantitatifkan data dari pengukuran suatu variabel. Dalam statistik, data merupakan karakteristik, simbol atau angka dari sebuah variabel yang diukur. Pengukuran hanya dilakukan terhadap variabel yang dapat didefinisikan seperti minat, kinerja ataupun sikap. Agar hasil penelitian tidak memberikan interpretasi yang berbeda maka definisi operasional terhadap variabel yang diteliti perlu dijelaskan terlebih dahulu. Dalam melakukan operasi pada statistik perbedaan data dan tujuan dari data yang tersaji tidak bisa dilakukan dalam dalam model skala pengukuran yang sama. Secara umum ada 4 tingkatan skala pengukuran yaitu skala nominal, skala ordinal, skala variabel dan skala rasio.

#### **A. Tipe Skala Pengukuran**

Skala pengukuran merupakan keseluruhan yang digunakan sebagai acuan untuk menentukan panjang pendeknya interval yang ada dalam alat ukur, sehingga alat ukur tersebut bila digunakan dalam pengukuran akan menghasilkan data kuantitatif.

Skala pengukuran atau aras pengukuran memiliki empat tipe, sebagaimana dikembangkan konsepnya oleh seorang psikolog bernama Stanley Smith Stevens pada artikel di majalah science berkepalanya *On the theory of scales of measurements*. Pada artikel ini, Stevens mengemukakan bahwa dalam sains dikenal empat tipe skala pengukuran yang masing-masing disebutnya sebagai skala nominal, ordinal, interval, dan rasio.

#### **1. Tipe skala**

Stevens (1946, 1951) mengajukan konsep bahwa skala pengukuran dapat dibedakan menjadi empat kelompok berdasarkan sifat bawaannya.

Sebagai akibat dari sifat itu, terdapat perbedaan dalam melakukan analisis terhadap pengamatan-pengamatan yang dilakukan dalam sains.

Tipe skala	Statistik yang layak	Transformasi skala yang diakui	Struktur matematis
nominal (juga kategoris)	modus, Khi-kuadrat	Pasangan satu-satu (equality (=))	himpunan yang tidak dapat diurutkan
Ordinal	median, percentile	Monotonic increasing (order (<))	totally ordered set
Interval	mean, standard deviation, correlation, regression, analysis of variance	Positive (affine)	linear affine line
Ratio	All statistics permitted for interval scales plus the following: geometric mean, harmonic mean, coefficient of variation, logarithms	Positive similarities (multiplication)	one-dimensional vector space

## 2. Skala Nominal

Dalam keempat tingkatan skala pengukuran, skala nominal merupakan tingkatan terendah karena skala ini hanya digunakan untuk membedakan satu objek dengan objek yang lainnya berdasarkan lambang yang diberikan. Sebelum memakai skala nominal biasanya data yang sudah diberi simbol dipisahkan dan dikelompokan berdasarkan jenis atau beberapa kategori pembeda anatara data tersebut. Biasanya lambang yang digunakan adalah suatu gambar yang mencirikan jenis data tersebut, namun terkadang simbol yang diberikan berupa angka atau sebarang bilangan (dengan catatan bilangan yang digunakan hanya digunakan sebagai lambang dari suatu kategori tidak memiliki arti numerik). Hal ini dimaksudkan pada angka atau sembarang bilangan tersebut tidak boleh melakukan operasi aritmatika (tidak boleh menjumlahkan, mengurangi, mengkalikan, membagi). Bilangan atau sembarang angka dalam hal ini hanya difungsikan sebagai lambang atau simbol saja dengan fungsi untuk membedakan satu data dengan data yang lainnya.

Contoh : Berikut merupakan data mengenai mahasiswa yang mendapatkan beasiswa di Universitas Pendidikan Ganesha. Mahasiswa yang tidak mendapatkan beasiswa dilambangkan atau diberi simbol berupa angka 0 (nol), sedangkan mahasiswa yang mendapatkan beasiswa diberi simbol angka 1 (satu). Dalam hal ini tidak bisa diartikan mahasiswa yang mendapatkan beasiswa lebih besar tingkatannya daripada mahasiswa yang tidak mendapatkan beasiswa, angka satu dalam hal ini hanyalah menyatakan lambang untuk mahasiswa yang mendapatkan beasiswa sedangkan angka nol hanyalah sebagai lambang untuk mahasiswa yang tidak mendapatkan beasiswa.

