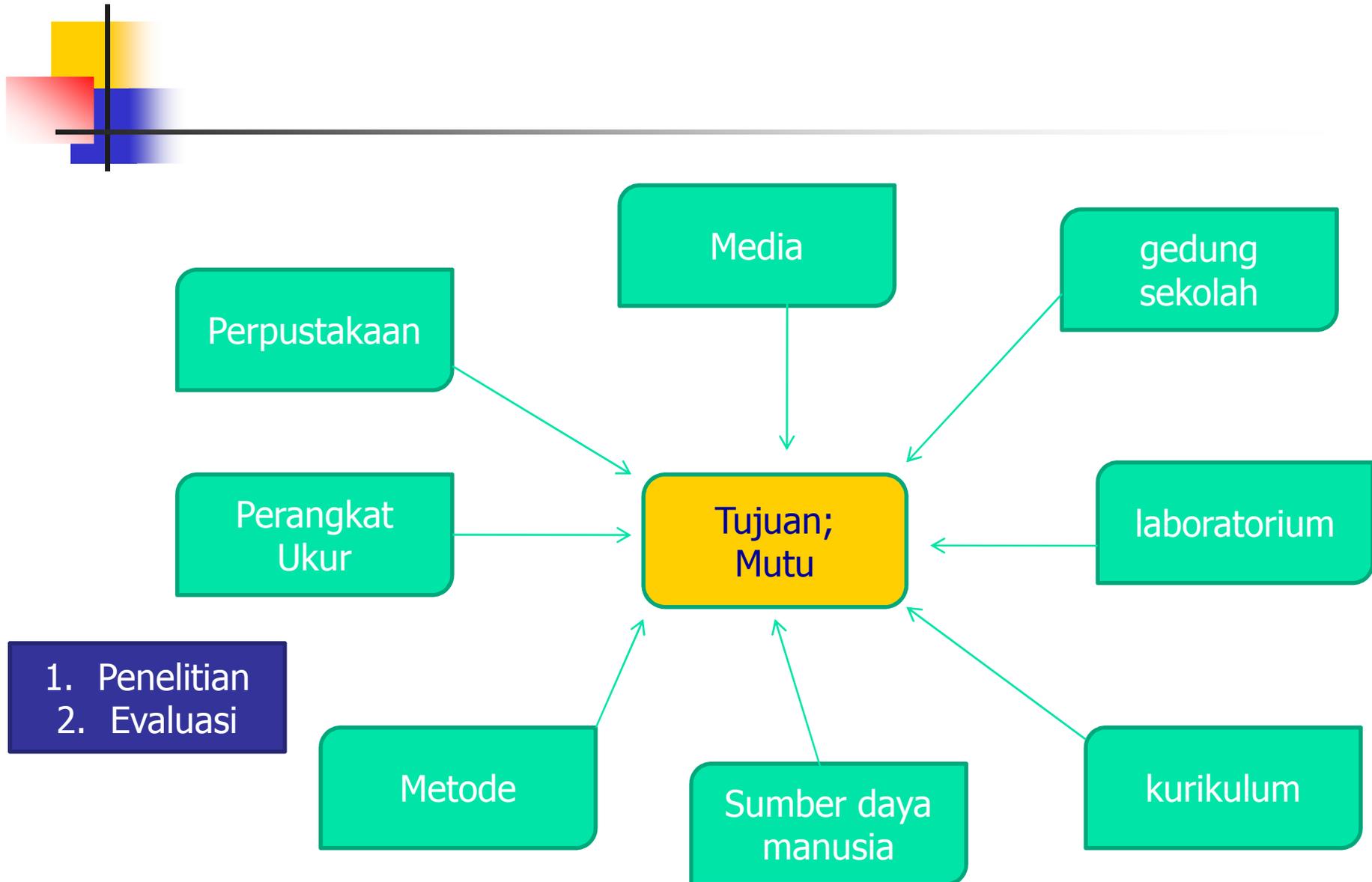


# **EVALUASI MUTU PENDIDIKAN DI SEKOLAH**

---



# Evaluasi sebagai satu sistem



# Standar Pendidikan (BNSP)



- Standar pengelolaan
- Standar kompetensi lulusan
- Standar proses
- Standar tenaga kependidikan
- Standar penilaian
- Standar sarana dan prasarana
- Standar isi
- Standar pembiayaan

