

PENGUKURAN TES DAN PSIKOMETRI

Univesitas Bhayangkara Jakarta Raya



Pengukuran dan Tes

Narasumber: Dr. Netty Merdiaty. MM., M. Si
Tanggal

PERTEMUAN 2



PENGUKURAN

- Ilmu pengukuran berasal dari statistika terapan untuk pengembangan alat tes
- Diuji kelayakannya dengan ilmu psikometri

definisi

- Prosedur pemberian angka pd atribut/ variabel sepanjang suatu kontinum

Contoh

- Panjang meja 50cm kontinum
- Tinggi badan 170 cm kontinum
- IQ 120 → ?

KONTINUM

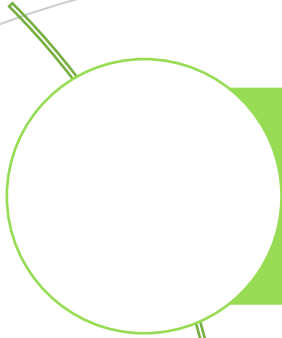
Kontinum fisik

- Diukur menggunakan skala fisik
- Contoh : berat, Panjang, dll

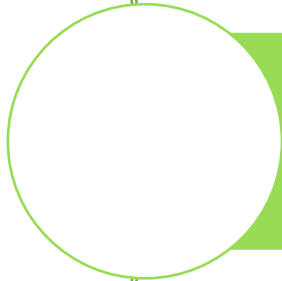
Kontinum Psikologis

- Diukur menggunakan skala psikologis
- Contoh : harga diri, kepribadian, dll

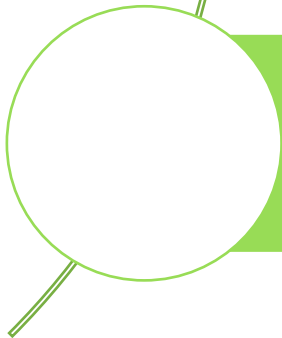
CIRI PENGUKURAN



Sebagai pembandingan antara atribut yang di ukur dengan alat ukurnya



Dinyatakan secara kuantitatif



Hasil bersifat deskriptif

Contoh ciri pengukuran

- Meja A memiliki Panjang 80 cm dan meja B memiliki Panjang 50 cm. Meja A lebih Panjang dibandingkan meja B
- Perbandingan : Panjang meja dan meteran
- Kuantitatif : 80 cm, dan 50 cm
- Hasil : meja A lebih Panjang dibandingkan meja B

SKALA PENGUKURAN

Skala Nominal

- Kategori, identifikasi
- Cth : nama, lokasi, dll

Skala Ordinal

- Jenjang kualitatif tidak menunjukkan perbedaan kuantitatif
- Cth : urutan buah yg enak, urutan buah yg wangi

SKALA INTERVAL

- Jenjang kuantitatif, 0 memiliki arti
- Cth : Angka, suhu, dll

SKALA RATIO

- Jenjang kualitatif dengan nilai 0 mutlak (tidak ada)
- Cth: berat, Panjang, dll

CONTOH

SKALA	CONTOH ATRIBUT	CONTOH ANGKA
Nominal	Kode jenis kelamin Kode lokasi No punggung	LK=1, Pr=2 Padang=1, Pariaman=2, BKT=3
Ordinal	Rangking kelas Status ekonomi	1,2,3 1=kaya, 2=menengah, 3=miskin
Interval	Angka derajat suhu tubuh Angka IQ	36 derajat 110
Ratio	Jarak tempuh Belanja perminggu penghasilan	56 km 500 ribu rupiah 300 ribu rupiah

CONTOH EVALUASI PENGUKURAN

1. Meja A lebih Panjang dibandingkan meja yang lain
2. Lampu B lebih terang
3. Nasi goreng lebih enak dibandingkan ayam goreng

Tes Psikologi

- ✓ Menyangkut pengukuran dan evaluasi
- ✓ Tes sebagai pengumpulan data dari sampel perilaku
- ✓ Atribut psikologi sebagai konstruk teoritis

Permasalahan Dalam Tes Psikologi

- Tidak ada pendekatan tunggal dalam pengukuran apapun yang dapat diterima secara universal
- Sampel perilaku yang jumlahnya terbatas misalnya. Mengukur kepribadian dengan 100 soal, apakah benar-benar bisa menggambarkan kepribadian.
- Selalu mungkin mengandung error
- Satuan dan skala pengukuran tidak dapat didefinisikan dengan baik.

PERFORMA

Performa maksimal

- Hasil terbaik yang identik dengan ranah kemampuan/kognitif

Performa tipikal

- Performa yang ditampakkan sebagai proyeksi kepribadian



THANK YOU
Universitas Bhayangkara Jakarta Raya

PENGUKURAN TES DAN PSIKOMETRI

Univesitas Bhayangkara Jakarta Raya



Model Penskalaan

Narasumber: Dr. Netty Merdiaty. MM., M. Si
Tanggal

PERTEMUAN 3



PENSKALAAN

pengertian

- prosedur untuk menempatkan karakteristik objek pada titik-titik sepanjang sebuah kontinum
- Penskalaan dalam psikologi adalah upaya untuk mengembangkan instrumen pengukuran terhadap penilaian individu

Tujuan

- Untuk akurasi estimasi penilaian individu yang bersifat subjektif

ELEMEN SKALA PSIKOLOGI

- ❑ SKALA → seperangkat nomor yang digunakan untuk menjelaskan konstruk psikologis
- ❑ INSTRUMENT → alat yg dipakai untuk menjalankan operasi pengukuran
- ❑ PROSEDUR → urutan Tindakan yang dilakukan Ketika dalam melakukan pengukuran

CONTOH PENGUKURAN

	Tinggi Badan	Kepribadian
Skala	Cm, Inchi	16pf, NEO
Instrumen	Penggaris, tongkat	Seperangkat pertanyaan
Prosedur	Berdiri didekat penggaris, berdiri di dekat tembok	Meminta subjek untuk mengisi pertanyaan

Kesalahan dalam penskalaan psikologi

Atribut psikologi sifatnya laten

Aitem psikologi ditulis berdasarkan indikator perilaku yang sifatnya terbatas.

- Respon yang diberikan oleh subjek terhadap stimulus dalam skala psikologi dipengaruhi variabel-variabel yg tidak relevan.
- Atribut psikologi yang terdapat dalam diri manusia stabilitasnya tidak tinggi.
- Interpretasi terhadap hasil ukur psikologi hanya dapat dilakukan secara normatif

Pendekatan utama pada penskalaan psikologi

- pendekatan dengan metode-metode yang berorientasi pada subjek,
- pendekatan dengan metode-metode yang berorientasi pada stimulus, dan
- pendekatan dengan metode-metode yang berorientasi pada respons

Atribut psikologi sebagai objek ukur

Atribut psikologi hanya diukur sampai dengan tingkat ordinal, karena tidak adanya titik nol yang sifatnya absolut.

