

## EVALUASI PEMBELAJARAN ABK KD 501, S-1, 2 SKS, SM 7

Mata kuliah Ini termasuk MKKP (kelompok mata kuliah keahlian profesi), pertemuan 16 kali. Tujuannya setelah selesai mengikuti perkuliahan ini mahasiswa diharapkan memiliki kemampuan konseptual maupun aplikatif dalam memahami konsep-konsep dasar pengukuran dan penilaian pendidikan, mengkonstruksi tes hasil belajar dan mampu menganalisisnya.

Komponen Evaluasi: UAS (35%), UTS (35%), Tugas (20%), dan Keaktifan (10%), Kehadiran (80%, Prasyarat UAS).

Dosen/Asisten: Budi Susetyo, M. Pd. (0918)  
Djadja Rahardja, M. Pd. ( )  
Zulkifli Sidiq, M. Pd. ( )  
**Iding Tarsidi, M. Pd. (1723 )**  
Dra. Oom Siti Homdijah ( )

## POKOK BAHASAN DAN MATERI PERKULIAHAN

- Pengukuran, Penilaian, dan Evaluasi Pendidikan (Konsep dasar, fungsi, jenis-jenis penilaian, dan alat penilaian pendidikan).
- Taxonomi Bloom dalam Penilaian Hasil Belajar (Kognitif, Afektif, dan Psikomotor).
- Pendekatan Interpretasi Hasil Belajar (PAN, PAP).
- Penilaian Kuantitatif dan Kualitatif, Tes Kinerja
- Tes dan Konstruksi Tes Hasil Belajar (Konsep, Jenis tes dan nontes, konstruksi tes obyektif dan esay).
- Persyaratan tes yang baik: Validitas, Reliabelitas, obyektivitas, Praktikabilitas. Validitas Tes Hasil Belajar (isi, kriteria, dan konstruksi).
- Reliabelitas (konsistensi internal, stabilitas, dan ekuivalensi).
- Analisis Butir Soal (tingkat kesukaran, daya pembeda, distraktor).
- Konversi Skor Tes Hasil Belajar/Skala Penilaian (skala huruf/angka, 1-10, Zskor, Tskor).
- Aplikasi komputer (Iteman, Anates, SPSS)
- Portofolio (pengertian, fungsi dan tujuan, karakteristik, tahapan, pelaporan)

## Buku Sumber

- Aiken, Lewis R. (1997). *Psychological Testing and Assesment*, Ninth Edition. Boston: Allyn and Bacon.
- Arikunto, Suharsimi (2002). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bina Aksara.
- Azwar, Saifuddin. (1997). *Reliabelitas dan Validitas*, Edisi ke-3. Yogyakarta: Pustaka Pelajar.
- Bloom, Benjamin S. (1956). *Taxonomy of Education Objectives*. New York: David Mc Kay Company, Inc.
- Cronbach, Lee J. (1970). *Essentials of Psychological Testing*. Third Edition. New York: Harper & Row Publisher.
- Gronlund, E. Norman. (1982). *Constructing Achievement Test*. Third Edition. Englewood Cliffs, New Jersey: Prentice Hall.
- Gronlund, Norman E. (1990). *Measurement and Evaluation in Teaching*. New York: Macmillan Publishing Company.

## KONSEP PENGUKURAN DAN PENILAIAN PENDIDIKAN

- Pengukuran (Measurement), ... as the assignment of numeral to objects or events according to rule (Steven, dalam Brown, 1970: 4).
- Pengukuran/mengukur adalah proses menentukan luas atau kuantitas (besar, banyak, jumlah) sesuatu objek dengan standar/skala/ukuran tertentu (meteran, timbangan) – jawaban dari How. (Wandt & Brown, 1977).
- Pengukuran adalah membandingkan sesuatu dengan sesuatu atas dasar ukuran tertentu – kuantitatif. Jenis-jenis pengukuran: (1) bukan untuk menguji, (2) untuk menguji, misal daya tahan lampu, (3) untuk menilai, dengan cara menguji (PBM).
- Penilaian adalah menilai sesuatu: yakni mengambil suatu keputusan terhadap sesuatu berdasarkan ukuran baik buruk, pandai bodoh. Bersifat kualitatif.
- Evaluation (measurement + value judgement). ... as systematic process of determining the extent to which educational objectives are achieved by pupils (Gronlund, 1976).
- Evaluation of educational ... is the estimation of the growth and progress of pupils toward objectives or values in the curriculum (Wrightstone, et. Al, 1956: 16).