# Pengukuran, Penilaian, dan Evaluasi

1. Pengukuran, adalah kegiatan membandingkan sesuatu dengan sesuatu atas dasar ukuran tertentu. Pengukuran sifatnya kuantitatif. Jenis-jenis pengukuran;
  - a. Pengukuran bukan untuk menguji; misal pengukuran yang dilakukan oleh penjahit
  - b. Pengukuran untuk menguji; misal pengukuran daya tahan lampu pijar.
  - c. Pengukuran untuk menilai, yang dilakukan dengan jalan menguji sesuatu; misal mengukur kemajuan belajar siswa dalam rangka mengisi rapor dengan tes.

2. Penilaian, adalah pemberian nilai terhadap sesuatu. Menilai berarti mengambil keputusan terhadap sesuatu dengan mendasarkan diri atau berpegang pada ukuran-ukuran tertentu, baik-buruk, sehat-sakit, pandai-bodoh. Penilaian sifatnya kualitatif.

Misal, suhu badan seseorang adalah  $36^{\circ}$  maka orang tersebut tergolong normal atau berbadan sehat.

3. Evaluasi, adalah kegiatan atau proses untuk menilai sesuatu oleh karena itu evaluasi mencakup pengukuran dan penilaian.

Pengukuran akan menghasilkan keterangan yang berupa angka yang diperoleh dari pengujian atau tes.

Penilaian merupakan penafsiran dari data kuantitatif.

## Beberapa pengertian evaluasi

- Proses yang menentukan sampai sejauhmana tujuan pendidikan dapat dicapai, (Tyler).
- Menyediakan informasi untuk pembuatan keputusan, (Cronbach, Stufflebeam, Alkin).
- Perbedaan apa yang ada dengan suatu standar untuk mengetahui apakah ada selisih (MacIcolm).
- Penelitian yang sistematis atau yang teratur tentang manfaat atau guna beberapa Objek, (joint comittee)

Evaluasi tidak menghakimi, karena bukan tugas evaluator menentukan apakah suatu program berguna atau tidak. Evaluator tidak bertindak sebagai wasit terhadap orang lain.

4. Evaluasi mutu pendidikan, adalah penilaian yang dilakukan dalam pendidikan berdasarkan hasil pengukuran yang diperoleh dari proses penyelenggaraan pendidikan dalam kurun waktu tertentu.

## Sasaran evaluasi

1. Objek evaluasi; siswa/mahasiswa, sarana dan prasarana, PBM, proyek atau program institusi
2. Aspek dan dimensi; dahulu berorientasi pada hasil, misal dari pendidikan, pelatihan aspek yang dievaluasi hasil yang telah dicapai peserta.

Akhir-akhir ini sasaran evaluasi menjadi luas banyak variabel yang dievaluasi.

- Model evaluasi Contack, Input, Proses, dan Product (CIPP) Stufflebeam mengemukakan empat aspek yaitu;
  - a. Konteks
  - b. Input
  - c. Proses implementasi,
  - d. Produk

#### Kreteria menilai objek

- Pencapaian Tujuan
- Indentifikasi kebutuhan yang potensial
- Nilai-nilai sosial
- Mutu, dan
- Efisiensi

## Siapa yang harus dilayani oleh evaluasi

- Agar evaluasi benar-benar bermanfaat, maka ada 3 audiensi yang tepat;
  - Evaluasi dapat diminta lebih dari satu audiensi
  - Masing-masing audien punya kebutuhan yang berbeda
  - Audien khusus kebutuhannya harus dirumuskan dengan jelas pada waktu memulai rencana

## Langkah dan Prosedur dalam evaluasi

- Memfokuskan evaluasi
- Mendesain evaluasi
- Mengumpulkan informasi
- Menganalisis informasi
- Melaporkan hasil evaluasi
- Mengelola evaluasi
- Mengevaluasi evaluasi

## Metode Evaluasi

Ada beberapa metode yang dapat digunakan dalam evaluasi yaitu;

1. Eksperimen, dan kuasi eksperimen (Cambell dan Stanley)
2. Metode naturalisme (Guba dan Lincoln)
3. Analisis sistem (Jury trials, Wolf)

# Evaluator

## Persyaratan/kompotensi evaluator

1. Mengetahui dan mengerti teknik pengukuran,
2. Memahami metode penelitian
3. Memahami kondisi sosial
4. Jujur,
5. Mengerti tentang evaluasi
6. Bertanggung jawab.