Kategori atribut psikologi :

- Atribut kemampuan (Kognitif)
- kemampuan yang bersifat potensial
- Kemampuan potensial umum (inteligensi)
- Kemampuan potensial khusus (bakat)
- Kemampuan yang bersifat actual (prestasi)
- Atribut bukan kemampuan

Karakteristik skala psikologi

- Stimulus atau aitem dalam skala psikologi berupa pertanyaan dan pernyataan yang tidak langsung (indikator perilaku yang bersangkutan)
- Atribut psikologi diungkap lewat indikator perilaku, indikator perilaku diterjemahkan dalam bentuk aitem.
- Respon subjek tidak diklasifikasikan sebagai jawaban “benar” atau “salah”

Perbedaan skala dan angket

Dimensi	Skala	Angket
Indikator psikologi yang diungkap arah pernyataan	Aspek diri yang dipersepsi subjek tidak langsung	Data fatual yang diketahui subjek langsung
Kesadaran pada tujuan ukur	Tidak sadar	Sadar
Penilaian respon	Prosedur penskalaan	Klasifikasi
Jumlah konstruk diungkap	Satu konstruk	Banyak konstruk
Pengujian realibilitas	Perlu diuji	Tidak perlu diuji
Pengujian validitas	Kejelasan konsep	Kejelasan tujuan
Jenis data	interval	ordinal

Langkah – Langkah konstruksi penskalaan

a. Sebelum melakukan pengukuran

Identifikasi tujuan ukur

- Pembatasan kawasan (domain) ukur berdasarkan kontrak yang didefinisikan oleh teori yang bersangkutan.

b. Sebelum membuat aitem

- Mengetahui indikator-indikator perilaku dari variabel dengan jelas
- Bentuk atau format stimulus yang hendak digunakan

TIPE SKALA PENGUKURAN

SKALA
LIKERT

SKALA
GUTTMAN

SKALA
RATING

SKALA
DIFFERENSIAL

SKALA LIKERT

Skala ini digunakan untuk mengatur sikap, pendapat dan persepsi, seseorang atau kelompok orang tentang fenomena atau gejala sosial yang terjadi.

	Sangat tidak setuju	Tidak setuju	Netral	Setuju	Sangat setuju
Perusahaan memberika layanan yang berkualitas					
Perusahaan bertempat ditempat yang strategis					
Jam operasional perusahaan tidak menyusahkan					

Skala Guttman

skala pengukuran untuk memperoleh/menginginkan tipe jawaban responden yang tegas, seperti : jawaban benar-salah, ya-tidak, pernah-tidak pernah, positif-negative, tinggi-rendah, baik-buruk, dan seterusnya.

1. Yakin atau tidakkah anda, pergantian presiden akan dapat mengatasi persoalan bangsa?
 - a. yakin
 - b. tidak
2. Apakah komentar anda jika menajer anda turun dari jabatannya?
 - a. setuju
 - b. tidak setuju
3. Pernahkah direktur saudara mengajak makan Bersama?
 - a. pernah
 - b. tidak pernah

RATING SCALE

Suatu konstruk teoritik yang ada secara hipotetis dan dikonsepsikan guna mendeskripsikan dimensi psikologis yang ada dalam diri individu

SKALA PENILAIAN (RATING SCALE) Keterampilan Membuat Preparat Basah Serbuk Sari Bunga Salak

Berilah centang (✓) di bawah skor 5 bila Anda anggap cara melakukan aspek keterampilan sangat tepat, skor 4 bila tepat, 3 bila agak tepat, 2 bila tidak tepat, dan skor 1 bila sangat tidak tepat untuk setiap aspek keterampilan di bawah ini!

No.	Aspek Keterampilan	Skor				
		1	2	3	4	5
1.	Memilih alat yang digunakan untuk mengambil serbuk sari : kuas kecil					
2.	Serbuk sari diambil dengan sapuan lembut dalam jumlah yang cukup banyak					
3.	Serbuk sari diletakkan dengan sapuan lembut di tengah-tengah kaca benda dalam jumlah yang cukup					
4.	Medium berupa air ditetaskan sebanyak 1 atau 2 tetes pada serbuk sari dengan menggunakan pipet tanpa tumpah dan meluber					
5.	Spekimen pengamatan yang telah diberi medium air ditutup dengan kaca penutup sedemikian rupa sehingga tidak terdapat gelembung udara					
	Jumlah Skor					

SKALA DIFFERENSIAL

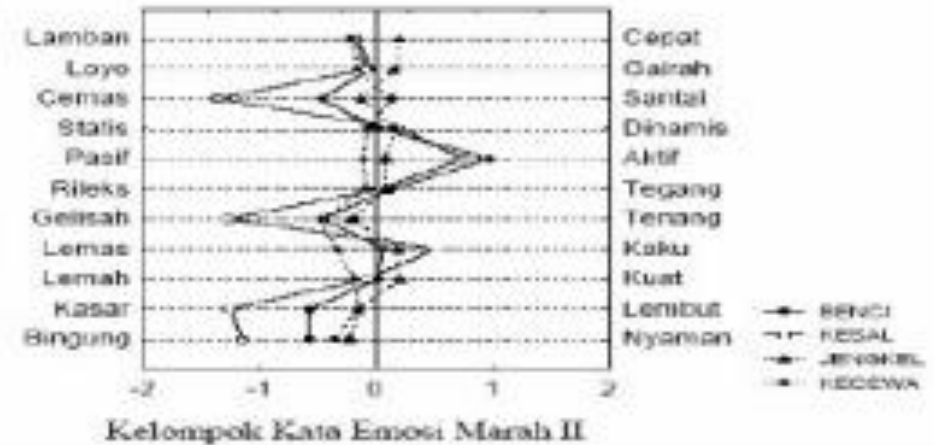
Skala differensial digunakan untuk mengatur sikap perbedaan simantik, responden untuk menjawab pernyataan dalam satu garis kontinum yang bertentangan yaitu positif negative. Data yang diperoleh biasanya data interval yang digunakan untuk mengukur sikap seseorang atau kelompok

Stimulus dapat berbentuk kata atau kalimat yang mengukur

- Perilaku
- Sikap
- Keyakinan
- Opini

Contoh Stimulus Semantik Diferensial

- Mengkritik Guru
- Keluarga Berencana
- Manajer Saya
- Televisi



SKALA SEMANTIC DIFFERENTIAL

Merupakan salah satu teknik self report untuk pengukuran sikap dimana subjek diminta memilih satu kata sifat atau frase dari sekelompok pasangan kata sifat ataupun pasangan frase yang disediakan yang paling mampu menggambarkan perasaan mereka terhadap suatu objek.

Contoh:

Pelayanan tidak berkualitas

:---:---:---:---:---:---: Pelayanan berkualitas

Lokasi tidak menyenangkan

:---:---:---:---:---:---: Lokasi menyenangkan

Jam kerja menyenangkan

:---:---:---:---:---:---: Jam kerja tidak menyenangkan

Suku bunga kredit tinggi

:---:---:---:---:---:---: Suku bunga kredit rendah

BOGARDUS SOCIAL SCALE

Suatu skala yang direncanakan untuk mengukur jarak social yang ingin dipertahankan oleh orang antara diri sendiri dengan para anggota suku lain atau kelompok etnis lain