## TUJUAN DAN FUNGSI EVALUASI PEMBELAJARAN

**Tujuan Umum:** (1). Memperoleh data sebagai petunjuk tingkat kemampuan dan keberhasilan siswa dalam pencapaian tujuan kurikuler setelah mengikuti PBM dalam kurun waktu tertentu, (2). Mengetahui efektivitas metode pengajaran, sarana, prasarana yang digunakan dalam PBM dalam kurun waktu tertentu.

**Tujuan Khusus:** (1). Memotivasi siswa dalam mengikuti program pembelajaran/ pendidikan-untuk memperbaiki diri dan meningkatkan prestasi belajarnya, (2). Menemukan faktor penyebab keberhasilan dan kegagalan siswa dlm mengikuti PBM.

**Fungsi Evaluasi Pembelajaran:**

- Memperoleh informasi hasil pembelajaran yang telah dicapai dalam rangka pelaksanaan program pembelajaran/pendidikan pada kurun waktu tertentu.
- Diketuainya relevansi antara program pembelajaran/pendidikan yang telah dirumuskan dengan tujuan yang hendak dicapai.
- Sebagai usaha perbaikan, penyempurnaan program pendidikan, sehingga lebih berdaya guna dan diperoleh hasil pembelajaran lebih baik.

Dengan demikian fungsi evaluasi pembelajaran, antara lain:

- Diagnostik – TPB hasilnya untuk deteksi kesulitan/kelemahan dalam belajar siswa
- Penempatan – TPB untuk klasifikasi dalam bidang/jurusan tertentu
- Formatif – TPB untuk melihat kemajuan belajar siswa & feedback kemajuan belajar
- Sumatif – untuk mengetahui penguasaan pelajaran, hasil akhir suatu PBM (UAS)
- Bimbingan – khususnya untuk siswa yang hasil belajarnya rendah, misalnya tentang cara belajar yang baik, cara membaca dan mendalami buku, dsb.

## TAXONOMI BLOOM DALAM PENDIDIKAN

- Menurut Scriven (1967, dalam Arikunto, 1987: 106), terdapat hubungan erat antara tujuan (yang dapat diukur) kurikulum dengan bahan pelajaran, bahan pelajaran dengan alat-alat evaluasi, dan tujuan kurikulum dengan alat-alat evaluasi.
- Menurut Ebel (1963, dalam Arikunto, 1987: 107), bahwa jika tujuan dirumuskan secara operasional, hasilnya dapat diukur. Tanda seseorang telah mencapai tujuannya terlihat dari perubahan tingkah laku.
- Tujuan pendidikan (di sekolah) dapat dirumuskan pada tiga tingkatan:
  1. Tujuan Instruksional Umum (TIU/TPU)
  2. Tujuan berbasis tingkah laku ~ taksonomi
  3. Tujuan yang jelas, operasional.
- Prinsip-prinsip yang digunakan Bloom dan Krathwohl dalam menyusun (taksonomi) menjadi suatu tingkatan menuju ke yang lebih sulit: metodologis, psikologis, logis, dan tujuan.
- Pada mulanya taksonomi Bloom terdiri dari dua bagian: cognitive & affektive domain (Bloom, 1956, dalam Arikunto, 109). Akhirnya Simpson melengkapinya dengan domain psikomotor (1966).
- Pemisahan domain hanya dibuat-buat, karena pada dasarnya manusia merupakan suatu kebulatan yang tidak dapat dipecah-pecah sehingga segala tindakannya juga merupakan suatu kebulatan.

## TAKSONOMI (BLOOM) DALAM PENDIDIKAN

- Pemisahan domain hanya rekayasa, karena pada dasarnya manusia merupakan suatu kebulatan yang tidak dapat dipecah-pecah sehingga segala tindakannya juga merupakan suatu kebulatan.
- Bloom, et.al., merumuskan tujuan pendidikan pada tiga tingkatan:
  1. Kategori tingkah laku yang masih verbal
  2. Perluasan kategori menjadi sederetan tujuan
  3. Tingkah laku konkrit yang terdiri dari tugas-tugas dalam pertanyaan sebagai ujian dan butir-butir soal.
- Taksonomi Bloom, et.al., meliputi domain/ranah: kognitif, affective, dan psychomotor

**Taxonomi**, skema hirarkhis untuk mengklasifikasikan perilaku. Tujuan pendidikan dapat dimasukkan pada salah satu atau lebih domain berikut:

1. **Domain Kognitif:** Intelektual/kemampuan berpikir, meliputi:
  - a. **Pengetahuan (Knowledge, recall):** proses psikologis dalam mengingat terhadap materi yang sudah dipelajari, meliputi:  
hal-hal yang khusus (simbol), istilah, fakta-fakta khusus (tanggal, peristiwa, tempat, nama orang), cara dan alat untuk melakukan hal khusus (urutan kronologis, menimbang), kecenderungan dan urutan (proses, arah, fenomena), klasifikasi/kategori, standard, metodologi, prinsip, teori.  
Kata Kerja Operasionalnya, antara lain: Mendefinisikan, menunjukkan, memberi nama, menyebutkan, menuliskan, memasang, memilih.