## Kriteria untuk menilai evaluasi

Standar untuk evaluasi dalam bidang pendidikan antara lain;

1. Bermanfaat dan praktis (utility)
2. Tepat sasaran (accuracy)
3. Realistik dan teliti (feasibility)
4. Dilakukan dengan legal dan etik (propriety)

---

## Model Evaluasi

Model evaluasi adalah model desain yang dibuat oleh ahli dan pakar evaluasi yang biasanya dinamakan sesuai dengan nama pembuat

1. Model CIPP (Context, Input, Process, dan Product) dari Stufflebeam.

Evaluasi sebagai suatu proses menggambarkan, memperoleh, dan menyediakan informasi yang berguna untuk menilai alternatif keputusan.

**Context**, berisi membantu merencanakan keputusan, menentukan kebutuhan yang akan dicapai oleh program dan merumuskan tujuan program.

---

---

**Input**, berisi pengaturan keputusan, menentukan sumber-sumber yang ada, alternatif yang diambil, rencana dan strategi untuk mencapai kebutuhan. (bagaimana prosedur kerja untuk mencapinya)

**Process**, berisi proses mengimplementasikan keputusan. Sejauhmana rencana telah diterapkan, apa harus direvisi. Jika pertanyaan sudah terjawab, prosedur dapat dimonitor, dikontrol, dan diperbaiki.

**Product**, berisi produk untuk keputusan selanjutnya. Apa hasil yang telah dicapai, apa yang harus dilakukan setelah program berjalan.

---

---

## 2. Model UCLA dari Alkin

Model ini hampir sama dengan model CIPP, menurut Alkin, evaluasi sebagai suatu proses menyakinkan keputusan, memilih informasi yang tepat, mengumpulkan data, dan menganalisis informasi sehingga dapat melaporkan data yang berguna bagi pembuat keputusan dalam memilih beberapa alternatif. Ada 5 macam evaluasi yaitu;

- **Sistem assessment**, memberikan informasi tentang keadaan atau posisi sistem
  - **Program Planning**, membantu pemilihan program tertentu yang mungkin akan berhasil memenuhi program.
-

- 
- **Program implementation**, menyiapkan informasi apakah program sudah diperkenalkan kepada kelompok tertentu yang tepat seperti yang direncanakan.
  - **Program improvement**, memberikan informasi bagaimana program berfungsi program bekerja atau berjalan. Apakah menuju pencapaian tujuan, adakah hal-hal masalah baru yang tidak terduga.
  - **Program cervication**, memberi informasi tentang nilai atau guna program.
-

---

### 3. Model Stake atau Model Countenance

- Dalam evaluasi ada dua dasar kegiatan yaitu, descriptions dan judgement, dan membedakan adanya tiga tahap dalam program pendidikan yaitu; anticendents (context), transaction (process), dan outcomes (output).
  - Matrix descriptions menunjukkan intents (goals) dan observations (effects) atau apa yang sebenarnya terjadi.
  - Judgements mempunyai dua aspek, yaitu standard dan judgement.
  - Stake mengatakan jika menilai suatu program pendidikan harus melakukan perbandingan yang relatif antara satu program dengan program lainnya, atau yang absolut (satu program dengan yang standar).
-

- 
- Evaluator membuat penilaian tentang program yang dievaluasi, meskipun description di satu pihak berbeda dengan judgement (menilai). Model ini antecedents (masukan), transaction (proses), dan outcomes (hasil) dibandingkan tidak hanya untuk menentukan apa ada perbedaan tujuan dengan keadaan yang sebenarnya, tetapi dibandingkan dengan standar yang absolut untuk menilai manfaat program.
-

---

## Istilah Pendekatan Evaluasi

Istilah pendekatan evaluasi sebagai pendapat tentang tugas evaluasi dan bagaimana dilakukan atau tujuan dan prosedur evaluasi (bagaimana memperoleh informasi yang tepat).