Mengenai perincian diatas dapat ditarik kesimpulan bahwa bilangan atau angka yang digunakan sebagai lambang atau simbol suatu kategori dalam skala nominal hanya berfungsi untuk membedakan data yang satu dengan yang lainnya serta dalam penggunaannya tidak berlaku operasi hitung atau aritmatika (menjumlahkan, mengurangkan, mengkalikan, ataupun membagi). Hubungan yang membatasi antara angka atau bilangan yang digunakan sebagai simbol hanyalah sama dengan dan tidak sama dengan.

Skala Nominal merupakan skala yang semata-mata hanya untuk memberikan indeks, atau nama saja dan tidak mempunyai makna yang lain. Contoh:

<b>Data</b>	<b>Kode (a)</b>	<b>Kode (b)</b>
Yuni	1	4
Desi	2	2
Ika	3	3
Astuti	4	1

Keterangan :

Kode 1 sampai dengan 4

(a) semata-mata hanyalah untuk memberi tanda saja, dan tidak dapat dipergunakan sebagai perbandingan antara satu data dengan data yang lain. Kode tersebut dapat saling ditukarkan sesuai dengan keinginan peneliti (menjadi alternatif b) tanpa mempengaruhi apa pun.

### 3. Skala Ordinal

Satu tingkat diatas skala pengukuran nominal adalah skala pengukuran ordinal. Ini dikarenakan skala pengukuran ordinal masih memiliki ciri pengukuran nominal yaitu membedakan data dalam berbagai kelompok menurut lambang yang diberikan pada populasi atau sampel, namun dalam skala pengukuran ordinal data yang dibedakan menurut lambang tersebut ditambah dengan pembeda yang lain yaitu memiliki pengertian lebih terhadap data yang lain (lebih bagus x lebih jelek, lebih tinggi x lebih rendah dan yang lainnya). Maka dalam skala pengukuran ordinal memungkinkan data untuk diranking.

Contoh : Dalam seri motoGP pembalap Valentino Rossi menduduki posisi terdepan, dan diikuti oleh Cassi Stoner sedangkan diurutan yang ketiga dan keempat masih tetap dipegang oleh Stoner Jambot dan Serul Van Ladret.

Maka dapat diranking yaitu :

1. Valentino Rossi
2. Cassi Stoner
3. Stoner jambot
4. Serul Van Ladret

Hal ini berbeda dengan skala pengukuran nominal, dimana angka hanya difungsikan sebagai lambang. Angka yang dideinisikan di skala ordinal tidak semata-mata hanya untuk sebagai lambang yang difungsikan untuk membedakan satu data dengan data yang lainnya namun juga bisa digunakan untuk mengurutkan data sesuai tingkatannya.

Maka dapat disimpulkan dalam pengukuran skala ordinal bilangan yang digunakan dapat difungsikan sebagai :

Lambang untuk membedakan antara satu data dengan data yang lainnya. Untuk mengurutkan data sesuai tingkatannya berdasarkan kualitas dari data tersebut dan yang telah ditentukan baik dari tingkat tinggi ke tempat yang lebih rendah maupun sebaliknya.

Dari data diatas dapat disimpulkan pada tingkat skala pengukuran skala ordinal kita bisa mengatakan suatu data lebih baik, lebih buruk, lebih besar ataupun lebih kecil dari data yang lainna berdasarkan informasi yang diberikan, Namun dalam hal ini kita tidak mengetahui berapa kali lebih besarnya ataupun lebih buruknya suatu data dari data yang lainnya.

Skala Ordinal merupakan skala ranking, di mana kode yang diberikan memberikan urutan tertentu pada data, tetapi tidak menunjukkan selisih yang sama dan tidak ada nol mutlak. Contoh :

<b>Data</b>	<b>Skala Kecantikan (a)</b>	<b>Skala Kecantikan (b)</b>
Yuni	4	10
Desi	3	6
Ika	2	5
Astuti	1	1

Skala kecantikan (a) di atas menunjukkan bahwa Yuni paling cantik (dengan skor tertinggi 4), dan Astuti yang paling tidak cantik dengan skor terendah (1). Akan tetapi, tidak dapat dikatakan bahwa Yuni adalah 4 kali lebih cantik dari pada Astuti. Skor yang lebih tinggi hanya menunjukkan skala pengukuran yang lebih tinggi, tetapi tidak dapat menunjukkan kelipatan. Selain itu, selisih kecantikan antara Yuni dan Desi tidak sama dengan selisih kecantikan antara Desi dan Ika meskipun keduanya mempunyai selisih yang sama (1). Skala kecantikan pada (a) dapat diganti dengan skala kecantikan (b) tanpa mempengaruhi hasil penelitian.

