

Arief Aulia Rahman, M.Pd  
Cut Eva Nasryah, M.Pd



# EVALUASI PEMBELAJARAN



# **EVALUASI PEMBELAJARAN**

**Arief Aulia Rahman, M.Pd**

**Cut Eva Nasryah, M.Pd**

**Uwais Inspirasi Indonesia**

# EVALUASI PEMBELAJARAN

**ISBN** : 978-623-227-231-6  
**Penulis** : Arief Aulia Rahman, M.Pd  
Cut Eva Nasryah, M.Pd  
**Tata Letak** : Funky  
**Design Cover** : Haqi

14,8 cm x 21 cm  
vi + 144 halaman  
Cetakan Pertama, Nopember 2019

Diterbitkan Oleh:

**Uwais Inspirasi Indonesia**

Anggota IKAPI Jawa Timur Nomor: 217/JTI/2019 tanggal 1 Maret 2019

## Redaksi:

Ds. Sidoarjo, Kec. Pulung, Kab. Ponorogo

Email: [Penerbituwais@gmail.com](mailto:Penerbituwais@gmail.com)

Website: [www.penerbituwais.com](http://www.penerbituwais.com)

Telp: 0352-571 892

WA: 0812-3004-1340/0823-3033-5859

Sanksi Pelanggaran Pasal 113 Undang-Undang Nomor 28 tahun 2014 tentang Hak Cipta, sebagaimana yang telah diatur dan diubah dari Undang-Undang nomor 19 Tahun 2002, bahwa:

### Kutipan Pasal 113

- (1) Setiap orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam pasal 9 ayat (1) huruf i untuk penggunaan secara komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp100.000.000,00 (seratus juta rupiah).
- (2) Setiap orang yang dengan tanpa hak dan/atau tanpa izin pencipta atau pemegang hak cipta melakukan pelanggaran hak ekonomi pencipta sebagaimana dimaksud dalam pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h, untuk penggunaan secara komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).
- (3) Setiap orang yang dengan tanpa hak dan/atau tanpa izin pencipta atau pemegang hak melakukan pelanggaran hak ekonomi pencipta sebagaimana dimaksud dalam pasal 9 ayat (1) huruf a, huruf b, huruf e, dan/atau huruf g, untuk penggunaan secara komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan/atau pidana denda paling banyak Rp1.000.000.000,00 (satu miliar rupiah).
- (4) Setiap orang yang memenuhi unsur sebagaimana dimaksud pada ayat (3) yang dilakukan dalam bentuk pembajakan, dipidana dengan pidana penjara paling lama 10 (sepuluh) tahun dan/atau pidana denda paling banyak Rp4.000.000.000,00 (empat miliar rupiah).

## PRAKATA

Alhamdulillah, rasa syukur kehadiran Allah SWT. Sesungguhnya atas berkat rahmat-Nya buku ini selesai disusun dan dapat diterbitkan. Shalawat beriring salam senantiasa disampaikan ke haribaan junjungan Rasulullah Muhammad SAW.

Undang-Undang menyatakan bahwa pendidik adalah tenaga profesional yang bertugas merencanakan dan melaksanakan proses pembelajaran, menilai hasil pembelajaran, melakukan bimbingan dan pelatihan, serta melakukan penelitian dan pengabdian kepada masyarakat. Dengan demikian, salah satu kompetensi yang mesti dimiliki seorang pendidik adalah mampu merancang dan melaksanakan evaluasi, baik terhadap proses maupun hasil pembelajaran.

Buku Evaluasi Pembelajaran ini disusun sebagai bahan ajar bagi mahasiswa di lembaga pendidikan tenaga kependidikan. Penguasaan terhadap materi buku ini diharapkan memberi mereka kemampuan dasar untuk melaksanakan evaluasi pembelajaran.

Ucapan terimakasih kepada orang tua kami tercinta, ayahanda Drs. Ahmad As'adi dan ibunda Dra. Aminah dan juga ayahanda Nasruddin dan ibunda Rawiyah Yusri, yang telah mendoakan dan selalu mendukung untuk terus berkarya. Dan kepada semua pihak yang telah terlibat dalam penyusunan hingga terbitnya buku ini kami haturkan terima kasih.

Kami menyadari masih terdapat kekurangan dalam buku ini untuk itu kritik dan saran terhadap penyempurnaan

buku ini sangat diharapkan. Kiranya karya ini memberi manfaat kepada pembaca, dan menorehkan secercah manfaat bagi perbaikan kualitas mahasiswa calon professional pendidikan.

Meulaboh, November 2019

Penyusun

# DAFTAR ISI

Prakata.....	iii
Daftar Isi.....	v
<b>BAB I KONSEP DASAR EVALUASI.....</b>	<b>1</b>
A. Pengertian Evaluasi, Penilaian, dan Pengukuran ...	2
B. Tujuan dan Fungsi Evaluasi Pembelajaran.....	8
C. Prinsip-Prinsip Evaluasi dan Penilaian Pembelajaran.....	13
D. Jenis Evaluasi Pembelajaran.....	19
<b>BAB II KARAKTERISTIK, MODEL DAN PENDEKATAN EVALUASI PEMBELAJARAN .....</b>	<b>33</b>
A. Karakteristik Instrumen Evaluasi .....	34
B. Model Evaluasi .....	37
C. Pendekatan Evaluasi.....	49
<b>BAB III INSTRUMEN EVALUASI JENIS TES .....</b>	<b>55</b>
A. Tes Tertulis Bentuk Uraian ( <i>Essay</i> ) .....	55
B. Tes Hasil Belajar Bentuk Objektif.....	59
C. Tes Tindakan ( <i>Performance Test</i> ).....	67
<b>BAB IV INSTRUMEN EVALUASI JENIS NON-TES .....</b>	<b>71</b>
A. Daftar Cek.....	71
B. Skala Rentang.....	72
C. Penilaian Sikap.....	73
D. Penilaian Proyek.....	78
E. Penilaian Produk.....	80
F. Penilaian Portofolio .....	82
G. Penilaian diri.....	85

<b>BAB V PENGUKURAN DOMAIN KOGNITIF, AFEKTIF, DAN PSIKOMOTORIK.....</b>	<b>91</b>
A. Pengukuran Ranah Kognitif.....	92
B. Pengukuran Ranah Afektif .....	96
C. Pengukuran Ranah Psikomotorik.....	107
<b>BAB VI ANALISIS INSTRUMEN PENILAIAN .....</b>	<b>113</b>
A. Analisis Logis/Rasional .....	113
B. Analisis Empirik .....	117
C. Reliabilitas tes.....	123
D. Taraf Kesukaran .....	130
E. Daya Pembeda.....	133
Daftar Pustaka .....	141
Tentang Penulis.....	143

# Bab I

## Konsep Dasar Evaluasi

Pendidik harus mampu mengetahui sejauh mana hasil belajar yang telah dilakukan, hasil dari proses pembelajaran dapat berupa kategori baik, tidak baik, bermanfaat, tidak bermanfaat dan lain-lain. Hal ini penting untuk diketahui karena hasil belajar menjadi tolak ukur keberhasilan seorang pendidik atau sebagai alat ukur sejauh mana proses pembelajaran yang ia terapkan dapat mengembangkan potensi peserta didik. Jika hasil belajar baik, maka proses pembelajaran dapat dikatakan berhasil dan demikian sebaliknya. Salah satu cara yang dilakukan untuk mengetahui hasil belajar yang dicapai oleh pendidik dalam proses pembelajaran yaitu mengevaluasi proses dan hasil belajar. Kemampuan mengevaluasi pembelajaran merupakan salah satu indikator yang harus dimiliki oleh seorang pendidik sebab kemampuan melakukan evaluasi pembelajaran merupakan kemampuan dasar yang wajib dimiliki setiap pendidik maupun calon pendidik.



## A. Pengertian Evaluasi, Penilaian, dan Pengukuran.

Secara umum, kebanyakan pendidik maupun calon pendidik mengidentifikasi bahwa kegiatan evaluasi sama halnya dengan melakukan penilaian dan pengukuran, karena aktifitas tersebut sudah terkandung dalam kegiatan evaluasi, namun pengukuran, penilaian dan evaluasi merupakan kegiatan bersifat hierarki dan tidak dapat dipisahkan, kegiatan tersebut harus dilakukan secara berurutan. Untuk lebih memahami tentang evaluasi, penilaian dan pengukuran maka akan disajikan contoh analogi sebagai berikut:

- Seseorang memberikan kepada anda 2 pensil dengan ukuran panjang yang berbeda dan anda diharuskan untuk memilih salah satunya, tentu anda akan memilih pensil yang lebih panjang dengan alasan dapat digunakan lebih lama dari pensil yang pendek.
- Fenomena jual beli yang terjadi di pasar. Pada umumnya anda akan memilih barang belanja seperti tomat, wortel, sayur dan lain-lain setelah membandingkan terlebih dahulu dengan barang-barang di kedai yang lain, biasanya pembeli melihat barang belanjannya dari segi warna yang cerah, kesegaran sayur/buah, bentuk yang bagus, dan harga yang lebih murah dari kedai yang lain sebelum membeli.

Dari contoh di atas, tanpa disadari anda telah melakukan **penilaian** sebelum menentukan pilihan, pada kasus 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 tomat mana yang akan kita beli berdasarkan

warna, kesegaran, bentuk, harga yang lebih murah. Sehingga kita dapat memperkirakan mana yang layak untuk dibeli.

Sebelum melakukan penilaian tersebut, tentu kita harus melakukan **pengukuran**. Pada kasus pertama, anda dapat mengukur pensil dengan menggunakan penggaris untuk menentukan mana pensil yang lebih panjang sebelum menentukan pilihan, sedangkan pada kasus kedua, anda memilih tomat yang terbaik lewat warna, kesegaran, bentuk dan harga. Hal itu juga diawali dengan proses pengukuran dimana kita membanding-bandingkan beberapa tomat yang ada sekalipun tidak menggunakan alat ukur yang paten tetapi berdasarkan pengalaman sebelum anda memilih tomat mana yang layak untuk dibeli. 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.

Evaluasi memiliki ruang lingkup yang lebih luas dari penilaian, jika yang ingin dinilai adalah sistem pembelajaran, maka ruang lingkup yang dinilai adalah seluruh komponen dalam pembelajaran, maka istilah yang tepat untuk menilai sistem pembelajaran adalah evaluasi, namun jika yang ingin dinilai adalah bagian/komponen pembelajaran, seperti hasil belajar maka istilah yang tepat digunakan adalah penilaian. Evaluasi dan penilaian bersifat kualitatif, sementara pengukuran bersifat kuantitatif (skor/angka) yang didapat melalui suatu alat ukur atau instrument yang baku. Dalam konteks kognitif, alat ukur atau instrument berupa tes, sedangkan afektif dapat berupa non-tes (Angket, wawancara, observasi, dll).

Evaluasi dan penilaian merupakan cara menentukan nilai sesuatu, namun berbeda dari segi ruang lingkup dan pelaksanaannya, evaluasi dan penilaian bersifat komprehensif yang meliputi pengukuran, sementara tes merupakan instrumen pengukurannya. Pengukuran dibatasi pada angka-angka tentang *learning proses*. Evaluasi dan penilaian pada hakikatnya suatu proses membuat keputusan tentang nilai suatu objek. Keputusan penilaian tidak hanya didasari dari hasil pengukuran (*quantitative description*), tetapi juga didasari pada hasil pengamatan dan wawancara (*qualitative description*).

Secara harafiah evaluasi berasal dari bahasa Inggris *evaluation* yang berarti penilaian atau penaksiran, evaluasi juga diartikan 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 merupakan penilaian terhadap data yang dikumpulkan melalui kegiatan asesmen. Sementara itu evaluasi adalah suatu keputusan tentang nilai berdasarkan hasil pengukuran. Sejalan dengan pengertian tersebut, 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. 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.

Penilaian (*assessment*) adalah penerapan berbagai cara dan penggunaan beragam alat penilaian untuk memperoleh informasi tentang sejauh mana hasil belajar peserta didik atau ketercapaian kompetensi (rangkaian 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 pendidik (guru) dapat mengetahui hasil pembelajaran yang telah dilakukan. Pendidik harus mengetahui sejauh mana peserta didik (*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.

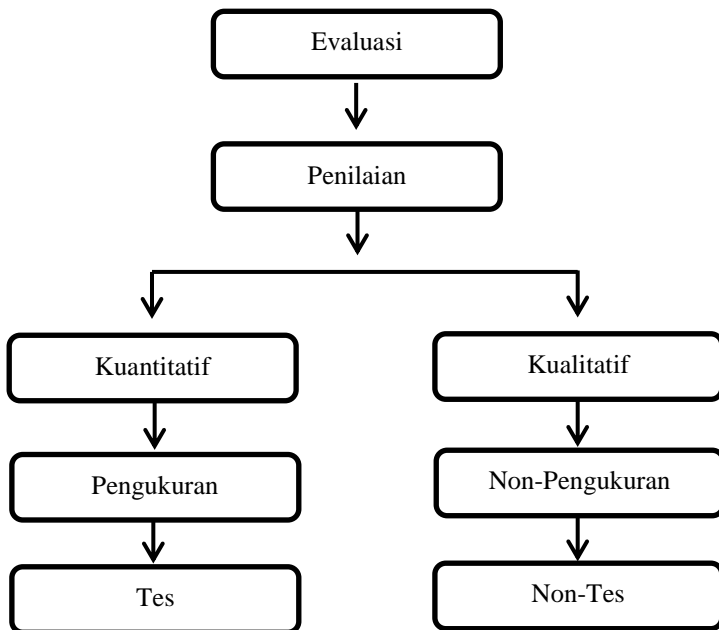
Pengukuran adalah penentuan besaran, dimensi, atau kapasitas, biasanya terhadap suatu standar atau satuan pengukuran. Pengukuran tidak hanya terbatas pada kuantitas

fisik, tetapi juga dapat diperluas untuk mengukur hampir semua benda yang bisa dibayangkan, seperti tingkat ketidakpastian, atau kepercayaan konsumen. 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 pendidik menaksir prestasi peserta didik dengan membaca atau mengamati apa saja yang dilakukan peserta didik, mengamati kinerja mereka, mendengar apa yang mereka katakan, dan menggunakan indera mereka seperti melihat, mendengar, menyentuh, mencium, dan merasakan.

Pengukuran memiliki dua karakteristik utama yaitu: 1) penggunaan angka atau skala tertentu; 2) menurut suatu aturan atau formula tertentu. Pengukuran merupakan proses yang mendeskripsikan *performance* peserta didik dengan menggunakan suatu skala kuantitatif sedemikian rupa sehingga sifat kualitatif dari *performance* siswa tersebut dinyatakan dengan angka-angka. 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. 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, pengertian pengukuran (*measurement*) sebagai kegiatan membandingkan suatu hal

dengan satuan ukuran tertentu sehingga sifatnya menjadi kuantitatif.

Perhatikan gambar berikut ini !



Gambar 1.2 Hubungan Evaluasi, Penilaian, pengukuran dan tes

Dengan demikian, istilah evaluasi, penilaian, pengukuran dan tes memiliki arti yang berbeda, namun dalam buku ini penulis menggunakan istilah evaluasi pembelajaran agar konsep dan pembahasan tidak hanya berkenaan dengan komponen prestasi dan hasil belajar saja, namun berkaitan dengan semua komponen dalam pembelajaran.

## B. Tujuan dan Fungsi Evaluasi Pembelajaran

Evaluasi pembelajaran bertujuan untuk mengetahui efisiensi dan efektifitas pembelajaran yang meliputi : tujuan, metode, konsep bahan ajar, media, sumber ajar, suasana belajar serta cara penilaian. Selain itu, evaluasi pembelajaran juga ditujukan untuk menilai efektifitas strategi pembelajaran, menilai dan meningkatkan efektifitas program kurikulum, menilai dan meningkatkan efektifitas pembelajaran, membantu belajar peserta didik, mengidentifikasi kekuatan dan kelemahan peserta didik, serta untuk menyediakan data yang membantu dalam membuat keputusan. Tujuan evaluasi diuraikan sebagai berikut:

1. *Keeping track*, yaitu untuk menelusuri dan melacak proses belajar peserta didik sesuai dengan rencana pelaksanaan pembelajaran yang telah ditetapkan. Untuk itu, guru harus mengumpulkan data dan informasi dalam kurun waktu tertentu melalui berbagai jenis dan teknik penilaian untuk memperoleh gambaran tentang pencapaian kemajuan belajar peserta didik.
2. *Checking-up*, yaitu untuk mengecek ketercapaian kemampuan peserta didik dalam proses pembelajaran dan kekurangan-kekurangan peserta didik selama mengikuti proses pembelajaran. Dengan kata lain, guru perlu melakukan penilaian untuk mengetahui bagian mana dari materi yang sudah dikuasai peserta didik dan bagian mana dari materi yang belum dikuasai.
3. *Finding-out*, yaitu untuk mencari, menemukan dan mendeteksi kekurangan kesalahan atau kelemahan peserta didik dalam proses pembelajaran, sehingga guru dapat dengan cepat mencari alternatif solusinya.

4. *Summing-up*, yaitu untuk menyimpulkan tingkat penguasaan peserta didik terhadap kompetensi yang telah ditetapkan. Hasil penyimpulan ini dapat digunakan guru untuk menyusun laporan kemajuan belajar ke berbagai pihak yang berkepentingan.

Dengan mengetahui makna penilaian ditinjau dari berbagai segi dalam sistem pendidikan, maka dengan cara lain dapat dikatakan bahwa tujuan atau fungsi penilaian ada beberapa hal:

1. Penilaian berfungsi selektif. Dengan cara mengadakan penilaian guru mempunyai cara untuk mengadakan seleksi atau penilaian terhadap peserta didiknya. Penilaian itu sendiri mempunyai beberapa tujuan, antara lain :
  - a. Untuk memilih peserta didik yang dapat diterima di sekolah tertentu.
  - b. Untuk memilih peserta didik yang dapat naik ke kelas atau tingkat berikutnya.
  - c. Untuk memilih peserta didik yang seharusnya mendapat beasiswa.
  - d. Untuk memilih peserta didik yang sudah berhak meninggalkan sekolah, dan sebagainya.
2. Penilaian berfungsi diagnostik. Apabila alat yang digunakan dalam penilaian cukup memenuhi persyaratan, maka dengan melihat hasilnya, guru akan mengetahui kelemahan peserta didik. Disamping itu diketahui pula sebab-sebab kelemahan itu. Jadi dengan mengadakan penilaian, sebenarnya guru mengadakan diagnosa kepada peserta didik tentang kebaikan dan kelemahannya. Dengan diketahui sebab-sebab kelemahan ini, maka akan lebih mudah dicari untuk cara mengatasinya.



3. Penilaian berfungsi sebagai penempatan. Sistem baru yang kini banyak dipopulerkan di negara Barat, adalah sistem belajar sendiri. Belajar sendiri dapat dilakukan dengan cara mempelajari sebuah paket belajar, baik itu berbentuk modul maupun paket belajar yang lain. Sebagai alasan dari timbulnya sistem ini adalah adanya pengakuan yang besar terhadap kemampuan individual. Setiap peserta didik sejak lahirnya telah membawa bakat sendiri sendiri sehingga pelajaran akan lebih efektif apabila disesuaikan dengan pembawaan yang ada. Akan tetapi disebabkan karena keterbatasan sarana dan tenaga, pendidikan, yang bersifat individual kadang-kadang sukar sekali dilaksanakan. Pendidikan yang bersifat melayani perbedaan kemampuan, adalah pengajaran secara kelompok. Untuk dapat menentukan dengan pasti dikelompok mana seorang peserta didik harus ditempatkan, digunakan suatu penilaian. Sekelompok peserta didik yang mempunyai hasil penilaian sama, akan berada dalam kelompok yang sama dalam belajar.
4. Penilaian berfungsi sebagai pengukur keberhasilan. Fungsi dari penilaian dimaksudkan untuk mengetahui sejauh mana suatu program berhasil diterapkan. Telah disinggung pada bagian sebelum ini, keberhasilan program ditentukan oleh beberapa faktor yaitu: guru, metode/strategi pembelajaran, media pembelajaran, kurikulum, sarana dan sistem administrasi.

Selain dari itu penilaian juga berguna bagi semua pihak pemangku kepentingan, mulai dari peserta didik, tenaga pengajar, sekolah dan juga masyarakat. Khusus bagi peserta

didik, guru dan sekolah penilaian memberikan manfaat sebagai berikut:

1. Peserta didik. Dengan diadakannya penilaian, maka peserta didik dapat mengetahui sejauh mana telah berhasil mengikuti pelajaran yang diberikan guru. Hasil yang diperoleh peserta didik dari pekerjaan menilai ini ada 2 kemungkinan:
  - a. Memuaskan. Jika peserta didik memperoleh hasil yang memuaskan, dan hal itu menyenangkan, tentu kepuasan itu ingin diperolehnya lagi pada kesempatan lain waktu. Akibatnya peserta didik akan mempunyai motivasi yang cukup besar untuk belajar yang lebih giat. Namun demikian, keadaan sebaliknya dapat terjadi, yakni peserta didik merasa sudah puas dengan hasil yang diperoleh dan usahanya kurang gigih lain kali.
  - b. Tidak memuaskan. Jika peserta didik tidak puas dengan hasil yang diperoleh ia akan berusaha agar lain kali keadaan itu tidak terulang lagi. Maka ia lalu bekerja giat. Namun demikian, keadaan sebaliknya dapat terjadi putus asa dengan hasil kurang memuaskan yang telah diterimanya.
2. Guru.
  - a. Dengan hasil penilaian yang diperoleh guru akan dapat mengetahui peserta didik mana yang sudah berhak meneruskan pelajarannya karena sudah berhasil menguasai bahan, maupun mengetahui peserta didik yang belum berhasil menguasai bahan. Dengan petunjuk ini guru dapat lebih memusatkan perhatiannya kepada peserta didik yang belum berhasil. Apa lagi jika

guru tahu akan sebab-sebabnya ia akan memberikan perhatian yang lebih teliti sehingga keberhasilan selanjutnya dapat diharapkan.

- b. Guru akan mengetahui apakah 'materi' yang diajarkan sudah tepat bagi peserta didik sehingga untuk memberikan pengajaran di waktu yang akan datang tidak perlu diadakan perubahan.
- c. Guru akan mengetahui apakah 'metode' yang digunakan sudah tepat atau belum. Jika sebagian besar dari peserta didik memperoleh angka jelek pada penilaian yang diadakan, mungkin hal ini disebabkan oleh pendekatan atau metode yang kurang tepat. Apabila demikian halnya, maka guru harus mawas diri dan mencoba mencari metode lain dalam belajar.

### 3. Sekolah.

- a. Apabila guru-guru mengadakan penilaian dan diketahui bagaimana hasil belajar peserta didik-peserta didiknya, dapat pula diketahui bahwa apakah kondisi belajar yang diciptakan oleh sekolah sudah sesuai dengan harapan atau belum. Hasil belajar merupakan cermin kualitas suatu sekolah.
- b. Informasi dari guru tentang tepat tidaknya kurikulum untuk sekolah itu dapat merupakan bahan pertimbangan bagi perencanaan sekolah untuk masa-masa yang akan datang.
- c. Informasi hasil penilaian yang diperoleh dari tahun ke tahun, dapat digunakan sebagai pedoman bagi sekolah, yang dilakukan oleh sekolah sudah memenuhi standar atau belum. Pemenuhan standar akan terlihat dari bagusnya angka-angka yang diperoleh peserta didik.

## **C. Prinsip-Prinsip Evaluasi dan Penilaian Pembelajaran**

### **1. Prinsip-Prinsip Evaluasi Pembelajaran**

Untuk mendapatkan hasil evaluasi yang baik, kegiatan evaluasi harus dilakukan berdasarkan prinsip-prinsip yang pasti, valid, komprehensif dll. Seperti penjelasan berikut ini:

#### **a. Kepastian dan Kejelasan.**

Dalam proses evaluasi maka kepastian dan kejelasan yang akan dievaluasi menduduki urutan pertama. Evaluasi akan dapat dilaksanakan apabila tujuan evaluasi tidak dirumuskan dulu secara jelas dalam definisi yang operational. Bila kita ingin mengevaluasi kemajuan belajar siswa maka pertama-tama kita identifikasi dan kita definisikan tujuan-tujuan instruksional pengajaran dan barulah kita kembangkan alat evaluasinya. Dengan demikian efektifitas alat evaluasi tergantung pada deskripsi yang jelas apa yang akan kita evaluasi. Pada umumnya alat evaluasi dalam pendidikan terutama pengajaran berupa test. Test ini dapat mencerminkan karakteristik aspek yang akan diukur. Kalau kita akan mengevaluasi tingkat intelegensi siswa, maka komponen-komponen intelegensi itu harus dirumuskan dengan jelas dan kemampuan belajar yang dicapai dirumuskan dengan tepat selanjutnya dikembangkan test sebagai alat evaluasi. Dengan demikian keberhasilan evaluasi lebih banyak ditentukan kepada kemampuan guru (evaluator) dalam mendefinisikan dengan jelas aspek-aspek individual ke dalam proses pendidikan.