1. Pendekatan eksperimen,

Evaluasi yang berorientasi pada penggunaan experimental science dalam program evaluasi.

Tujuannya evaluator memperoleh kesimpulan umum tentang dampak suatu program tertentu dengan mengontrol faktor lain ikut mempengaruhi program.

---

---

## 2. Pendekatan berorientasi pada tujuan (goal oriented approach)

Pendekatan ini memakai tujuan program sebagai kriteria untuk menentukan keberhasilan. Evaluator hanya mengukur sampai dimana pencapaian tujuan telah dicapai. Pendekatan evaluasi yang semacam ini merupakan pendekatan yang amat wajar dan praktis untuk desain dan pengembangan program.

Pengembang model ini perlu menjelaskan hubungan antara kegiatan khusus yang ditawarkan dan hasil yang akan dicapai.

---

---

### 3. Pendekatan yang berfokus pada keputusan

Pendekatan ini menekankan pada peranan informasi yang sistematis untuk pengelolaan program dalam menjalankan tugasnya.

Pengumpulan data dan laporan dibuat untuk menambah efektifitas pengelola program, maka program sering berubah selama beroperasi dari awal sampai akhir. Kebutuhan pemegang keputusan juga berubah, dan evaluasi harus disesuaikan dengan keadaan.

---

---

#### 4. Pendekatan bebas tujuan (goal free evaluation)

Tujuan pendidikan tak dapat dikatakan sebagai pemberian, seperti tujuan lain, ia harus dievaluasi. Tujuan umumnya hanya formalitas dan jarang menunjukkan tujuan yang sebenarnya dari proyek atau tujuan berubah.

Banyak hasil program penting tidak sesuai dengan tujuan program, misal membangun pusat vocational pendidikan baru akan menciptakan lapangan kerja baru, tetapi tak pernah mencapai tujuan proyek yang nyata. Oleh karena itu pendekatan ini bertujuan untuk mengurangi bias dan menambah objektivitas.

---

---

## 5. Formatif-Sumatif

- Evaluasi formatif adalah evaluasi yang dilakukan pada waktu program masih berjalan dan merupakan umpan balik dalam kemajuan program.
  - Evaluasi sumatif adalah evaluasi yang dilakukan setelah program selesai dan hasil evaluasi merupakan akhir dari suatu program.
-

---

## Fungsi Evaluasi Mutu

- a. Memperoleh data yang dapat memberikan petunjuk sampai dimana tingkat kemampuan dan keberhasilan dalam mencapai kompetensi tertentu.
  - b. Mengetahui efektivitas metode pengajaran, sarana, prasarana yang dipergunakan dalam proses pembelajaran.
-

---

# Sasaran evaluasi mutu pendidikan di sekolah

Program kegiatan untuk peningkatan mutu yang dirancang sekolah dalam rencana kegiatan jangka pendek dan jangka panjang dalam segala aspek.

---

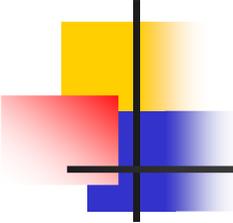
# EVALUASI MUTU PEMBALAJARAN (PBM)

- Kurikulum adalah seperangkat rencana dan pengaturan mengenai; tujuan, isi, dan bahan pelajaran serta cara yang dipergunakan sebagai pedoman penyelenggaraan kegiatan pembelajaran untuk mencapai tujuan.
- KTSP adalah kurikulum operasional yang disusun dan dilaksanakan oleh masing-masing satuan pendidikan dengan mengacu pada Standar Nasional Pendidikan yang disusun oleh BSNP terutama berkaitan dengan Standar Isi (SI) dan Standar Kompetensi Lulusan (SKL).
- KTSP diberlakukan di seluruh jenjang pendidikan

- Pada pembelajaran alat ukur berfungsi sebagai alat untuk membantu mengungkap kemampuan laten yang berada dalam diri peserta didik yang disebut dengan tes hasil belajar.
- Hasil pengukuran merupakan input yang memberikan gambaran mengenai kemampuan peserta didik dan berfungsi sebagai indikator keberhasilan dalam mencapai tujuan dalam kurikulum.