## TAXONOMI BLOOM DALAM TUJUAN PENDIDIKAN (KOGNITIF)

### **Persyaratan TIK/TPK:**

**Audience (obyek): siswa**

**Behavior (tingkah laku), dalam bentuk kata kerja operasional**

**Condition: bagaimana tingkah laku itu berlangsung**

**Degree (standard minimal/kriteria keberhasilan)**

### **b. Comprehension (Pemahaman)**

Mengertinya seseorang tentang apa yang sedang dikomunikasikan dan dapat diserap materi yang sudah dipelajarinya. Misalnya, Menterjemahkan materi dari satu ke bentuk ke bentuk lain (kata ke angka). Menginterpretasikan materi (menjelaskan, meringkaskan), mengklasifikasikan, meramalkan arah/akibat sesuatu, membedakan.

**KKOnya, antara lain: Membedakan, menjelaskan, mengubah, menguraikan, meringkaskan, menghitung luas.**

### **c. Application (Penerapan): kemampuan menggunakan apa yang telah dipelajari dalam situasi konkrit yang baru. Misal: peraturan, konsep, metode, prinsip, hukum, teori.**

**KKOnya: menghitung, mendemonstrasikan, mempergunakan, menunjukkan, menghubungkan, memecahkan (persoalan).**

### **d. Analysis, kemampuan menguraikan suatu materi ke dalam bagian-bagian, sehingga struktur organisasinya dapat dipahami. Unsur, hubungan, dan prinsip-prinsip yang terorganisir.**

## TAXONOMI BLOOM (SYNTHESIS & EVALUATION)

**KKOnya:** Memerinci/menguraikan, membedakan, mengidentifikasi, memisahkan, memilih, membuat diagram.

- e. **Synthesis**, kemampuan untuk menggabungkan bagian-bagian, membentuk keseluruhan yang baru. Hasilnya dapat berupa: tulisan, rencana/mechanisme, hubungan abstraksi.

**KKOnya:** Menyusun, mengkombinasikan, memodifikasi, mengarang, menciptakan, mengorganisir, merencanakan, merekonstruksi, merevisi, menceritakan, menyimpulkan.

- f. **Evaluation**, kemampuan untuk mempetimbangkan/menilai suatu materi (pernyataan, laporan, puisi, dsb) berdasarkan kriteria tertentu.

**KKOnya:** Menduga, mempertimbangkan, menilai, mengkritik, menyimpulkan, membenarkan, menafsirkan, membandingkan.

### 2. **Domain Afektif**

Segala sesuatu yang berkaitan dengan aspek sikap, minat dan nilai (Krathwohl, 1974). Tingkat kompetensinya, meliputi:

**Menerima/memperhatikan**, adalah kepekaan/kesadaran seseorang dalam menerima rangsangan dari luar dalam bentuk masalah, situasi, gejala, dll., atau kemauan untuk memperhatikan suatu kegiatan obyek.

**Contoh**, siswa menyadari bahwa disiplin harus ditegakkan

## AFFECTIVE DOMAIN

- **Menanggapi**, adalah kemauan yang dimiliki seseorang untuk tunduk, ikut secara aktif dalam suatu kondisi tertentu, berbuat sukarela/kehendak sendiri, dan memberikan reaksi, merasa puas/menyenangkan.  
Misal: siswa tumbuh minat untuk mempelajari materi lebih mendalam lagi
- **Menilai/Menghargai**, adalah memberikan nilai terhadap suatu kegiatan atau obyek, menerima nilai-nilai, memegang teguh nilai/mengabdikan diri kepada keyakinannya, jika ada kegiatan yang tidak dikerjakan dirasakan akan membawa kerugian atau penyesalan.  
Contoh: tumbuhnya kemauan yang kuat dari siswa untuk bertindak disiplin baik di sekolah di rumah maupun dalam kehidupan masyarakat.
- **Mengatur/mengorganisasi**, adalah konseptualisasi nilai, mempertemukan perbedaan nilai sehingga terbentuk nilai baru yang lebih universal, yang membawa ke perbaikan umum, mengorganisasikan sistem nilai secara teratur, harmonis, dan konsisten. Misal: siswa mendukung penegakkan disiplin nasional sebagaimana yang telah dicanangkan oleh pemerintah.
- **Characterization by a value or value complex**, adalah keterpaduan semua sistem nilai yang telah dimiliki seseorang, yang mempengaruhi pola kepribadian dan tingkah lakunya secara permanen (philosophy of life).