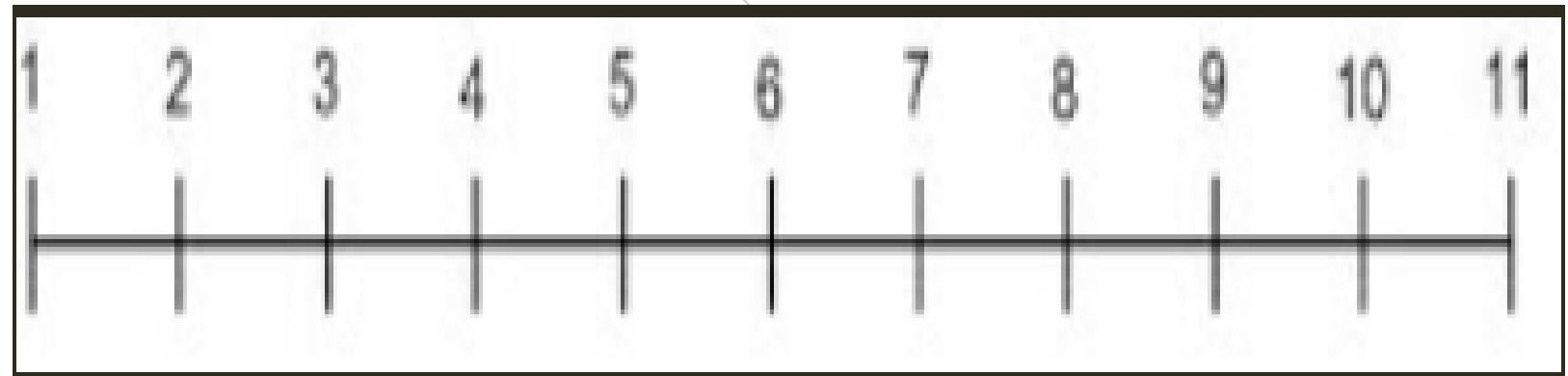
Bogardus Social Scale

Suatu skala yang direncanakan untuk mengukur jarak social yang ingin dipertahankan oleh orang antara diri sendiri dengan para anggota suku lain atau kelompok etnis lain

	To close kinship by marriage	To try club as personal charms	To try street as neighbors	To employment in my occupation in my country	To citizenship in my country	As visitors only to my country	Would exclude from my country
Canadians	1	1	3	4	5	6	7
Chinese	1	1	3	4	5	6	7
English	1	1	3	4	5	6	7
French	1	1	3	4	5	6	7
German	1	1	3	4	5	6	7
Hindus	1	1	3	4	5	6	7
			(etc.)				

SKALA THURSTONE

Skala Thurstone adalah skala yang disusun dengan memilih butir yang berbentuk skala interval. Setiap butir memiliki kunci skor dan jika diurut, kunci skor menghasilkan nilai yang berjarak sama.



METODE PENSKALAAN

Metode penskalaan adalah pemberian nilai-nilai kedalam variabel-variabel yang sesuai dengan skalanya. Terdapat dua macam metode penskalaan.



**SKALA
RATING**



**SKALA
RANGKING**

JENIS PENSKALAAN

JENIS	DESKIPSI
MERANKING	Subjek mengurutkan stimulus berdasarkan kesesuaiannya dengan kondisi dirinya
MENILAI (RATING)	Subjek menilai stimulus berdasarkan kesesuaiannya dengan kondisi dirinya
MENKATEGORIKAN	Subjek meletakkan stimulus pada kategori yang sesuai dengan kondisi dirinya
MEMBANDINGKAN	Subjek memilih pasangan stimulus yang sesuai dengan kondisi dirinya
MENGESTIMASI	Subjek mengestimasi dengan memberikan penilaian pada atribut yang sesuai dengan kondisi dirinya
MEMETAKAN KEMIRIPAN	Subjek memetakan kesamaan antar stimulus pada sebuah peta dimensi stimulus



THANK YOU
Universitas Bhayangkara Jakarta Raya

PENGUKURAN TES DAN PSIKOMETRI

Univesitas Bhayangkara Jakarta Raya



Validitas

Narasumber: Dr. Netty Merdiaty.MM.,M.Si
Tanggal

PERTEMUAN 4



Validitas

- Validitas berarti ketepatan atau kecermatan
- Validitas merupakan sejauh mana alat ukur benar-benar mengukur apa yang akan diukur

Ciri-ciri validitas :

1. Bersifat relative
2. Bukan merupakan sifat yang menetap pada alat ukur
3. Merupakan sejauh mana alat itu valid mengukur apa yang ingin diukur, bukan masalah valid atau tidak

KOEFISIEN VALIDITAS

- Hasil estimasi validitas suatu pengukuran dinyatakan secara empiris dengan koefisien validitas
- Koefisien validitas merupakan korelasi antara distribusi skor tes yang bersangkutan dengan distribusi skor kriteria.
- Koefisien validitas hanya bermakna ketika mempunyai harga positif.
- Koefisien validitas dikatakan tinggi apabila mendekati

TIPE VALIDITAS

- a) Content validity
- b) Criterion related validity
- c) Construct validity

CONTENT VALIDITY

- Bertujuan : menetapkan sejauh mana seseorang telah menguasai kemampuan tertentu.
- Apakah alat ukur benar-benar mengukur apa yang secara keseluruhan ingin diukur.
- Berkaitan dengan relevansi item-item alat ukur, per aitem maupun secara keseluruhan.
- Validitas banyak digunakan untuk tes prestasi

MENENTUKAN CONTENT VALIDITY

- Judgement ahli
- Menguji item yang ditulis dengan blue print yang dibuat sebelum penulisan item

Face Validity : tidak menunjuk pada apa yang alat ukur ingin ukur, namun hanya sekedar penampakan yang sifatnya common sense.

CRITERION – RELATED VALIDITY

- Bertujuan : membuat prediksi posisi seseorang di waktu yang akan datang pada suatu variable tertentu.
- Diperoleh dengan mengkorelasikan skor dari alat ukur (yang ingin diuji validitasnya) dengan skor yang diperoleh dari kriteria pada saat ini dan pada waktu yang akan datang.
- Kriteria merupakan ukuran lain yang mengukur hal yang sama dengan alat ukur validitasnya.

Dua bentuk Criterion-Related Validity

- Predictive Validity

Mengkorelasikan skor alat ukur dengan skor kriteria yang diperoleh pada waktu yang akan datang.

- Concurrent Validity

Mengkorelasikan skor alat ukur dengan skor kriteria yang diperoleh pada saat ini (tidak ada interval waktu pengukuran antara alat ukur dan kriteria).

Permasalahan dalam Criterion Related Validity

a) Bagaimana menentukan kriterion yang tepat?

b) Two way correction for attenuation

Yaitu koreksi penuh yang melibatkan alat ukur dan kriterion. Dilakukan bila keduanya memiliki reliabilitas yang rendah.

Permasalahan dalam Criterion

- One way Correction for attenuation

Yaitu koreksi yang hanya melibatkan kriterion.

Dilakukan bila kriterion memiliki reliabilitas yang jauh lebih rendah daripada alat ukur

CONSTRUCT VALIDITY

- Bertujuan: menentukan sejauh mana seseorang memiliki beberapa sifat seperti yang diukur oleh alat ukur.
- Construct Validity adalah sejauh mana alat ukur bisa dikatakan mengukur sebuah konstruk atau sifat teoritisnya

Langkah – Langkah Construct Validity

- Menemukan ukuran konstruk
- Menentukan korelasi antara ukuran-ukuran konstruk dengan alat ukur lain yang sejenis.
- Menentukan ukuran-ukuran yang benar-benar mengukur konstruk yang diukur

Cara lain

- Jika Memiliki korelasi yang tinggi dengan alat ukur lain yang mengukur atribut yang sama (Convergent Validity).
- Jika Memiliki korelasi yang rendah dengan alat ukur lain yang mengukur atribut yang berbeda (discriminant Validity)

Faktor-faktor yang mempengaruhi validitas

- Panjang alat ukur
- Variabilitas kemampuan kelompok
- Instruksi tes yang ambigu
- Perbedaan sosio kultural
- Penambahan item-item yang tidak tepat

Interpretasi Koefisien Validitas

- Bersifat relative
- Kesepakatan umum, koefisien validitas $r_{xy} \geq 0,30$ dianggap memuaskan



THANK YOU

Universitas Bhayangkara Jakarta Raya

PENGUKURAN TES DAN PSIKOMETRI

Univesitas Bhayangkara Jakarta Raya



Reliabilitas

Narasumber: Dr. Netty Merdiaty. MM., M. Si
Tanggal

PERTEMUAN 5



Reliabilitas

- Reliabilitas berarti keajegan atau bisa disebut dengan
- konsistensi.
- Reliabilitas merupakan sejauh mana suatu pengukuran dapat dipercaya.
- Term untuk menyatakan apakah alat ukur memberikan hasil sama apabila digunakan mengukur subjek yang sama pada waktu yang berbeda.
- Mengasumsikan tidak ada realchange(perubahan yang berarti) dalam diri subjek.