#### **b. Teknik Evaluasi**

Teknik evaluasi yang dipilih sesuai dengan tujuan evaluasi. Hendaklah diingat bahwa tidak ada teknik evaluasi yang

cocok untuk semua keperluan dalam pendidikan. Tiap-tiap tujuan (pendidikan) yang ingin dicapai dikembangkan teknik evaluasi tersendiri yang cocok dengan tujuan tersebut. Kecocokan antara tujuan evaluasi dan teknik yang digunakan perlu dijadikan pertimbangan utama.

c. Komprehensif.

Evaluasi yang komprehensif memerlukan teknik bervariasi yaitu teknik evaluasi tunggal yang mampu mengukur tingkat kemampuan siswa dalam belajar, meskipun hanya dalam satu pertemuan jam pelajaran. Sebab dalam kenyataannya tiap-tiap teknik evaluasi mempunyai keterbatasan-keterbatasan tersendiri. *Test obyektif* misalnya akan memberikan bukti obyektif tentang tingkat kemampuan siswa. Tetapi hanya memberikan informasi sedikit dari siswa tentang apakah ia benar-benar mengerti tentang materi tersebut. Lebih-lebih pada *test subyektif* yang penilaiannya lebih banyak tergantung pada subyektivitas evaluatornya. Atas dasar prinsip inilah maka seyogyanya dalam proses belajar-mengajar, untuk mengukur kemampuan belajar siswa digunakan teknik evaluasi yang bervariasi.

d. Kesadaran adanya kesalahan pengukuran.

Evaluator harus menyadari keterbatasan dan kelemahan dalam teknik evaluasi yang digunakan. Atas dasar kesadaran ini, maka dituntut untuk lebih hati-hati dalam kebijakan-kebijakan yang diambil setelah melaksanakan evaluasi. Evaluator menyadari bahwa dalam pengukuran yang dilaksanakan, hanya mengukur sebagian (sampel) saja dari suatu kompleksitas yang seharusnya diukur, lagi pula pengukuran dilakukan hanya pada saat tertentu saja.

Maka dapat terjadi salah satu aspek yang sifatnya menonjol yang dimiliki siswa tidak termasuk dalam sampel pengukuran. Inilah yang disebut *sampling error* dalam evaluasi. Sumber kesalahan (*error*) yang lain terletak pada alat/instrument yang digunakan dalam proses evaluasi. Penyusunan alat evaluasi tidak mudah, lebih-lebih bila aspek yang diukur sifatnya kompleks. Dalam laporan hasil evaluasi, evaluator perlu melaporkan adanya kesalahan pengukuran ini. Pengukuran dengan test, kesalahan pengukuran dapat ditunjukkan dengan koefisien kesalahan pengukuran.

- e. Evaluasi adalah alat, bukan tujuan.  
Evaluator menyadari sepenuhnya bahwa tiap-tiap teknik evaluasi digunakan sesuai dengan tujuan evaluasi. Hasil evaluasi yang diperoleh tanpa tujuan tertentu akan membuang waktu dan uang, bahkan merugikan peserta didik. Maka dari itu yang perlu dirumuskan lebih dahulu ialah tujuan evaluasi, baru dari tujuan ini dikembangkan teknik yang akan digunakan dan selanjutnya disusun test sebagai alat evaluasi. Jangan sampai terbalik, sebab tanpa diketahui tujuan evaluasi data yang diperoleh akan sia-sia. Atas dasar pengertian tersebut di atas maka kebijakan-kebijakan pendidikan yang akan diambil dirumuskan dulu dengan jelas sebelumnya dipilih prosedur evaluasi yang digunakan dengan demikian.

## **2. Prinsip-Prinsip Penilaian Pembelajaran**

- a. Berorientasi kepada pencapaian kompetensi.  
Berarti penilaian mencakup semua aspek kompetensi dengan menggunakan berbagai teknik yang sesuai untuk memantau perkembangan kemampuan peserta didik.

Penilaian hasil belajar oleh pendidik didasarkan pada ukuran pencapaian kompetensi yang ditetapkan. Penilaian hasil belajar oleh pendidik dapat dipertanggung jawabkan, baik dari segi teknik, prosedur, maupun hasilnya. Contoh : guru matematika dapat menjelaskan secara benar kepada pihak terkait, tentang proses penilaian, teknik penilaian, prosedur, dan hasil yang sesuai dengan kenyataan kemampuan hasil belajar peserta didiknya.

b. Valid

Penilaian valid berarti menilai apa yang seharusnya dinilai dengan menggunakan alat yang sesuai untuk mengukur kompetensi, sehingga penilaian tersebut menghasilkan informasi yang akurat tentang aktivitas belajar. Penilaian hasil belajar oleh pendidik harus mengukur pencapaian kompetensi yang ditetapkan dalam standar isi (standar kompetensi dan kompetensi dasar) dan standar kompetensi lulusan. Misalnya apabila pembelajaran menggunakan pendekatan eksperimen maka kegiatan eksperimen harus menjadi salah satu obyek yang di nilai.

Contoh : Dalam pelajaran penjaskes, guru menilai kompetensi permainan badminton siswa, penilaian dianggap valid jika menggunakan test praktek langsung, jika menggunakan tes tertulis maka tes tersebut tidak valid.

c. Adil

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. Peserta didik berhak memperoleh nilai secara adil, penilaian hasil belajar tidak menguntungkan atau merugikan peserta didik karena

berkebutuhan khusus serta perbedaan latar belakang agama, suku, budaya, adat istiadat, status sosial, ekonomi, fisik, dan gender.

Contoh : guru penjaskes laki-laki hendaknya tidak memandang fisik dan rupa dari murid perempuan yang cantik kemudian memberi perlakuan khusus, semua murid berhak diperlakukan sama saat kegiatan belajar mengajar maupun dalam pemberian nilai. Nilai yang diberikan sesuai dengan kenyataan hasil belajar siswa tersebut.

d. Obyektif

Penilaian yang bersifat objektif tidak memandang dan membeda-bedakan latar belakang peserta didik, namun melihat kompetensi yang dihasilkan oleh peserta didik tersebut, bukan atas dasar siapa dirinya. Penilaian harus dilaksanakan secara objektif dan tidak dipengaruhi oleh subyektivitas penilai.

Contoh : Guru memberi nilai 85 untuk materi volley pada si A yang merupakan tetangga dari guru tersebut, namun si B, yang kemampuannya lebih baik, mendapatkan nilai hanya 80. Ini adalah penilaian yang bersifat subyektif dan tidak disarankan. Pemberian nilai haruslah berdasarkan kemampuan siswa tersebut.

e. Berkesinambungan

Pelaksanaan penilaian hasil belajar dilakukan secara terencana, bertahap, dan terus menerus untuk memperoleh gambaran tentang perkembangan belajar peserta didik.

Contoh : guru matematika melakukan kegiatan belajar mengajar secara terencana, guru menjelaskan materi tiap pertemuan, memberikan tugas, mengadakan ulangan



harian, ujian tengah semester, serta ujian akhir semester, semua dilaksanakan secara terus menerus dan bertahap, dan dari setiap tahap tersebut, guru mengumpulkan informasi yang akan diolah untuk menghasilkan nilai. Berkesinambungan yaitu penilaian dilakukan secara terencana, bertahap, terus menerus, untuk memperoleh gambaran pencapaian kompetensi peserta didik dalam kurun waktu tertentu.

f. Menyeluruh.

Penilaian diambil dengan mencakup seluruh aspek kompetensi peserta didik dan menggunakan berbagai teknik penilaian yang sesuai, termasuk mengumpulkan berbagai bukti aktivitas belajar peserta didik. Penilaian meliputi pengetahuan (*cognitif*), keterampilan (*psycho-motor*), dan sikap (*affectif*).

Contoh : Dalam penilaian hasil akhir belajar, guru Seni Budaya mengumpulkan berbagai bukti aktivitas siswa dalam catatan sebelumnya, penilaian yang dikumpulkan mulai dari pengetahuan tentang seni budaya, keterampilan menari, menggambar, bermusik, kehadiran dalam kegiatan belajar mengajar, dan penilaian sikap peserta didik, semua hal tersebut digabungkan menjadi satu dan menghasilkan nilai. Menyeluruh yaitu mencakup seluruh domain yang tertuang pada setiap kompetensi dasar.

g. Terbuka

Penilaian harus bersifat transparan dan pihak yang terkait harus tau bagaimana pelaksanaan penilaian tersebut, dari aspek apa saja nilai tersebut didapat, dasar pengambilan keputusan, dan bagaimana pengolahan nilai tersebut sampai hasil akhirnya tertera, dan dapat diterima.

Contoh : pada tahun ajaran baru, guru kimia menerangkan tentang kesepakatan pemberian nilai dengan bobot masing-masing aspek, misal, Partisipasi kehadiran diberi bobot 20%, Tugas individu dan kelompok 20%, Ujian tengah semester 25%, ujian akhir semester 35%. Sehingga disini terjadi keterbukaan penilaian antara murid dan guru.

h. Bermakna

Penilaian hasil belajar oleh pendidik memiliki arti, makna, dan manfaat yang dapat ditindaklanjuti oleh pihak lain, terutama pendidik, peserta didik, orang tua, dan masyarakat.

Contoh : bagi guru, hasil penilaian dapat bermakna untuk melihat seberapa besar keberhasilan metode pembelajaran yang digunakan, sebagai evaluasi untuk perbaikan kedepan, serta memberikan pengukuran prestasi belajar kepada siswa.

## **D. Jenis Evaluasi Pembelajaran**

### **1. Bentuk Tes sebagai Instrument Evaluasi**

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

Secara harfiah, kata “tes” berasal dari bahasa Perancis Kuno *testum* dengan arti piring untuk menyisahkan logam-logam mulia, dalam bahasa Inggris ditulis dengan *test* yang dalam bahasa Indonesia diterjemahkan dengan “tes”, “ujian”, atau “percobaan”. *Testing* berarti saat dilaksanakannya atau peristiwa berlangsungnya pengukuran dan penilaian. *Tester* adalah orang yang melaksanakan tes atau pembuat tes. *Testee* adalah pihak yang dikenai tes (peserta tes).

Dari segi istilah, yang dimaksud dengan tes adalah alat pengukur yang mempunyai standar yang obyektif sehingga dapat digunakan secara meluas, serta dapat betul-betul digunakan untuk mengukur dan membandingkan keadaan psikis atau tingkah laku individu.

Dalam dunia evaluasi pendidikan, yang dimaksud dengan tes adalah cara atau prosedur dalam pengukuran dan penilaian di bidang pendidikan, yang berbentuk pemberian tugas atau serangkaian tugas, baik berupa pertanyaan-pertanyaan yang harus dijawab, atau perintah-perintah oleh *testee*, sehingga dapat dihasilkan nilai yang melambangkan tingkah laku atau prestasi *testee*, nilai mana dapat dibandingkan dengan nilai-nilai yang dicapai oleh *testee* lainnya, atau dibandingkan dengan nilai standar tertentu.

Sehingga kami dapat menyimpulkan bahwa tes merupakan suatu teknik atau cara yang digunakan dalam rangka melaksanakan kegiatan pengukuran, yang di dalamnya terdapat berbagai pertanyaan, pernyataan, atau serangkaian tugas yang harus dikerjakan atau dijawab oleh peserta didik untuk mengukur aspek perilaku peserta didik.

Banyak alat atau instrument yang dapat digunakan dalam kegiatan evaluasi. Salah satunya adalah tes. Di sekolah juga sering disebut dengan tes prestasi belajar. Tes banyak digunakan untuk mengukur prestasi belajar peserta didik dalam bidang kognitif, seperti pengetahuan, pemahaman, aplikasi, analisis, dan evaluasi. Penggunaan tes dalam dunia pendidikan sudah dikenal sejak dahulu kala, sejak orang mengenal pendidikan itu sendiri.

## 2. Fungsi Tes

### a. Fungsi untuk Kelas:

- 1) Mengadakan diagnosis terhadap kesulitan belajar siswa.
- 2) Mengevaluasi celah antara bakat dengan pencapaian.
- 3) Menaikkan tingkat prestasi.
- 4) Mengelompokkan siswa dalam kelas pada waktu metode kelompok.
- 5) Merencanakan kegiatan proses belajar-mengajar untuk siswa secara perseorangan.
- 6) Menentukan siswa mana yang memerlukan bimbingan khusus.
- 7) Menentukan tingkat pencapaian untuk setiap anak.

### b. Fungsi untuk Bimbingan:

- 1) Menentukan arah pembicaraan dengan orang tua tentang anak-anak mereka.
- 2) Membantu siswa dalam menentukan pilihan.
- 3) Membantu siswa mencapai tujuan pendidikan dan jurusan.
- 4) Memberi kesempatan kepada pembimbing, guru, dan orang tua dalam memahami kesulitan anak.

### c. Fungsi untuk Administrasi:

- 1) Memberi petunjuk dalam pengelompokan siswa.
- 2) Penempatan siswa baru.
- 3) Membantu siswa memilih kelompok.
- 4) Menilai kurikulum.
- 5) Memperluas hubungan masyarakat (*public relation*).

- 6) Menyediakan informasi untuk badan-badan lain di luar sekolah.

### 3. Langkah-langkah dalam Penyusunan Tes

Urutan langkah yang dilakukan dalam penyusunan tes adalah:

- a. Menentukan tujuan mengadakan tes.
- b. Mengadakan pembatasan terhadap bahan yang akan dijadikan tes.
- c. Merumuskan tujuan instruksional khusus dari tiap bagian bahan.
- d. Menderetkan semua indikator dalam tabel persiapan yang memuat pula aspek tingkah laku yang terkandung dalam indikator itu.

### 4. Bentuk-bentuk tes

- a. Berdasarkan Fungsinya  
Ditinjau dari segi fungsi yang dimiliki oleh tes sebagai alat pengukur perkembangan belajar peserta didik, tes ini dapat dibedakan menjadi enam golongan:

#### 1) Tes seleksi

Tes seleksi sering dikenal dengan istilah “ujian saringan” atau “ujian masuk”. Tes ini dilaksanakan dalam rangka penerimaan calon siswa baru, dimana hasil tes digunakan untuk memilih calon peserta didik yang tergolong paling baik dari sekian banyak calon yang mengikuti tes. Materi pada tes seleksi ini merupakan materi prasyarat untuk mengikuti program pendidikan yang akan diikuti oleh calon.

Sesuai dengan sifatnya, yaitu menyeleksi atau melakukan penyaringan, maka materi tes seleksi terdiri atas butir-butir soal yang cukup sulit, sehingga hanya calon-calon yang tergolong memiliki kemampuan tinggi sajalah yang dimungkinkan dapat menjawab butir-butir soal tes dengan betul. Tes seleksi dapat dilaksanakan secara lisan, secara tertulis, dengan tes perbuatan, dan dapat pula dilaksanakan dengan mengombinasikan ketiga jenis tes tersebut secara serempak. Sebagai tindak lanjut dari hasil tes seleksi, maka para calon yang dipandang memenuhi batas persyaratan minimal yang telah ditentukan dinyatakan sebagai peserta tes yang lulus dan dapat diterima sebagai siswa baru, dinyatakan tidak lulus dan karenanya tidak dapat diterima sebagai siswa baru.

## 2) Tes awal

Tes awal sering dikenal dengan istilah *pre-test*. Tes jenis ini dilaksanakan dengan tujuan untuk mengetahui sejauh manakah materi atau bahan pelajaran yang akan diajarkan telah dapat dikuasai oleh para peserta didik. Jadi tes awal adalah tes yang dilaksanakan sebelum bahan pelajaran diberikan kepada peserta didik. Karena itu maka butir-butir soalnya dibuat yang mudah-mudah. Setelah tes awal berakhir, maka sebagai tindak lanjutnya adalah :

- a) Jika dalam tes awal itu semua materi yang ditanyakan dalam tes sudah dikuasai dengan baik oleh peserta didik, maka materi yang telah

ditanyakan dalam tes awal itu tidak diajarkan lagi,

b) Jika materi yang dapat dipahami oleh peserta didik baru sebagian saja, maka yang diajarkan adalah materi pelajaran yang belum cukup dipahami oleh para peserta didik tersebut.

### 3) Tes akhir

Tes akhir sering dikenal dengan istilah *post-test*. Tes akhir dilaksanakan dengan tujuan untuk mengetahui apakah semua materi pelajaran yang tergolong penting sudah dapat dikuasai dengan sebaik-baiknya oleh para peserta didik.

### 4) Tes diagnostik

Tes diagnostik (*diagnostic test*) adalah tes yang digunakan untuk menentukan secara tepat, jenis kesukaran yang dihadapi oleh peserta didik dalam suatu mata pelajaran tertentu.

### 5) Tes formatif

Tes formatif adalah tes hasil belajar yang bertujuan untuk mengetahui, sudah sejauh manakah peserta didik “telah terbentuk” setelah mereka mengikuti proses pembelajaran dalam jangka waktu tertentu. Tes formatif ini biasanya dilaksanakan di tengah-tengah perjalanan program pengajaran, yaitu dilaksanakan pada setiap kali satuan pelajaran atau subpokok bahasan berakhir atau dapat diselesaikan. Di sekolah-sekolah tes formatif ini biasa dikenal dengan istilah “ulangan harian”.

Tindak lanjut yang perlu dilakukan setelah diketahuinya hasil tes formatif adalah:

- a) Jika materi yang diteskan itu telah dikuasai dengan baik, maka pembelajaran dilanjutkan dengan pokok bahasan yang baru.
- b) Jika ada bagian-bagian yang belum dikuasai, maka sebelum dilanjutkan dengan pokok bahasan baru, terlebih dahulu diulangi atau dijelaskan lagi bagian-bagian yang belum dikuasai oleh peserta didik.

6) Tes sumatif

Tes sumatif adalah tes hasil belajar yang dilaksanakan setelah sekumpulan satuan program pengajaran selesai diberikan. Tes sumatif dilaksanakan secara tertulis, agar semua siswa memperoleh soal yang sama. Butir-butir soal yang dikemukakan dalam tes sumatif ini pada umumnya juga lebih sulit atau lebih berat daripada butir-butir soal tes formatif. Yang menjadi tujuan utama tes sumatif adalah untuk menentukan nilai yang melambangkan keberhasilan peserta didik setelah mereka menempuh proses pembelajaran dalam jangka waktu tertentu.

b. Berdasarkan aspek psikis yang ingin diungkap

Ditilik dari aspek kejiwaan yang ingin diungkap, tes setidak-tidaknya dapat dibedakan menjadi lima golongan, yaitu:

- 1) Tes intelegensi, yakni tes yang dilaksanakan dengan tujuan mengungkap atau mengetahui tingkat kecerdasan seseorang.



- 2) Tes kemampuan, yakni tes yang dilaksanakan dengan tujuan untuk mengungkap kemampuan dasar atau bakat khusus yang dimiliki oleh *testee*.
- 3) Tes sikap, yakni salah satu jenis tes yang dipergunakan untuk mengungkap predisposisi atau kecenderungan seseorang untuk melakukan suatu respon tertentu terhadap dunia sekitarnya, baik berupa individu-individu maupun obyek-obyek tertentu.
- 4) Tes kepribadian, yakni tes yang dilaksanakan dengan tujuan mengungkap ciri-ciri khas dari seseorang yang banyak sedikitnya bersifat lahiriah.
- 5) Tes hasil belajar, yang juga sering dikenal dengan istilah tes pencapaian, yakni tes yang biasa digunakan untuk mengungkap tingkat pencapaian atau prestasi belajar.

c. Penggolongan lain-lain

Ditilik dari banyaknya orang yang mengikuti tes, dapat dibedakan menjadi 2 golongan, yaitu:

- 1) Tes individual, yaitu tes dimana tester hanya berhadapan dengan satu orang *testee* saja.
- 2) Tes kelompok, yaitu tes dimana tester berhadapan dengan lebih dari satu orang *testee*.

Ditilik dari segi waktu yang disediakan bagi *testee* untuk menyelesaikan tes, yaitu:

- 1) *Power test*, yakni tes dimana waktu yang disediakan buat *testee* untuk menyelesaikan tes tersebut tidak dibatasi.

- 2) *Speed test*, yaitu tes dimana waktu yang disediakan buat *testee* untuk menyelesaikan tes tersebut dibatasi.

Ditilik dari segi bentuk responnya, tes dapat dibedakan menjadi dua golongan yaitu:

- 1) *Verbal test*, yaitu suatu tes yang menghendaki respon (jawaban) yang tertuang dalam bentuk ungkapan kata-kata atau kalimat, baik secara lisan ataupun tertulis.
- 2) *Nonverbal test*, yakni tes yang menghendaki respon (jawaban) dari *testee* bukan berupa ungkapan kata-kata atau kalimat, melainkan berupa tindakan atau tingkah laku.

Akhirnya, apabila ditinjau dari segi cara mengajukan pertanyaan dan cara memberikan jawabannya, tes dapat dibedakan menjadi dua golongan, yaitu:

- 1) Tes tertulis, yakni jenis tes dimana tester dalam mengajukan butir-butir pertanyaan atau soalnya dilakukan secara tertulis dan *testee* memberikan jawabannya juga secara tertulis.
- 2) Tes lisan, yakni tes dimana tester di dalam mengajukan pertanyaan-pertanyaan atau soalnya dilakukan secara lisan, dan *testee* memberikan jawabannya secara lisan juga.

## **2. Bentuk Non Tes Sebagai Instrumen Evaluasi**

Nontes adalah cara penilaian hasil belajar peserta didik yang dilakukan tanpa menguji peserta didik tetapi dengan melakukan pengamatan secara sistematis. Teknik evaluasi non

tes berarti melaksanakan penilaian dengan tidak menggunakan tes. Teknik penilaian ini umumnya untuk menilai kepribadian anak secara menyeluruh meliputi sikap, tingkah laku, sifat, sikap sosial, dan lain-lain. Yang berhubungan dengan kegiatan belajar dalam pendidikan, baik secara individu maupun secara kelompok.

Pada prinsipnya, setiap melakukan evaluasi pembelajaran, kita dapat menggunakan teknik tes dan nontes, sebab hasil belajar atau aspek-aspek pembelajaran bersifat aneka ragam. Perlu diketahui bahwa tes bukanlah satu-satunya cara untuk melakukan evaluasi hasil belajar siswa, teknik lain yang dapat dilakukan adalah teknik non tes. Dengan teknik ini evaluasi hasil belajar peserta didik dilakukan tanpa menguji peserta didik tersebut, melainkan dilakukan dengan pengamatan secara sistematis (*observation*), melakukan wawancara (*interview*), penyebaran angket (*questionnaire*), memeriksa atau meneliti dokumen-dokumen (*documentary analysis*). Teknik non tes ini memegang peranan penting terutama dalam rangka evaluasi hasil belajar peserta didik dalam ranah sikap hidup (*affective domain*) dan ranah keterampilan (*psychomotoric domain*), sedangkan teknik tes sering digunakan untuk mengevaluasi hasil belajar peserta didik dari segi ranah berfikirnya (*cognitive domain*). Berikut ini adalah beberapa jenis evaluasi non tes:

Ada beberapa teknik non tes, yaitu Skala bertingkat (*rating scale*), Kuesioner (*questionnaire*), Daftar cocok (*check-list*), Wawancara (*interview*), Pengamatan (*observation*) dan Riwayat hidup. Adapun penjelasannya adalah sebagai berikut :

a. Skala bertingkat (*rating scale*)

Skala menggambarkan nilai yang berbentuk angka terhadap suatu hasil pertimbangan. Biasanya angka-angka yang digunakan diterangkan pada skala dengan jarak yang sama. Meletakkannya secara bertingkat dari yang rendah ke yang tinggi.

Kita dapat menilai hampir segala sesuatu dengan skala. Dengan maksud agar pencatatannya dapat objektif, maka penilaian terhadap penampilan atau penggambaran kepribadian seseorang disajikan dalam bentuk skala.

b. Kuesioner

Kuesioner juga sering dikenal sebagai angket. Pada dasarnya kuesioner adalah Sebuah daftar pertanyaan yang harus diisi oleh orang yang akan diukur (responden). Dengan kuesioner ini orang dapat diketahui tentang keadaan atau data diri, pengalaman, pengetahuan sikap atau pendapatnya dan lain-lain.