### 3. PSYCHOMOTOR DOMAIN

**Adalah segala aspek motorik yang berhubungan dengan keterampilan atau kemampuan bertindak setelah siswa menerima pengalaman belajar tertentu.**

**Menurut Harrow, ranah psikomotor meliputi: 1) Gerakan refleks, 2) Gerakan fundamental/dasar, 3) Kemampuan perseptual VAKT, 4) Kemampuan fisik, 5) Gerakan terampil, dan 6) Komunikasi nondiskursif (mimik, tarian).**

**Dari 6 jenjang kemampuan tersebut, terdapat 3 kelompok utama kemampuan dalam ranah psikomotor, yaitu:**

- 1. Keterampilan motorik: memperlihatkan gerak, menunjukkan hasil, menampilkan, melompat, dsb.**
- 2. Manipulatif benda-benda: menyusun, membentuk, memindahkan, menggeser, mereparasi, dsb.**
- 3. Koordinasi neuromuscular: memasang, memadukan, menghubungkan, mengamati, memotong.**
- KKO ranah psikomotor merupakan aktualisasi kata-kata (dalam tingkah laku) yang dapat diamati (diukur), misalnya: mempertontonkan, melompat, menampilkan, membentuk, dan memasang.**

## ACUAN PENILAIAN (PAP DAN PAN)

- **PAP (Criterion Refferenced Evaluation)**, yaitu penilaian berdasarkan standard absolut atau kriteria kompetensi yang telah dtetapkan terlebih dahulu. Caranya membandingkan skor mentah hasil tes dari masing-masing siswa dengan skor maksimal ideal (SMi) yang harus dikuasai siswa.
- **Nilai = skor mentah / skor maksimal ideal x 100 (cocok untuk tes formatif)**

### Contoh

Nama	Skor Mentah	Nilai	Keterangan:
A	60	$60 / 120 \times 100 = 50$	SMi = 120
B	40	$40 / 120 \times 100 = 33$	
C	80	$80 / 120 \times 100 = 67$	

**Kriteria: Nilai 80 ke atas = A; 66 – 79 = B; 56 – 65 = C; 46 – 55 = D; 45 ke bawah = E**

- **PAN (Norm Refferenced Evaluation)**, yaitu penilaian berdasarkan standard relatif atau kompetensi yang tidak ditetapkan lebih dahulu, tetapi ditetapkan kemudian berdasarkan penguasaan kelompok (norma). Caranya; membandingkan skor mentah hasil tes dari masing-masing siswa dengan skor rata-rata kelompok (cocok untuk tes sumatif). Dalam hal ini perlu dicari rerata dan simpangan baku.
- **$Xbar = \Sigma X/N$  atau  $\Sigma fX/N$  ;  $S = \sqrt{\Sigma X^2 / N - (\Sigma X)^2 / N}$  atau  $S = \sqrt{\Sigma fX^2 / N - \Sigma fX^2 / N}$**

## JENIS EVALUASI PEMBELAJARAN (TES & NONTES)

**Test is a systematic procedure for observing a person's behavior and describing it with the aid of a numerical scale or a category system (Cronbach, L.J. 70).**

**Tes:**

- **Kognitif, mencakup: 1). Abilitas potensial (potensial umum/intelegensi) dan potensial khusus/bakat, dan 2). Abilitas aktual/prestasi.**
- **Non kognitif (aspek afektif dan kepribadian).**

**Secara garis besar tes mengukur:**

- **Performansi maksimum (Abilitas aktual/prestasi), untuk mengungkap seluruh kemampuan siswa dan seberapa baik mereka dapat melakukannya. Stimulus yang disajikan harus jelas struktur dan tujuannya, sehingga siswa tahu arah jawaban yang diinginkan.**
- **Performansi tipikal (Abilitas potensial), untuk mengungkap kecenderungan reaksi atau perilaku individu ketika berada dalam situasi tertentu. Maka tujuan ukurnya adalah bukan kemampuan siswa, melainkan mengukur kecenderungan yang dapat dilakukan siswa.**