Stability Reliability

- Reliability over time : Apakah hasilnya sama kalau diaplikasikan pada waktu berbeda.
- Test retest methods : bila alat ukur reliabel maka korelasi antar dua pengukuran, subjek yang sama sebaiknya tidak lebih kecil dari 0.6

Koefisien reliabilitas

- a) Tinggi rendahnya reliabilitas ditunjukkan dengan koefisien reliabilitas.
- b) Koefisien reliabilitas awalnya merupakan korelasi antara dua distribusi alat ukur

Pendekatan reliabilitas

- Tes Ulang
- Tes Sejajar/Paralel Form
- Konsistensi Internal
 - a. Formula Spearman Brown
 - b. Formula Rulon
 - c. Formula Alpha
 - d. Formula Kuder Richardson 20

Types of reliability

- Inter rater/inter observer reliability
- Test retest reliability
- Parallel forms reliability
- Internal consistency reliability:
 1. Average inter item correlation
 2. Average item total correlation
 3. Split half reliability
 4. Cronbach's alpha
 5. Formula spearman Brown
 6. Formula Rulon
 7. Formula Kuder Rihardson

Tes ulang

- Menggunakan satu alat ukur yang dikenakan dua kali pengukuran pada subjek yang sama
- Antara pengukuran pertama dan yang kedua ada jeda waktu
- Koefisien reliabilitasnya adalah korelasi dari hasil pengukuran pertama dengan yang kedua
- Kurang praktis dan memungkinkan carry over effect

Tes Sejajar

- Membutuhkan dua alat ukur yang dianggap memenuhi asumsi paralel
- Koefisien reliabilitasnya adalah korelasi dari hasil pengukuran alat ukur 1 dengan alat ukur 2.
- Sulit mencari alat ukur yang paralel dan tidak menghilangkan kemungkinan carry over effect.

Konsistensi Internal

- a) • Membutuhkan satu alat ukur yang dikenakan sekali pada sekelompok subjek.
- b) Komputasi koefisien reliabilitas dilakukan dengan membelah alat ukur menjadi beberapa bagian.
- c) Bentuk dan sifat alat ukur serta banyaknya belahan menentukan teknik perhitungan koefisien reliabilitasnya.

BELAHAN TES PADA KONSISTEN INTERNAL

2 bagian

3 bagian

MULTI-bagian

1. KOEF ALPHA CRONBACH

Dapat dibelah menjadi beberapa bagian

Dibelah menjadi bagian yang sama panjang

2. KOEF RELIABILITAS HORST

Dibelah beberapa bagian

Tidak perlu sama panjang

Jika sama Panjang maka identic dengan alpha

3. KOEF RELIABILITAS FELDT

Tidak dapat dibelah

Panjang tidak sama

Semua hitungan digunakan tanpa belahan

EFEK PERUBAHAN PANJANG TES

Perubahan Panjang tes fokus pada banyaknya aitem

Bertambah atau berkurang aitem tidak akan mempengaruhi mean dan varian

Bertambah atau berkurang aitem mempengaruhi reliabilitas dan validitas

EFEK PERUBAHAN PANJANG TES

Semakin Panjang tes maka reliabilitas akan bertambah

Tapi tidak akan mencapai 1,0

Bisa mencapai 1,0 asal aitem bersifat homogen

EFEK HETEROGENITAS

Heterogen pd kelompok diidentifikasi dengan besarnya variasi

Semakin banyak variasi maka kelompok disebut heterogen

Heterogen akan mempengaruhi reliabilitas dan validitas

EFEK HETEROGENITAS

Semakin heterogen maka semakin tinggi skor reliabilitas

Pada sampel yang heterogeny dapat terjadi overestimasi

Sedanglan sampel yang homogen dapat terjadi underestimasi

Equivalence Reliability

- Split Half Method : Dividing the items of composite system into two parts, and then correlating the two parts, if the correlations is high, reliability should be high.
- Cronbach Alpha : Average Correlation among items. High Average correlation high equivalence

Formula Spearman Brown

- Jumlah item genap
- Alat ukur dibelah menjadi dua bagian (ganjil-genap atau atas bawah)

$$r_{xx'} = \frac{2(r_{y_1y_2})}{(1 + r_{y_1y_2})}$$

◆ $r_{y_1y_2}$ = Korelasi belahan 1 dan 2

Formula alpha

Dapat dibelah menjadi beberapa bagian yang seimbang

$$\alpha = \left(\frac{k}{k-1} \right) \left(1 - \frac{S^2_1 + S^2_2 + \dots + S^2_n}{S^2_x} \right)$$

- ◆ α = Koefisien reliabilitas alpha
- ◆ k = Banyaknya belahan
- ◆ n = banyaknya belahan
- ◆ S^2_j = varian skor belahan (j)
- ◆ S^2_x = varian skor tes (X)

Interpretasi Koefisien Reliabilitas

- Bersifat Relatif
- Kesepakatan umum, koefisien reliabilitas harus setinggi mungkin, 0,9 dianggap memuaskan
- Dengan reliabilitas dapat diketahui error standar dalam pengukuran.

$$S_e = S_x \sqrt{(1 - r_{xx'})}$$

- ◆ S_x = Standar deviasi skor tes
- ◆ $r_{xx'}$ = Koefisien reliabilitas tes

Faktor-faktor yang mempengaruhi reliabilitas

1. Faktor Internal : Faktor yang melekat pada alat ukur itu sendiri
2. Faktor eksternal : Faktor yang terdapat di luar alat ukur itu sendiri.

Faktor Internal

- Banyaknya butir, makin banyak butir makin reliabel.
- Range skor total, makin besar range skor total, alat ukur makin reliabel, karena menunjukkan bahwa subjek uji coba heterogen.
- Homogenitas aitem, makin homogen aitem, makin reliabel (jangan sampai tujuan untuk mengukur matematika, di dalamnya terdapat butir kimia).

Faktor Eksternal

- Variabilitas Kelompok, makin homogen kelompok testee, reliabilitas makin rendah, semakin heterogen testee, reliabilitas semakin tinggi.
- Terkaan testee saat menjawab. Terkaan memuat keakuratan hasil ukur kurang bisa dipercaya (menimbulkan kesalahan). Makin besar kesalahan, alat ukur semakin tidak reliabel.
- Fluktuasi keadaan sesaat pada diri subjek kondisi subjek yang berubah-ubah akan mempengaruhi reliabilitas.



THANK YOU
Universitas Bhayangkara Jakarta Raya

PENGUKURAN TES DAN PSIKOMETRI

Univesitas Bhayangkara Jakarta Raya



Norma Pengukuran

Narasumber: Dr. Netty Merdiaty. MM., M. Si
Tanggal

PERTEMUAN 6



APA ITU NORMA ?