Skala nominal dan skala ordinal biasanya mempergunakan analisis statistik non parametrik, contoh : Korelasi Kendall, Korelasi Rank Spearman, Chi Square dll.

#### **4. Skala Interval**

Dua tingkatan diatas skala nominal atau satu tingkatan diatas skala ordinal adalah skala interval. Skala interval merupakan yang sifatnya hampir mirip dengan skala pengukuran nominal maupun ordinal. Namun dalam skala interval terdapat sifat tambahan yang membedakannya dengan skala nominal dan skala ordinal, yaitu dalam skala interval selain kita dapat membedakan data yang satu dengan yang lainnya dengan menggunakan lambang serta dapat merangkingnya, kita juga dapat mengukur perbedaan atau interval atau jarak antara data yang satu dengan yang lainnya.

Contoh : Pada pengukuran suhu dari 5 buah cairan yang berbeda dilaboratorium yang dilakukan oleh seorang profesor didapatkan hasil sebagai berikut :

Cairan A bersuhu 20° celcius, cairan B bersuhu 40° celcius, cairan C bersuhu 55° celcius, cairan D bersuhu 65° celcius dan cairan E bersuhu 70° celcius.

Dari data diatas dapat kita lihat antara satu data dengan data yang lainnya memiliki interval. Data tersebut selain bisa dibedakan, diranking berdasarkan tingkatannya juga dapat diketahui interval dari masing-masing data. Interval merupakan jarak antara satu data dengan data sebelum dan sesudahnya. Kita bisa lihat interval antara data D dengan data E memiliki interval 6° celcius. Namun dalam skala interval tidak bisa kita lakukan suatu perbandingan antara satu data dengan data yang lainnya. Kita tidak bisa mengatakan bahwa suhu cairan benda B yang bersuhu 40° celcius 2 kali lebih panas daripada suhu cairan benda A yang bersuhu 20° celcius. Hal ini dikarenakan skal interval tidak memiliki titik nol mutlak. Titik nol mutlak yang dimaksudkan adalah benda yang bersuhu 0° celcius bukan berarti tidak memiliki panas (titik nolnya tidak bernilai kosong). Titik nol yang ditetapkan adalah berdasarkan perjanjian.

Maka dapat disimpulkan bilangan yang digunakan dalam skala interval memiliki 3 fungsi, yakni sebagai berikut :

- Sebagai lambang yang berfungsi sebagai pembeda antara satu data dengan data yang lainnya.
- Untuk mengurutkan data sesuai tingkatannya atau meranking suatu data menurut tingkatannya.
- Untuk mengetahui interval antara satu data dengan data yang lainnya ( jarak antara satu data dengan data yang lainnya).

Skala Interval merupakan skala pengukuran yang mempunyai selisih sama antara satu pengukuran dengan pengukuran yang lain, tetapi tidak memiliki nilai nol mutlak.

Contoh :

<b>Data</b>	<b>Nilai Mata Kuliah (a)</b>	<b>Skor Nilai Mata Kuliah (b)</b>
Yuni	A	4
Desi	B	3
Ika	C	2
Astuti	D	1

Tabel di atas menunjukkan bahwa nilai A setara dengan 4, B setara dengan 3, C setara dengan 2 dan D setara dengan 1. Selisih antara nilai A dan B adalah sama dengan selisih antara B dan C dan juga sama persis

dengan selisih antara nilai C dan D. Akan tetapi, tidak boleh dikatakan bahwa Yuni adalah empat kali lebih pintar dibandingkan Astuti, atau Ika dua kali lebih pintar dari pada Astuti. Meskipun selisihnya sama, tetapi tidak mempunyai nilai nol mutlak.