**Achievement Tes; yang disusun secara terencana untuk mengungkap performansi maksimum (abilitas aktual) siswa dalam menguasai materi yang sudah diajarkan.**

1. **Standardized Test.**
2. **Teacher-made-test, mencakup: 1). Essay type test dan 2). Objective type test, mencakup: a). Completion type test (completion, fill in); b). Selection type test (true false, multiple choice, matching); dan c. Tes praktek (Skill test).**

## NON TES

Penilaian terhadap hasil belajar yang dilakukan dengan tidak menguji siswa.  
Jenis-jenisnya:

- a. Observasi, mengumpulkan data dengan mengamati dan mencatat secara sistematis fenomena sasaran. Misal untuk menilai tingkah laku individu, proses terjadinya sesuatu kegiatan buatan maupun sebenarnya.
  - b. Wawancara, mengumpulkan data dengan tanya jawab lisan dan tatap muka, dengan arah dan tujuan yang telah ditentukan.
  - c. Angket, cara mengumpulkan data secara tertulis, bukan untuk mengukur kemampuan (kognitif), melainkan afektif. Bentuknya (pilihan ganda, skala).
  - d. Check list, terdiri dari daftar item dengan cara menandai (v) pada kolom (ya) atau (tidak).
  - e. Rating scale, menggambarkan suatu nilai terhadap sesuatu hasil pertimbangan secara bertingkat, misal, menggambarkan prestasi atau kecenderungan sesuatu.
- Tahapan observasi, meliputi: pedoman observasi tentang aspe-aspek yang akan diamati sesuai tujuan, menentukan jenis perilaku yang diamati, memahami prosedur, interpretasi, pencatatan, kuantifikasi dan cara melaporkan.

## PERBEDAAN TES OBYEKTIF DAN ESSAY

### TES OBYEKTIF

Baik untuk mengukur aspek pengetahuan, pemahaman, aplikasi dan analisis.

Jumlah item cukup banyak, merepresentasikan bahan ajar

Persiapan menulis soal sukar dan memakan waktu

Penskorannya obyektif, sederhana, reliabilitas tinggi.

Mendorong siswa untuk mengingat, menginterpretasi, atau analisis ide orang lain.

### TES ESSAY

Sangat baik untuk mengukur aspek sintesis dan evaluasi.

Jumlah item sedikit, mencakup bahan yang terbatas.

Persiapannya relatif lebih mudah, meskipun untuk item yang lebih baik adalah sukar.

Penskorannya subyektif, sukar dan kurang reliabel.

Mendorong siswa organisasikan, dan mengintegrasikan ide-ide orang lain.

## PERSYARATAN TES YANG BAIK

### ■ Validitas (isi, konstruksi, kriteria)

Ketepatan dan ketelitian suatu alat pengukur mengukur gejala yang hendak diukur (timbangan untuk mengukur berat, literan untuk volume, meteran untuk mengukur panjang). ~ Validitas isi.

### ■ Reliabelitas (stabilitas, konsist: paralel test-retest, split-half method)

Keajegan/ketetapan/kemantapan. Dapat mengukur dengan hasil yang ajeg mengenai gejala yang diukur. Menekankan ketetapan hasil (tes)

### ■ Objektivitas, tidak ada unsur subyektif yang mempengaruhi. Menekankan ketetapan sistem skoring. Objektivitas suatu tes dapat dipengaruhi oleh: a. Bentuk tes (objektif atau esay), b. Penilai (subyektif: ada kesan penilai terhadap siswa, tulisan, bahasa, waktu mengadakan tes, kelelahan, dsb.)

### ■ Praktikabilitas (praktis, mudah dilaksanakan, mudah pemeriksaannya, dilengkapi dengan petunjuk-petunjuk yang jelas).

**KONSTRUKSI TES** Merupakan serangkaian pertanyaan, latihan yang dibangun/disusun untuk mengukur perubahan tingkah laku

### ■ Prosedur standard dalam perencanaan dan penyusunan THB:

Identifikasi tujuan dan kawasan ukur, Batasan perilaku dan kompetensi, Spesifikasi tes/Kisi-kisi, Menulis item-Mereview, Uji coba awal (Uji validitas dan reliabelitas), Analisis item, Merakit dan menyusun instruksi tes, Uji coba 2, Tes final (siap pakai).