- Secara empiris norma dapat ditetapkan dengan mengetahui apa yang sesungguhnya dilakukan oleh individu di dalam kelompoknya
- Norma merupakan suatu batas
- Tes → Skor : Hasil yang diperoleh subjek dari suatu tes Skor mentah : hasil/nilai sebenarnya yang didapatkan subjek dari suatu tes
- Untuk menilai dengan tepat posisi individu yang mengacu pada sampel yang terstandar, maka skor mentah diubah menjadi ukuran relatif.

Norma

- ✓ Penyebaran skor-skor dari suatu kelompok standar yang menjadi pasangan untuk memberi makna/arti pada skor-skor mentah yang diperoleh individu
- ✓ Patokan untuk membandingkan skor mentah subjek yang satu dengan yang lain untuk mencapai pengertian/makna skor mentah tersebut (distandarisasi dari suatu kelompok)
- ✓ Contoh : Norma pada tes intelegens

Tujuan norma

1. Menginterpretasikan suatu skor tes agar skor tes tersebut bermakna
2. Menunjukkan posisi relatif individu dalam suatu kelompok yang terstandarisasi
3. Sebagai dasar pembandingan performa setiap individu dalam suatu kelompok yang terstandarisasi
4. Membandingkan performa individu pada dua jenis tes yang berbeda

Hal – hal yang perlu diingat dalam mentransformasi skor mentah ke dalam norma

1. Bukan mengubah skor mentah yang diperoleh individu, tetapi menampilkannya ke dalam unit lain yang lebih sederhana
2. Perhitungan yang dilakukan mengarah pada sekumpulan informasi yang tidak terdapat pada skor mentah
3. Dalam scaled score, skor mentah individu dapat tampil dalam unit yang lebih normatif dan lebih dapat diinterpretasikan

Fungsi Norma

- Pedoman untuk memisahkan sejumlah individu ke dalam beberapa bagian dengan di dasarkan pada kenyataan atau data
- Memberikan kategorisasi atau tempat bagi individu di mana sesungguhnya posisi dia di dalam kelompoknya
- Menggambarkan kedudukan individu terhadap orang lain

Kurva normal

- Dikatakan normal, jika kurva memperlihatkan seperti lonceng.
- Pada dasarnya, kurve ini mengindikasikan bahwa jumlah terbesar kasus pada pusat kisaran (tengah) dan jumlahnya menurun secara bertahap pada kedua belah bagian
- Kebanyakan distribusi sifat manusia membentuk kurve normal
- Semakin banyak jumlah subjeknya, akan semakin mirip distribusinya dengan kurve normal teoritis.

Pembuatan Kurva Normal

- Ada beberapa konsep statistik yang perlu dicari terlebih dahulu, antara lain:
- Mean (M)
- Standar Deviasi/ Simpangan Baku
- Persentase kasus-kasus yang berada pada simpang baku tertentu

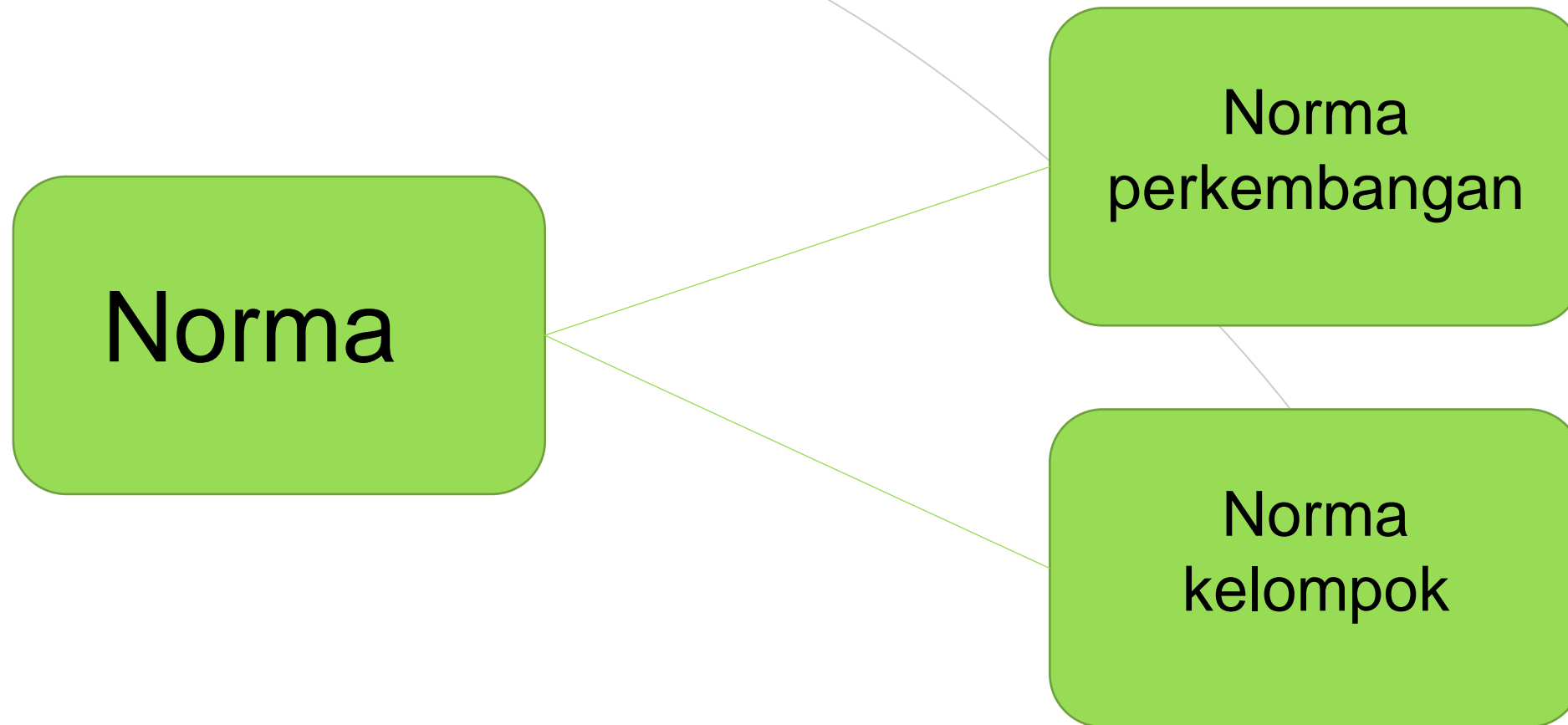
Norma Kelas

- a) Norma kelas didapat dengan menghitung skor mentah rata-rata yang didapatkan oleh anak-anak dalam tiap kelas, kemudian ditentukan normanya.
- b) Jika metode yang diberikan untuk masing-masing kelas berbeda, maka cara ini cukup baik untuk diberlakukan, namun tidak dapat diperbandingkan dengan kelas yang lain karena normanya berbeda

Sampel Normatif

- Norma apapun dibatasi oleh populasi normatif tertentu tempat norma itu diturunkan
- Norma-norma tes psikologis sama sekali tidak absolut, universal ataupun permanen
- Sampel harus mewakili populasi yang dimaksud oleh tujuan penelitian
- Faktor-faktor selektif samar-samar yang mungkin membuat sampel tidak representatif seharusnya diselidiki dengan hati-hati

Jenis – jenis norma



Norma perkembangan

- ❑ Suatu cara untuk menafsirkan skor tes/pengukuran dengan memberikan indikasi tentang tingkat perkembangan individu
- ❑ Skor mentah yang dihubungkan dengan suatu norma perkembangan memperlihatkan sejauh mana perkembangan individu
- ❑ Cenderung kasar secara psikometris
- ❑ Lebih digunakan untuk tujuan deskriptif pada studi klinis individual

Beberapa jenis norma perkembangan

1. Mental Age (MA) / Usia Mental

- a) Yang menjadi norma adalah nilai rata-rata yang diperoleh kelompok usia sampel standar pada tes tersebut
- b) Terdapat pada skala Binet – Simon (item dikelompokkan menurut tingkat usia)
- c) Skor yang dicapai seorang anak dapat menunjukkan tingkat usia perkembangannya

ex : Nilai rata-rata kelompok usia 8 tahun pada tes $X = 23$, maka skor 23 memiliki mental age = 8. Artinya kemampuan subjek pada tes setara dengan kemampuan anak usia 8 tahun.