## 5. Skala Rasio

Tiga tingkat diatas skala nominal atau dua tingkat diatas skala ordinal atau satu tingkat diatas skal interval merupakan skala rasio. Skala rasio memiliki ketiga sifat yang dimiliki oleh skala nominal, skala ordinal, skala interval. Selain kita bisa membedakan satu data dengan data yang lainnya karena lambang atau simbol yang diberikan, mengurutkan data berdasarkan tingkatannya dan mengetahui interval antara satu data dengan data yang lainnya dalam skala rasio kita juga bisa membandingkan anntara satu data dengan data yang lainnya berdasarkan kuantitatif nilai yang dimiliki oleh data.

Contoh : Dalam suatu keluarga terdapat Ayah, Ibu, Kakak, dan Adik. Ayah memiliki tinggi badan 172 sentimeter sedangkan Ibu hanya 165 sentimeter serta tinggi badan kakak dan adik masing-masing adalah 162 sentimeter dan 86 sentimeter.

Dalam skala rasio kita bisa menyimpulkan tinggi badan Ayah adalah dua kali lipat dari tinggi badan adik. Berbeda dalam skala interval, pada skala pengukuran rasio sudah memiliki titik nol mutlak.

Maka dapat disimpulkan pada skala pengukuran rasio, angka yang berperan didalamnya memiliki fungsi sebagai berikut :

- Sebagai lambang untuk membedakan antara data satu dengan data yang lainnya.
- Untuk mengurutkan peringkat, misal, makin besar bilangannya, peringkat makin tinggi atau sebaliknya.
- Untuk dapat memperlihatkan jarak atau perbedaan antara data yang satu dengan data yang lainnya.
- Rasio (perbandingan) antara satu data dengan data yang lainnya dapat diketahui dan mempunyai arti.
- Titik nol yang digunakan dalam skala rasio merupakan suatu titik mutlak (tidak memiliki nilai) atau bukan berdasarkan perjanjian.
- Skala Fatio merupakan skala pengukuran yang paling tinggi di mana selisih tiap pengukuran adalah sama dan mempunyai nilai nol mutlak.

Contoh :

<b>Data</b>	<b>Tinggi Badan</b>	<b>Berat badan</b>
Yuni	170	60
Desi	160	50
Ika	150	40
Astuti	140	30

Tabel di atas adalah menggunakan skala rasio, artinya setiap satuan pengukuran mempunyai satuan yang sama dan mampu mencerminkan kelipatan antara satu pengukuran dengan pengukuran yang lain. Sebagai contoh : Yuni mempunyai berat badan dua kali lipat berat Astuti, atau, Desi mempunyai tinggi 14,29% lebih tinggi dari pada Astuti.

Skala pengukuran interval dan rasio biasanya dikenai alat statistik parametrik.

## **B. Macam-Macam Data**

### **1. Data Nominal (diskrit),**

Data nominal (diskrit) adalah data yang hanya digolong-golongkan secara terpisah, secara diskrit, atau kategori. Contoh : misalnya dalam suatu kelas setelah dihitung terdapat 50 mahasiswa, terdiri dari 30 pria dan 20 wanita. data nominal/diskrit bukan kontinum.

### **2. Data Kontinum**

Data kontinum adalah data yang bervariasi menurut tingkatan dan ini diperoleh dari hasil pengukuran.

- a. Data Ordinal** : adalah yang berbetuk Ranking atau peringkat.  
contoh : Juara I, II, dan III.

Dalam skala maka jaraknya tidak sama I II III IV V VI 98 93  
76 70 50 70.

- b. Data Interval** : adalah data yang jaraknya sama tetapi tidak mempunyai nilai nol (0) absolut/mutlak.

Contoh : Skala thermometer, walaupun ada nilai 0°C, tetapi tetap ada nilainya. -2 -1 0 1 2 3 4 5 6 7 8 9 10 11 12.

- c. **Data Ratio** : adalah data yang jaraknya sama, dan mempunyai nilai nol mutlak.  
 Contoh : Data tentang berat, panjang dan volume.

### C. Macam-Macam Skala

#### 1. Skala Likert (Skala Sikap)

Skala ini digunakan untuk mengukur sikap, pendapat, dan persepsi. Dengan skala likert, maka variabel akan dijabarkan menjadi dimensi, dan indikator variabel. Indikator variabel sebagai titik tolak untuk menyusun item-item instrumen. Jawaban setiap item instrumen yang menggunakan skala likert mempunyai gradasi yang sangat positif sampai sangat negatif yang dapat berupa kata-kata.