## Lingkup Evaluasi Pembelajaran KTSP di SD/MI

- a. **Kognitif (*Cognitive*)**, meliputi aspek; pengetahuan, pemahaman, penerapan, analisis, sintesis, dan evaluasi.
- b. **Sikap (*Attitude*)**, menunjuk kepada kecenderungan bertindak (predisposisi) seseorang, meliputi aspek-aspek: penerimaan, merespon, menghargai, pengorganisasian atau pengintegrasian, pengkarakterisasian.
- c. **Keterampilan (*Psikomotor*)**, berkaitan dengan kemampuan pergerakan syaraf-otot, meliputi aspek-aspek: persepsi (*perception*), kesiap-sediaan (*mental set*), respon/gerakan dipimpin/terbimbing (*guided respons*), gerakan kebiasaan-mekanisme (*mechanism*), gerakan khas/kompleks, yang menunjukkan taraf keterampilan tertentu (*skillful*) serta profisiensi (koordinatif), dan gerakan penyesuaian (*adaptation*).



# Jenis Tes

---

Tes yang biasa digunakan adalah tes objektif dan tes uraian khusus untuk mengukur kognitif.

Berdasarkan Cara Menjawab

- Butir tes yang ada pilihan jawabannya (*selected response test items*).
- Butir tes yang tidak ada pilihan jawaban (*constructed response test item*)

# Model Tes Hasil Belajar

- **Teori Ujian Klasik**, biasa dibuat oleh guru, memiliki ketergantungan antara butir dengan kemampuan kelompok peserta didik.
- **Validitas**, berfungsi menilai ketepatan dengan sasaran yang diukur pada setiap butir. Jenis; validitas isi, kriteria (serentak dan prediktif), dan konstruk.
- **Reliabilitas**, berfungsi menilai mutu perangkat tes yang terdiri atas sejumlah butir. Jenis; reliabilitas stabilitas, konsistensi internal, dan ekuivalen.
- **Analisis butir** untuk mencari butir yang sesuai dengan kemampuan kelompok tertentu.

- 
- Teori Ujian Modern atau Teori Responsi Butir (*Item Response Theory/IRT*), tes yang dibuat standar.
  - Memiliki ciri-ciri ; **unidimensi, invariansi, dan independensi lokal**, serta kecocokan model logistik.
  - IRT terdiri dari parameter kemampuan dan parameter butir. Model IRT; L1P, L2P, dan L3P.
  - Analisis butir pada teori responsi butir berguna untuk mencari butir tes atau sejumlah butir yang baik sebagai perangkat tes dengan menggunakan fungsi informasi butir.
  - Informasi butir merupakan kecocokan antara parameter kemampuan dengan parameter butir dan dipengaruhi oleh probabilitas menjawab benar.
-

# Konstruksi Teori Ujian Klasik

- Tahapan pembuatan tes adalah; menentukan tujuan tes, menetapkan standar kompetensi dan kompetensi dasar, penyusunan tabel spesifikasi berdasarkan SK-KD, penyusunan butir tes, ujicoba butir, seleksi butir, pengujian reliabilitas, pengujian validitas, analisis butir, dan butir tes siap digunakan.

# Konstruksi Teori Ujian Modern

- Sama dengan teori ujian klasik dan dilanjutkan dengan pengujian persyaratan model logistik
- Syarat-syarat teori ujian modern adalah; independensi lokal, invariansi kelompok, unidimensi, dan kecocokan model logistik
- Nilai informasi butir merupakan gambaran mengenai mutu setiap butir tes yang terdapat dalam suatu perangkat tes, mengenai kecocokan kemampuan peserta tes, tingkat kesukaran, daya beda, dan faktor tebakan.
- Perangkat tes yang bermutu masing-masing butir tesnya memiliki nilai informasi yang tinggi atau nilai variansinya kecil.