### ■ Keunggulan dan Keterbatasan bentuk tes dapat ditinjau dari: (1). Hasil belajar yang diukur, (2). Representasi isi, (3). Persiapan soal, (4). Penyekoran, dan (5). Pengaruhnya pada pengetahuan/belajar

## KONSTRUKSI TES

- Keunggulan Tes bentuk B-S: relatif mudah dalam membuat dan penyekorannya serta representatif. Kelemahannya, spekulatif, membatasi penjawab, dan lebih menekankan aspek ingatan.
- Keunggulan dan Keterbatasan tes PG: relatif dapat mengukur semua jenjang kognitif, variatif dalam bentuk/jawabannya. Keterbatasannya, sukar membuat distraktor, perlu waktu menyusun.
- Hal-hal yang perlu diperhatikan dalam penulisan item tes: tingkat kesukaran setiap item sesuai dengan tujuan dan fungsi tes serta tingkat pendidikan siswa. Penulis item tes hendaknya menempatkan diri sebagai testee, sehingga dapat menghasilkan item dengan TK sedang. Kreativitas dan skill penulis item tes dapat mempengaruhi kualitas tes yang dibuatnya.
- Kaidah Penulisan Item Tes PG: mengukur hasil belajar penting, masalah tunggal atau terdiri dari satu pertanyaan atau pernyataan (stem), stem bersifat positif, hindari pengulangan materi dalam option, atau petunjuk yang mengarahkan ke jawaban yang betul, terdapat beberapa option, kunci dan distraktor, posisi jawaban benar sebaiknya acak. Contoh: ketepatan suatu alat ukur dalam mengukur gejala yang hendak diukur, merujuk kepada makna: a. Reliabelitas, b. Objektivitas, c. Validitas, d. praktikabilitas.

## KAIDAH PENULISAN ITEM TES

- **Kaidah Penulisan Item Tes B-S:** hanya mencakup satu ide, tegas B/S, bahasanya sederhana, singkat, jelas, jika mengemukakan pendapat maka harus disertai sumbernya. Contoh: Menurut Thorndike, agar bahan pelajaran tidak terlupakan, kita harus sering mengulangnya.
- **Item tes Jawaban Pendek/Singkat:** berupa kalimat pernyataan yang belum sempurna, atau yang harus dijawab secara singkat (1 atau 2 kata), dan hanya ada satu jawaban yang pasti. Contoh: TMI terletak di kota ..., tidak menggunakan kata/kalimat langsung dikutip dari buku. Contoh: Jika tiga orang siswa mempunyai 270 kelereng, maka rata-rata kelereng siswa adalah..
- **Kaidah Penulisan Item Tes Matching:** materi/premis dan respons homogen, premis/stimulus di kiri dan respon/option di kanan (dalam jumlah yang tidak sama, respon lebih banyak), keduanya dalam 1 halaman, berisi kata/kalimat pendek, ada petunjuk pemasangan yang jelas.
- **Keunggulan dan Keterbatasan Matching:** Keunggulannya: mengukur kemampuan asosiatif (istilah, peristiwa, definisi). Keterbatasannya, sukar menemukan pasangan homogen benar.
- **Kaidah Penulisan Item Tes Essay:** item meliputi ide-ide pokok dan komprehensif, rumusan soal jelas dan mudah dipahami, soal disusun dari yang mudah ke yang sukar, ada pedoman penilaiannya, pertanyaan/tugas menunjukkan penguasaan yang penting (menuntut penerapan atau analisis) siswa, pertanyaan dimulai dengan bandingkan, berikan alasan, jelaskan, mengapa; sebaiknya tidak menanyakan sikap/pendapat (karena ini bukan menguji benar-salah). Contoh: jelaskan keunggulan dan kelemahan tes bentuk objektif dan esai?
- **Banyaknya butir soal tes ditentukan oleh:** waktu yang tersedia dan bentuk soal tes yang digunakan.

**TABEL SPESIFIKASI KONSTRUKSI TES  
(KISI – KISI)**

**Aspek yang diukur**

<b>Pokok Materi</b>	<b>Ingatan (50%)</b>	<b>Pemahaman (25%)</b>	<b>Aplikasi (25%)</b>	<b>Jumlah (100%)</b>
<b>Pengertian Evaluasi (15%)</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>8</b>
<b>Fungsi Evaluasi (20%)</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>9</b>
<b>Jenis dan Bentuk Tes (35%)</b>	<b>G</b>	<b>H</b>	<b>I</b>	<b>18</b>
<b>Persyaratan Tes yang baik (30%)</b>	<b>J</b>	<b>K</b>	<b>L</b>	<b><u>15</u> 50</b>