Contoh :

Pernyataan Positif Negatif

- |                              |   |   |
|------------------------------|---|---|
| 1. Sangat Setuju (SS)        | 5 | 1 |
| 2. Setuju (S)                | 4 | 2 |
| 3. Ragu-ragu (RR)            | 3 | 3 |
| 4. Tidak Setuju (TS)         | 2 | 4 |
| 5. Sangat Tidak Setuju (STS) | 1 | 5 |

Contoh:

Jika instrumen tersebut diberikan kepada 100 orang karyawan yang diambil secara random. Setelah dilakukan analisis, misalnya:

25 orang menjawab SS;

40 orang menjawab ST;

5 orang menjawab RR;

20 orang menjawab TS;

10 orang menjawab STS.

Berdasarkan data tersebut 65 orang (65%) karyawan menjawab setuju dan sangat setuju.

Dihitung berdasarkan skoring:

Jumlah skor untuk 25 orang (SS) =  $25 \times 5 = 125$

Jumlah skor untuk 40 orang (ST) =  $40 \times 4 = 160$

Jumlah skor untuk 5 orang (RR) =  $5 \times 3 = 15$

Jumlah skor untuk 20 orang (TS) =  $20 \times 2 = 40$

Jumlah skor untuk 10 orang (STS) =  $10 \times 1 = 10$

Jumlah = 350

Jumlah skor kriterium yang tertinggi =  $5 \times 100 = 500$

Jumlah skor kriterium yang terendah =  $1 \times 100 = 100$

Tingkat persetujuan terhadap metode kerja baru =  $TP = (HS:ST) \times 100\%$   
 $TP = (350:500) \times 100\% = 70\%$ .

Secara kontinum dapat digambarkan sebagai berikut : STS TS RR ST SS  
100 200 300 350 400 500

## 2. Skala Guttman

Skala pengukuran dengan tipe ini, akan didapat jawaban yang tegas yaitu : “ya – tidak”, “benar – salah”, “positif – negatif”, dll. Data yang diperoleh adalah data interval dan rasio dikhotomi (dua alternatif). Pada skala ini ada 2 interval, yaitu setuju dan tidak setuju.

Contoh : Bagaimana pendapat anda, bila orng itu menjabat Kepala Sekolah? a. Setuju b. Tidak Setuju Bentuk instrumen – pilihan ganda dan checklist.

## 3. Semantic Deferential (Osgood)

Skala ini juga digunakan untuk mengukur sikap, hanya bentuknya tidak pilihan ganda atau checklist, tetapi tersusun dalam suatu garis kontinum yang jawabannya sangat positif terletak di bagian kiri garis, atau sebaliknya. Data yang diperoleh adalah interval. Responden dapat memberikan, pada rentang jawaban yang positif sampai dengan negatif. Hal itu tergantung kepada persepsi responden kepada yang dimiliki.

Contoh :

Beri Nilai Gaya Kepemimpinan Kepala Sekolah Anda  
1. Bersahabat 5 4 3 2 1 Tidak bersahabat  
2. Tepat Janji 5 4 3 2 1 Lupa janji  
3. Bersahabat 5 4 3 2 1 Memusuhi  
4. Bersahabat 5 4 3 2 1 Suka marah  
5. Memberikan 5 4 3 2 1 Mendominasi bawahan kepercayaan pada bawahan.

## 4. Rating Scale

*Rating scale* adalah data mentah yang diperoleh berupa angka, kemudian ditafsirkan dalam pengertian kualitatif, sedangkan Likert, Guttman, dan Semantic deferential data kualitatif ke kuantitatif. Responden menjawab, senang atau tidak senang, setuju atau tidak setuju, pernah atau tidak pernah adalah merupakan data kualitatif. Dalam skala model Rating Skala, responden tidak akan menjawab salah satu dari jawaban kualitatif yang telah disediakan, tetapi menjawab salah satu dari jawaban kuantitatif yang telah disediakan. Rating Skala ini lebih fleksibel, tidak terbatas untuk pengukuran sikap saja, tetapi untuk mengukur persepsi,

status skala ekonomi, kelembagaan, pengetahuan, kemampuan, proses kegiatan dan lain-lain.

Contoh :

Seberapa baik tata ruang kantor SMUN VI ?