Ditinjau dari segi cara menjawabnya maka dibedakan menjadi dua, yaitu kuesioner tertutup dan kuesioner terbuka. Kuesioner tertutup adalah kuesioner yang disusun dengan menyediakan pilihan jawaban langkah sehingga pengisi hanya tinggal memberi tanda pada jawaban yang dipilih. Sedangkan kuesioner terbuka adalah kuesioner yang disusun sedemikian rupa sehingga para pengisi bebas mengemukakan pendapatnya. Kuesioner terbuka disusun apabila macam jawaban pengisi belum terperinci dengan jelas sehingga jawabannya akan beraneka ragam.

c. Daftar cocok (*check list*)

Yang dimaksud dengan daftar cocok adalah deretan pernyataan (yang biasanya singkat-singkat), dimana responden yang dievaluasi tinggal membubuhkan tanda cocok ( $\checkmark$ ) di tempat yang sudah disediakan. Ada pendapat yang mengatakan bahwa sebenarnya skala bertingkat dapat digolongkan ke dalam daftar cocok karena dalam skala bertingkat, responden juga diminta untuk memberikan tanda cocok pada pilihan yang tepat.

d. Wawancara (*interview*)

Wawancara adalah suatu cara yang digunakan untuk mendapatkan jawaban dari responden dengan jalan tanya jawab sepihak. Dikatakan sepihak karena dalam wawancara ini responden tidak diberi kesempatan sama sekali untuk mengajukan pertanyaan. Pertanyaan hanya diajukan oleh subjek evaluasi.

Ada dua macam *interview*, yaitu *interview* bebas dan *interview* terpimpin. *Inteview* bebas merupakan di mana responden mempunyai kebebasan untuk mengutarakan pendapatnya, tanpa dibatasi oleh patokan-patokan yang telah dibuat oleh subjek evaluasi. Sedangkan *interview* terpimpin adalah *interview* yang dilakukan oleh subjek evaluasi dengan cara mengajukan pertanyaan-petanyaan yang udah disusun terlebih dahulu. Jadi dalam hal ini responden pada waktu menjawab pertanyaan tinggal memilih jawaban yang sudah dipersiapkan oleh penanya. Pertanyaan itu kadang-kadang bersifat sebagai pemimpin, mengarahkan dan penjawab sudah dipimpin oleh sebuah daftar cocok, sehingga dalam menuliskan jawaban, ia

tinggal membubuhkan tanda cocok di tempat yang sesuai dengan keadaan responden.

e. Pengamatan (*observation*)

Pengamatan adalah suatu teknik yang dilakukan dengan cara mengadakan pengamatan secara teliti serta pencatatan secara sistematis. Ada 3 macam observasi :

- 1) Observasi partisipan, yaitu observasi yang dilakukan oleh pengamat, tetapi dalam pada itu pengamat memasuki dan mengikuti kegiatan kelompok yang sedang diamati. Observasi partisipan dilaksanakan sepenuhnya jika pengamat betul-betul mengikuti kegiatan kelompok, bukan hanya pura-pura. Dengan demikian ia dapat menghayati dan merasakan seperti apa yang dirasakan orang-orang dalam kelompok yang diamati.
- 2) Observasi sistematis, yaitu observasi di mana faktor-faktor yang diamati sudah diatur menurut kategorinya. Berbeda dengan observasi partisipan, maka dalam observasi ini pengamat berada di luar kelompok. Dengan demikian maka pengamat tidak dibingungkan oleh situasi yang melingkungi dirinya.
- 3) Observasi eksperimental, pengamatan ini terjadi jika pengamatan tidak berpartisipasi dalam kelompok. Dalam hal ini ia dapat mengendalikan unsur-unsur penting dalam situasi sedemikian rupa sehingga situasi itu dapat diatur sesuai dengan tujuan evaluasi.

f. Riwayat hidup

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

## Bab II

# Karakteristik, Model dan Pendekatan Evaluasi Pembelajaran

**T**es maupun non-tes merupakan instrumen yang digunakan pendidik dalam mengukur keefektifan proses pembelajaran yang telah diterapkan kepada peserta didik, instrumen tersebut berupa observasi, wawancara, angket, tes dan lain-lain. Instrumen sebagai alat ukur yang begitu penting tentu harus memiliki syarat-syarat yang menunjukkan karakteristik instrumen yang layak. Pendidik sering kali membuat instrumen tanpa mengikuti aturan-aturan yang berlaku. Instrumen yang akan digunakan tidak dapat langsung diambil dari buku atau bahan ajar yang dijual di pasaran, instrumen juga tidak dapat diambil dari soal-soal yang telah lama dan belum diketahui kualitasnya. Maka dari itu perlu ada pemahaman terkait perancangan instrumen evaluasi yang baik agar alat ukur yang dibuat sesuai dengan apa yang ingin diukur pendidik.



## A. Karakteristik Instrumen Evaluasi

Evaluasi dilakukan untuk menemukan kekurangan-kekurangan dalam proses belajar mengajar, sehingga kualitas pembelajaran dapat ditingkatkan. Evaluasi dapat menilai sejauh mana keefektifan pembelajaran yang diterapkan pendidik. Semua komponen dalam pembelajaran dapat diketahui apakah dapat berfungsi sebagaimana mestinya atau tidak. Pendidik dapat mengetahui sejauh mana perkembangan peserta didik, baik hasil belajar, sikap, kemampuan individu dan kelompok, kemampuan psikomotor dan lain-lain.

Proses evaluasi dan penilaian harus didasari pada proses pengukuran yang melibatkan instrumen sebagai alat ukur, instrumen dapat berupa tes maupun non-tes, kualitas instrumen mempengaruhi hasil pengukuran, maka dari itu penting untuk memahami syarat-syarat dan kaidah dalam penyusunan instrumen. Alat ukur yang baik dapat mengukur data secara akurat sesuai dengan fungsinya, tentu karakteristik instrumen yang baik memiliki dasar valid, reliabel, relevan, representatif, praktis, deskriminatif, spesifik dan proporsional.

### 1. Valid

Valid berarti kecocokan, maka instrumen yang valid dapat mengukur apa yang hendak diukur secara akurat, misalkan, alat ukur hasil belajar, maka alat ukur tersebut harus benar-benar mampu mengukur hasil belajar, seperti halnya mengukur keefektifan media, model, metode, strategi dan komponen pembelajaran lain. Uji yang digunakan dalam mengetahui valid atau tidaknya suatu alat ukur disebut validitas. Validitas dapat ditinjau dari berbagai segi, yaitu validitas isi (*Content Validity*), validitas

konstruk (*construct validity*), validitas ramalan (*predictive validity*) dan validitas bandingan (*concurrent validity*).

2. Reliabel

Instrumen baik juga harus reliabel, artinya memiliki hasil yang konsisten, misalkan pendidik memberikan tes kepada sekelompok peserta didik untuk mengukur kemampuan mereka dalam memecahkan suatu masalah, kemudian tes tersebut diberikan lagi pada kelompok peserta didik yang berbeda, ternyata hasilnya sama atau tidak jauh berbeda, maka instrumen tersebut memiliki tingkat reliabel yang tinggi.

3. Relevan

Relevan berarti sesuai, artinya instrumen yang dirancang harus sesuai dengan indikator, standar kompetensi dan kompetensi dasar yang telah ditetapkan, misalnya dalam menilai hasil belajar, maka instrumen yang dibuat harus berdasarkan domain kognitif, jangan sampai ingin mengukur domain kognitif namun menggunakan wawancara atau angket. Hal ini tentu tidak relevan.

4. Representatif.

Representatif artinya materi alat ukur harus betul-betul mewakili dari seluruh materi yang disampaikan. Hal ini dapat dilakukan bila guru menggunakan silabus sebagai acuan pemilihan materi tes. Guru juga harus memperhatikan proses seleksi materi, mana materi yang bersifat aplikatif dan mana yang tidak, mana yang penting dan mana yang tidak.

5. **Praktis**  
Praktis artinya mudah digunakan. Jika alat ukur itu sudah memenuhi syarat tetapi sukar digunakan, berarti tidak praktis. Kepraktisan ini bukan hanya dilihat dari pembuat alat ukur (guru), tetapi juga bagi orang lain yang ingin menggunakan alat ukur tersebut.
6. **Deskriminatif**  
Deskriminatif artinya adalah alat ukur itu harus disusun sedemikian rupa, sehingga dapat menunjukkan perbedaan-perbedaan yang sekecil apapun. Semakin baik suatu alat ukur, maka semakin mampu alat ukur tersebut menunjukkan perbedaan secara teliti. Untuk mengetahui apakah suatu alat ukur cukup deskriminatif atau tidak, biasanya didasarkan atas uji daya pembeda alat ukur tersebut.
7. **Spesifik**  
Spesifik artinya suatu alat ukur disusun dan digunakan khusus untuk objek yang diukur. Jika alat ukur tersebut menggunakan tes, maka jawaban tes jangan menimbulkan ambivalensi atau spekulasi.
8. **Proporsional**  
Proporsional artinya suatu alat ukur harus memiliki tingkat kesulitan yang proporsional antara sulit, sedang dan mudah. Begitu juga ketika menentukan jenis alat ukur, baik tes maupun non-tes.

## B. Model-model Evaluasi

Pada tahun 1949, Tyler pernah mengemukakan model evaluasi *black box*. Model ini banyak digunakan oleh orang-orang yang melakukan kegiatan evaluasi. Studi tentang evaluasi belum begitu menarik perhatian orang banyak, karena kurang memiliki nilai praktis. Baru sekitar tahun 1960-an studi evaluasi mulai berdiri sendiri menjadi salah satu program studi di perguruan tinggi, tidak hanya di jenjang sarjana (S.1) dan magister (S.2) tetapi juga pada jenjang doktor (S.3). Sekitar tahun 1972, model evaluasi mulai berkembang. Taylor dan Cowley, misalnya, berhasil mengumpulkan berbagai pemikiran tentang model evaluasi dan menerbitkannya dalam suatu buku. Model evaluasi yang dikembangkan lebih banyak menggunakan pendekatan positivisme yang berakar pada teori psikometrik. Dalam model tersebut, pengukuran dan tes masih sangat dominan, sekalipun tidak lagi diidentikkan dengan evaluasi. Penggunaan desain eksperimen seperti yang dikemukakan Campbell dan Stanley menjadi ciri utama dari model evaluasi. Berkembangnya model evaluasi pada tahun 70-an tersebut diawali dengan adanya pandangan alternatif dari para *expert*. Pandangan alternatif yang dilandasi sebuah paradigma fenomenologi banyak menampilkan model evaluasi.

Dari sekian banyak model-model evaluasi yang dikemukakan, tes dan pengukuran tidak lagi menempati posisi yang menentukan. Penggunaannya hanya untuk tujuan-tujuan tertentu saja, bukan lagi menjadi suatu keharusan, seperti ketika model pertama ditampilkan. Tes dan pengukuran tidak lagi menjadi parameter kualitas suatu studi evaluasi yang dilakukan. Perkembangan lain yang menarik dalam model evaluasi ini adalah adanya suatu upaya untuk bersikap eklektik

dalam penggunaan pendekatan positivisme maupun fenomenologi yang oleh Patton (1980) disebut *paradigm of choice*. Walaupun usaha ini tidak melahirkan model dalam pengertian terbatas tetapi memberikan alternatif baru dalam melakukan evaluasi.

Dalam studi tentang evaluasi, banyak sekali dijumpai model-model evaluasi dengan format atau sistematika yang berbeda, sekalipun dalam beberapa model ada juga yang sama. Hal ini dapat dikelompokkan model evaluasi sebagai berikut :

- a. Model evaluasi kuantitatif, yang meliputi : model Tyler, model teoritik Taylor dan Maguire, model pendekatan sistem Alkin, model *Countenance Stake*, model CIPP, model ekonomi mikro.
- b. Model evaluasi kualitatif, yang meliputi : model studi kasus, model iluminatif, dan model responsif

Ada juga model evaluasi dikelompokkan yang membagi model evaluasi menjadi ke dalam empat model utama, yaitu "*measurement, congruence, educational system, dan illumination*". Dari beberapa model evaluasi di atas, beberapa diantaranya akan dikemukakan secara singkat sebagai berikut :

#### 1. Model Tyler

Nama model ini diambil dari nama pengembangnya yaitu Tyler. Dalam buku *Basic Principles of Curriculum and Instruction*, Tyler banyak mengemukakan ide dan gagasannya tentang evaluasi. Salah satu bab dari buku tersebut diberinya judul *how can the the effectiveness of learning experience be evaluated ?* Model ini dibangun atas dua dasar pemikiran. *Pertama*, evaluasi ditujukan kepada

tingkah laku peserta didik. *Kedua*, evaluasi harus dilakukan pada tingkah laku awal peserta didik sebelum melaksanakan kegiatan pembelajaran dan sesudah melaksanakan kegiatan pembelajaran (hasil). Dasar pemikiran yang kedua ini menunjukkan bahwa seorang evaluator harus dapat menentukan perubahan tingkah laku apa yang terjadi setelah peserta didik mengikuti pengalaman belajar tertentu, dan menegaskan bahwa perubahan yang terjadi merupakan perubahan yang disebabkan oleh pembelajaran.

Penggunaan model Tyler memerlukan informasi perubahan tingkah laku terutama pada saat sebelum dan sesudah terjadinya pembelajaran. Istilah yang populer dikalangan guru adalah tes awal (*pre-test*) dan tes akhir (*post-test*). Model ini mensyaratkan validitas informasi pada tes akhir. Untuk menjamin validitas ini maka perlu adanya kontrol dengan menggunakan disain eksperimen. Model Tyler disebut juga model *black box* karena model ini sangat menekankan adanya tes awal dan tes akhir. Dengan demikian, apa yang terjadi dalam proses tidak perlu diperhatikan. Dimensi proses ini dianggap sebagai kotak hitam yang menyimpan segala macam teka-teki.

Menurut Tyler, ada tiga langkah pokok yang harus dilakukan, yaitu :

- a. Menentukan tujuan pembelajaran yang akan dievaluasi.
- b. Menentukan situasi dimana peserta didik memperoleh kesempatan untuk menunjukkan tingkah laku yang berhubungan dengan tujuan.
- c. Menentukan alat evaluasi yang akan dipergunakan untuk mengukur tingkah laku peserta didik.

## 2. Model yang Berorientasi pada Tujuan

Sebelum KBK 2004, Anda mungkin pernah mengenal adanya tujuan pembelajaran umum dan tujuan pembelajaran khusus. Model evaluasi ini menggunakan kedua tujuan tersebut sebagai kriteria untuk menentukan keberhasilan. Evaluasi diartikan sebagai proses pengukuran hinggamana tujuan pembelajaran telah tercapai. Model ini banyak digunakan oleh guru-guru karena dianggap lebih praktis untuk menentukan hasil yang diinginkan dengan rumusan yang dapat diukur. Dengan demikian, terdapat hubungan yang logis antara kegiatan, hasil dan prosedur pengukuran hasil. Tujuan model ini adalah membantu Anda merumuskan tujuan dan menjelaskan hubungan antara tujuan dengan kegiatan. Jika rumusan tujuan pembelajaran dapat diobservasi (*observable*) dan dapat diukur (*measurable*), maka kegiatan evaluasi pembelajaran akan menjadi lebih praktis dan simpel.

Model ini dapat membantu Anda menjelaskan rencana pelaksanaan pembelajaran dengan proses pencapaian tujuan. Instrumen yang digunakan bergantung kepada tujuan yang ingin diukur. Hasil evaluasi akan menggambarkan tingkat keberhasilan tujuan program pembelajaran berdasarkan kriteria program khusus. Kelebihan model ini terletak pada hubungan antara tujuan dengan kegiatan dan menekankan pada peserta didik sebagai aspek penting dalam program pembelajaran. Kekurangannya adalah memungkinkan terjadinya proses evaluasi melebihi konsekuensi yang tidak diharapkan.

### 3. Model Pengukuran

Model pengukuran (*measurement model*) banyak mengemukakan pemikiran-pemikiran dari R.Thorndike dan R.L.Ebel. Sesuai dengan namanya, model ini sangat menitikberatkan pada kegiatan pengukuran. Pengukuran digunakan untuk menentukan kuantitas suatu sifat (*atribute*) tertentu yang dimiliki oleh objek, orang maupun peristiwa, dalam bentuk unit ukuran tertentu. Anda dapat menggunakan model ini untuk mengungkap perbedaan-perbedaan individual maupun kelompok dalam hal kemampuan, minat dan sikap. Hasil evaluasi digunakan untuk keperluan seleksi peserta didik, bimbingan, dan perencanaan pendidikan. Objek evaluasi dalam model ini adalah tingkah laku peserta didik, mencakup hasil belajar (kognitif), pembawaan, sikap, minat, bakat, dan juga aspek-aspek kepribadian peserta didik. Untuk itu, instrumen yang digunakan pada umumnya adalah tes tertulis (*paper and pencil test*) dalam bentuk tes objektif, yang cenderung dibakukan. Oleh sebab itu, dalam menganalisis soal sangat memperhatikan *difficulty index* dan *index of discrimination*. Model ini menggunakan pendekatan Penilaian Acuan Norma (*norm-referenced assessment*).

### 4. Model Kesesuaian (Ralph W.Tyler, John B.Carrol, and Lee J.Cronbach)

Menurut model ini, evaluasi adalah suatu kegiatan untuk melihat kesesuaian (*congruence*) antara tujuan dengan hasil belajar yang telah dicapai. Hasil evaluasi dapat Anda gunakan untuk menyempurnakan sistem bimbingan



peserta didik dan untuk memberikan informasi kepada pihak-pihak yang memerlukan. Objek evaluasi adalah tingkah laku peserta didik, yaitu perubahan tingkah laku yang diinginkan (*intended behaviour*) pada akhir kegiatan pendidikan, baik yang menyangkut aspek kognitif, afektif maupun psikomotor. Teknik evaluasi yang dapat Anda gunakan tidak hanya tes (tulisan, lisan, dan perbuatan), tetapi juga non-tes (observasi, wawancara, skala sikap, dan sebagainya). Model evaluasi ini memerlukan informasi perubahan tingkah laku pada dua tahap, yaitu sebelum dan sesudah kegiatan pembelajaran. Berdasarkan konsep ini, Anda perlu melakukan *pre and post-test*. Adapun langkah-langkah yang harus ditempuh dalam model evaluasi ini adalah merumuskan tujuan tingkah laku (*behavioural objectives*), menentukan situasi dimana peserta didik dapat memperlihatkan tingkah laku yang akan dievaluasi, menyusun alat evaluasi, dan menggunakan hasil evaluasi. Oleh sebab itu, model ini menekankan pada pendekatan penilaian acuan patokan (PAP).

5. *Educational System Evaluation Model* (Daniel L.Stufflebeam, Michael Scriven, Robert E.Stake, dan Malcolm M.Provus)

Menurut model ini, evaluasi berarti membandingkan *performance* dari berbagai dimensi (tidak hanya dimensi hasil saja) dengan sejumlah kriteria, baik yang bersifat mutlak/interen maupun relatif/eksteren. Model yang menekankan sistem sebagai suatu keseluruhan ini sebenarnya merupakan penggabungan dari beberapa model, sehingga objek evaluasinya pun diambil dari

beberapa model, yaitu (1) model *countenance* dari Stake, yang meliputi : keadaan sebelum kegiatan pembelajaran berlangsung (*antecedents*), kegiatan yang terjadi dan saling mempengaruhi (*transactions*), hasil yang diperoleh (*outcomes*), (2) model CIPP dari Stufflebeam, yang meliputi *Context, Input, Process, dan Product*, (3) model Scriven yang meliputi *instrumental evaluation and consequential evaluation*, (4) model Provus yang meliputi : *design, operation program, interim products, dan terminal products*. Dari keempat model yang tergabung dalam *educational system model*, akan dijelaskan secara singkat tentang dua model, yaitu model *countenance* dan model CIPP.

Model Stake menitikberatkan evaluasi pada dua hal pokok, yaitu *description* dan *judgement*. Setiap hal tersebut terdiri atas tiga dimensi, seperti telah dijelaskan di atas, yaitu *antecedents (context), transaction (process), dan outcomes (output)*. *Description* terdiri atas dua aspek, yaitu *intents (goals)* dan *observation (effects)* atau yang sebenarnya terjadi. Sedangkan *judgement* terdiri atas dua aspek, yaitu *standard* dan *judgement*. Dalam model ini, evaluasi dilakukan dengan membandingkan antara satu program dengan program lain yang dianggap standar. Stake mengatakan *description* berbeda dengan *judgement* atau menilai. Dalam ketiga dimensi di atas (*antecedents, transaction, outcomes*), perbandingan data tidak hanya untuk menentukan apakah ada perbedaan tujuan dengan keadaan yang sebenarnya tetapi juga dibandingkan dengan standar yang absolut untuk menilai manfaat program. Menurut Stake, suatu hasil penelitian tidak dapat diandalkan jika tidak dilakukan evaluasi.

Model CIPP berorientasi kepada suatu keputusan (*a decision oriented evaluation approach structured*). Tujuannya membantu kepala madrasah dan guru di dalam membuat keputusan. Evaluasi diartikan sebagai suatu proses menggambarkan, memperoleh dan menyediakan informasi yang berguna untuk menilai alternatif keputusan. Sesuai dengan nama modelnya, model ini membagi empat jenis kegiatan evaluasi, yaitu :

- a. *Context evaluation to serve planning decision*, yaitu konteks evaluasi untuk membantu merencanakan keputusan, menentukan kebutuhan yang akan dicapai oleh program pembelajaran, dan merumuskan tujuan program pembelajaran.
- b. *Input evaluation, structuring decision*. Kegiatan evaluasi bertujuan untuk membantu mengatur keputusan, menentukan sumber-sumber yang ada, alternatif apa yang akan diambil, apa rencana dan strategi untuk mencapai tujuan, dan bagaimana prosedur kerja untuk mencapainya.
- c. *Process evaluation, to serve implementing decision*. Kegiatan evaluasi ini bertujuan untuk membantu melaksanakan keputusan. Pertanyaan yang harus Anda jawab adalah hinggamana suatu rencana telah dilaksanakan, apakah rencana tersebut sesuai dengan prosedur kerja, dan apa yang harus diperbaiki.
- d. *Product evaluation, to serve recycling decision*. Kegiatan evaluasi ini bertujuan untuk membantu keputusan selanjutnya. Pertanyaan yang harus Anda jawab adalah hasil apa yang telah dicapai dan apa yang dilakukan setelah program berjalan.

Proses evaluasi tidak hanya berakhir dengan suatu deskripsi mengenai keadaan sistem yang bersangkutan, tetapi harus sampai pada *judgment* sebagai simpulan dari hasil evaluasi. Model ini menuntut agar hasil evaluasi digunakan sebagai *input* untuk *decision making* dalam rangka penyempurnaan sistem secara keseluruhan. Pendekatan yang digunakan adalah penilaian acuan norma (PAN) dan penilaian acuan patokan (PAP).

6. *Illuminative Model* (Malcolm Parlett dan Hamilton)

Jika model *measurement* dan *congruence* lebih berorientasi pada evaluasi kuantitatif-terstruktur, maka model ini lebih menekankan pada evaluasi kualitatif-terbuka (*open-ended*). Kegiatan evaluasi dihubungkan dengan *learning milieu*, dalam konteks madrasah sebagai lingkungan material dan psiko-sosial, dimana guru dan peserta didik dapat berinteraksi. Tujuan evaluasi adalah untuk mempelajari secara cermat dan hati-hati terhadap pelaksanaan sistem pembelajaran, faktor-faktor yang mempengaruhinya, kelebihan dan kekurangan sistem, dan pengaruh sistem terhadap pengalaman belajar peserta didik. Hasil evaluasi lebih bersifat deskriptif dan interpretasi, bukan pengukuran dan prediksi. Model ini lebih banyak menggunakan *judgment*. Fungsi evaluasi adalah sebagai input untuk kepentingan pengambilan keputusan dalam rangka penyesuaian dan juga penyempurnaan sistem pembelajaran yang sedang dikembangkan.