2. Grade Equivalents (GE) / Kesetaraan Kelas

- a) Biasa digunakan pada tes prestasi pendidikan
- b) Norma kelas didapatkan dengan menghitung skor rata-rata yang didapatkan oleh siswa di kelas tersebut.

ex : Rata-rata soal yang dapat dipecahkan oleh anak kelas 4 dalam sampel standar = 23. Maka

3. Ordinal Scales (OS) / Skala Ordinal

- a) Berasal dari riset dalam psikologi anak
- b) Untuk mengidentifikasi tahap yang dicapai anak dalam perkembangan fungsi-fungsi perilaku tertentu
- c) Perkembangan anak berlangsung secara bertahap dengan urutan yang teratur. Tahap yang lebih rendah harus dilalui dulu sebelum tahap yang lebih tinggi dapat dicapai (bisa dilihat pada Gesell Developmental Schedules)
- d) Penting karena merupakan deskripsi perilaku anak (apa sesungguhnya bisa dilakukan seorang oleh anak)

Ex : Perilaku motorik, bahasa

Norma dalam kelompok

- Suatu cara menafsirkan skor tes dengan membandingkannya dengan performa sekelompok subjek pada tes tersebut,
- Kelompok subjek = kelompok standarisasi
- Penilaian dilakukan dengan melihat kedudukan subjek dalam penyebaran skor kelompok
- Kedudukan subjek dalam pengukuran kelompok standar menggambarkan kedudukan subjek dalam penyebaran populasi

Beberapa jenis norma dalam kelompok

Persentil

- Menunjukkan posisi relatif seseorang dalam sampel standarisasi
- Dianggap sebagai urutan dalam kedudukan suatu kelompok yang terdiri dari 100 orang, persentil yang lebih tinggi menunjukkan nilai yang lebih tinggi
- P_{50} = median, $P_{>50}$ = performansi di atas rata-rata, $P_{<50}$ = performansi di bawah rata-rata
- Kelebihan persentil : mudah dihitung dan dimengerti oleh awam, dapat digunakan secara universal, sesuai untuk semua jenis tes dan kelompok sampel

2. Standard Score (Z-score)

- Skor mentah diubah menjadi bentuk lain agar posisi atau kedudukan subjek terlihat
- Mengungkapkan jarak individu dari nilai rata-rata (mean) dalam satuan simpang baku/standar deviasi (SD)
- Rumus :

$$Z = \frac{(x-m)}{s}$$

m = mean

S = standar deviasi

T-scale

- Membandingkan skor individu pada dua tes yang berbeda
- Tidak memakai nilai minus (-)
- Ketentuan :

Skor 50 = rata-rata

Skor 60 = 1 SD di atas rata-rata

Rumus :

$$***T = 50 + 10Z***$$

Cara dalam menginterpretasi nilai tes

Dalam menginterpretasi nilai tes, mengacu pada 2 cara, yaitu :

1. Criterion Referenced Testing (CRT)

- ❖ Dikenal pula sebagai tes berujuk dominan
- ❖ Fokus pada apa yang dapat dilakukan dan yang diketahui individu bukan bagaimana dibandingkan dengan orang lain (biasanya untuk mengukur keterampilan tertentu)

2. Norm Referenced Testing (NRT)

- ❖ Melihat kemampuan dengan membandingkannya dengan orang lain (menggunakan norma)

BEBERAPA KESALAHAN YANG MUNGKIN TERJADI DALAM MENGINTERPRETASIKAN NORMA

1. Norma yang digunakan tidak dibuat berdasarkan sampel yang mewakili populasi yang akan diperbandingkan
2. Kebanyakan tes yang menggunakan skala besar tidak dilakukan penormaan ulang karena proses pengembangan norma memakan waktu dan membutuhkan biaya yang mahal
3. Digunakan jumlah sampel yang kecil dalam membuat norma



THANK YOU

Universitas Bhayangkara Jakarta Raya

PENGUKURAN TES DAN PSIKOMETRI

Univesitas Bhayangkara Jakarta Raya



Konstruksi Tes

Narasumber: Dr. Netty Merdiaty. MM., M. Si
Tanggal

PERTEMUAN 9



DEFINISI TES

TES

Instrumen / alat dalam pengukuran

TES

Prosedur sistematis yg disusun berdasarkan cara dan aturan tertentu yang bertujuan untuk Mengukur sesuatu

TES DALAM PSIKOLOGI

Subjek tes :
manusia

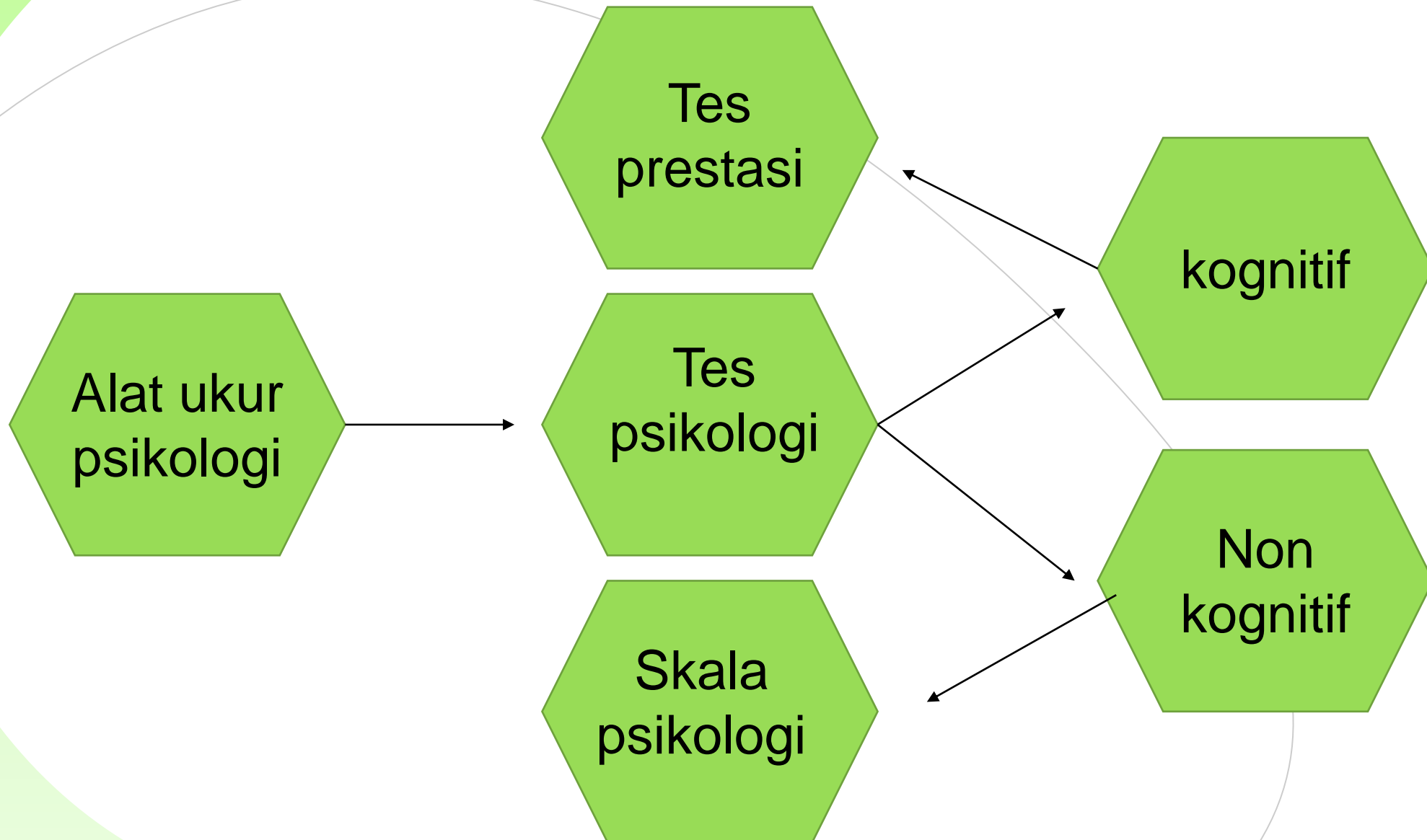
Bentuk : paper &
pencil test untuk
klasikal/ masal