Berilah jawaban angka 4 bila tata ruang itu sangat baik 3 bila tata ruang itu cukup baik 2 bila tata ruang itu kurang baik 1 bila tata ruang itu sangat tidak baik

Jawaban dengan melingkari nomor jawaban Interval No Pernyataan Tentang Tata Ruang Kantor, Jawaban :

1. Penataan meja kerja sehingga arus kerja menjadi pendek 4 3 2 1
2. Pencahayaan dalam tiap ruangan 4 3 2 1
3. Pencahayaan buatan/listrik tiap ruang sesuai dengan 4 3 2 1
4. Warna lantai sehingga tidak menimbulkan pantulan cahaya 4 3 2 1
5. Sirkulasi udara setiap ruangan 4 3 2 1
6. Keserasian warna alat-alat kantor, perabotan dengan ruangan 4 3 2 1
7. Penempatan lemari arsip 4 3 2 1
8. Penempatan ruang pimpinan 4 3 2 1
9. Meningkatkan keakraban sesama pegawai 4 3 2 1
10. Kebersihan ruangan 4 3 2 1.

Bila instrumen tersebut digunakan sebagai angket dan diberikan pada 30 responden, sebelum dianalisis maka tabulasi data dapat dilakukan seperti dalam halaman berikut.

Jumlah skor kriterium (bila setiap butir mendapat skor tertinggi) =  $4 \times 10 \times 30 = 1200$ . Untuk ini skor tertinggi = 4, jumlah butir = 10 dan jumlah responden = 30. Jumlah skor hasil pengumpulan data = 818. Dengan demikian kualitas tata ruang kantor SMAN VI menurut persepsi 30 responden itu  $818:1200=68\%$  dari kriteria yang ditetapkan. Hal ini secara kontinum dapat dibuat kategori sebagai berikut : 300 600 900 1200 818 Sangat/ tidak/ Kurang/ Cukup /Sangat Baik : Nilai interval 818 termasuk dalam kategori interval kurang baik dan cukup baik, tetapi lebih mendekati cukup baik.

Contoh : Seberapa tinggi pengetahuan anda terhadap mata pelajaran-mata pelajaran berikut sebelum dan sesudah mengikuti pendidikan dan latihan. 0 = bila sama sekali belum tahu 1 = telah mengetahui sampai dengan 25% 2 = telah mengetahui sampai dengan 50% 3 = telah mengetahui sampai

dengan 75% 4 = telah mengetahui 100% (semuanya) Mohon dijawab dengan cara melingkari nomor sebelum dan sesudah latihan pengetahuan.

0 1 2 3 4 Komunikasi 0 1 2 3 4

0 1 2 3 4 Tata Ruang Kantor 0 1 2 3 4

0 1 2 3 4 Pengambilan Keputusan 0 1 2 3 4

0 1 2 3 4 Sistem Pembuatan Laporan 0 1 2 3 4

0 1 2 3 4 Pemasaran 0 1 2 3 4

0 1 2 3 4 Akuntansi 0 1 2 3 4

0 1 2 3 4 Statistik 0 1 2 3 4.

Dengan diketahui pengetahuan sebelum dan sesudah mengikuti diklat, maka pengaruh pendidikan dan latihan dalam menambah pengetahuan para pegawai yang mengikuti diklat dapat diketahui. Data untuk tambahan pengetahuan bisa ratio. Instrumen untuk menjaring data nominal.

Contoh :

a. Berapakah jumlah pegawai di tempat anda bekerja? ..... Pegawai.

b. Berapakah orang yang dapat berbahasa Belanda? ..... Orang.

c. Berapa orangkah pemimpin yang anda sukai? ..... Orang.

d. Berapakah jumlah komputer yang dapat digunakan di lembaga anda?  
..... komputer

e. Darimana anda mengetahui tata kerja yang baru?

*Instrumen untuk menjaring data ordinal*

Contoh :

Berilah ranking ke sepuluh barang berikut yang laris di pasaran. Nama

Barang Ranking Nomor A ..... B ..... C .....

D ..... E ..... F ..... G ..... H

..... I ..... J .....

Misalnya barang D paling laris di pasaran, maka barang tersebut diberi ranking 1.