# Skor

Skor adalah harga kuantitatif suatu jawaban terhadap butir dalam tes. Dengan memberikan skor, dapat diperoleh gambaran tentang kemampuan siswa dalam memahami sesuatu yang tertuang dalam tes dan skor tes merupakan unsur dalam memberikan nilai.

## Skoring Tes Objektif

Pemberian skor tes obyektif pada umumnya menggunakan 1 untuk jawaban yang benar dan 0 untuk jawaban yang salah (dikotomi). Skor siswa ditentukan oleh banyaknya jawaban benar pada satu ujian. Skoring tes obyektif perlu adanya hukuman bagi siswa yang menjawab salah, "correction for guessing".

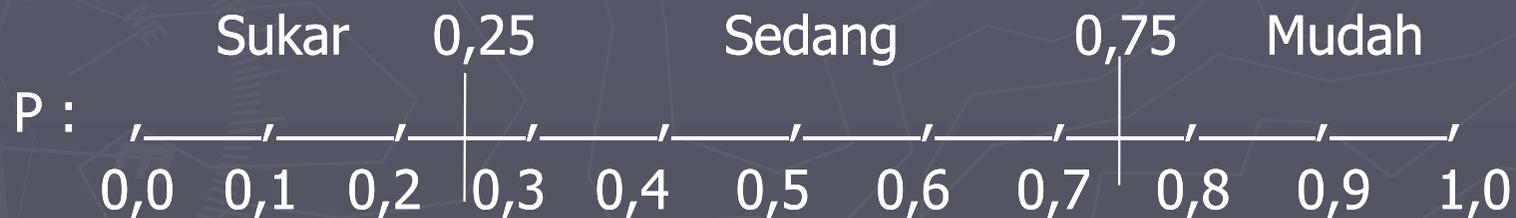
## Skoring Tes Uraian

skoring tes uraian pada umumnya menggunakan rentangan skor, 1, 2, 3, 4 dst. (politomi)

# Analisis Butir (Tingkat Kesukaran, Daya Beda, dan Distraktor)

## 1. Tingkat Kesukaran

Tingkat kesukaran (difficulty index) adalah derajat kesukaran atau taraf kesukaran butir soal dalam suatu tes dan umumnya dinyatakan dengan  $p$  (porporasi). Taraf kesukaran butir soal merupakan rasio antara peserta tes yang benar dengan jumlah butir soal. Secara teoritis bahwa  $p$  adalah probabilitas empiris untuk lulus butir soal tertentu bagi kelompok peserta tes. Pembagian derajat kesukaran menurut Witherington sbb;



## 2. Daya Beda

Daya beda (D) butir soal adalah kemampuan suatu butir tes hasil belajar untuk dapat membedakan (diskriminasi) antara peserta tes yang berkemampuan tinggi dengan berkemampuan rendah.

Mengetahui daya beda sangat penting bagi penyusunan butir, karena ada anggapan bahwa setiap individu memiliki perbedaan dalam hasil belajar, sehingga butir soal yang dibuat juga harus dapat memberikan gambaran adanya perbedaan. Daya beda dapat diketahui melalui besar kecilnya angka indek diskriminasi dari setiap butir soal.

Daya beda dapat diketahui melalui besar kecilnya angka indeks diskriminasi dari setiap butir soal.

Interpretasi daya beda (D) menurut Ebel (1965) sebagai berikut;

- ▶  $0,70 \leq D \leq 1,0$  memiliki daya beda baik sekali
- ▶  $0,40 \leq D \leq 0,69$  memiliki daya beda cukup baik
- ▶  $0,30 \leq D \leq 0,39$  memerlukan revisi sedikit atau tidak
- ▶  $0,20 \leq D \leq 0,29$  memerlukan revisi atau disisihkan
- ▶  $D \leq 0,19$  disisihkan atau direvisi total

### 3. Analisis distraktor

Distraktor berjalan dengan baik jika dapat merangsang peserta tes untuk memilihnya atau telah dipilih oleh peserta yang berkemampuan rendah minimal 5 % dan bagi peserta yang perkemampuan tinggi tidak akan terjebak.