Objek evaluasi model ini mencakup latar belakang dan perkembangan sistem pembelajaran, proses pelaksanaan

sistem pembelajaran, hasil belajar peserta didik, kesukaran-kesukaran yang dialami dari perencanaan sampai dengan pelaksanaan, termasuk efek samping dari sistem pembelajaran itu sendiri. Pendekatan yang digunakan lebih menyerupai pendekatan yang diterapkan dalam bidang antropologi sosial, psikiatri, dan sosiologi. Cara-cara yang digunakan tidak bersifat *standard*, melainkan bersifat fleksibel dan selektif. Berdasarkan tujuan dan pendekatan evaluasi dalam model ini, maka ada tiga fase evaluasi yang harus Anda tempuh, yaitu : *observe, inquiry further, dan seek to explain.*

## 7. Model Responsif

Sebagaimana model illuminatif, model ini juga menekankan pada pendekatan kualitatif-naturalistik. Evaluasi tidak diartikan sebagai pengukuran melainkan pemberian makna atau melukiskan sebuah realitas dari berbagai perspektif orang-orang yang terlibat, berminat dan berkepentingan dengan program pembelajaran. Tujuan evaluasi adalah untuk memahami semua komponen program pembelajaran melalui berbagai sudut pandangan yang berbeda. Sesuai dengan pendekatan yang digunakan, maka model ini kurang percaya terhadap hal-hal yang bersifat kuantitatif. Instrumen yang digunakan pada umumnya mengandalkan observasi langsung maupun tak langsung dengan interpretasi data yang impresionistik. Langkah-langkah kegiatan evaluasi meliputi kegiatan observasi, merekam hasil wawancara, mengumpulkan data, mengecek pengetahuan awal (*preliminary understanding*) peserta didik dan mengembangkan disain atau model. Berdasarkan langkah-

langkah ini, evaluator mencoba responsif terhadap orang-orang yang berkepentingan pada hasil evaluasi. Hal yang penting dalam model responsif adalah pengumpulan dan sintesis data.

Kelebihan model ini adalah peka terhadap berbagai pandangan dan kemampuannya dalam mengakomodasi pendapat yang ambigu serta tidak fokus. Sedangkan kekurangannya antara lain (1) pembuat keputusan sulit menentukan prioritas atau penyederhanaan informasi (2) tidak mungkin menampung semua sudut pandangan dari berbagai kelompok (3) membutuhkan waktu dan tenaga. Evaluator harus dapat beradaptasi dengan lingkungan yang diamati. Untuk mempelajari lebih jauh tentang model ini, silahkan Anda membaca buku Stake (1975) atau Lincoln dan Guba (1985).

Setelah Anda mempelajari berbagai model evaluasi, model mana yang akan digunakan dalam pembelajaran ? Jawabannya tentu sangat bergantung kepada tujuan evaluasi yang ditetapkan. Namun demikian, perlu juga Anda pahami bahwa keberhasilan suatu evaluasi pembelajaran secara keseluruhan bukan hanya dipengaruhi penggunaan yang tepat pada sebuah model evaluasi melainkan juga dipengaruhi oleh berbagai faktor. *Pertama*, tujuan pembelajaran, baik tujuan pembelajaran umum maupun tujuan pembelajaran khusus (*instructional objective*). Seringkali kedua tujuan pembelajaran ini saling bertentangan satu sama lain dilihat dari kebutuhan madrasah, kurikulum, guru, peserta didik, lingkungan, dan sebagainya. Bahkan, kadang-kadang guru sendiri mempunyai tujuan sendiri-sendiri. Semuanya harus dipertimbangkan agar terdapat keseimbangan dan keserasian.

*Kedua*, sistem madrasah. Faktor ini perlu dipertimbangkan dengan matang dan hati-hati karena melibatkan berbagai komponen yang saling berinteraksi dan ketergantungan. Mengingat kompleksnya sistem madrasah, maka fungsi madrasah juga menjadi ganda. Di satu pihak, madrasah ingin mewariskan kebudayaan masa lampau dengan sistem norma, nilai dan adat yang dianggap terbaik untuk generasi muda. Di pihak lain, madrasah berkewajiban mempersiapkan peserta didik menghadapi masa depan, memperoleh keterampilan dan kemampuan untuk berinovasi, bahkan menghasilkan perubahan. Jadi, madrasah sekaligus bersikap konservatif-radikal serta reaksioner-progresif. Oleh sebab itu, peranan evaluasi menjadi sangat penting. Tujuannya adalah untuk melihat dan mempertimbangkan hal-hal apa yang perlu diberikan di madrasah. Begitu juga bentuk kurikulum dan silabus mata pelajaran sangat bergantung pada evaluasi yang dilaksanakan oleh guru-guru di madrasah, sehingga timbul masalah lainnya yaitu teknik evaluasi apa yang akan digunakan untuk mencapai tujuan itu.

*Ketiga*, pembinaan guru. Banyak program pembinaan guru yang belum menyentuh secara langsung tentang evaluasi. Program pembinaan guru lebih banyak difokuskan kepada pengembangan kurikulum dan metodologi pembelajaran. Hal ini pula yang menyebabkan perbaikan sistem evaluasi pembelajaran menjadi kurang efektif. Guru juga sering dihadapkan dengan beragam kegiatan, seperti membuat persiapan mengajar, mengikuti kegiatan ekstra kurikuler, penyesuaian diri, dan kegiatan administratif lainnya. Artinya, bagaimana mungkin kualitas sistem evaluasi pembelajaran di madrasah dapat ditingkatkan, bila fokus pembinaan guru hanya menyentuh domain-domain tertentu saja, ditambah lagi

dengan kesibukan-kesibukan guru di luar tugas pokoknya sebagai pengajar.

### **C. Pendekatan Evaluasi**

Pendekatan merupakan sudut pandang seseorang dalam mempelajari sesuatu. Dengan demikian, pendekatan evaluasi merupakan sudut pandang seseorang dalam menelaah atau mempelajari evaluasi. Dilihat dari komponen pembelajaran, pendekatan evaluasi dapat dibagi dua, yaitu pendekatan tradisional dan pendekatan sistem. Dilihat dari penafsiran hasil evaluasi, pendekatan evaluasi dibagi menjadi dua, yaitu *criterion-referenced evaluation* dan *norm-referenced evaluation*.

#### **1. Pendekatan tradisional**

Pendekatan ini berorientasi kepada praktik evaluasi yang telah berjalan selama ini di sekolah yang ditujukan kepada perkembangan aspek intelektual peserta didik. Aspek-aspek keterampilan dan pengembangan sikap kurang mendapat perhatian yang serius. Peserta didik hanya dituntut untuk menguasai mata pelajaran. Kegiatan-kegiatan evaluasi juga lebih difokuskan kepada komponen produk saja, sementara komponen proses cenderung diabaikan. Hasil kajian Spencer cukup memberikan gambaran betapa pentingnya evaluasi pembelajaran. Ia mengemukakan sejumlah isi pendidikan yang dapat dijadikan dasar pertimbangan untuk merumuskan tujuan pendidikan secara komprehensif dan pada gilirannya menjadi acuan dalam membuat perencanaan evaluasi. Namun demikian, tidak sedikit guru mengalami kesulitan untuk mengembangkan sistem evaluasi di sekolah karena



bertentangan dengan tradisi yang selama ini sudah berjalan. Misalnya, ada tradisi bahwa target kuantitas kelulusan setiap sekolah harus di atas 95%, begitu juga untuk kenaikan kelas. Ada juga tradisi bahwa dalam mata pelajaran tertentu nilai peserta didik dalam buku rapot harus minimal enam. Seharusnya, kebijakan evaluasi lebih menekankan kepada target kualitas yaitu kepentingan dan kebermaknaan pendidikan bagi anak.

## 2. Pendekatan sistem

Sistem adalah totalitas dari berbagai komponen yang saling berhubungan dan ketergantungan. Jika pendekatan sistem dikaitkan dengan evaluasi, maka pembahasan lebih difokuskan kepada komponen evaluasi, yang meliputi : komponen kebutuhan dan *feasibility*, komponen input, komponen proses, dan komponen produk. Dalam bahasa Stufflebeam disingkat CIPP, yaitu *context*, *input*, *process* dan *prudent*. Komponen-komponen ini harus menjadi landasan pertimbangan dalam evaluasi pembelajaran secara sistematis. Berbeda dengan pendekatan tradisional yang hanya menyentuh komponen produk saja, yaitu perubahan perilaku apa yang terjadi pada peserta didik setelah mengikuti proses pembelajaran. Pendekatan ini tentu tidak salah, hanya tidak sistematis. Padahal, Anda juga tahu bahwa hasil belajar tidak akan ada bila tidak melalui proses, dan proses tidak bisa berjalan bila tidak ada masukan dan guru yang melaksanakan.

Dalam literatur modern tentang evaluasi, terdapat dua pendekatan yang dapat digunakan untuk menafsirkan hasil evaluasi, yaitu penilaian acuan patokan (*criterion-referenced evaluation*) dan penilaian acuan norma (*norm-referenced*

*evaluation*). Artinya, setelah Anda memperoleh skor mentah dari setiap peserta didik, maka langkah selanjutnya adalah mengubah skor mentah menjadi nilai dengan menggunakan pendekatan tertentu.

a. Penilaian Acuan Patokan (PAP)

Pendekatan ini sering juga disebut penilaian norma absolut. Jika Anda ingin menggunakan pendekatan ini, berarti Anda harus membandingkan hasil yang diperoleh peserta didik dengan sebuah patokan atau kriteria yang secara absolut atau mutlak telah ditetapkan oleh guru. Anda juga dapat menggunakan langkah-langkah tertentu untuk menggunakan PAP, seperti menentukan skor ideal, mencari rata-rata dan simpangan baku ideal, kemudian menggunakan pedoman konversi skala nilai. Pendekatan ini cocok digunakan dalam evaluasi atau penilaian formatif yang berfungsi untuk perbaikan proses pembelajaran. Umumnya, seorang guru yang menggunakan PAP sudah dapat menyusun pedoman konversi skor menjadi skor standar sebelum kegiatan evaluasi dimulai. Oleh sebab itu, hasil pengukuran dari waktu ke waktu dalam kelompok yang sama atau berbeda dapat dipertahankan *keajegannya*. PAP dapat menggambarkan prestasi belajar peserta didik secara objektif apabila alat ukur yang digunakan adalah alat ukur yang standar.

b. Penilaian Acuan Norma (PAN)

Salah satu perbedaan PAP dengan PAN adalah penggunaan tolak ukur hasil/skor sebagai pembanding. Pendekatan ini membandingkan skor setiap peserta didik dengan teman satu kelasnya. Makna nilai dalam bentuk angka maupun kualifikasi memiliki sifat relatif. Artinya, jika Anda sudah

menyusun pedoman konversi skor untuk suatu kelompok, maka pedoman itu hanya berlaku untuk kelompok itu saja dan tidak berlaku untuk kelompok yang lain, karena distribusi skor peserta didik sudah berbeda.

c. Pendekatan Deskript

Dalam pendekatan ini, istilah *diskret* digunakan untuk menggambarkan dua aspek yang berbeda dalam tes bahasa, yakni (1) isi atau tugas, dan (2) model jawaban dan penyekoran jawaban.

Dari segi isi atau tugas, tes dengan pendekatan ini menyangkut satu aspek kebahasaan saja pada satu kesempatan pengetesan, misalnya aspek fonologi, morfologi, sintaksis, atau kosa-kata saja. Tiap satu butir soal hanya dimaksudkan untuk mengukur satu aspek kebahasaan saja. Dari segi model jawaban, tes dengan pendekatan ini berupa penjodohan (*matching*), benar-salah (*true-flase*), pilihan ganda (*multiple choiche*), atau mengisi kotak kosong yang disediakan dengan jawaban yang sudah tersedia pada kolom lain. Dari segi penyekoran jawaban, model jawaban yang seperti itu sangat memudahkan guru atau korektor dalam memberikan penilaian. Penyekoran berdasarkan model jawaban seperti itu memiliki reliabilitas yang sangat tinggi. Dengan bantuan komputer misalnya, penyekoran jawaban hampir 100% tidak diragukan lagi keakuratannya.

d. Pendekatan Integratif

Jika dalam pendekatan diskret, aspek-aspek kebahasaan dan kemampuan berbahasa itu diperlakukan secara terpisah, maka dalam pendekatan integratif aspek-aspek

bahasa dan kemampuan berbahasa itu dicakup secara bersamaan.

Jika dalam tes diskret hanya diujikan satu aspek kebahasaan saja pada satu waktu, maka dalam tes integratif berusaha diukur beberapa aspek kebahasaan secara bersamaan. Prinsip ini sesuai dengan pandangan psikologi Gestalt yang intinya “bahwa tingkah laku itu dipelajari sebagai kesatuan yang tidak terpisahkan atau *“gestalts”*”.

Berdasarkan pandangan ini, maka tes integratif tidak secara khusus mengeteskan salah satu aspek kebahasaan seperti fonologi, morfologi, sintaksis, atau kosa kata, atau salah satu dari kemampuan berbahasa seperti membaca, menulis, berbicara, atau menyimak, melainkan sebuah tes dalam satu waktu meliputi beberapa aspek kebahasaan dan kemampuan berbahasa sekaligus.

e. Pendekatan Pragmatik

Pendekatan pragmatik pada awalnya digunakan dalam kaitannya dengan teori tentang kemampuan memahami berdasarkan kemampuan tata bahasa pragmatik (*pragmatic expectancy grammar*). Kemampuan itu merupakan kemampuan untuk memahami teks atau wacana, tidak hanya dalam konteks linguistic melainkan juga dengan memanfaatkan kemampuan pemahaman unsur-unsur ekstra linguistik (seluk beluk bidang yang dibahas dalam teks bacaan).

f. Pendekatan Komunikatif

Tes bahasa komunikatif adalah tes yang melibatkan konsep kompetensi komunikatif. Kompetensi komunikatif adalah suatu kompetensi yang melihat kemampuan

pelajar tidak hanya kemampuan membentuk kalimat yang benar tetapi juga menggunakannya secara tepat. Tes bahasa secara komunikatif bertujuan untuk mengukur bagaimana orang yang diuji mampu menggunakan bahasa di dalam situasi kehidupan nyata.

## Bab III

### Instrumen Evaluasi Jenis Tes

**T**es merupakan instrumen evaluasi yang paling umum dipakai dalam dunia pendidikan sebagai alat ukur untuk domain kognitif. Tes memiliki jenis yang beragam sesuai dengan fungsinya, seperti tes prestasi belajar (*achievement test*), tes penguasaan (*proficiency test*), tes bakat (*aptitude test*), tes diagnostik (*diagnostic test*), dan tes penempatan (*placement test*). Jika dilihat dari bentuk jawaban peserta didik, maka tes dapat dibagi menjadi tiga jenis, yaitu tes tertulis, tes lisan, dan tes perbuatan. Tes tertulis ada dua bentuk, yaitu bentuk uraian (*essay*) dan bentuk objektif (*objective*).

#### A. Tes Tertulis Bentuk Uraian (Essay).

Dilihat dari penamaannya, tes bentuk uraian merupakan tes yang menuntut penerima tes mengeluarkan jawaban-jawaban berbentuk uraian, baik secara bebas maupun secara terbatas. Tes bentuk uraian secara bebas artinya jawaban uraian peserta didik yang menuntut kemampuan peserta didik dalam menyusun, mengorganisasikan dan merumuskan

jawaban menggunakan kata-kata sendiri serta mampu mengukur kecakapan peserta didik untuk berfikir tingkat tinggi.

Sementara tes uraian terbatas tepat digunakan untuk mengukur kemampuan siswa dalam menjelaskan hubungan sebab akibat, menerapkan suatu prinsip atau teori, memberikan alasan yang relevan, merumuskan hipotesis, membuat kesimpulan yang tepat, menjelaskan suatu prosedur, dan sebagainya.

Berikut adalah contoh masing-masing pertanyaan tes uraian terbatas (*restricted respons items*) dan uraian bebas (*extended respons items*):

1. Tes uraian dalam bentuk bebas atau terbuka.

Contoh:

Coba jelaskan fungsi dan tujuan belajar Matematika dalam kehidupan dan berikan contohnya.

2. Tes uraian dalam bentuk uraian terbatas.

Contoh:

Andi memiliki 18 kelereng merah dan 22 kelereng putih lalu dimasukkan kedalam kotak. Tiap kotak berisi kelereng merah yang sama banyak dan kelerengn putih yang sama banyak pula. Berapa banyak kotak yang diperlukan?. Berapa kelereng merah dan kelereng putih dalam setiap kotak?

Tes uraian sebagaimana dicontohkan di atas memiliki berbagai karakteristik, yaitu:

1. Tes tersebut bentuk pertanyaan atau perintah yang menghendaki jawaban berupa uraian atau paparan kalimat yang pada umumnya cukup panjang.
2. Bentuk pertanyaan atau perintah itu menuntut kepada tester untuk memberikan penjelasan, komentar, penafsiran, membandingkan, membedakan, dan sebagainya.
3. Jumlah soal butir uraiannya terbatas yaitu berkisar lima sampai dengan sepuluh butir.
4. Pada umumnya butir-butir soal uraian diawali dengan kata-kata, "*uraikan*",.... "*Mengapa*" ,...."*Terangkan*", ...."*Jelaskan*".

Untuk penyusunan jenis tes bentuk uraian ada beberapa langkah yang dapat dipedomani sebagai berikut:

1. Dalam menyusun butir-butir soal tes uraian diusahakan agar soal tersebut dapat mencakup ide-ide pokok dari materi pelajaran yang telah diajarkan.
2. Untuk menghindari tumbuhnya perbuatan curang oleh tester misalnya, menyontek dan bertanya kepada tester yang lainya hendaknya sesuatu kalimat pada soal berlawanan dengan buku pelajaran.
3. Dalam menyusun butir-butir soal tes uraian hendaknya diusahakan agar pertanyaan-pertanyaan itu jangan dibuat seragam melainkan bervariasi.

Contohnya:

*Jelaskan perbedaan antara ... dengan .. dan kemukakan alasannya... mengapa...*

4. Kalimat soal yang disusun hendaklah ringkas dan padat.



5. Sebelum tester mengerjakan soal hendaklah seorang tester mengemukakan cara mengerjakannya, contoh, "Jawaban soal harus ditulis di atas lembaran jawaban dan sesuai dengan urutan nomor.

Sebagaimana jenis tes lainnya, tes uraian juga memiliki beberapa kelebihan dan kekurangan. Kelebihan tes uraian diantaranya adalah:

- Bagi guru, menyusun tes tersebut sangat mudah dan tidak memerlukan waktu yang lama.
- Si penjawab mempunyai kebebasan dalam menjawab dan mengeluarkan isi hati dan buah pikirannya.
- Melatih mengeluarkan pikiran dalam bentuk kalimat atau bahasa yang teratur.
- Lebih ekonomis, hemat karena tidak memerlukan kertas terlalu banyak untuk membuat soal tes, dapat didektekan atau ditulis dipapan tulis.

Sedangkan kelemahan tes uraian yakni:

- Tidak atau kurang dapat digunakan untuk mengetes pelajaran yang luas atau banyak sehingga kurang dapat menilai isi pengetahuan siswa yang sebenarnya.
- Kemungkinan jawaban dan keterangan sifatnya menyulitkan penjelasan pengetesan dalam menemukannya.
- Baik buruknya tulisan dan panjang pendeknya jawaban yang sama mudah menimbulkan evaluasi dan perskoran (*scoring*) yang kurang objektif.

## **B. Tes Hasil Belajar Bentuk Objektif**

Tes objektif disebut objektif karena cara pemeriksaannya yang seragam terhadap semua peserta didik yang mengikuti sebuah tes. Tes objektif juga dikenal dengan istilah tes jawaban pendek (*short answer test*), dan salah satu tes hasil belajar yang terdiri dari butir-butir soal (*items*) yang dapat dijawab oleh tester dengan jalan memilih salah satu (atau lebih), di antara beberapa kemungkinan jawaban yang telah dipasangkan pada masing-masing items atau dengan jalan menuliskan jawabannya berupa kata-kata atau simbol-simbol tertentu pada tempat-tempat yang disediakan untuk masing-masing butir yang bersangkutan. Terdapat beberapa jenis tes bentuk objektif, misalnya: bentuk melengkapi (*completion test*), pilihan ganda (*multiple choice*), menjodohkan (*matching*), bentuk pilihan benar-salah (*true false*). Lebih jelasnya diuraikan subbagi berikut.

### **1. Melengkapi (*Completion test*).**

Completion test adalah dikenal dengan istilah melengkapi atau menyempurnakan. Salah satu jenis objektif yang hampir mirip sekali dengan tes objektif fill in. Letak perbedaannya ialah pada tes objektif bentuk fill in bahan yang dites itu merupakan satu kesatuan. Sedangkan pada tes objektif bentuk completion tidak harus demikian.

Contoh:

*Isilah titik-titik dibawah ini dengan jawaban yang benar dan tepat. Faktor prima dari bilangan 15 adalah .....*

Test completion memiliki kelebihan yakni:

- a) Test ini amat mudah dalam penyusunannya.
- b) Jika dibanding dengan tes objektif bentuk *fill in*, tes objektif ini lebih menghemat tempat (kertas).
- c) Karena bahan yang disajikan dalam tes ini cukup banyak dan beragam.
- d) Test ini juga dapat digunakan untuk mengukur berbagai taraf kompetensi dan tidak sekedar mengungkapkan taraf pengenalan atau hapalan saja.

Kekurangan tes completion yakni :

- a) Pada umumnya tester cenderung menggunakan tes model ini untuk mengungkapkan daya ingat atau aspek hapalan saja.
- b) Dapat terjadi bahwa butir-butir item dari tes model ini kurang relevan untuk disajikan.
- c) Karena pembuatannya mudah, maka tester sering kurang hati-hati dalam membuat soal-soal.

## **2. Test objektif bentuk *multifile choice test* (pilihan berganda)**

*Test multifile choice*, tes pilihan ganda merupakan tes objektif dimana masing-masing tes disediakan lebih dari kemungkinan jawaban, dan hanya satu dari pilihan-pilihan tersebut yang benar atau yang paling benar. Penyusunan tes dalam bentuk *multifile choice*:

- a) Hendaknya antara pernyataan dalam soal dengan alternatif jawaban terdapat kesesuaian.
- b) Kalimat pada tiap-tiap butir soal hendaknya dapat disusun dengan jelas.

- c) Sebaiknya soal hendaknya disusun menggunakan bahasa yang mudah dipahami.
- d) Setiap butir pertanyaan hendaknya hanya mengandung satu masalah, meskipun masalah itu agak kompleks.

Contoh: Hasil pembagian  $\frac{3}{4} : \frac{1}{2}$  adalah:

- a.  $1 \frac{1}{2}$
- b.  $2 \frac{1}{2}$
- c.  $3 \frac{1}{2}$
- d.  $4 \frac{1}{2}$

Menurut Sumadi Surya Brata, merinci tes *multiple choice* ada beberapa macam yaitu:

1. Jenis jawaban benar. Contoh: Hasil penjumlahan  $-8 + 3$  adalah:
  - a. 6
  - b. 5
  - c. 4
  - d. 3
  
2. Jawaban yang sesuai yang paling tepat pertanyaan yang diikuti dengan alternatif. Contoh: membaca ayat al-Quran bertujuan untuk:
  - a. Mendapat pahala.
  - b. Melaksanakan perintahnya.
  - c. Mengingat zikir kepada Allah.
  - d. Selamat dunia akhirat.

3. Jawaban tidak sesuai. Contoh: Diantara makhluk Allah yang diciptakan yakni manusia dan ..
  - a. Hewan.
  - b. Laut.
  - c. Bumi.
  - d. Tanah.
  
4. Jawaban negatif dalam suatu soal bentuk *multifile choice* peserta didik diberi pernyataan yang disediakan alternatif jawaban. Sebagian besar dari alternatif tersebut merupakan jawaban yang benar, kecuali ada satu yang salah. Contoh: Manakah diantara Rasul-Rasul di bawah ini yang tidak termasuk ulul azmi.
  - a. Adam.
  - b. Ibrahim.
  - c. Musa.
  - d. Isa.

### 3. Test objektif bentuk *matching* (menjodohkan)

Test bentuk ini sering dikenal dengan istilah tes menjodohkan, tes mencari pandangan, tes menyesuaikan, tes mencocokkan. Ciri-ciri tes ini adalah :

- a) Test terdiri dari satu seri pertanyaan dan satu seri jawaban.
- b) Tugas tes adalah mencari dan menetapkan jawaban-jawaban yang telah bersedia sehingga sesuai dengan atau cocok atau merupakan pasangan, atau merupakan “jodoh” dari pertanyaan.

Contoh sebagai berikut :

1. Cabang Ilmu yang mempelajari tentang iklim....	a. Geomorfologi
2. Cabang Ilmu yang mempelajari tentang cuaca....	b. Hidrologi
3. Cabang Ilmu yang mempelajari bentuk muka bumi....	c. Klimatologi
4. Cabang Ilmu yang mempelajari perairan di darat....	d. Meteorologi
	e. Pedologi

Test bentuk *matching* memiliki kelebihan dan kekurangan. Kelebihan dari tes ini adalah.

- Pembuatan mudah.
- Dapat dinilai dengan mudah, cepat dan objektif.
- Apabilas tes jenis ini dibuat dengan baik, maka faktor merubah praktis dapat dihilangkan.
- Test ini sangat berguna untuk menilai berbagai hal.