## LANJUTAN KISI-KISI

- Tabel di atas adalah untuk materi yang homogen, yakni bahwa antara pokok materi yang satu dengan materi yang lain mempunyai kesamaan dalam imbangan aspek tingkah laku yang diukur, misalnya 50% ingatan, 25% pemahaman dan 25% aplikasi. Maka banyaknya butir soal untuk setiap sel (A s/d L) diperoleh dengan cara menghitung persentase dari banyaknya soal bagi tiap pokok materi yang sudah tertulis di kolom paling kanan.
- Sel A =  $50/100 \times 8 = 4$  Soal
- Sel B =  $25/100 \times 8 = 2$  Soal
- Sel C =  $25/100 \times 8 = 2$  Soal

Catatan: Kecuali dengan cara menentukan jumlah butir soal untuk tiap-tiap pokok materi, ada cara lain, yaitu memulai dari pengisian sel-sel kemudian baru diperoleh jumlah soal tiap pokok materi. Misalnya berdasarkan waktu yang telah ditentukan diperkirakan akan disusun 40 butir soal. Maka tiap sel diperoleh imbangan jumlah soal sbb:

## LANJUTAN KISI - KISI

Aspek yang diukur

	Ingatan (50%)	Pemahaman (30%)	Aplikasi (20%)	Jumlah (100%)
Pokok Materi				
Pokok Bahasan I (40%)	A	B	C	
Pokok Bahasan II (30%)	D	E	F	
Pokok Bahasan III (30%)	G	H	I	
J u m l a h (100%)				40

## LANJUTAN KISI – KISI

Cara lain, misalnya berdasarkan waktu yang telah ditentukan diperkirakan akan disusun 40 butir soal. Maka tiap sel diperoleh imbangan jumlah soal sbb:

- Sel A =  $50/100 \times 40/100 \times 40 = 8$  Soal
- Sel B =  $30/100 \times 40/100 \times 40 = 4,8$  (5 soal)
- Sel C =  $20\% \times 40\% \times 40 = 3,2$  (3 soal)
- Sel D =  $50/100 \times 30/100 \times 40 = 6$  soal

Selanjutnya idem

- Untuk pokok-pokok materi yang tidak homogen, tidak perlu mencantumkan angka persentase imbangan aspek tingkah laku yang diukur di kepala kolom. Pemberian imbangan dilakukan tiap pokok materi didasarkan atas banyaknya soal untuk pokok materi itu dan imbangan yang dikehendaki oleh penilai menurut sifat pokok materi yang bersangkutan.

## TABEL SPESIFIKASI MATERI YANG TIDAK HOMOGEN

Aspek yang diukur	Aspek yang diukur			Jumlah
	Ingatan ( I )	Pemahaman ( P )	Aplikasi ( A )	
<b>Pokok Materi</b>				
<b>Pokok Bahasan I (25%)</b>	<b>A</b>	<b>B</b>	<b>C</b>	<b>10</b>
<b>Pokok Bahasan II (40%)</b>	<b>D</b>	<b>E</b>	<b>F</b>	<b>16</b>
<b>Pokok Bahasan III (35%)</b>	<b>G</b>	<b>H</b>	<b>I</b>	<b>14</b>
<b>J u m l a h (100%)</b>				<b>40</b>

Misal: untuk PB I aspek (I) = 60%, (P) = 30%, dan aspek (A) = 10%, maka:  
sel A =  $60/100 \times 10 = 6$  soal; sel B =  $30/100 \times 10 = 3$  soal; dan sel C =  $10/100 \times 10 = 1$  soal. Untuk PB II dan PB III dalam mengisi sel-sel selanjutnya idem.

## TABEL SPESIFIKASI MATERI YANG TIDAK HOMOGEN (LANJUTAN)

- Misalnya, PB I mayoritas hafalan, PB II mayoritas pemahaman, dan PB III mayoritas aplikasi. Makaimbangan aspek tingkah laku tidak dapat dituliskan pada kepala kolom. Penentuan angka yang menunjukkan banyaknya butir soal pada tiap sel, ditentukan per Pokok Bahasan (PB).
- Misalnya, PB I, ingatan 50%, pemahaman 30% dan aplikasi 20%, maka:
  - Sel A =  $50/100 \times 10 = 5$  soal
  - Sel B =  $30/100 \times 10 = 3$  soal
  - Sel C =  $20/100 \times 10 = 2$  soal

**KISI-KISI SOAL ULANGAN UMUM CAWU: I, II, III, EBTA  
TAHUN .....**

MATA PELAJARAN: ..... KELAS: ..... NAMA PENYUSUN: .....