Bentuk : individual
face to face oleh
psikolog

Administrasi tes :
persiapan,
pelaksanaan, skoring,
interpretasi

Semakin berkualitas
alat tes maka hasil
semakin akurat

ASAL MULA TES



JENIS TES

Tes Kognitif

- Abilitas potensial umum (intelegensi)
- Abilitas potensial khusus (bakat)
- Abilitas aktual (prestasi)

Tes Non Kognitif

- Afektif
- Kepribadian

TES NON KOGNITIF

Afektif

Untuk mengetahui sikap ataupun perilaku individu

Kepribadian

Untuk mendeskripsikan perilaku melalui observasi, inventori, ataupun proyektif

Contoh tes

Kepribadian, kepercayaan diri, optimisme, stress, dll

JENIS TES KOGNITIF

Tes
Intelegensi
(IQ)

Tes Bakat

Tes Aktual /
prestasi

TES KOGNITIF

Tes berkaitan dengan atribut kognitif

Mebutuhkan analisa

Jawaban yang diberikan benar dan salah

INTELEGENSI

Kecerdasan individu secara umum

tingkat perkembangan kognitif umum individu

IQ

BAKAT

mengetahui kemampuan individu dalam bidang-bidang khusus

meramalkan untuk kerja di masa yang akan datang

Jenis : verbal, numerik, spasial, perceptual, reasoning, mekanik, memori, clerical (admin)

PRESTASI

Mengetahui penguasaan pada subyek yang telah dipelajari sebelumnya.

Menggambarkan hasil belajar

Membagi kawasan tujuan pendidikan mejadi tiga bagian, yaitu kawasan kognitif, kawasan afektif, dan kawasan psikomotorik

KEGUNAAN ALAT UKUR/TES

Membantu untuk riset sosial

Membantu untuk mengambil kesimpulan

Untuk proses seleksi

Untuk menilai aspek psikologis

KARAKTERISTIK TES YANG BAIK

Mampu
menghasilkan
data

Memberikan
informasi yg
akurat

Hasil objektif

Ada prosedur
tes

Mudah untuk
digunakan

Valid dan
reliabel

VALIDITAS ALAT TES

Mengukur sesuai
apa yg diukur

Memberikan
informasi yg akurat

- Timbangan untuk mengukur berat badan
- Jika timbangan untuk mengukur tinggi badan maka alat ukur tidak valid
- Alat dapat memberikan hasil sesuai keadaan yang sebenarnya
- Dikatakan tidak valid jika skor tidak mencerminkan keadaan yg sebenarnya

RELIABILITAS TES

Mengacu pada konsistensi alat ukur

Jika pengetasan beberapa kali hasilnya tidak jauh berbeda maka alat tersebut reliable

Biasanya, tes yg reliabilitas nya rendah maka validitasnya pun rendah

Sumber : aitem yg buruk, kesalahan administrasi, subjek/ testee

FUNGSI TES

Fungsi deskripsi

Menggambarkan keadaan yang sebenarnya

Fungsi prediksi

Untuk memprediksi peluang keberhasilan yang akan datang

Fungsi seleksi

Untuk memilih individu yang sesuai dengan kualifikasi yang diharapkan

FUNGSI TES

Fungsi klasifikasi

Untuk mengelompokkan individu dalam kelompok yang sejenis

Fungsi evaluasi

Untuk mengevaluasi tindakan/ treatment yang telah dilakukan pada individu

Fungsi hipotesis

Untuk menguji hipotesis dan asumsi yang dikemukakan pada penelitian ilmiah

JENIS TES

TES STANDAR

- Tes yang dilakukan secara baku dari segi materi, administratif, cara skoring dan interpretasi
- Dirancang untuk digunakan untuk jangka Panjang
- Alat tes memiliki hak cipta
- Contoh : Tes IQ, UN, SNMPTN

JENIS TES

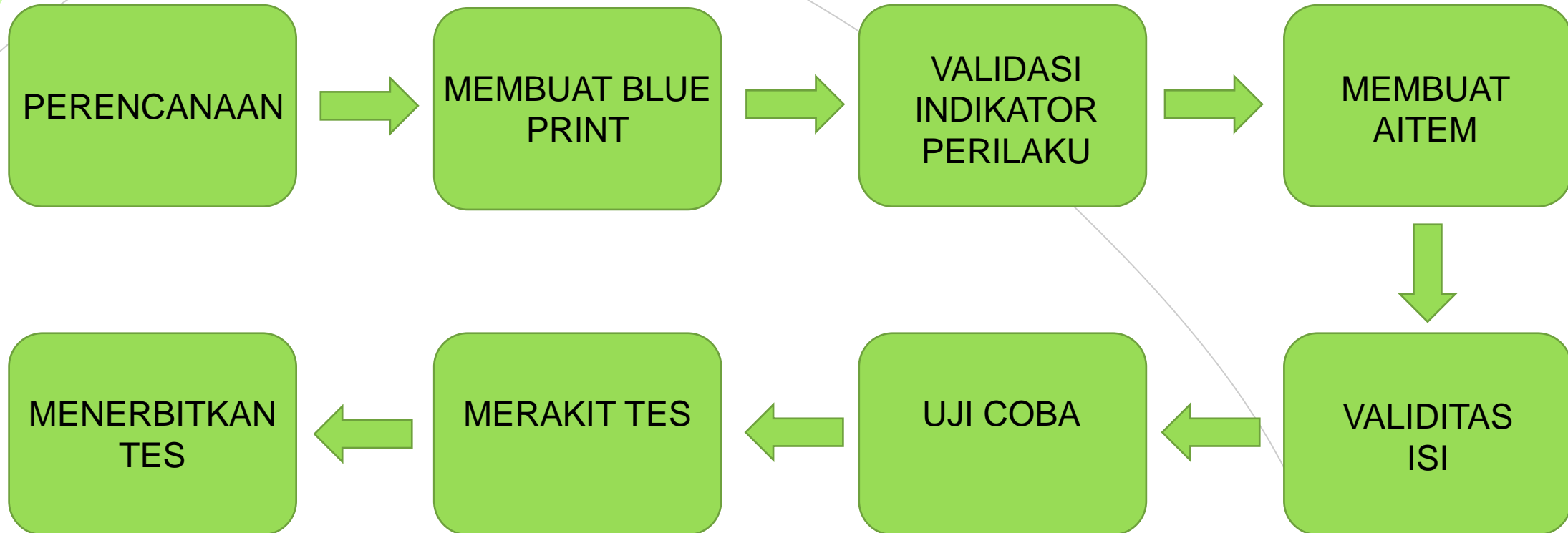
TES NON STANDAR

- Dijumpai pada tes prestasi belajar
- Digunakan untuk keperluan prediksi keberhasilan belajar
- Kualitas lebih rendah daripada tes standar
- Cth: UTS, UAS