Kelemahan dari test *matching* yakni:

- *Matching* test cenderung lebih banyak mengungkap aspek hapalan atau daya ingat.
- Karena mudah disusun, maka tes jenis ini kurang baik acap kali dijadikan “pelarian” bagi pengajaran, yaitu kalau pengajar tidak sempat lagi untuk membuat tes bentuk lain.

- Karena jawaban yang pendek, maka tes ini kurang baik untuk mengevaluasi pengertian dan kemampuan membuat tafsiran.

Adapun cara menyusunnya.

- Hendaknya butir-butir dari soal yang dituangkan dalam bentuk meching test ini jumlahnya tidak kurang dari 10 dan tidak lebih dari 15 soal.
- Daftar yang berada disebelah kiri hendaknya dibuat lebih panjang ketimbang daftar yang disebelah kanan, agar jawaban dapat dengan cepat dicari dan ditemukan oleh tester.
- Sekalipun kadang-kadang sulit dilaksanakan, usahakanlah agar petunjuk tentang cara mengerjakan soal dibuat ringkas dan setengah mungkin.

## 5. Test objektif bentuk fill in (isian).

Test objektif bentuk fill in ini biasanya berbentuk cerita atau karangan. Test objektif fill ini memiliki kelebihan dan kelemahan. Kelebihannya ialah:

- a) Dengan menggunakan tes objektif bentuk fill in maka masalah yang diwujudkan tertuang secara keseluruhan dalam konteksnya.
- b) Cara penyusunannya mudah.

Adapun kekurangannya adalah:

- a) Test objektif fill ini cenderung lebih banyak mengungkapkan aspek pengetahuan atau pengenalan saja.
- b) Test ini juga sifatnya konfrensif, sebab hanya dapat mengungkapkan sebahagian saja dari bahan yang seharusnya ditekankan.

Cara penyusunan tes objektif bentuk fill in:

- a) Agar tes ini dapat digunakan secara efisien sebaiknya jawaban yang harus diisikan ditulis pada lembar jawaban atau pada tempat yang terpisah.
- b) Ungkapan cerita yang dijadikan bahan tes hendaknya disusun ringkas mungkin demi menghemat tempat atau kertas serta waktu penyesuaiannya.
- c) Apabila jenis mata pelajaran yang akan disajikan itu memungkinkan pengajaran atau pengujian soal juga dapat dituangkan dalam bentuk gambar.

## 6. Test objektif bentuk *True False* (benar salah).

Test ini juga sering dikenal dengan tes objektif bentuk “Ya-Tidak” tes objektif bentuk true false adalah salah satu bentuk tes, dimana ada yang benar dan ada yang salah. Contohnya adalah:

- a) (B)-(S). Rasulullah dilahirkan pada tahun 571 H bertepatan dengan tahun Gajah.
- b) (B)-(S). Rasulullah dijuluki dengan “Al-Amin” karena beliau tidak pernah bohong.

Kelebihan dan kekurangan *test true-false*, kelebihanya ialah :

- a) Pembuatan mudah dapat dipergunakan berulang kali.
- b) Dapat mencakup bahan pelajaran yang luas.
- c) Tidak terlalu banyak memakan kertas.
- d) Bagi tester cara mengerjakannya mudah.

Adapun kekurangannya adalah:

- a) Test objektif bentuk true false membuka peluang bagi tester untuk berspekulasi dalam memberikan jawaban.



- b) Sifatnya awal terbatas dalam arti bahwa tes tersebut hanya dapat mengungkapkan daya ingat dan pergerakan kembali saja.
- c) Dapat terjadi bahwa butir-butir soal tes objektif, jenis ini tidak dapat dijawab dengan dua kemungkinan saja yakni benar atau salah.

Contohnya:

- a) B-S Test objektif lebih baik dari pada tes subjektif.
- b) B-S IPS lebih berguna untuk dipelajari ketimbang IPA.

Adapun cara penyusunan *test true false* adalah:

- a) Seyogianya membuat petunjuk yang jelas, bagaimana mengerjakan soal tes, agar anak tidak bingung.
- b) Jangan membuat pernyataan yang masih dapat dipersoalkan antara benar dan salahnya, pernyataan sudah benar atau salah.
- c) Setiap soal supaya mengandung satu pengertian saja, jangan membuat soal yang banyak mengandung pengertian.
- d) Dalam membuat soal jangan ada kata-kata yang meragukan misalnya dengan kata “Kadang” “Barang kali”. Sekarang ini bentuk true false tidak diperlukan lagi untuk tes hasil belajar karena bentuk ini dianggap kurang tepat untuk mengukur tingkat kemajuan belajar anak.

Berikut ditampilkan kisi-kisi instrumen penilaian tes tertulis pilihan ganda dan uraian sebagai berikut:

Mata pelajaran :  
Kelas/Semester :  
Alokasi waktu :  
Jumlah Soal :  
Bentuk Soal : Pilihan ganda/Uraian

No	KD	Materi	Indikator	No urut Soal

### C. Tes Tindakan (*Performance Test*)

Tes tindakan adalah tes yang menuntut jawaban peserta didik dalam bentuk perilaku, tindakan, atau perbuatan di bawah pengawasan penguji yang akan mengobservasi penampilannya dan membuat keputusan tentang kualitas hasil belajar yang dihasilkannya atau ditampikannya. Peserta didik bertindak sesuai dengan apa yang diperintahkan dan ditanyakan.

Tes tindakan dapat digunakan untuk menilai kualitas suatu pekerjaan yang telah selesai dikerjakan oleh peserta didik, termasuk juga keterampilan dan ketepatan menyelesaikan suatu pekerjaan, kecepatan dan kemampuan merencanakan suatu pekerjaan. Tindakan atau unjuk kerja yang dapat dinilai seperti: memainkan alat musik, bernyanyi,

membaca puisi atau deklamasi, menggunakan peralatan laboratorium, dan mengoperasikan suatu alat.

Penilaian ini cocok digunakan untuk menilai ketercapaian kompetensi yang menuntut peserta didik menunjukkan unjuk kerja. Cara penilaian ini dianggap lebih otentik daripada tes tertulis karena apa yang dinilai lebih mencerminkan kemampuan peserta didik yang sebenarnya.

Contoh tes tindakan:

*Coba tunjukkan di depan kelas bagaimana cara mengajar dengan menggunakan model pembelajaran aktif tipe jigsaw.*

Tes jenis ini sangat bermanfaat untuk memperbaiki kemampuan/perilaku peserta didik, karena secara objektif kesalahan-kesalahan yang dibuat oleh peserta didik dapat diamati dan diukur, sehingga menjadi dasar pertimbangan untuk praktik selanjutnya. Sebagaimana jenis tes yang lain, tes tindakan pun mempunyai kelebihan dan kekurangan.

Kelebihan tes tindakan adalah:

1. satu-satunya teknik tes yang dapat digunakan untuk mengetahui hasil belajar dalam bidang keterampilan, seperti keterampilan membaca al-Qur'an berdasarkan ilmu tajwid.
2. sangat baik digunakan untuk mencocokkan kesesuaian antara pengetahuan teori dengan keterampilan praktik, sehingga hasil penilaian menjadi lengkap.
3. dalam pelaksanaannya tidak memungkinkan peserta didik untuk saling menyontek.

4. guru dapat lebih mengenal karakteristik masing-masing peserta didik sebagai dasar tindak lanjut hasil penilaian, seperti pembelajaran remedial.

Adapun kelemahan/kekurangan tes tindakan adalah:

1. memakan waktu yang lama.
2. dalam hal tertentu membutuhkan biaya yang besar.
3. cepat membosankan.
4. jika tes tindakan sudah menjadi sesuatu yang rutin, maka ia tidak mempunyai arti apa-apa lagi.
5. memerlukan syarat-syarat pendukung yang lengkap, baik waktu, tenaga maupun biaya. Jika syarat-syarat tersebut tidak terpenuhi, maka hasil penilaian tidak dapat dipertanggungjawabkan dengan baik.

Contoh :

Format Penilaian Tindakan Dalam Praktik Pelajaran ...

Nama Sekolah :  
Mata Pelajaran :  
Nama Peserta Didik :  
Kelas/Semester :  
Hari/Tanggal :  
Tujuan :

Pertunjuk: Berilah penilaian dengan menggunakan tanda cek (V) pada setiap aspek yang tertera di bawah ini sesuai dengan tingkat penguasaan peserta didik.

Keterangan nilai :

SB = Sangat Baik

B = Baik

C = Cukup

K = Kurang

SK = Sangat Kurang

No	Aspek yang di amati	SB	B	C	K	SK

Guru ybs,

.....

# Instrumen Evaluasi Jenis Non-Tes

**H**asil dari satu proses pembelajaran mencakup tidak hanya aspek kognitif, tapi juga aspek afektif dan psikomotorik. Sehingga hasil dari proses pembelajaran dapat berupa pengetahuan teoritis, keterampilan dan sikap. Pengetahuan teoritis dapat diukur dengan menggunakan teknik tes. Keterampilan dapat diukur dengan menggunakan tes perbuatan. Sedangkan hasil belajar berupa perubahan sikap hanya dapat diukur dengan teknik non-tes. Instrumen evaluasi jenis non-tes dapat digunakan jika kita ingin mengetahui kualitas proses dan produk dari suatu pembelajaran yang berkenaan dengan domain afektif, seperti sikap, minat, bakat, motivasi, dan lain-lain. Termasuk jenis instrumen evaluasi jenis non-tes adalah observasi, wawancara, skala sikap, dan lain-lain.

### A. Daftar Cek.

Penilaian unjuk kerja dapat dilakukan dengan menggunakan daftar cek (ya - tidak). Pada penilaian unjuk

kerja yang menggunakan daftar cek, peserta didik mendapat nilai apabila kriteria penguasaan kemampuan tertentu dapat diamati oleh penilai. Jika tidak dapat diamati, peserta didik tidak memperoleh nilai. Kelemahan cara ini adalah penilai hanya mempunyai dua pilihan mutlak, misalnya benar-salah, dapat diamati tidak dapat diamati. Dengan demikian tidak terdapat nilai tengah. Berikut contoh daftar cek.

Contoh *checklists*

### Format Penilaian Kegiatan Siswa

Nama peserta didik: \_\_\_\_\_

Kelas: \_\_\_\_

No	Aspek yang diamati	Ya	Tidak
1.	Mengamati		
2.	Bertanya		
3.	Menarik kesimpulan		
..	...		
Skor yang dicapai			

## B. Skala Rentang.

Penilaian unjuk kerja yang menggunakan skala rentang memungkinkan penilai memberi nilai penguasaan kompetensi tertentu karena pemberian nilai secara kontinum di mana pilihan kategori nilai lebih dari dua.

Penilaian sebaiknya dilakukan oleh lebih dari satu penilai agar faktor subjektivitas dapat diperkecil dan hasil penilaian lebih akurat. Berikut contoh skala rentang:

## Format Penilaian Kegiatan Siswa

Nama Siswa: \_\_\_\_\_

Kelas: \_\_\_\_\_

No	Aspek yang diamati	Nilai				
		1	2	3	4	5
1.	Mengamati					
2.	Bertanya					
3.	Menarik kesimpulan					
..	...					
Skor yang dicapai						

Kriteria Penskoran: semakin baik penampilan siswa semakin tinggi skor yang diperoleh.

### C. Penilaian Sikap

Berangkat dari perasaan (suka atau tidak suka) yang terkait dengan kecenderungan bertindak seseorang dalam merespon sesuatu/objek. Sikap juga sebagai ekspresi dari nilai-nilai atau pandangan hidup yang dimiliki oleh seseorang. Sikap dapat dibentuk untuk terjadinya perilaku atau tindakan yang diinginkan.

Sikap terdiri dari tiga komponen, yakni: komponen afektif, komponen kognitif, dan komponen konatif. Komponen afektif adalah perasaan yang dimiliki oleh seseorang atau penilaiannya terhadap sesuatu objek. Komponen kognitif adalah kepercayaan atau keyakinan seseorang mengenai objek. Adapun komponen konatif adalah kecenderungan untuk berperilaku atau berbuat dengan cara-cara tertentu berkenaan dengan kehadiran objek sikap.



Secara umum, objek sikap yang perlu dinilai dalam proses pembelajaran berbagai mata pelajaran adalah sebagai berikut:

- Sikap terhadap materi pelajaran. Peserta didik perlu memiliki sikap positif terhadap materi pelajaran. Dengan sikap positif dalam diri peserta didik akan tumbuh dan berkembang minat belajar, akan lebih mudah diberi motivasi, dan akan lebih mudah menyerap materi pelajaran yang diajarkan.
- Sikap terhadap guru/pengajar. Peserta didik perlu memiliki sikap positif terhadap guru. peserta didik yang tidak memiliki sikap positif terhadap guru akan cenderung mengabaikan hal-hal yang diajarkan. Dengan demikian, peserta didik yang memiliki sikap negatif terhadap guru/pengajar akan sukar menyerap materi pelajaran yang diajarkan oleh guru tersebut.
- Sikap terhadap proses pembelajaran. Peserta didik juga perlu memiliki sikap positif terhadap proses pembelajaran yang berlangsung. Proses pembelajaran di sini mencakup suasana pembelajaran, strategi, metodologi, dan teknik pembelajaran yang digunakan. Proses pembelajaran yang menarik, nyaman dan menyenangkan dapat menumbuhkan motivasi belajar peserta didik, sehingga dapat mencapai hasil belajar yang maksimal.
- Sikap berkaitan dengan nilai-nilai atau norma-norma tertentu berhubungan dengan suatu materi pelajaran. Misalnya kasus atau masalah lingkungan hidup, berkaitan dengan materi Biologi atau Geografi. peserta didik juga perlu memiliki sikap yang tepat, yang

dilandasi oleh nilai-nilai positif terhadap kasus lingkungan tertentu (kegiatan pelestarian/kasus perusakan lingkungan hidup). Misalnya, peserta didik memiliki sikap positif terhadap program perlindungan satwa liar. Dalam kasus yang lain, peserta didik memiliki sikap negatif terhadap kegiatan ekspor kayu glondongan ke luar negeri.

- Sikap berhubungan dengan kompetensi afektif lintas kurikulum yang relevan dengan mata pelajaran.
- Penilaian sikap dapat dilakukan dengan beberapa cara atau teknik. Teknik-teknik tersebut antara lain: observasi perilaku, pertanyaan langsung, dan laporan pribadi. Teknik-teknik tersebut secara ringkas dapat diuraikan sebagai berikut.

### **1. Observasi perilaku**

Perilaku seseorang pada umumnya menunjukkan kecenderungan seseorang dalam sesuatu hal. Misalnya orang yang biasa minum kopi dapat dipahami sebagai orang kecenderungannya yang senang kepada kopi. Oleh karena itu, guru dapat melakukan observasi terhadap peserta didik yang dibinanya. Hasil observasi dapat dijadikan sebagai umpan balik dalam pembinaan. Observasi perilaku di sekolah dapat dilakukan dengan menggunakan buku catatan khusus tentang kejadian-kejadian berkaitan dengan peserta didik selama di sekolah. Berikut contoh format buku catatan harian. Contoh halaman sampul Buku Catatan Harian:

Contoh halaman sampul Buku Catatan Harian:

BUKU CATATAN HARIAN PERILAKU SISWA

(Nama Sekolah)

Mata Pelajaran : .....

Nama Guru : .....

Tahun Pelajaran : .....

2019

Contoh isi Buku Catatan Harian :

No	Hari/Tanggal	Nama Siswa	Kejadian (Positif/negatif)

Catatan dalam lembaran buku tersebut, selain bermanfaat untuk merekam dan menilai perilaku peserta didik sangat bermanfaat pula untuk menilai sikap peserta didik serta dapat menjadi bahan dalam penilaian perkembangan peserta didik secara keseluruhan. Selain itu, dalam observasi perilaku dapat juga digunakan daftar cek (*Checklist*) yang memuat perilaku-perilaku tertentu yang diharapkan muncul dari peserta didik pada umumnya atau dalam keadaan tertentu. Berikut contoh format Penilaian Sikap.

## Contoh Format Penilaian sikap dalam diskusi pelajaran

No	Nama	Perilaku				Nilai	Ket
		Bekerjasama	Berinisiatif	Penuh Perhatian	Sistematis		

Catatan: Kolom perilaku diisi dengan angka yang sesuai:

1 = sangat kurang

2 = kurang

3 = sedang

4 = baik

5 = amat baik

## 2. Pertanyaan langsung.

Kita juga dapat menanyakan secara langsung tentang sikap seseorang berkaitan dengan sesuatu hal. Misalnya, bagaimana tanggapan peserta didik tentang kebijakan yang baru diberlakukan di sekolah mengenai “Peningkatan Ketertiban”.

Berdasarkan jawaban dan reaksi lain yang tampil dalam memberi jawaban dapat dipahami sikap peserta didik itu terhadap objek sikap. Dalam penilaian sikap peserta didik di sekolah, guru juga dapat menggunakan teknik ini dalam menilai sikap dan membina peserta didik.

### **3. Laporan pribadi.**

Melalui penggunaan teknik ini di sekolah, peserta didik diminta membuat ulasan yang berisi pandangan atau tanggapannya tentang suatu masalah, keadaan, atau hal yang menjadi objek sikap. Misalnya, peserta didik diminta menulis pandangannya tentang “Kerusuhan Antaretnis” yang terjadi akhir-akhir ini di Indonesia. Dari ulasan yang dibuat oleh peserta didik tersebut dapat dibaca dan dipahami kecenderungan sikap yang dimilikinya.

### **D. Penilaian Proyek**

Penilaian proyek merupakan kegiatan penilaian terhadap suatu tugas yang harus diselesaikan dalam periode/waktu tertentu. Tugas tersebut berupa suatu investigasi sejak dari perencanaan, pengumpulan data, pengorganisasian, pengolahan dan penyajian data.

Penilaian proyek dapat digunakan, diantaranya untuk mengetahui pemahaman dan pengetahuan dalam bidang tertentu, kemampuan peserta didik mengaplikasikan pengetahuan tersebut dalam penyelidikan tertentu, dan kemampuan peserta didik dalam menginformasikan subyek tertentu secara jelas.

Dalam penilaian proyek setidaknya ada 3 (tiga) hal yang perlu dipertimbangkan yaitu:

- Kemampuan pengelolaan.  
Kemampuan peserta didik dalam memilih topik dan mencari informasi serta dalam mengelola waktu pengumpulan data dan penulisan laporan.

- Relevansi.  
Kesesuaian dengan mata pelajaran, dalam hal ini mempertimbangkan tahap pengetahuan, keterampilan, dan pemahaman dalam pembelajaran.
- Keaslian.  
Proyek yang dilakukan peserta didik harus merupakan hasil karyanya, dengan mempertimbangkan kontribusi guru pada proyek peserta didik, dalam hal ini petunjuk atau dukungan.

Penilaian proyek dapat dilakukan mulai perencanaan, proses selama pengerjaan tugas, dan terhadap hasil akhir proyek. Dengan demikian guru perlu menetapkan hal-hal atau tahapan yang perlu dinilai, seperti penyusunan disain, pengumpulan data, analisis data, kemudian menyiapkan laporan tertulis. Laporan tugas atau hasil penelitiannya juga dapat disajikan dalam bentuk poster. Pelaksanaan penilaian ini dapat menggunakan alat/instrumen penilaian berupa daftar cek (*checklist*) ataupun skala rentang (*rating scale*).

- a) Beberapa contoh kegiatan peserta didik dalam penilaian proyek: penelitian sederhana tentang perilaku terpuji dan tidak terpuji ditemui dalam kehidupan sehari-hari.
- b) Penelitian sederhana tentang pelaksanaan zakat di desanya. Contoh format penilaian proyek sebagai berikut:

## PENILAIAN PROYEK

Mata pelajaran :  
Kelas/Semester :  
Alokasi waktu :  
Jumlah Soal :  
Standar kompetensi :  
Kompetensi Dasar :

No.	Aspek yang dinilai	Skor

### E. Penilaian Produk

Penilaian produk adalah penilaian terhadap keterampilan dalam membuat suatu produk dan kualitas produk tersebut. Penilaian produk tidak hanya diperoleh dari hasil akhir saja tetapi juga proses pembuatannya. Penilaian produk meliputi penilaian terhadap kemampuan peserta didik membuat produk-produk teknologi dan seni, seperti: makanan, pakaian, hasil karya seni (patung, lukisan, gambar), barang-barang terbuat dari kayu, keramik, plastik, dan logam. Pengembangan produk meliputi 3 (tiga) tahap dan dalam setiap tahapan perlu diadakan penilaian yaitu:

- Tahap persiapan, meliputi: menilai kemampuan peserta didik merencanakan, menggali, dan mengembangkan gagasan, dan mendesain produk.

- Tahap pembuatan (produk), meliputi: menilai kemampuan peserta didik menyeleksi dan menggunakan bahan, alat, dan teknik.
- Tahap penilaian (*appraisal*), meliputi: menilai kemampuan peserta didik membuat produk sesuai kegunaannya dan memenuhi kriteria keindahan.

Penilaian produk biasanya menggunakan cara holistik atau analitik.

- Cara holistik, yaitu berdasarkan kesan keseluruhan dari produk, biasanya dilakukan pada tahap appraisal.
- Cara analitik, yaitu berdasarkan aspek-aspek produk, biasanya dilakukan terhadap semua kriteria yang terdapat pada semua tahap proses pengembangan. Contoh format kisi-kisi instrumen penilaian produk sebagai berikut:

#### PENILAIAN PRODUK

Mata pelajaran :  
 Kelas/Semester :  
 Alokasi waktu :  
 Jumlah Soal :  
 Standar kompetensi :  
 Kompetensi Dasar :

No.	Aspek yang dinilai	Skor



## **F. Penilaian Portofolio.**

Penilaian portofolio merupakan penilaian berkelanjutan yang didasarkan pada kumpulan informasi yang menunjukkan perkembangan kemampuan peserta didik dalam satu periode tertentu. Informasi perkembangan peserta didik tersebut dapat berupa karya peserta didik (hasil pekerjaan) dari proses pembelajaran yang dianggap terbaik oleh peserta didiknya, hasil tes (bukan nilai), piagam penghargaan atau bentuk informasi lain yang terkait dengan kompetensi tertentu dalam satu mata pelajaran. Berdasarkan informasi perkembangan tersebut, guru dan peserta didik sendiri dapat menilai perkembangan kemampuan peserta didik dan terus melakukan perbaikan. Dengan demikian, portofolio dapat memperlihatkan perkembangan kemajuan belajar peserta didik melalui karya peserta didik, antara lain: karangan, puisi, surat, komposisi, musik.

Hal-hal yang perlu diperhatikan dan dijadikan pedoman dalam penggunaan portofolio di sekolah, antara lain:

- Saling percaya antara guru dan peserta didik.  
Dalam proses penilaian guru dan peserta didik harus memiliki rasa saling percaya, saling memerlukan dan saling membantu sehingga terjadi proses pendidikan berlangsung dengan baik.
- Kerahasiaan bersama antara guru dan peserta didik.  
Kerahasiaan hasil pengumpulan informasi perkembangan peserta didik perlu dijaga dengan baik dan tidak disampaikan kepada pihak-pihak yang tidak berkepentingan sehingga memberi dampak negatif proses pendidikan.

- Milik bersama (*joint ownership*) antara peserta didik dan guru.  
Guru dan peserta didik perlu mempunyai rasa memiliki berkas portofolio sehingga peserta didik akan merasa memiliki karya yang dikumpulkan dan akhirnya akan berupaya terus meningkatkan kemampuannya.
- Kepuasan.  
Hasil kerja portofolio sebaiknya berisi keterangan dan atau bukti yang memberikan dorongan peserta didik untuk lebih meningkatkan diri.
- Kesesuaian.  
Hasil kerja yang dikumpulkan adalah hasil kerja yang sesuai dengan kompetensi yang tercantum dalam kurikulum.
- Penilaian proses dan hasil.  
Penilaian portofolio menerapkan prinsip proses dan hasil. Proses belajar yang dinilai misalnya diperoleh dari catatan guru tentang kinerja dan karya peserta didik.
- Penilaian dan pembelajaran  
Penilaian portofolio merupakan hal yang tak terpisahkan dari proses pembelajaran. Manfaat utama penilaian ini sebagai diagnostik yang sangat berarti bagi guru untuk melihat kelebihan dan kekurangan peserta didik.

Teknik penilaian portofolio di dalam kelas memerlukan langkah-langkah sebagai berikut:

- Jelaskan kepada peserta didik maksud penggunaan portofolio, yaitu tidak semata-mata merupakan kumpulan hasil kerja peserta didik yang digunakan oleh guru untuk penilaian, tetapi digunakan juga oleh

peserta didik sendiri. Dengan melihat portofolionya peserta didik dapat mengetahui kemampuan, keterampilan, dan minatnya. Proses ini tidak akan terjadi secara spontan, tetapi membutuhkan waktu bagi peserta didik untuk belajar meyakini hasil penilaian mereka sendiri.