=====  
TPU ! PB/SPB/ ! BAHAN ! JML ! URAIAN ! INDIKATOR ! NO. !JG! TKS ! BNTK  
TEMA PB SOAL MATERI (TPK) SOAL SOAL  
=====

**KISI – KISI NASKAH EVALUASI HASIL BELAJAR/EBTA  
SLB ..... TAHUN ..... MATA PELAJARAN .....**

=====  
NOMOR ASPEK YANG DIUKUR  
INGATAN PEMAHAMAN APLIKASI JML ! KET  
NO.! PB ! SPB ! TKS ! BS! PG! IS ! MJD ! PG ! ES! PG ! MJD ! ES!  
=====

## MODEL SILABUS

**SATUAN PENDIDIKAN: SDLB TUNAGRAHITA SEDANG**  
**PROGRAM KHUSUS: BINA DIRI**  
**STANDARD KOMPETENSI: 1. MAMPU MERAWAT DIRI**  
**2. MAMPU MENGURUS DIRI**

NO	KOMPETENSI DASAR	MATERI POKOK	KEGIATAN PEMBELAJARAN	INDIKATOR PENILAIAN	JENIS TES	ALOKASI WAKTU	BAHAN/ALAT	SUMBER
1	MELAKUKAN MAKAN DAN MINUM	MAKAN MENGGUNAKAN SENDOK	MENGENAL ALAT-ALAT MAKAN  MELAKUKAN MAKAN	MENYIAPKAN SEBELUM MAKAN  MELAKUKAN MAKAN BENAR	LISAN, PERBUATAN, TULISAN	4 X 30 MENIT	SENDOK, PIRING, NASI	BUKU RUANGANN KHUS CERMIN

## PERSYARATAN TES HASIL BELAJAR YANG BAIK

### ■ Validitas Isi

Menunjukkan sejauhmana item dalam tes mencakup seluruh kawasan isi (materi dan tujuan secara komprehensif & relevan) yang hendak diukur oleh tes itu. Teknik analisisnya dengan statistik dan analisis rasional.

### ■ Validitas Konstruk

Menunjukkan sejauhmana alat ukur/tes dapat mengukur trait (konstruk teoretik) yang hendak diukur. Teknik analisisnya melalui analisis faktor.

### ■ Validitas Kriteria

Menunjukkan sejauhmana alat ukur tersebut dapat berhubungan dengan kriteria tertentu yang telah ditetapkan. (1). Validitas Prediksi: sejauhmana hasil dari alat ukur yang telah dibuat dapat digunakan untuk meramalkan kondisi yang berkaitan dengan perilaku belajar dimasa yang akan datang, analisisnya melalui korelasi; (2). Validitas Konkuren/serentak: menunjukkan sejauhmana alat ukur yang telah dibuat berhubungan dengan kriteria (alat ukur standard), dan pengesanan dilakukan secara bersamaan. Analisisnya melalui korelasi.

### ■ Reliabelitas

Jika tes tersebut senantiasa dapat memberikan hasil yang tetap apabila dilakukan pengukuran berkali-kali kepada subyek.

(1). Reliabelitas Stabilitas: Hasil tes yang diperoleh siswa menunjukkan kondisi tetap apabila dilakukan pengukuran pada waktu tertentu dan dilakukan pengukuran kembali pada waktu tertentu. Model tes – retest, melalui korelasi.

## VALIDITAS DAN RELIABELITAS TES

### ■ Reliabelitas Konsistensi Internal

■ Ini menggunakan satu tes/alat ukur pada subyek. Tes reliabel jika terdapat kaitan antara item yang terdapat dalam alat ukur tersebut. Analisisnya menggunakan korelasi K – R 20, dll.

### ■ Reliabelitas Ekuivalen

Ini menggunakan dua alat ukur/tes yang setara diberikan kepada subyek secara berturut-turut. Analisisnya menggunakan korelasi.

### ■ Obyektif

Jika tes tersebut disusun, dilaksanakan, dan diberi skor apa adanya. Penyusunan tes berdasarkan materi dan TIK, dilaksanakan secara baik tanpa rekayasa, dan dinilai oleh siapapun menunjukkan hasil yang sama.

### ■ Praktis

THB yang dibuat dapat dilaksanakan dengan mudah, tidak berbelit-belit, ekonomis, serta tidak memakan waktu, biaya dan tenaga yang banyak. (1). Sederhana; tidak memerlukan peralatan yang banyak atau tidak sulit didapat, (2). Lengkap, tes dilengkapi dengan petunjuk cara mengerjakannya, kunci jawaban, pedoman penilaian, dan penentuan nilainya.