Nomor	Alternatif					Keterangan
	A	B	C	D	E	
1	4	6	5	30*	5	* kunci
2	1	44*	2	1	2	

Butir soal nomor 1 kunci jawaban D

Pengecoh A dipilih 4 peserta dengan demikian = 8%

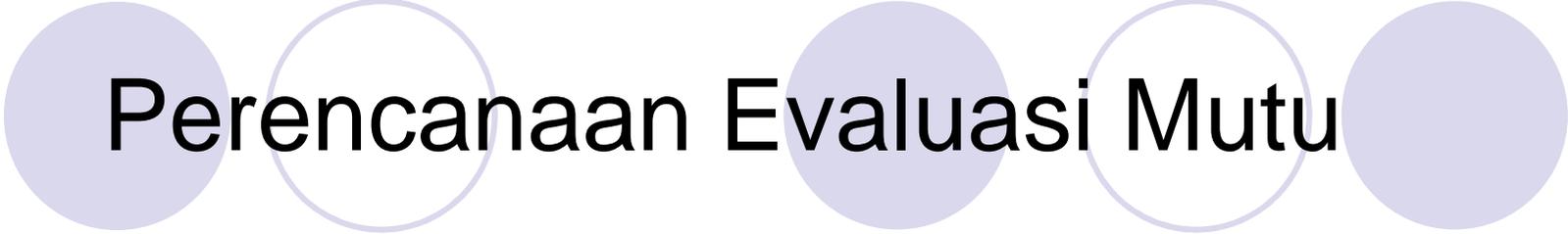
Pengecoh B dipilih 6 peserta dengan demikian = 12%

Pengecoh A dipilih 5 peserta dengan demikian = 10%

Pengecoh E dipilih 5 peserta dengan demikian = 10%

► Analisis Butir dengan Program komputer  
banyak program analisis butir untuk teori klasik maupun teori modern yang dapat digunakan antara lain;

1. Bigstep
2. Iteman
3. Ascal
4. NCSS
5. Param
6. Anates
7. Rashcal
8. dll.



# Perencanaan Evaluasi Mutu

## Analisis Kebutuhan

Kegiatan analisis kebutuhan akan menghasilkan kesenjangan antara kondisi nyata dengan kondisi yang diharapkan. Sasaran analisis kebutuhan; siswa, kelas, sekolah dengan tujuan utama menentukan dan mendekatkan jarak kesenjangan.



## Menyusun proposal evaluasi

- ❖ Bagian Pendauluan terdiri dari;
- ❖ Latar belakang masalah
- ❖ Identifikasi masalah dan pembatasan masalah
- ❖ Pertanyaan masalah
- ❖ Tujuan umum dan khusus
- ❖ Manfaat hasil evaluasi



- **Menyusun Instrumen Evaluasi**

- Mengidentifikasi komponen mutu dan indikatornya
- Membuat kisi-kisi, dari indikator, sumber data metode pengumpulan data
- Menyusun butir-butir instrumen
- Menyusun kriteria penilaian
- Menyusun pedoman pengerjaan