- Tentukan bersama peserta didik sampel-sampel portofolio apa saja yang akan dibuat. Portofolio antara peserta didik yang satu dan yang lain bisa sama bisa berbeda. Misalnya, untuk kemampuan menulis peserta didik mengumpulkan karangan-karangannya. Sedangkan untuk kemampuan menggambar, peserta didik mengumpulkan gambar-gambar buatannya.
- Kumpulkan dan simpanlah karya-karya tiap peserta didik dalam satu map atau folder.
- Berilah tanggal pembuatan pada setiap bahan informasi perkembangan peserta didik sehingga dapat terlihat perbedaan kualitas dari waktu ke waktu.
- Tentukan kriteria penilaian sampel-sampel portofolio peserta didik beserta pembobotannya bersama para peserta didik agar dicapai kesepakatan. Diskusikan dengan para peserta didik bagaimana menilai kualitas karya mereka. Contoh; untuk kemampuan menulis karangan, kriteria penilaiannya misalnya: penggunaan tata bahasa, pemilihan kosa-kata, kelengkapan gagasan, dan sistematika penulisan. Sebaiknya kriteria penilaian suatu karya dibahas dan disepakati bersama peserta didik sebelum peserta didik membuat karya tersebut. Dengan demikian, peserta didik mengetahui harapan (standar) guru dan berusaha mencapai harapan atau standar itu.

- Mintalah peserta didik menilai karyanya secara berkesinambungan. Guru dapat membimbing peserta didik tentang bagaimana cara menilai dengan memberi keterangan tentang kelebihan atau kekurangan karya tersebut dan bagaimana cara memperbaikinya. Hal ini dapat dilakukan pada saat membahas portofolio.
- Setelah suatu karya dinilai dan ternyata nilainya belum memuaskan, kepada peserta didik dapat diberi kesempatan untuk memperbaiki lagi. Namun, antara peserta didik dan guru perlu dibuat “kontrak” atau perjanjian mengenai jangka waktu perbaikan, misalnya setelah 2 minggu karya yang telah diperbaiki harus diserahkan kepada guru.
- Bila perlu, jadwalkan pertemuan untuk membahas portofolio. Jika dianggap perlu, undanglah orang tua peserta didik untuk diberi penjelasan tentang maksud dan tujuan portofolio sehingga orangtua dapat membantu dan memotivasi anaknya.

Contoh format kisi-kisi instrumen portofolio sebagai berikut:

No.	Jenis Tagihan/Tanggal	Nilai	Keterangan	KD/Indikator yang dinilai

### G. Penilaian Diri.

Penilaian diri (self assessment) adalah suatu teknik penilaian, di mana subjek yang ingin dinilai diminta untuk menilai dirinya sendiri berkaitan dengan, status, proses dan

tingkat pencapaian kompetensi yang dipelajarinya dalam mata pelajaran tertentu.

Teknik penilaian diri dapat digunakan dalam berbagai aspek penilaian, yang berkaitan dengan kompetensi kognitif, afektif dan psikomotor. Dalam proses pembelajaran di kelas, berkaitan dengan kompetensi kognitif, misalnya: peserta didik dapat diminta untuk menilai penguasaan pengetahuan dan keterampilan berpikir sebagai hasil belajar dalam mata pelajaran tertentu, berdasarkan kriteria atau acuan yang telah disiapkan. Berkaitan dengan kompetensi afektif, misalnya, peserta didik dapat diminta untuk membuat tulisan yang memuat curahan perasaannya terhadap suatu objek sikap tertentu. Selanjutnya, peserta didik diminta untuk melakukan penilaian berdasarkan kriteria atau acuan yang telah disiapkan.

Berkaitan dengan kompetensi psikomotorik, peserta didik dapat diminta untuk menilai kecakapan atau keterampilan yang telah dikuasainya sebagai hasil belajar berdasarkan kriteria atau acuan yang telah disiapkan. Penggunaan teknik ini dapat memberi dampak positif terhadap perkembangan kepribadian seseorang. Keuntungan penggunaan teknik ini dalam penilaian di kelas antara lain sebagai berikut.

- Dapat menumbuhkan rasa percaya diri peserta didik, karena mereka diberi kepercayaan untuk menilai dirinya sendiri;
- Peserta didik menyadari kekuatan dan kelemahan dirinya, karena ketika mereka melakukan penilaian, harus melakukan introspeksi terhadap kekuatan dan kelemahan yang dimilikinya;

- Dapat mendorong, membiasakan, dan melatih peserta didik untuk berbuat jujur, karena mereka dituntut untuk jujur dan objektif dalam melakukan penilaian.

Penilaian diri dilakukan berdasarkan kriteria yang jelas dan dengan cara yang objektif. Oleh karena itu, penilaian diri oleh peserta didik di kelas perlu dilakukan melalui langkah-langkah sebagai berikut.

- Menentukan kompetensi atau aspek kemampuan yang akan dinilai.
- Menentukan kriteria penilaian yang akan digunakan.
- Merumuskan format penilaian, dapat berupa pedoman penskoran, daftar tanda cek, atau skala rentang.
- Meminta peserta didik untuk melakukan penilaian diri.
- Guru mengkaji sampel hasil penilaian secara acak, untuk mendorong peserta didik supaya senantiasa melakukan penilaian diri secara cermat dan objektif.
- Menyampaikan umpan balik kepada peserta didik berdasarkan hasil kajian terhadap sampel hasil penilaian yang diambil secara acak.

Perlu dicatat bahwa tidak ada satu pun alat penilaian yang dapat mengumpulkan informasi prestasi dan kemajuan belajar peserta didik secara lengkap. Penilaian tunggal tidak cukup untuk memberikan gambaran/informasi tentang kemampuan, keterampilan, pengetahuan dan sikap seseorang. Lagi pula, interpretasi hasil tes tidak mutlak dan abadi karena anak terus berkembang sesuai dengan pengalaman belajar yang dialaminya.

Alat penilaian tertulis seperti pilihan ganda yang mengarah kepada hanya satu jawaban yang benar (*convergent*

*thinking*), tidak mampu menilai keterampilan/ kemampuan lain yang dimiliki peserta didik. Hal ini amat menghambat penguasaan beragam kompetensi yang tercantum pada kurikulum secara utuh. Alat penilaian pilihan ganda kurang mampu memberikan informasi yang cukup untuk dijadikan umpan-balik guna mendiagnosis atau memodifikasi pengalaman belajar. Karena itu, guru hendaknya mengembangkan alat-alat penilaian yang membedakan antara jenis-jenis kompetensi yang berbeda dari tiap tingkat pencapaian. Hasil penilaian dapat menghasilkan rujukan terhadap pencapaian peserta didik dalam domain kognitif, afektif, dan psikomotor, sehingga hasil tersebut dapat menggambarkan profil peserta didik secara lengkap.

Contoh format penilaian diri sebagai berikut:

Format 1

Nama :

Kelas :

No.	Indikator	Penilaian		
		0	1	2
1.	Partisipasi dalam kelompok			
2.	Teamwork			
3.	Memberi ide/gagasan			
4.	Memberi pertanyaan			
...	.....			

Kriteria:

0 = Tidak pernah/jelek

1 = Jarang/cukup

2 = Sering/baik

## Format 2

Nama :

Kelas :

No.	Indikator	Penilaian	
		Ya	Tidak
1.	Partisipasi dalam kelompok		
2.	Teamwork		
3.	Memberi ide/gagasan		
4.	Memberi pertanyaan		
...	.....		

## Format 3

Nama :

Kelas :

Anggota Kelompok :

Kegiatan Kelompok :

Untuk pernyataan dibawah ini masing-masing penilaiannya dengan huruf A, B, atau C sesuai dengan pendapatmu

A = selalu

B = jarang

C = tidak pernah

1. \_\_\_\_\_ Selama diskusi saya memberikan saran-saran kepada kelompok untuk didiskusikan.



2. \_\_\_\_ Ketika kami berdiskusi, setiap anggota memberikan masukan untuk di diskusikan.
3. \_\_\_\_ Semua anggota kelompok harus melakukan sesuatu dalam kegiatan kelompok.
4. \_\_\_\_ Setiap anggota kelompok mengerjakan kegiatannya sendiri dalam kegiatan kelompok.
5. Selama kegiatan kelompok saya:
  - \_\_\_\_ mendengarkan
  - \_\_\_\_ bertanya
  - \_\_\_\_ mengajukan gagasan/pendapat
  - \_\_\_\_ mengendalikan kelompok
  - \_\_\_\_ mengganggu kelompok
  - \_\_\_\_ tidur

Format 4

Nama :

Kelas :

Komentar Siswa :

---

---

---

---

---

---

---

---

---

---

## Bab V

# Pengukuran Domain Kognitif, Afektif, Dan Psikomotorik

Setiap kompetensi yang telah dipelajari oleh siswa dalam proses pembelajaran harus dinilai melalui penilaian otentik. Kompetensi yang dinilai mencakup tiga ranah, yaitu ranah kognitif, afektif dan psikomotorik. Dalam membuat instrument penilaian yang akan dilakukan perlu memperhatikan ranah atau domain pembelajaran, apakah penilaian dilakukan untuk menganalisis kemampuan berfikir, otak, akal, mental, atau menganalisis kemampuan bersikap, berakhlak, berperilaku, atau menganalisis kemampuan skill atau kinerja. Benyamin S. Bloom mengembangkan suatu metode pengklasifikasian tujuan pendidikan yang disebut dengan taksonomi (*taxonomy*). ia berpendapat bahwa taksonomi tujuan pembelajaran harus senantiasa mengacu kepada tiga jenis domain atau ranah, yaitu ranah proses berfikir (*kognitif*); ranah nilai atau sikap (*afektif*); dan ranah keterampilan (*psikomotor*).

## A. Pengukuran Ranah Kognitif

Ranah kognitif adalah ranah yang mencakup kegiatan berfikir/akal/otak. Bloom mengelompokkan ranah kognitif ke dalam enam kategori dari yang sederhana sampai kepada yang paling kompleks dan diasumsikan bersifat hirarkis, yang berarti tujuan pada level yang tinggi dapat dicapai apabila tujuan pada level yang rendah telah dikuasai. Tingkat kompetensi tersebut dapat dilihat pada gambar berikut



Gambar 5.1 Domain Kognitif Menurut Bloom

Tingkatan pengetahuan ialah kemampuan mengingat kembali, misalnya, pengetahuan mengenai istilah-istilah, pengetahuan mengenai klasifikasi dan sejenisnya. Jadi, tingkatan pengetahuan mencakup ingatan akan hal-hal yang pernah dipelajari dan disimpan dalam ingatan. Singkatnya dapat dikatakan bahwa pengetahuan yang disimpan dalam ingatan itu, dapat digali kembali pada saat dibutuhkan melalui

bentuk ingatan (*recall*) atau mengingatkan kembali (*recognition*).

contoh:

*Siswa dapat mendeskripsikan Model Pembelajaran Berbasis Masalah.*

Tingkatan pemahaman yaitu kemampuan menggunakan informasi dalam situasi yang tepat, mencakup kemampuan untuk membandingkan, menunjukkan persamaan dan perbedaan, mengidentifikasi karakteristik, menganalisis dan menyimpulkan. Kata-kata operasional yang biasa digunakan ialah: mengklasifikasi, menjelaskan, mengikhtisarkan, membedakan dan yang sejenis.

contoh:

*Siswa mampu menjelaskan kelebihan dan kelemahan metode ceramah dalam proses pembelajaran.*

Tingkatan penerapan mencakup kemampuan untuk menggunakan atau menerapkan informasi yang telah dipelajari ke dalam situasi atau konteks yang lain, yaitu mampu mengaplikasikan atas pengetahuan dan pemahaman yang telah dimiliki sebagai hasil dari proses pembelajaran. Kata-kata operasional yang biasa digunakan ialah: mendemonstrasikan, menghitung, menyelesaikan, menyesuaikan, mengoperasikan, menghubungkan, menyusun dan yang sejenis.

Contoh:

*Siswa dapat mengoperasikan software program excel 2000, untuk menghitung central tendency atas data yang terdapat pada tabel III, tanpa kesalahan.*

Tingkatan analisis yaitu mengenal kembali unsur-unsur, hubungan-hubungan dan susunan informasi atau masalah, misalnya: menganalisis hubungan-hubungan meliputi kemampuan untuk mengidentifikasi, memisahkan atau membedakan komponen atau elemen suatu fakta, konsep, pendapat, asumsi, hipotesis atau kesimpulan dan memeriksa setiap komponen tersebut untuk melihat ada tidaknya konstruksi. Kata-kata operasional yang biasa digunakan ialah: menemukan perbedaan, memisahkan, membuat diagram, membuat estimasi, menjabarkan ke dalam bagian-bagian, menyusun urutan dan yang sejenis.

Contoh:

*Siswa dapat membuat perbedaan suatu hasil matematika.... dengan menggunakan berbagai cara penyelesaian.*

Tingkatan sintesis yaitu mengkombinasikan kembali bagian-bagian dari pengalaman yang lalu dengan bahan yang baru menjadi suatu keseluruhan yang baru dan terpadu, misalnya membuat suatu rencana atau menyusun usulan kegiatan dengan suatu kesatuan atau pola baru. Bagian-bagian dihubungkan satu sama lain sehingga tercipta suatu bentuk baru. Adanya kemampuan ini dinyatakan dalam membuat rencana seperti penyusunan satuan pelajaran atau proposal penelitian. Kata-kata operasional yang biasa digunakan ialah: menggabungkan, menciptakan, merumuskan, merancang, membuat komposisi, dan yang sejenis.

Contoh:

*Setelah menggunakan berbagai metode, model, pendekatan dan strategi pembelajaran, guru dapat merumuskan atau merangkum satu metode mengajar yang baru dengan*

*meminimalisir kelemahan metode tersebut melalui gabungan dari beberapa metode belajar.*

Tingkatan evaluasi yaitu menggunakan kriteria untuk mengukur nilai suatu gagasan, karya dan sebagainya, misalnya menimbang-nimbang dan memutuskan mencakup kemampuan untuk membuat penelitian dan keputusan tentang nilai suatu gagasan, metode, produk, atau benda dengan menggunakan kriteria tertentu. Kata-kata operasional yang biasa digunakan ialah: menimbang, mengkritik, membandingkan, memberi alasan, menyimpulkan, memberi dukungan, dan yang sejenis.

Contoh:

*Setelah membaca karya al-Manfaluthi dan karya Hamka dalam novelnya 'Tenggelamnya Kapal Vanderwijk', siswa dapat mengemukakan sekurang-kurangnya 3 alasan bahwa novel Hamka itu bukan plagiat.*

## **RANAH KOGNITIF**

<b>Tingkat Kompetensi</b>	<b>Contoh Kata Kerja Operasional</b>
Pengetahuan ( <i>Knowledge</i> )	Mengenali, mendeskripsikan, menanamkan, memasang, membuat daftar, memilih.
Pemahaman ( <i>Comprehension</i> )	Mengklasifikasi, menjelaskan, mengikhtisarkan, membedakan
Penerapan ( <i>Aplication</i> )	Mendemonstrasikan, menghitung, menyelesaikan, menyesuaikan, mengoperasikan, menghubungkan, menyusun

Analisis ( <i>Analysis</i> )	Menemukan perbedaan, memisahkan, membuat diagram, membuat estimasi, menjabarkan ke dalam bagian-bagian, menyusun urutan
Sintesis ( <i>Synthesis</i> )	Menggabungkan, menciptakan, merumuskan, merancang, membuat komposisi
Evaluasi ( <i>Evaluation</i> )	Menimbang, mengkritik, membandingkan, memberi alasan, menyimpulkan, memberi dukungan

Berkenaan dengan pengukuran terhadap ranah kognitif ini banyak dijumpai, dan hampir sebagian besar contoh-contoh yang dikemukakan dalam buku ini adalah berkenaan dengan hal itu. Berbeda halnya dengan ranah afektif seperti yang akan dibahas berikut ini, yang bentuk pertanyaannya berbeda dengan ranah kognitif. Untuk mengukur kognitif dapat dilakukan dengan tes, yaitu: tes lisan di kelas, pilihan berganda, uraian obyektif, uraian non obyektif, jawaban singkat, menjodohkan, unjuk karya dan portofolio.

## **B. Pengukuran Ranah Afektif**

Ranah afektif adalah ranah yang berkaitan dengan sikap dan nilai. Sikap adalah salah satu istilah bidang psikologi yang berhubungan dengan persepsi dan tingkah laku. Istilah sikap dalam bahasa Inggris disebut *attitude*. *Attitude* adalah suatu cara bereaksi terhadap suatu perangsang. Suatu kecenderungan untuk bereaksi terhadap suatu perangsang atau situasi yang dihadapi. Sikap melibatkan beberapa pengetahuan tentang situasi, namun aspek yang paling esensial dalam sikap adalah adanya perasaan atau emosi,

kecenderungan terhadap perbuatan yang berhubungan dengan pengetahuan.

Sikap melibatkan pengetahuan tentang situasi. Situasi di sini dapat digambarkan sebagai suatu obyek yang pada akhirnya akan mempengaruhi emosi, kemudian memungkinkan munculnya reaksi atau kecenderungan untuk berbuat. Dalam beberapa hal, sikap adalah penentuan yang paling penting dalam tingkah laku manusia. Sebagai reaksi maka sikap selalu berhubungan dengan dua alternatif senang dan tidak senang untuk melaksanakan atau menjauhinya. Perasaan senang meliputi sejumlah perasaan yang lebih spesifik seperti rasa puas, sayang, dan lain-lain, perasaan tidak senang meliputi sejumlah rasa yang spesifik pula yaitu rasa takut, gelisah, cemburu, marah, dendam, dan lain-lain. Sikap juga diartikan sebagai "suatu konstruk untuk memungkinkan terlihatnya suatu aktivitas".

Pengertian sikap itu sendiri dapat dipandang dari berbagai unsur yang terkait seperti sikap dengan kepribadian, motif, tingkat keyakinan, dan lain-lain. Namun dapat diambil pengertian yang memiliki persamaan karakteristik, dengan demikian sikap adalah tingkah laku yang terkait dengan kesediaan untuk merespon obyek sosial yang membawa dan menuju ke tingkah laku yang nyata dari seseorang. Hal itu berarti tingkah laku dapat diprediksi apabila telah diketahui sikapnya.

Tiap orang mempunyai sikap yang berbeda-beda terhadap suatu objek. Ini berarti bahwa sikap itu dipengaruhi oleh berbagai faktor yang ada pada diri masing-masing seperti perbedaan bakat, minat, pengalaman, pengetahuan, intensitas



perasaan dan juga situasi lingkungan. Demikian juga sikap seseorang terhadap suatu yang sama mungkin saja tidak sama.

Krathwohl, Bloom dan Masria mengembangkan taksonomi ini yang berorientasi kepada perasaan atau afektif. Taksonomi ini menggambarkan proses seseorang di dalam mengenali dan mengadopsi suatu nilai dan sikap tertentu yang menjadi pedoman baginya dalam bertindak laku.

Domain afektif, Krathwohl membaginya atas lima kategori/ tingkatan yaitu; Pengenalan (*receiving*), pemberian respon (*responding*), penghargaan terhadap nilai (*valuing*), pengorganisasian (*organization*) dan pengamalan (*characterization*).



Gambar 5.2 Domain afektif Menurut Krathwohl dkk

Menurut A.J. Nitko Jenjang Afektif sama dengan pendapat Kratwohl hanya saja uraiannya lebih terperinci pada masing-masing tingkatan. Pembagian ini bersifat hierarkhis, pengenalan tingkat yang paling rendah dan pengamalan sebagai tingkat yang paling tinggi, seseorang memiliki kompetensi pengamalan jika sudah memiliki kompetensi pengenalan, pemberian respon, penghargaan terhadap nilai pengorganisasian.

Pengenalan/penerimaan mencakup kemampuan untuk mengenal, bersedia menerima dan memperhatikan berbagai stimulasi. Dalam hal ini peserta didik bersikap pasif, sekedar mendengarkan atau memperhatikan saja. Contoh kata kerja operasional pada tingkat ini adalah : mendengarkan, menghadiri, melihat dan memperhatikan.

Pemberian respon mencakup kemampuan untuk berbuat sesuatu sebagai reaksi terhadap suatu gagasan, benda atau sistem nilai, lebih dari sekedar pengenalan. Dalam hal ini mahasiswa diharapkan untuk menunjukkan perilaku yang diminta, misalnya berpartisipasi, patuh atau memberikan tanggapan secara sukarela, misalnya berpartisipasi, patuh atau memberikan tanggapan secara sukarela bila diminta. Contoh hasil belajar dalam tingkat ini berpartisipasi dalam kebersihan kelas, berlatih membaca al-Qur'an, dan lain-lain. Kata kerja operasionalnya meliputi: mengikuti, mendiskusikan, berlatih, berpartisipasi, dan mematuhi.

Penghargaan terhadap nilai merupakan perasaan, keyakinan atau anggapan bahwa suatu gagasan, benda atau cara berfikir tertentu mempunyai nilai. Dalam hal ini mahasiswa secara konsisten berperilaku sesuai dengan suatu

nilai meskipun tidak ada pihak lain yang meminta atau mengharuskan. Nilai ini dapat saja dipelajari dari orang lain misalnya dosen, teman atau keluarga. Dalam proses belajar mengajar, peserta didik tidak hanya menerima nilai yang ajarkan tetapi telah tidak mampu untuk memilih baik atau buruk jenjang ini mulai dari hanya sekedar penerimaan sampai ketinggian komitmen yang lebih tinggi (menerima tanggung jawab untuk fungsi kelompok yang lebih efektif. Kata kerja operasionalnya adalah : memilih, meyakinkan, bertindak dan mengemukakan argumentasi.

Pengorganisasian menunjukkan saling berhubungan antara nilai-nilai tertentu dalam suatu sistem nilai, serta menentukan nilai mana yang mempunyai prioritas lebih tinggi daripada nilai yang lain. Dalam hal ini mahasiswa menjadi committed terhadap suatu sistem nilai. Dia diharapkan untuk mengorganisasikan berbagai nilai yang dipilihnya ke dalam suatu sistem nilai dan menentukan hubungan diantara nilai-nilai tersebut. Kata kerja operasional pada tingkat pengorganisasian adalah: memilih, memutuskan, memformulasikan, membandingkan dan membuat sistematisasi.

Pengamalan (*characterization*) berhubungan dengan pengorganisasian dan pengintegrasian nilai-nilai kedalam suatu sistem nilai pribadi. Hal ini diperlihatkan melalui perilaku yang konsisten dengan sistem nilai tersebut. Ini adalah merupakan tingkatan afektif tertinggi, karena sikap batin peserta didik *philosophy of life* yang mapan. Contoh hasil belajar pada tingkat ini adalah: siswa memiliki kebulatan sikap untuk menjadikan surat Al-Ashr sebagai pegangan hidup dalam disiplin waktu baik di sekolah, di rumah maupun di tengah masyarakat. Kata kerja operasional pada tingkat ini adalah:

menunjukkan sikap, menolak, mendemonstrasikan dan menghindari.

### **RANAH AFEKTIF**

<b>Tingkatan Kompetensi</b>	<b>Contoh Kata Kerja Operasional</b>
Pengenalan	Mendengarkan, menghindari, memperhatikan
Pemberian Respon	Mengikuti, mendiskusikan, berpartisipasi, mematuhi
Penghargaan terhadap nilai	Memilih, meyakinkan, bertindak, mengemukakan argumentasi
Pengorganisasian	Memilih, memutuskan, memformulasikan, membandingkan, membuat sistematisasi
pengalaman	Menunjukkan sikap, menolak, mendemonstrasikan, menghindari

Afektif yang harus dikembangkan oleh guru dalam proses belajar tentunya sangat tergantung kepada mata pelajaran dan jenjang kelas, namun yang pasti setiap mata pelajaran memiliki indikator afektif dalam kurikulum hasil belajar.

Pengukuran ranah afektif tidaklah semudah mengukur ranah Afektif karena tidak dapat dilakukan setiap selesai menyajikan materi pelajaran. Perubahan sikap seseorang memerlukan waktu yang relatif lama, demikian juga pengembangan minat dan penghargaan serta nilai-nilai.

Pengukuran afektif berguna untuk mengetahui sikap dan minat siswa ataupun untuk mengetahui tingkat pencapaian kompetensi afektif pada setiap tingkat (level). Pada mata pelajaran tertentu, misalnya seorang siswa mendapatkan nilai tertinggi pada mata pelajaran tertentu belum tentu menyenangi mata pelajaran tersebut.