### **Rangkuman**

- Pengukuran merupakan penggunaan aturan untuk menetapkan bilangan pada obyek atau peristiwa. Dengan kata lain, pengukuran memberikan nilai-nilai variabel dengan notasi bilangan.
  - Skala pengukuran merupakan aturan-aturan yang diperlukan untuk mengkuantitatifkan data dari pengukuran suatu variabel.
  - Dalam statistik, data merupakan karakteristik, simbol atau angka dari sebuah variabel yang diukur. Pengukuran hanya dilakukan terhadap variabel yang dapat didefinisikan seperti minat, kinerja ataupun sikap.
  - Agar hasil penelitian tidak memberikan interpretasi yang berbeda maka definisi operasional terhadap variabel yang diteliti perlu dijelaskan terlebih dahulu.
  - Dalam melakukan operasi pada statistik perbedaan data dan tujuan dari data yang tersaji tidak bisa dilakukan dalam dalam model skala pengukuran yang sama.
  - Secara umum ada 4 tingkatan skala pengukuran yaitu skala nominal, skala ordinal, skala variabel dan skala rasio.
-

## DAFTAR PUSTAKA

- Alwasilah, et al. 1996. *Glossary of educational Assessment Term*. Jakarta: Ministry of Educational and Culture.
- Anas, Sudijono. 1996. *Pengantar Evaluasi Pendidikan*. Jakarta: PT. Raja. Grafindo. Persada.
- Ancok, Djamaludin. 1993. *Teknik Penyusunan Skala Pengukur*. Gadjah Mada University Press.
- Anderson W. Lorin, Krathwohl R. David, 2014. *Kerangka Landasan Untuk Pembelajaran, Pengajaran, dan Asesmen*. Yogyakarta: Pustaka Pelajar.
- Arikunto, S. dan Jabar. 2004. *Evaluasi Program Pendidikan*. Jakarta: Bumi Aksara.
- Arikunto, Suharsimi. 2007. *Dasar-Dasar Evaluasi Pendidikan edisi revisi*. Jakarta: PT. Bumi Aksara.
- Arikunto, Suharsimi. 2003. *Manajemen Penelitian*. Jakarta: Rineka Cipta.
- Bloom, Benjamin S. et. all. 1971. *Handbook on Formative and Summative Evaluation of Student Learning*, Ditempatkan di Perpustakaan FPS IKIP Malang.
- Cangelossi, S. James. 1991. *Evaluating Classroom Instruction*. New York: Longman Publishing Group.
- Depdikbud. 1994. *Kurikulum Pendidikan Dasar (GBPP)*. Depdikbud: Jakarta.
- Djemari, Mardapi. 2004. *Pengembangan Sistem Penilaian Berbasis Kompetensi, Seminar Nasional Rekayasa Sistem Penilaian Dalam Rangka Meningkatkan Kualitas Pendidikan*. HAPY Yogyakarta, 26 s/d 27 Maret 2004.

- Djemari, Mardapi. 1999. *Pengukuran, Penilaian, Dan Evaluasi*. Yogyakarta.
- Ebel, R. L., & Frisbie, D. A. 1986. *Essentials of Educational Measurement*. New Jersey: Prentice Hall Inc. P. 230.
- Echols, Jhon M. Dan Hassan Shadily. 1983. *Kamus Inggris Indonesia*. Jakarta: Gramedia.
- Elliot et al. 2000. *Educational Psychology: Effective Teaching, Effective Learning, 3rd edition*. United States of America: Mc Graw Hill Companies.
- Elliot, J. Andrew and Dweck, S. Carol. 2005. *Handbook of Competence And Motivation*. New York: The Guilford Press.
- Estina Ekawati dan Sumaryanta. 2011. *Pengembangan Instrumen Penilaian Pelajaran Matematika SD/SMP*. Jakarta: Pusat Pengembangan dan Pemberdayaan Pendidik dan Tenaga Kependidikan (PPPPTK) Matematika.
- Gronlund, N.E & R.L Linn. 1990. *Measurement and Evaluation in Teaching*. 6th. Ed. New York: MacMillan Publishing Company.
- H. M. Sukardi,. 2008. *Evaluasi Pendidikan Prinsip dan Oporasionalnya*. Jakarta: PT. Bumi Aksara.
- Herman, J. L. 1997. *Large-scale assessment in support of school reform: Lessons learned in the search for alternative measures. International Journal of Educational Research*, 27, 395-413.
- Herman, J.L., P.R. Aschbacher, dan L. Winters. 1992. *A Paractical Guide to Alternative Assessment*. Alexandria, Va.: Associoation for Supervision and Curriculum Development.
- Kumano, Y. 2001. *Authentic Assessment and Portofolio Assessment-Its Theory and Practice*. Japan: Shizuoka University.