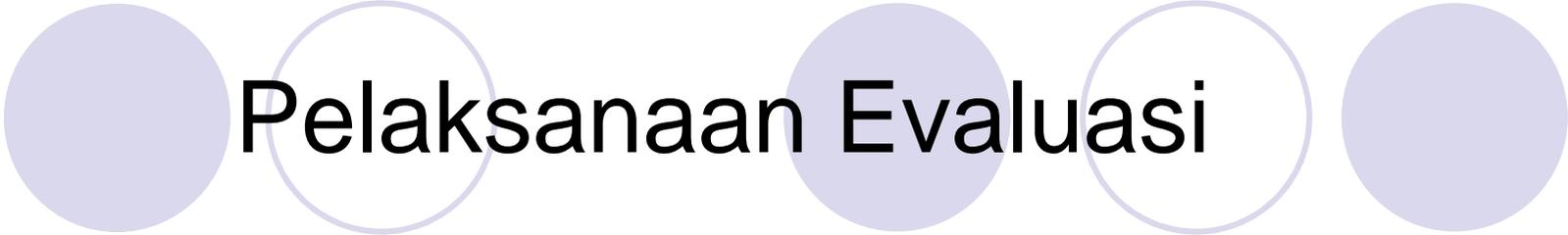
- **Syarat instrumen**

valid

reliabel

praktis

ekonomis



# Pelaksanaan Evaluasi

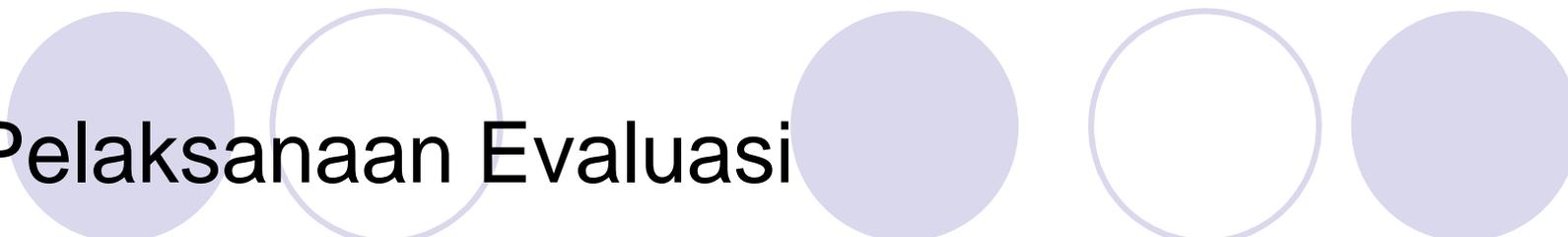
- **Persiapan Evaluasi**

- penyusunan evaluasi

- validasi instrumen

- menentukan jumlah sampel

- penyamaan persepsi antar evaluator sebelum mengambil data



- **Pelaksanaan Evaluasi**

pengambilan data dengan tes

pengambilan data dengan observasi

pengambilan data dengan angket

pengambilan data dengan wawancara

pengambilan data dengan dokumentasi



# ● Monitoring Pelaksanaan Evaluasi

- Fungsi pemantauan
- Sasaran pemantauan
- Teknik dan alat pemantauan
- Perencanaan pemantauan
- Pemanfaatan hasil pemantauan



## Analisis Data Evaluasi

- **Tabulasi data**
  - data narasi berpotensi tabulasi
  - data narasi non tabulasi
- **Pengolahan data secara manual**
  - Analisis data kuantitatif dengan statistika deskriptif dan statistika inferensial.
- Analisis data kualitatif
- **Analisis dengan komputer**

# Menyusun Laporan Evaluasi

- Permasalahan Evaluasi
- Metodologi Evaluasi
- Hasil Hasil Evaluasi
- Kesimpulan Hasil Evaluasi

## Sistematika

- Bab I Pendahuluan (latar belakang, Rumusan masalah, tujuan evaluasi, manfaat, dan batasan konsep/variabel evaluasi)
- Bab II Pembahasan Kepustakaan
- Bab III Metodologi Evaluasi
- Bab IV Hasil-hasil Evaluasi
- Bab V Pembahasan Hasil Evaluasi
- Bab VI Kesimpulan dan Rekomendasi

Biasanya ada ringkasan eksekutif (executive summary)

The slide features five decorative circles in a light purple color. Two are solid, and three are hollow with a thin purple outline. They are arranged horizontally across the top of the slide, with the title 'Perlu diskusi kelompok' centered between the second and fourth circles.

# Perlu diskusi kelompok

- Membuat rancangan evaluasi mutu di sekolah  
Dengan sistematika sebagai berikut;

- PENDAHULUAN

1. Latar belakang masalah,
2. Rumusan masalah,
3. Tujuan dan manfaat evaluasi
4. Batasan konsep/variabel evaluasi

- a. Definisi Program

- b. Batasan Program /aspek yang akan dievaluasi

- METODOLOGI/PENDEKATAN DAN DISAIN EVALUASI

- RENCANA KERJA

- JADWAL PELAKSANAAN

- ANGGARAN

- DAFTAR PUSTAKA