Ada beberapa bentuk skala yang dapat digunakan untuk mengukur sikap (afektif) yaitu: (1) Skala likert, (2) Skala pilihan ganda, (3) Skala thurstone, (4) Skala guttman, (5) Skala differential, dan (6) Pengukuran minat.

## **1. Skala likert**

Skala likert digunakan untuk mengukur sikap seseorang terhadap sesuatu, misalnya pada mata pelajaran al-Qur'an Hadits siswa menunjukkan sikap dan perilaku gemar melafalkan ayat-ayat al-Qur'an, siswa menunjukkan sikap hormat pada orang tua dll. Skala likert terdiri dari dua unsur yaitu pernyataan dan alternatif jawaban. Pernyataan ada dua bentuk yaitu pernyataan positif dan negatif, sedangkan alternatif jawaban terdiri dari: sangat setuju, setuju, netral, kurang setuju dan tidak setuju.

Langkah-langkah untuk membuat skala likert untuk menilai afektif antara lain adalah: (1) pilih variabel afektif yang akan diukur, (2) buat pernyataan positif terhadap variabel yang diukur, (3) minta pertimbangan kepada beberapa orang tentang pernyataan positif dan negatif yang dirumuskan, (4) tentukan alternatif jawaban yang digunakan, (5) tentukan penskorannya dan, (6) tentukan dan hilangkan pernyataan yang tidak berfungsi dengan pernyataan lainnya.

Contoh :

Saya membaca al-Qur'an setiap selesai shalat Magrib

- a. sangat setuju
- b. setuju
- c. netral
- d. kurang setuju
- e. tidak setuju

## 2. Skala pilihan ganda

Skala ini bentuknya seperti soal bentuk pilihan ganda yaitu suatu pernyataan yang diikuti oleh sejumlah alternatif pendapat.

Contoh:

Dalam melaksanakan shalat fardhu, saya merasa:

- a. senang karena dapat berdialog dengan Allah
- b. mudah untuk melakukan konsentrasi
- c. tidak begitu sulit untuk berkonsentrasi
- d. dapat berkonsentrasi tetapi mudah terganggu
- e. sulit untuk berkonsentrasi

## 3. Skala thurstone

Skala ini mirip dengan skala likert karena merupakan instrumen yang jawabannya menunjukkan adanya tingkatan, thurstone menyarankan pernyataan yang diajukan + 10 item

Contoh :

1	2	3	4	5	6	7	8	9	10	11
Very Favouereble			Netral				Very unfavouereble			

#### 4. Skala Guttman

Skala ini sama dengan skala yang disusun Bogardus yaitu pernyataan yang dirumuskan empat atau tiga pernyataan. Pernyataan tersebut menunjukkan tingkatan yang berurutan, apabila responden setuju persyaratan 2, diduga setuju pernyataan 1, selanjutnya setuju pernyataan 3 diduga setuju pernyataan 1 dan 2 dan apabila setuju pernyataan 4 diduga setuju pernyataan 1,2 dan 3.

Contoh afektif yang indikatornya hormat pada orang tua

1. Saya permissi kepada orang tua bila bermain ke tetangga
2. Saya permissi kepada orang tua bila pergi kemana saja
3. Saya permissi kepada orang tua bila pergi kapan saja dan kemana saja
4. Saya tidak pergi kemana saja tanpa permissi kepada orang tua

#### 5. Skala differential

Skala ini bertujuan untuk mengukur konsep-konsep untuk tiga dimensi. Dimensi yang akan diukur dalam kategori :

baik – tidak baik

kuat – lemah cepat – lambat

aktif – pasif

Contoh :

Baik	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	Tidak baik
Berguna	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	Tidak berguna
Aktif	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>7</u>	Pasif

## 6. Pengukuran Minat

Untuk mengetahui/mengukur minat siswa terhadap mata pelajaran terlebih dahulu ditentukan indikatornya misalnya : kehadiran di kelas, keaktifan bertanya, tepat waktu mengumpulkan tugas, kerapian. Catatan, mengerjakan latihan, mengulang pelajaran dan mengunjungi perpustakaan dan lain-lain. Untuk mengukur minat ini lebih tepat digunakan kuesioner skala likert dengan skala lima yaitu; sangat sering, sering, netral, jarang dan tidak pernah.

Tabel  
Contoh Format Penilaian Minat Siswa Terhadap Mata pelajaran

No	Pernyataan	Skala					JLH
		SS	S	N	J	TP	
1	Saya Senang mengikuti pelajaran ini						
2	Saya selalu hadir pada mata pelajaran ini						
3	Saya bertanya jika ada yang saya tidak pahami kepada guru						
4	Saya mencatat semua yang dijelaskan oleh guru						
5	Saya mempelajari kembali semua yang diajarkan oleh guru dirumah						
Jumlah.....							



Jawaban sangat sering diberi skor 5, sering diberi skor 4, netral diberi skor 3, jarang skor 2, dan tidak pernah skor 1. Selanjutnya tehnik penskoran minat siswa terhadap mata pelajaran dengan item pernyataan 5 butir maka skor terendah 5 dan skor tertinggi 25, jika dibagi menjadi tiga kategori maka skala 5 sampai 11 termasuk minat rendah, 12 sampai 18 berminat dan 19 sampai 25 sangat berminat, maka dapat dikomfersi ke pengukuran kualitatif karena penilaian afektif dilakukan secara kualitatif, maka 5 - 11 = C, 12 - 18 = B, 19 - 25 = A.

Paling tidak ada dua komponen afektif yang penting untuk dinilai setiap mata pelajaran yaitu sikap dan minat. Sikap terhadap mata pelajaran bisa positif, netral dan negatif. Tentu diharapkan sikap siswa terhadap semua mata pelajaran positif sehingga akan muncul minat yang tinggi untuk mempelajarinya, karena minat belajar yang besar cenderung menghasilkan prestasi yang tinggi sebaliknya minat belajar yang kurang akan menghasilkan prestasi yang rendah.

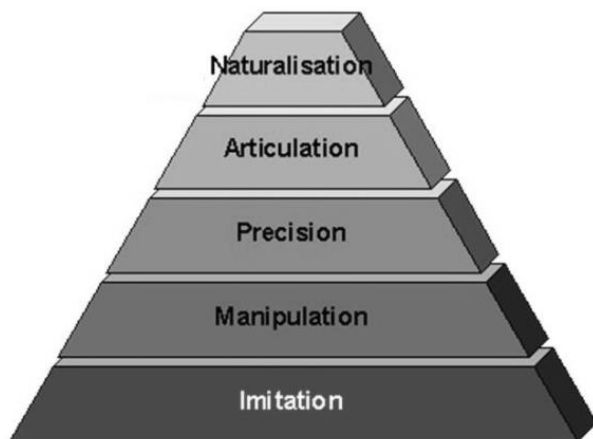
Apabila dari sekian banyak siswa ternyata tidak berminat dan bersikap baik dengan substansi mata pelajaran pendidikan agama maka guru harus mencari sebab-sebabnya, perlu dikaji dan dilihat kembali secara menyeluruh hal yang terkait dengan pelajaran mata pelajaran tersebut atau guru belum menyampaikan diawal pembelajaran indikator yang dimiliki oleh siswa, oleh karenanya guru seharusnya menyampaikan kepada siswa kompetensi dasar yang harus dicapai siswa sekaligus indikator-indikator yang mesti dimiliki siswa.

### C. Pengukuran Ranah Psikomotorik

Ranah psikomotorik menurut Dave's adalah: (a) imitasi, (b) manipulasi, (c) ketepatan, (d) artikulasi, dan (e) naturalisasi. Imitasi: mengamati dan menjadikan perilaku orang lain sebagai pola. Apa yang ditampilkan mungkin kualitas rendah. Contoh: menjiplak hasil karya seni. Manipulasi: mampu menunjukkan perilaku tertentu dengan mengikuti instruksi dan praktek. Contoh: membuat hasil karya sendiri setelah mengikuti pelajaran, ataupun membaca mengenai hal tersebut. Ketepatan: meningkatkan metode supaya lebih tepat. Beberapa kekeliruan tampak jelas. Contoh: bekerja dan melakukan sesuatu kembali, sehingga menjadi "cukup baik." Artikulasi: mengkoordinasikan serangkaian tindakan, mencapai keselarasan dan internal konsistensi. Contoh: memproduksi film video yang menampilkan musik, drama, warna, suara dsb. Naturalisasi: telah memiliki tingkatperformance yang tinggi sehingga menjadi alami, dalam melakukan tidak perlu berpikir banyak. Misalkan: Michael Jordan bermain basket, Nancy Lopez memukul bola golf.

Penyusunan tujuan psikomotor secara hierarkhis dalam lima tingkat sebagai berikut: (1) Meniru. Tujuan pembelajaran pada tingkat ini diharapkan peserta didik dapat meniru suatu perilaku yang dilihatnya, (2) Manipulasi. Tujuan pembelajaran pada tingkat ini menuntut peserta didik untuk melakukan suatu perilaku tanpa bantuan visual, sebagaimana pada tingkat meniru. Tetapi diberi petunjuk berupa tulisan atau instruksi verbal, (3) Ketepatan Gerakan. Tujuan pembelajaran pada level ini peserta didik mampu melakukan suatu perilaku tanpa menggunakan contoh visual maupun petunjuk tertulis, dan melakukannya dengan lancar, tepat, seimbang dan akurat, (4)

Artikulasi. Tujuan pembelajaran pada level ini peserta didik mampu menunjukkan serangkaian gerakan dengan akurat, urutan yang benar, dan kecepatan yang tepat, dan (5) Naturalisasi. Tujuan pembelajaran pada tingkat ini peserta didik mampu melakukan gerakan tertentu secara spontan tanpa berpikir lagi cara melakukannya dan urutannya.



Gambar 5.3, Gambar Ranah Psikomotorik Menurut Harrow dkk

Meniru (*imitation*), pada pada tingkat ini diharapkan peserta didik untuk dapat meniru suatu perilaku yang dilihatnya. Manipulasi (*manipulation*), pada tingkat ini peserta didik diharapkan untuk melakukan suatu perilaku tanpa bantuan visual, sebagaimana pada tingkat meniru. Peserta didik diberi petunjuk berupa tulisan atau instruksi verbal, dan diharapkan melakukan tindakan (perilaku) yang diminta. Contoh kata kerja yang digunakan sama dengan untuk kemampuan meniru. Ketetapan gerakan (*precision*), pada tingkat ini peserta didik diharapkan melakukan suatu perilaku tanpa menggunakan Contoh visual maupun petunjuk tertulis, dan melakukannya dengan lancar, tepat dan akurat. Artikulasi

(*artikulation*), pada tingkat ini peserta didik diharapkan untuk menunjukkan serangkaian gerakan dengan akurat, urutan yang benar, dan kecepatan yang tepat. Naturalisasi (*naturalization*) Pada tingkat ini peserta didik diharapkan melakukan gerakan tertentu secara spontan atau otomatis. Peserta didik melakukan gerakan tersebut tanpa berfikir lagi cara melakukannya dan urutannya.

### **RANAH PSIKOMOTORIK**

<b>Tingkat Kompetensi</b>	<b>Contoh Kata Kerja Operasional</b>
Meniru	Mengulangi, mengikuti, memegang, menggambar, mengucapkan, melakukan
Memanipulasi	Mengulangi, mengikuti, memegang, menggambar, mengucapkan, melakukan, (tidak melihat contoh/tidak mendengar suara)
Ketepatan Gerakan	Mengulangi, mengikuti, memegang, menggambar, mengucapkan, melakukan, (tepat, lancar tanpa kesalahan)
Artikulasi	Menunjukkan gerakan, akurat benar, kecepatan yang tepat, sifatnya: selaras, stabil dan sebagainya.
Naturalisasi	Gerakan spontan/otomatis, tanpa Berpikir melakukan dan urutannya

Pengukuran ranah psikomotorik merupakan merupakan pengukuran yang dilakukan dengan mengamati kegiatan peserta didik dalam melakukan sesuatu. Penilaian ini cocok digunakan untuk menilai ketercapaian kompetensi yang menuntut peserta didik menunjukkan unjuk kerja. Cara

penilaian ini dianggap lebih otentik daripada tes tertulis karena apa yang dinilai lebih mencerminkan kemampuan peserta didik yang sebenarnya. Unjuk kerja yang dapat diamati seperti: bermain peran, memainkan alat musik, bernyanyi, membaca puisi/deklamasi, menggunakan peralatan laboratorium, dan mengoperasikan suatu alat. Pengukuran ranah psikomotorik perlu memperhatikan hal-hal berikut:

- a. Langkah-langkah kinerja yang diharapkan dilakukan peserta didik untuk menunjukkan kinerja dari suatu kompetensi.
- b. Kelengkapan dan ketepatan aspek yang akan dinilai dalam kinerja tersebut.
- c. kemampuan-kemampuan khusus yang diperlukan untuk menyelesaikan tugas.
- d. Upayakan kemampuan yang akan dinilai tidak terlalu banyak, sehingga semua dapat diamati.
- e. kemampuan yang akan dinilai diurutkan berdasarkan urutan yang akan diamati

Bentuk-bentuk teknik pengukuran pada ranah psikomotorik antara lain:

### **1. Daftar Cek**

Pengukuran ranah psikomotorik dapat dilakukan dengan menggunakan daftar cek (ya-tidak). Pada pengukuran ranah psikomotorik yang menggunakan daftar cek, peserta didik mendapat nilai apabila kriteria penguasaan kemampuan tertentu dapat diamati oleh penilai. Jika tidak dapat diamati, peserta didik tidak memperoleh nilai. Kelemahan cara ini adalah penilai hanya mempunyai dua pilihan mutlak, misalnya

benar-salah, dapat diamati-tidak dapat diamati. Dengan demikian tidak terdapat nilai tengah.

Berikut contoh daftar cek dalam mengukur ranah psikomotorik:

Format Penilaian Praktek Sholat  
(Menggunakan Daftar Tanda Cek)

Nama Peserta didik :

Kelas :

No	Aspek yang dinilai	Ya	Tidak
1	Niat		
2	Berdiri		
3	Takbiratul Ihram		
4	Membaca Surat Alfatihah		
5	Rukuk dengan Tumakninah		
6	Iktidal		
7	Sujud dua kali dengan tumakninah		
8	Duduk antara dua Sujud		
9	Tasyahud awal		
10	Tasyahud akhir		
11	Membaca shalawat		
12	Salam		
13	Tertib		
Skor yang dicapai			
Skor maksimum			

## 2. Skala Rentang

Pengukuran ranah psikomotorik yang menggunakan skala rentang memungkinkan penilai memberi nilai penguasaan

kompetensi tertentu karena pemberian nilai secara kontinum di mana pilihan kategori nilai lebih dari dua. Penilaian sebaiknya dilakukan oleh lebih dari satu penilai agar faktor subjektivitas dapat diperkecil dan hasil penilaian lebih akurat.

Berikut contoh skala rentang:

**Format Penilaian Praktek Sholat  
(Menggunakan Skala Rentang)**

Nama Peserta didik :

Kelas :

No	Aspek yang dinilai	Nilai			
		1	2	3	4
1	Niat				
2	Berdiri				
3	Takbiratul Ihram				
4	Membaca Surat Alfatihah				
5	Rukuk dengan Tumakninah				
6	Iktidal				
7	Sujud dua kali dengan tumakninah				
8	Duduk antara dua Sujud				
9	Tasyahud awal				
10	Tasyahud akhir				
11	Membaca shalawat				
12	Salam				
13	Tertib				
Jumlah					
Skor Maksimum					

Kriteria Penskoran: semakin baik penampilan peserta didik semakin tinggi skor yang diperoleh.

## Bab VI

### Analisis Instrumen Penilaian

**A**nalisis instrumen penilaian dikaji segi analisis logis/rasional dan analisis empirik. Analisis logis/rasional meliputi ranah materi, ranah konstruksi dan ranah bahasa. Sedangkan analisis empirik meliputi seperti validitas, reliabilitas, tingkat kesukaran dan daya beda tes.

#### A. Analisis Logis/Rasional

Analisis logis/rasional meliputi analisis materi, konstruksi dan bahasa. Analisis materi dimaksudkan sebagai penelaahan yang berkaitan dengan substansi keilmuan yang ditanyakan dalam soal serta tingkat kemampuan yang sesuai dengan soal. Analisis konstruksi dimaksudkan sebagai penelaahan yang umumnya berkaitan dengan teknik penulisan soal. Analisis bahasa dimaksudkan sebagai penelaahan soal yang berkaitan dengan penggunaan bahasa Indonesia yang baik dan benar. Berikut ditampilkan analisis logis yang terhadap bentuk soal uraian dan bentuk soal pilihan yang diadopsi dari



Pengembangan Sistem Penilaian yang dirancang oleh Departemen Pendidikan Nasional sebagai berikut:

### Telaah Butir Soal Uraian

No	Kriteria	Nomor Soal				
		1	2	3	4	5
A	Ranah Materi					
	Butir soal sesuai dengan indikator					
	Batasan pertanyaan dan jawaban yang diharapkan, jelas					
	Isi materi yang ditanyakan sesuai dengan jenjang, jenis sekolah dan tingkat kelas					
B	Ranah Konstruksi					
	Rumusan kalimat dalam bentuk kalimat Tanya atau perintah yang menuntut jawaban terurai					
	Ada petunjuk yang jelas cara mengerjakan soal					
	Ada pedoman penskorannya					
	Butir soal tidak bergabung pada butir soal sebelumnya					
C	Ranah Bahasa					
	Rumusan kalimat komunikatif					
	Kalimat menggunakan bahasa yang baik dan benar					
	Rumusan kalimat tidak menimbulkan penafsiran ganda atau salah pengertian					
	Menggunakan bahasa yang umum					
	Rumusan soal tidak mengandung kata-kata yang dapat menyinggung perasaan siswa					

Keterangan :

Beri tanda cek (V) jika menurut saudara sesuai dengan kriteria, dan beri tanda silang (X) jika menurut saudara tidak sesuai dengan kriteria

### Telaah Butir Soal Pilihan Ganda

No	Kriteria	Nomor Soal									
		1	2	3	4	5	6	7	8	9	10
A	Ranah Materi										
	Butir soal sesuai dengan indikator										
	Batasan pertanyaan dan jawaban yang diharapkan, jelas										
	Isi materi yang ditanyakan sesuai dengan jenjang, jenis sekolah dan tingkat kelas										
B	Ranah Konstruksi										
	Rumusan kalimat dalam bentuk kalimat Tanya atau perintah yang menuntut jawaban terurai										
	Ada petunjuk yang jelas cara mengerjakan soal										
	Ada pedoman penskorannya										

	Butir soal tidak bergabung pada butir soal sebelumnya										
C	Ranah Bahasa										
	Rumusan kalimat komunikatif										
	Kalimat menggunakan bahasa yang baik dan benar										
	Rumusan kalimat tidak menimbulkan penafsiran ganda atau salah pengertian										
	Menggunakan bahasa yang umum										
	Rumusan soal tidak mengandung kata-kata yang dapat menyinggung perasaan siswa										

Keterangan :

Beri tanda cek (V) jika menurut saudara sesuai dengan kriteria, dan beri tanda silang (X) jika menurut saudara tidak sesuai dengan kriteria

## **B. Analisis Empirik**

Analisis empirik terhadap instrumen/soal dilakukan dengan melakukan menguji validitas, reliabilitas, taraf kesukaran dan daya pembeda.

### **1. Validitas Tes**

#### *a. Pengertian Validitas Tes*

Valid artinya sah atau tepat. Jadi tes yang valid berarti tes tersebut merupakan alat ukur yang tepat untuk mengukur suatu objek. Berdasarkan pengertian ini, maka validitas tes pada dasarnya berkaitan dengan ketepatan dan kesesuaian antara tes sebagai alat ukur dengan objek yang diukur. Mengukur berat badan tentu tidak valid menggunakan meteran. Di kilang padi, ada timbangan yang valid untuk mengukur berat beras, akan tetapi timbangan ini tidak valid untuk mengukur berat emas dengan bentuk cincin.

Mengukur keterampilan siswa, misalnya mengukur unjuk kerja siswa, tentu tidak valid menggunakan tes pilihan ganda. Jadi, tes yang digunakan perlu disesuaikan dengan karakteristik hasil belajar yang diukur.

#### *b. Cara-cara Menentukan Validitas Tes*

Pada garis besarnya, cara-cara menentukan validitas tes dibedakan kepada dua, yaitu validitas rasional/logis dan validitas empiris atau validitas berdasarkan pengalaman.

Validitas rasional dapat dicapai dengan menjawab pertanyaan berikut ini:

1. Apakah tes benar-benar mengukur kompetensi atau hasil belajar yang akan diukur ?
2. Apakah bentuk tes sesuai digunakan untuk mengukur hasil belajar siswa ?

Untuk menentukan validitas instrumen secara empiris, peneliti harus melakukan uji coba (try out). Uji coba dilakukan kepada sebahagian sebagian siswa. Kemudian hasil uji coba tersebut diuji validitasnya. Banyak cara yang dapat kita tempuh untuk menguji validitas tes secara empiris. Pada makalah ini akan diperkenalkan tiga cara yang lazim digunakan.

### 1. Validitas eksternal

Validitas eksternal dilakukan dengan cara mengkorelasikan skor hasil uji coba instrumen yang dibuat guru dengan instrumen yang sudah baku.

Misalnya seorang guru Fikih membuat tes ujian semester genap kelas III tingkat Aliyah. Untuk menguji validitas eksternal tes yang dibuat guru, dapat dibandingkan dengan tes yang sudah baku, misalnya Tes Toufel.

Test kemampuan berbahasa Inggris yang dibuat guru dapat diuji validitas eksternal dengan cara:

- a. Mengujicobakan secara bersamaan tes yang dibuat guru dan tes toufel yang telah baku.
- b. Memberi skor-skor tes buatan dan tes toufel.

- c. Mencari angka korelasi antara skor-skor tes buatan dengan skorskor tes toulfe. Teknik korelasi yang digunakan adalah teknik korelasi Product Moment.
- d. Menguji signifikansi angka korelasi yang diperoleh pada langkah ketiga. jika angka korelasi yang diperoleh ternyata signifikan, berarti tes yang dibuat guru dapat dianggap VALID.

## 2. Validitas Internal

Validitas Internal dapat dilakukan dengan dua cara, yaitu analisis faktor dengan analisis butir.

- a. Analisis Faktor. Analisis faktor dilakukan dengan cara mengkorelasikan skor faktor dengan skor total. Teknik korelasi yang digunakan adalah teknik korelasi Product Moment. Jika terdapat korelasi positif dan signifikan, berarti item-item pada faktor tersebut dianggap valid.
- b. Analisis Butir Analisis butir dilakukan dengan cara mengkorelasikan skor-skor item dengan skor total. Korelasi dilakukan dengan teknik korelasi product moment. Jika terdapat korelasi positif dan signifikan antara skor item dengan skor total berarti item tersebut dianggap valid.

$$g_{pbi} = \frac{M_p - M_t}{s_t} \sqrt{\frac{p}{q}}$$

Keterangan :

$g_{pbi}$  = Koefisien Korelasi Biserial

$M_p$  = rerata skor dari subjek yang menjawab benar per item

$M_t$  = rerata skor total

$s_t$  = standar deviasi dari skor total

$p$  = proporsi peserta didik yang menjawab benar

$$p = \frac{\text{banyaknya siswa yang menjawab benar}}{\text{jumlah seluruh siswa}}$$

$q$  = proporsi peserta didik yang menjawab salah

$$q = 1 - p$$

Contoh penggunaannya ;

Guru memberikan skor kepada anak didiknya dengan ketentuan setiap item tes yang yang dijawab benar diberikan skor 1 dan bila salah diberi skor 0. Datanya tertera pada tabel berikut :

No	Nama	Butir Soal / Item										Skor Total
		1	2	3	4	5	6	7	8	9	10	
1	Feberianus	1	0	1	0	1	1	1	1	1	1	8
2	Dwy	0	0	1	0	1	0	0	1	1	1	5
3	Istiqamah	0	1	0	0	0	1	0	1	0	1	4
4	Yunna	1	1	0	0	1	1	0	0	1	0	5
5	Yessy	1	1	1	1	1	1	0	0	0	0	6
6	Tira	1	0	1	0	1	0	1	0	0	0	4
7	Nana	1	1	1	1	1	1	1	0	0	0	7
8	Nurhayati	0	1	0	1	1	1	1	1	1	1	8

Pertanyaan hitung validitas butir test nomor 1:

Langkah-langkah penyelesaian:

1. Buat tabel persiapan menghitung validitas item sebagai berikut;

No	Nama	X	Xt
1	Feberianus	1	8
2	Dwy	0	5
3	Istiqamah	0	4
4	Yunna	1	5
5	Yessy	1	6
6	Tira	1	4
7	Nana	1	7
8	Nurhayati	0	8
Jumlah		5	47

2. Hitung harga  $M_p$

$$M_p = \frac{8 + 5 + 6 + 4 + 7}{5}$$

$$M_p = 6$$

3. Hitung harga  $M_t$

$$M_t = \frac{8 + 5 + 4 + 5 + 6 + 4 + 7 + 8}{8}$$

$$M_t = 5,87$$



4. Hitung harga  $S_t$  (standar deviasi total)

$$S_t^2 = \frac{n \sum X_t^2 - (\sum X_t)^2}{n(n-1)}$$

$$S_t^2 = \frac{8(295) - (47)^2}{8(8-1)}$$

$$S_t^2 = 2,69$$

Dengan demikian dapat diketahui harga standar deviasi total dengan menarik akar dari varians total di atas yaitu 1,64.