- Lehman, H. 1990. *The System Approach To Education*, Innotech. Publication-Vol 20 NO. 05, Manila.
- Lerner, T. C. 1988. *Learning Disabilities: Theories, Diagnosis, and Teaching Strategies*, Boston: Houghton Mufflin.
- Lidz, Carol. 2003. *Early Childhood Assessment*. New Jersey: Jhon Wiley & Sons.
- M. Ngalim Purwanto. 2002. *Prinsip-Prinsip Dan Teknik Evaluasi Pengajaran*. Bandung: PT Remaja Rosdakarya.
- Mardapi, Djemari. 2004. *Penyusunan Tes Hasil Belajar*. Yogyakarta: Program Pascasarjana Universitas Negeri Yogyakarta.
- Masidjo, Ign., 1995. *Penilaian Pencapaian Hasil Belajar Siswa di Sekolah*. Yogyakarta: Kanisius.
- Masri Singarimbun & Sofian Effendi. 2006. *Metode Penelitian Survei*. Jakarta: Pustaka LP3ES.
- Mehrens, W.A & Lehmann, I. J. 1984. *Measurement and evaluation in education and psychology*, Third edition. New York: Holt, Rinehart and Winston.
- Moh Nazir. 2003. *Metode Penelitian*. Jakarta : Ghalia Indonesia.
- Popham. J. 1995. *Classroom assessment: What teacher need to know?*. New York: Allyn & Bacon A Viacom Company Needham Hights, MA 021194.
- Santoso, Singgih, dan Fandy Tjiptono. 2002. *Riset Pemasaran : Konsep dan Aplikasinya dengan SPSS*. Jakarta: PT Elex Media Computindo Kelompok Gramedia.
- Steven, S. S. 1946. *On the theory of measurement scales*. *Science*, 103, 677-680.

- Steven, S. S. 1951. *Mathematics, measurement, and psychophysics*. Dalam S. Stevens (ed). *Handbook of experimental psychology*. New York: Jhon Wiley (pp. 1-49).
- Stufflebeam, dkk. 1971. *The Relevance of The Cipp Evaluation Model For Educational Accountability*. Ohio State Univ. Colombus: Evaluation Center.
- Sugiyono. 2001. *Metode Penilaian*. Bandung: Alfabeta.
- Sugiyono. 2005. *Memahami Penelitian Kualitatif*. Bandung: CV. Alfabeta.
- Sukadji, Soetarlinah. 2000. *Penyusunan dan Mengevaluasi Laporan Penelitian*. Universitas Indonesia Press: Jakarta.
- Sunardi & Sunaryo. 2006. *Intervensi Dini Anak Berkebutuhan Khusus*. Bandung: Jurusan PLB FIP UPI.
- Supranto. 1997. *Pengukuran tingkat kepuasan pelanggan untuk menaikkan pangsa pasar*. Jakarta: Rineka Cipta.
- Suryabrata, Sumadi. 1987. *Psikologi Pendidikan*. Jakarta: Rajawali.
- Tayibnapis, Farida Yusuf. 2000. *Evaluasi Program*. Jakarta: PT. Rineka Cipta.
- Zainul & Nasution. 2001. *Penilaian Hasil Belajar*. Jakarta: Dirjen Dikti.

## Tentang Penulis



### **Anetha Lyta Flora Tilaar**

lahir di Minahasa pada tanggal 27 September 1960. Lulus S1 pada tahun 1983 di Institut Keguruan Ilmu Pendidikan Manado Jurusan Pendidikan Matematika. Pada tahun 1997 meraih gelar Magister Sains dari Universitas Samratulangi Manado.

Gelar Doktor diperoleh pada tahun 2009 di Universitas Negeri Jakarta. Saat ini penulis aktif sebagai staf pengajar tetap di Fakultas Matematika dan Ilmu Pengetahuan Alam Jurusan Matematika baik di jenjang S1, S2 maupun S3. Selain itu, penulis juga dipercayakan sebagai Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Manado. Adapun penulis telah banyak mempublikasikan hasil-hasil penelitian di berbagai Jurnal Internasional maupun Nasional terakreditasi.

**Penerbit:**

Percetakan Unima Press  
Kampus Unima di Tondano

ISBN 978-602-14611-3-6



9 786021 461136