5. Hitung harga  $p$

$$P = 5/8$$

$$P = 0,625$$

6. Hitung harga  $q$

$$q = 1 - 0,625$$

$$q = 0,375$$

sehingga diperoleh:

$$g_{pbi} = \frac{6 - 5,87}{1,64} \sqrt{\frac{0,625}{0,375}}$$

$$g_{pbi} = \frac{6 - 5,87}{1,64} \sqrt{\frac{0,625}{0,375}}$$

$$g_{pbi} = 0,08 \times 1,29$$

$$g_{pbi} = 0,10$$

Selanjutnya untuk menerima apakah butir tes yang dicari tersebut valid atau invalid, maka harga yang diperoleh tersebut dibandingkan dengan harga kritik korelasi. Secara umum, jika koefisien korelasi lebih besar dari 0,3, maka butir instrument tersebut sudah dikategorikan valid.

### **C. Reliabilitas Tes**

Menurut arti kata reliabel berarti dapat dipercaya. Berdasarkan arti kata tersebut, maka instrumen yang reliabel adalah instrumen yang hasil pengukurannya dapat dipecahkan. Salah satu kriteria instrumen yang dapat dipercaya jika instrumen tersebut digunakan secara berulang-ulang, hasil pengukurannya tetap. Mistar dapat dipercaya sebagai alat ukur, karena berdasarkan pengalaman jika mistar digunakan dua kali atau lebih mengukur panjang sebuah benda, maka hasil pengukuran pertama dan selanjutnya terbukti tidak berbeda. Sebuah tes dapat dikatakan reliabel jika tes tersebut digunakan secara berulang terhadap peserta didik yang sama hasil pengukurannya relatif tetap sama.

Banyak rumus-rumus yang dapat digunakan untuk mengetahui tingkat reliabilitas. Akan tetapi pada pembahasan ini diperkenalkan hanya dua buah rumus, yaitu rumus KR 21 dan rumus Alpha.

#### **1) Menentukan tingkat reliabilitas instrumen dengan rumus KR 21.**

Rumus KR 21 digunakan apabila alternatif jawaban pada instrumen bersifat dikotomi, misalnya benar-salah dan pemberian skor = 1 dan 0. Contoh penggunaan rumus KR 21.

Langkah pertama tes hasil uji coba diberi skor-skor, kemudian didistribusikan ke dalam tabel kerja sebagai berikut:

No	Nomor Butir																				Skor (X)
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
1	1	0	1	1	1	1	0	1	0	1	1	1	0	1	1	1	1	0	0	1	14
2	1	1	1	1	1	0	1	1	0	0	1	1	1	1	1	0	1	1	1	1	16
3	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	1	0	0	1	1	11
4	1	1	0	0	1	1	0	1	1	0	1	1	0	0	1	1	0	0	0	0	10
5	1	1	1	1	1	1	0	0	1	1	0	0	1	1	1	0	0	1	0	0	12
6	1	1	1	1	1	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	9
7	1	0	1	0	1	0	1	0	1	0	1	0	0	1	1	0	1	0	1	0	10
8	0	1	1	1	0	0	0	1	0	0	1	0	0	1	0	0	1	0	1	0	8
9	1	1	1	1	1	1	0	0	1	1	1	0	0	0	0	1	1	1	1	1	14
10	0	0	1	0	0	1	0	0	0	0	0	0	0	1	1	0	0	1	0	1	7
Jumlah	8	6	9	6	8	6	4	4	5	3	8	3	4	7	6	4	6	3	6	5	111

Langkah kedua menghitung varians skor total ( $S_t^2$ ) dengan rumus:

$$S_t^2 = \frac{n \sum X_t^2 - (\sum X_t)^2}{n(n-1)}$$

$$S_t^2 = \frac{10(1307) - (111)^2}{10(10-1)}$$

$$S_t^2 = 8,32$$

Langkah ketiga menghitung reliabilitas instrumen dengan rumus KR 21

$$r_{11} = \left(\frac{n}{n-1}\right)\left(1 - \frac{M(n-M)}{nS_t^2}\right)$$

Keterangan :

$r_{11}$  = reliabilitas instrument

$n$  = banyaknya butir soal

$M$  = mean/rata-rata skor

$nS_t^2$  = varians total

Jika dimasukkan ke rumus maka perhitungannya:

$$r_{11} = \left(\frac{20}{20-1}\right)\left(1 - \frac{11,1(20-11,1)}{20 \times 8,32}\right)$$

$$r_{11} = \left(\frac{20}{20-1}\right)\left(1 - \frac{11,1(20-11,1)}{20 \times 8,32}\right)$$

$$r_{11} = 1,0526 - 0,5932$$

$$r_{11} = 0,4594$$

## 2) Menentukan tingkat reliabilitas tes dengan rumus Alpa.

Rumus Alpa digunakan untuk mencari reliabilitas instrumen yang skornya merupakan rentangan 0 - 10, 0 - 100 atau berbentuk skala 1 - 3, 1 - 5 atau 1 - 10.

Rumus alpha:

$$r_{11} = \left[ \frac{k}{k-1} \right] \left[ 1 - \frac{\sum S_b^2}{S_t^2} \right]$$

Keterangan:

$r_{11}$  = reliabilitas instrument

$k$  = jumlah butir tes

$\sum S_b^2$  = jumlah varians butir

$S_t^2$  = varians total

Sebagai contoh perhitungan berikut ini disajikan tabel analisis 5 butir pertanyaan atau butir soal dari 10 orang peserta didik.

No	Nomor butiran					Skor total	Kuadrat skor total
	1	2	3	4	5		
1	10	8	9	10	8	45	2025
2	8	7	8	9	7	39	1521
3	6	5	6	8	7	32	1024
4	5	4	3	0	2	14	196
5	9	10	8	7	6	40	1600
6	7	5	3	4	7	26	676
7	3	4	4	5	6	22	484
8	4	3	5	5	5	22	484
9	6	2	2	2	3	15	225
10	7	6	1	5	4	23	529
Jlh	65	54	49	55	55	278	8764
Jlh Kuadrat	465	344	309	389	337	1844	

Sebelum dicari angka reliabilitasnya, perlu terlebih dahulu dicari varians butir dan varians skor total dengan rumus:

$$\sigma^2 = \frac{\sum X^2 - \frac{(X)^2}{N}}{N}$$

Untuk memperoleh jumlah varians butir dicari dulu varians setiap butir, kemudian dijumlahkan.

$$\sigma^2(1) = \frac{465 - \frac{(65)^2}{10}}{10}$$

$$\sigma^2(1) = \frac{465 - 422,5}{10}$$

$$\sigma^2(1) = 4,25$$

$$\sigma^2(2) = \frac{344 - \frac{(54)^2}{10}}{10}$$

$$\sigma^2(2) = \frac{344 - 291,6}{10}$$

$$\sigma^2(2) = 5,24$$

$$\sigma^2(3) = \frac{309 - \frac{(49)^2}{10}}{10}$$

$$\sigma^2(3) = \frac{309 - 240,1}{10}$$

$$\sigma^2(3) = 6,89$$

$$\sigma^2(4) = \frac{389 - \frac{(55)^2}{10}}{10}$$

$$\sigma^2(4) = \frac{389 - 302,5}{10}$$

$$\sigma^2(4) = 8,65$$

$$\sigma^2(5) = \frac{337 - \frac{(55)^2}{10}}{10}$$

$$\sigma^2(5) = \frac{337 - 302,5}{10}$$

$$\sigma^2(5) = 3,50$$

Dengan demikian diperoleh total varian butir adalah:

$$\sigma^2 = 4,25 + 5,24 + 6,89 + 8,65 + 3,50$$

$$\sigma^2 = 28,8$$

Sedangkan varians total dihitung sebagai berikut

$$\sigma^2(t) = \frac{8764 - \frac{(278)^2}{10}}{10}$$

$$\sigma^2(t) = \frac{8764 - 7728,4}{10}$$

$$\sigma^2(t) = 103,56$$

Selanjutnya harga-harga yang diperoleh dimasukkan ke dalam rumus Alpha sebagai berikut:

$$r_{11} = \left[ \frac{k}{k-1} \right] \left[ 1 - \frac{\sum S_b^2}{S_t^2} \right]$$

$$r_{11} = \left[ \frac{5}{5-1} \right] \left[ 1 - \frac{28,8}{103,56} \right]$$

$$r_{11} = \left[ \frac{5}{4} \right] [1 - 0,27]$$

$$r_{11} = 1,25 \times 0,73$$

$$r_{11} = 0,91$$

Dengan demikian diperoleh koefisien reliabilitas tes sebesar 0,91. Selanjutnya dengan merujuk Sudijono (2002) suatu tes dikatakan reliabel apabila koefisien  $\geq 0,70$ . Dengan demikian tes tersebut reliabel.





Adapun rumus mencari P adalah

$$P = \frac{B}{JS}$$

Dimana:

$P$  = indeks kesukaran

$B$  = banyaknya siswa yang menjawab soal itu dengan betul

$JS$  = jumlah seluruh siswa peserta tes

Misalnya :

Ada 20 orang dengan nama kode A-T yang mengajarkan tes yang terdiri dari 20 soal. Jawaban tesnya dianalisa dan jawabannya tertera seperti dibawah ini.

(1= jawaban benar, 0= jawaban salah)

Siswa	Nomor Soal																				Sekor Siswa
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
A	1	1	0	0	1	0	1	1	0	1	1	0	1	0	1	1	1	1	0	1	13
B	0	1	0	0	1	0	1	1	0	0	1	1	1	1	0	0	0	1	1	1	11
C	1	1	0	0	1	1	1	0	0	1	0	1	1	0	1	1	0	1	1	1	12
D	0	1	0	0	1	1	0	1	1	0	0	1	1	1	0	0	0	1	1	0	9
E	1	1	0	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	0	14
F	0	0	0	1	1	1	0	0	0	0	1	1	1	0	0	0	1	0	0	0	8
G	1	0	0	1	0	0	1	1	0	1	0	1	1	1	1	1	1	0	1	1	13
H	0	0	0	1	0	0	1	1	0	0	1	1	0	1	0	1	0	1	0	1	9
I	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	0	1	1	1	17
J	0	1	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	0	13
K	1	1	0	1	0	0	1	0	0	1	0	1	1	0	0	1	1	0	1	0	10
L	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0	4
M	1	0	0	0	1	0	1	1	1	1	1	1	1	0	1	0	1	1	0	1	13
N	0	1	1	0	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	16
O	1	1	0	0	1	0	1	0	1	1	0	1	1	0	1	0	1	1	1	0	12
P	0	1	0	0	1	1	1	0	0	0	0	1	1	1	0	0	0	0	1	1	10
Q	1	0	0	0	0	0	0	1	0	1	1	1	1	0	0	0	0	1	1	1	9
R	0	1	0	1	1	0	1	1	0	1	1	1	1	0	0	0	0	1	0	1	11
S	1	1	0	1	1	0	1	1	1	0	0	1	1	1	1	0	1	1	0	1	14
T	0	1	0	1	1	0	1	1	0	0	0	1	1	1	0	0	0	1	0	1	10
J L H	10	14	4	9	15	6	18	17	7	11	10	18	20	10	9	7	10	14	13	13	

Dari tabel yang disajikan di atas dapat ditafsirkan bahwa :

- Soal nomor 1 mempunyai taraf kesukaraan  $\frac{10}{20} = 0,5$
- Soal nomor 13 adalah soal yang paling mudah karena seluruh siswa peserta tes dapat menjawab :

$$\text{Indeks kesukarannya} = \frac{20}{20} = 1,0$$

Menurut ketentuan yang sering diikuti, indeks kesukaraan sering diklasifikasikan sebagai berikut:

- Soal dengan P 0,00 sampai 0,30 adalah soal sukar
- Soal dengan P 0,30 sampai 0,70 adalah soal sedang
- Soal dengan P 0,70 sampai 1,00 adalah soal mudah

Walaupun demikian itu yang berpendapat bahwa: soal-soal yang dianggap baik, yaitu soal-soal sedang, adalah soal-soal yang mempunyai indeks kesukaraan 0,30 sampai dengan 0,70.

### E. Daya Pembeda

Daya pembeda soal adalah kemampuan sesuatu soal untuk membedakan antara siswa yang pandai (berkemampuan tinggi) dengan siswa yang tidak pandai (berkemampuan rendah). Angka yang menunjukkan besarnya beda pembeda disebut indeks Diskriminasi, disingkat D. Seperti halnya indeks kesukaraan, indeks diskriminasi (daya pembeda) ini berkisar antara 0,00 sampai 1,00 hanya bedanya indeks kesukaraan tidak mengenal tanda negatif. Tanda negatif pada indeks diskriminasi digunakan jika sesuatu soal “terbalik” menunjukkan kualitas tester yaitu anak pandai disebut tidak pandai dan anak tidak pandai disebut pandai. Dengan demikian ada tiga titik pada daya pembeda yaitu :

-1,00      0,00      1,00  
Daya pembeda      Daya pembeda      Daya pembeda  
negative      rendah      tinggi (positif)

Bagi sesuatu soal dapat dijawab benar oleh siswa pandai maupun siswa tidak pandai, maka soal itu tidak baik karena tidak mempunyai daya pembeda. Demikian pula jika semua baik yang pandai maupun yang tidak pandai tidak dapat menjawab dengan benar, soal tersebut tidak baik, juga karena tidak mempunyai daya pembeda. Soal yang baik adalah soal yang dijawab benar oleh siswa-siswa yang pandai saja. Seluruh pengikut tes dikelompokkan menjadi 2 kelompok yaitu kelompok pandai atau kelompok atas (*upper group*) dan kelompok tidak pandai atau kelompok bawah (*lower group*).

### Cara menentukan daya pembeda (nilai D)

Untuk ini perlu dibedakan antara kelompok kecil (kurang dari 100) dan kelompok besar (100 orang ke atas).

- a. Untuk Kelompok Kecil Seluruh kelompok tester dibagi dua sama besar, 50% kelompok atas dan 50% kelompok bawah. Contoh :

Siswa	Sekor	
A	9	
B	8	
C	7	
D	7	
E	6	Kelompok atas (JA)
F	5	
G	5	
H	4	
I	4	
J	3	Kelompok Bawah (JB)

Seluruh pengikut tes, dideretkan mulai dari skor teratas sampai terbawah, lalu dibagi dua.

- b. Untuk Kelompok Besar Mengingat biaya dan waktu untuk menganalisa, maka untuk kelompok besar biasanya hanya diambil kedua kutubnya saja, yaitu 27% skor teratas sebagai kelompok atas (JA) dan 27% skor terbawah sebagai kelompok bawah (JB).

JA = Jumlah kelompok atas

JB = Jumlah kelompok bawah

Contoh :

9	
9	
8	
8	
8	27 % sebagai JA
.	
-	
.	
.	
.	27 % sebagai JB
2	
1	
1	
1	
0	

Rumus Mencari D

Rumus untuk menentukan indeks diskriminasi adalah:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = PA - PB$$

Dimana

J = Jumlah Peserta tes

JA = Banyaknya peserta kelompok atas

JB = Banyaknya peserta kelompok bawah

BA = Banyaknya peserta kelompok atas yang menjawab soal benar

BB = Banyaknya peserta kelompok bawah menjawab soal benar

PA =  $\frac{BA}{JA}$  = banyaknya peserta kelompok atas yang menjawab benar (ingat p sebagai simbol indeks kesukaran).

PB =  $\frac{BB}{JB}$  = proporsi peserta kelompok bawah yang menjawab benar

Contoh Perhitungan :

Dari hasil analisa tes yang terdiri dari 10 butir soal yang dikerjakan oleh 20 orang siswa, terdapat dalam tabel sebagai berikut:

Tabel analisa 10 butir soal, 20 orang siswa.

Siswa	Kelom Pok	Nilai Sosial										Skor Siswa
		1	2	3	4	5	6	7	8	9	10	
A	B	1	0	1	0	0	0	1	1	1	0	5
B	A	0	1	1	1	1	1	0	0	1	1	7
C	A	1	0	1	0	1	1	1	1	1	1	8
D	B	0	0	1	0	0	1	1	1	1	1	5
E	A	1	1	1	1	1	1	1	1	1	1	10
F	B	0	1	0	0	0	1	1	1	1	1	6
G	B	0	1	0	0	0	1	1	1	1	1	6
H	B	0	1	1	0	0	1	0	1	1	1	6
I	A	1	1	1	0	0	1	1	1	1	1	8
J	A	1	1	1	1	0	0	1	0	1	1	7
K	A	1	1	1	0	0	1	1	1	1	0	7
L	B	0	1	0	1	1	0	0	1	1	0	5
M	B	0	1	0	0	0	0	0	1	1	0	3
N	A	0	0	1	0	1	1	1	1	1	1	7
O	A	1	1	0	1	1	1	1	1	1	1	9
P	B	0	1	0	0	0	1	0	0	1	0	3
Q	A	1	1	0	1	0	1	1	1	1	1	8
R	A	1	1	1	1	0	1	1	1	1	0	8
S	B	1	0	1	0	0	1	1	1	1	0	6
T	B	0	1	0	1	0	1	1	1	1	0	6
Jumlah		11	15	12	8	6	16	15	17	20	10	

Berdasarkan nama-nama siswa tersebut dapat kita peroleh skor-skor sebagai berikut :

A=5    F=6    K=7    P=3  
 B=7    G=6    L=5    Q=8  
 C=8    H=6    M=3    R=8  
 D=5    I=8    N=7    S=6  
 E=10    J=7    O=0    T=6



Dari angka-angka yang belum teratur kemudian dibuat array (uraian penyebaran), dari skor yang paling tinggi ke skor yang paling rendah.

Kelompok Atas	Kelompok Bawah
10	6
9	6
8	6
8	6
8	6
8	5
7	5
7	5
7	3
7	3
10 orang	10 orang

Array ini sekaligus menunjukkan adanya kelompok atas (JA) dan kelompok bawah (JB) dengan pemilikannya sebagai berikut :

Kelompok (JA)	Kelompok (JB)
B = 7	A = 5
C = 8	D = 5
E = 10	F = 6
I = 8	G = 6
J = 7	H = 6
K = 7	L = 5
N = 7	M = 3
O = 9	P = 3
Q = 8	S = 6
R = 8	T = 6
———— 10 Orang	———— 10 Orang

Mari kita perhatikan lagi tabel analisa, khusus untuk butir soal nomor 1.

- Dari kelompok atas yang menjawab betul 8 orang.
- Dari kelompok bawah yang menjawab betul 3 orang

Kita terapkan dalam rumus indeks diskriminasi :

$$JA = 10$$

$$JB = 10$$

$$P = 0,8$$

$$PB = 0,3$$

$$BA = 8$$

$$BB = 3$$

$$\text{Maka } D = PA - PB$$

$$= 0,8 - 0,3 = 0,5$$

Butir soal ini jelek karena lebih banyak dijawab benar oleh kelompok bawah dibandingkan dengan jawaban kelompok atas. Ini berarti bahwa untuk menjawab soal dengan benar dapat dilakukan dengan menebak:

Butir- butir soal yang baik adalah butir-butir soal yang mempunyai indeks diskriminasi 0,4 sampai 0,7

## Klasifikasi Daya Pembeda

D : 0,00 – 0,20 : jelek (poor)

D : 0,20 – 0,40 : cukup (satisfactory)

D : 0,40 – 0,70 : baik (good)

D : 0,70 – 1,00 : baik sekali (excellent)

D : Negatif, semuanya tidak wajib, jadi semua butir soal yang mempunyai nilai D negatif sebaiknya dibuang saja.

## DAFTAR PUSTAKA

- Arifin, Z. *Evaluasi Instruksional: Prinsip, Teknik, Prosedur*, Bandung: PT. Remaja Rosdakarya, 2013.
- Arikunto, S. *Dasar-Dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara, 2005.
- Departemen Pendidikan Nasional Badan Standar Nasional Pendidikan, *Pedoman Model Penilaian Kelas Kurikulum Tingkat Satuan Pendidikan*, Jakarta: 2006/2007.
- Departemen Pendidikan Nasional, *Pengembangan Sistem Penilaian*, Jakarta: Direktorat Pendidikan Menengah Umum, 2004.
- Departemen Pendidikan, *Model Penilaian Kurikulum Tingkat Satuan Pendidikan*, Jakarta, 1995.
- Guba, E.G. & Lincoln, Y.S. *Effective Evaluation*, San Francisco: Jossey Bass Pub, 1985.
- Joni, T.R. *Pengukuran dan Penilaian Pendidikan*, Surabaya: Karya Anda, 1984.
- Mardapi, D. *Pengembangan Sistem Penilaian Kurikulum Berbasis Kompetensi*, Jakarta, 2004.
- Mohrens, W.A. *Measurement and Evaluation in Education and Psychology*, New York: Rinchart and Wionston, 1984.

- Purwanto, M.N. *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*, Bandung: Remaja Rosda Karya, 1990.
- Sax, G. *Principles of Educational and Psychological Measurement and Evaluation*, Belmont California: Wads Worth Pub.Co, 1980.
- Stambeek, C.S. *Prinsip Dan Teknik Pengukuran Dan Penilaian Didalam Dunia Pendidikan*, Cet II, Mutiara S. Wijaya, Jakarta, 1986.
- Sudijono, A. *Pengantar Evaluasi Pendidikan*, Jakarta: Raja Grafindo Persada, 2004.
- Surapranata, S. & Hatta, M. *Penilaian Portofolio Impelementasi Kurikulum 2004*, Bandung: Remaja Rosda Karya, 2005.
- Surapranata, S. *Panduan Penulisan Tes Tertulis Implementasi Kurikulum 2004*, Bandung: Remaja Rosda Karya, 2005.
- Sutomo, *Teknik Penilaian Pendidikan*, Surabaya: Bina Ilmu, 1995.
- Thoah, M.C. *Teknik Evaluasi Pendidikan*, Cet I, Rajawali Pers, Jakarta, 1991.

## TENTANG PENULIS



Arief Aulia Rahman, lahir di Langsa, 11 Oktober 1991, anak kedua dari bapak Drs. Ahmad As'adi dan ibu Dra. Aminah Sulaiman. Telah menyelesaikan sekolah dasar, sekolah menengah pertama dan sekolah menengah atas di kota Langsa, Provinsi Nanggroe Aceh Darussalam.

Setelah itu melanjutkan pendidikan pada program sarjana pendidikan Matematika di Universitas Muhammadiyah Sumatera Utara (UMSU) selama 3 tahun 8 bulan, pada tahun 2014 mengambil program master pendidikan matematika di Universitas Negeri Medan (UNIMED) selama 1 tahun 5 bulan. Pernah menjadi guru bidang Matematika di MAN 2 Model Medan dan Sekarang aktif sebagai dosen pendidikan matematika di Sekolah Tinggi Keguruan dan Ilmu Pendidikan Bina Bangsa Meulaboh (STKIP BBM), menjadi salah satu anggota *The Indonesian Mathematical Society (indoMS)* dan Forum Dosen Indonesia (FDI). Dua kali mendapat penghargaan pada tahun 2017 dalam rangka pengabdian kepada masyarakat di Aceh Tamiang melalui pelatihan membuat media ajar berbasis pendekatan realistik untuk guru-guru di Sekolah Dasar dan pentingnya strategi pembelajaran untuk menumbuhkan minat belajar siswa.



Cut Eva Nasryah, Lahir di Medan, 11 Januari 1992. Anak dari bapak Nasruddin dan ibu rawiyah yusri. Menyelesaikan sekolah dasar hingga sekolah menengah atas di Sumatera Utara. Melanjutkan pendidikan pada program S-1 pendidikan matematika selama 4 tahun dan S-2 pendidikan Matematika selama 2 tahun di Universitas Negeri Medan (UNIMED). Sekarang aktif sebagai dosen di STKIP Bina Bangsa Meulaboh (STKIP BBM). Penelitian-penelitian yang dilakukan terkait pengembangan perangkat pembelajaran, *realistic mathematics education*, kecerdasan emosional, dan etnomatematika.