PERBEDAAN

	STANDAR	NON-STANDAR
TUJUAN	Kebutuhan secara umum	Kebutuhan secara khusus
ISI	Tidak boleh diubah	Sesuai dengan kelas
ADMINISTRATIF	Harus diikuti	Ditentukan oleh guru
NORMA	Dibuat oleh penerbit	Dibuat oleh guru
EVALUASI	Dilakukan penerbit	Dilakukan guru

LANGKAH DALAM KONSTRUKSI TES





THANK YOU

Universitas Bhayangkara Jakarta Raya

PENGUKURAN TES DAN PSIKOMETRI

Univesitas Bhayangkara Jakarta Raya



Tes Kognitif

Narasumber: Dr. Netty Merdiaty. MM., M. Si
Tanggal

PERTEMUAN 10



TES KOGNITIF

TES KOGNITIF

- Abilitas potensial umum (intelegensi)
- Abilitas potensial khusus (bakat)
- Abilitas aktual (prestasi)

TES NON KOGNITIF

- Aspek afektif dan kepribadian

Apakah ujian = tes?

Ujian Psikologi **TIDAK SAMA DENGAN** Tes Psikologi

- Individu yang menjalani ujian biasanya sudah tahu mengenai apa yang harus dikerjakannya dan apa saja yang harus dipersiapkan karena sudah jelas batas-batasnya dan sudah diketahui pula prosedur pengerjaannya.
- Individu yang menjalani tes, akan merasakan ketidakpastian mengenai apa yang akan ditanyakan.
- Seperti contoh di atas, apabila individu akan menjalani tes psikologi, maka individu akan merasakan “ketidakpastian” mengenai apa yang akan dikerjakan. Sedangkan apabila individu akan menjalani ujian psikologi, maka individu mampu memiliki gambaran apa yang akan dikerjakan.
- Namun dalam dunia pendidikan, kedua kata ini sering memiliki konotasi yang sama (contoh : UAS dan THB)

Agar tidak semua bentuk ujian dapat diklasifikasikan ke dalam tes, maka berikut merupakan definisi tes secara mendasar.

- Anne Anastasi (1976) : Tes merupakan suatu pengukuran yang objektif dan standar terhadap **sampel perilaku**.
- Frederick G. Brown (1976) : Tes adalah prosedur yang **sistematik** (objektif, standar, dll) guna mengukur **sampel perilaku** seseorang.
- Lee J. Cronbach (1970) : Prosedur sistematis untuk melihat **perilaku** seseorang dan menggambarkannya melalui bantuan skala numerik.

KARAKTERISTIK TES YANG BAIK

1. **Valid** → akurasi dan kecermatan hasil ukur
 - Instrumen ukur atau tes disebut sebagai tidak valid apabila?
 - Tes yg valid → sebelumnya dilakukan uji validitas → validitas isi dan validitas logis

2. **Reliabel** → konsistensi hasil ukur tes

Tes yang tidak reliabel → tidak valid

Reliabilitas dipengaruhi oleh:

- Kualitas aitem
- Proses pengukuran
- Subjek yg dites

3. **Objektivitas** → dalam pemberian skor, penafsiran, pemaknaan hasil tes

Objektivitas di dalam tes psikologi → mencakup prosedur administrasi dan penafsiran hasil

4. **Praktis dan ekonomis**

- Penyajian
- Pemberian skor
- Interpretasi hasil
- Tidak rumit pelaksanaannya
- Tidak banyak memakan waktu
- Meminimalisir biaya

TES KEMAMPUAN KOGNITIF



TES PRESTASI DAN KOMPETENSI

1. Potensial

- Potensial umum
- Potensial khusus

2. Aktual

- Prestasi

KARAKTERISTIK TES POTENSIAL

- Berkaitan dengan abilitas
- Memerlukan stimulasi untuk berkembang
- Dapat dioptimalkan
- Tidak mudah berubah
- Pemaknaan fokus pada masadepan

KARAKTERISTIK TES PRESTASI

- Hasil dari pembelajaran/ pelatihan
- Berkembang sejalan dengan proses pembelajaran
- Dibatasi oleh potensi
- Dapat berubah secara signifikan

Tujuan pengukuran tes kemampuan (kognitif)

Tes Kemampuan Potensial

- Bertujuan untuk memprediksi peluang keberhasilan yang akan datang

Tes Kemampuan Aktual

- Bertujuan untuk menggambarkan prestasi yang dicapai sebagai hasil dari usaha pembelajaran/ pelatihan

BENTUK KEPUTUSAN

KEPUTUSAN DIDAKTIK

- Untuk memenuhi kebutuhan pengajaran seperti pembuatan kurikulum

KEPUTUSAN ADMINISTRATIF

- Untuk memenuhi kebutuhan administratif seperti nilai kelulusan

KEPUTUSAN BIMBINGAN PENYULUHAN

- Untuk memberikan bimbingan dlm penjurusan dan karir

Tujuan hasil tes prestasi

1. TUJUAN PENEMPATAN

- Untuk mengklasifikasi individu dalam bidang jurusan yg sesuai dgn kemampuannya

2. TUJUAN NORMATIF

- Untuk melihat kemampuan belajar yang telah dicapai siswa

3. TUJUAN DIAGNOSTIK

- Untuk mendiagnosa kesulitan dalam belajar, mendeteksi kelemahan siswa yg dpt diperbaiki segera

4. TUJUAN SUMATIF

- Pengukuran akhir yang digunakan untuk memperoleh informasi mengenai penguasaan pelajaran (lulus / tdk lulus)

Rancangan Tes Prestasi

1. Menentukan wilayah pengukuran

- Atribut kognitif : prestasi, intelegensi, atau bakat

2. Menentukan dasar konseptual

- Definisi, faktor, proses belajar, dll

3. Menentukan subjek

- Siswa TK, SD, SMP, SMA

Lanjutan

4. Menentukan tujuan tes

- Sesuai tujuan hasil tes belajar

5. Menentukan materi tes

- Materi proyektif utk atribut non kognitif
- Materi non proyektif utk atribut kognitif

6. Menentukan tipe soal

- Sblm menentukan tipe soal harus mempertimbangkan tujuan tes, cara skoring, pencetakan tes, & kegiatan penyelenggaraan

7. Menentukan jumlah soal

- Sblm menentukan jumlah soal harus melihat blue print yg telah di tentukan, reliabilitas, dan waktu pelaksanaan tes

8. Menentukan taraf kesulitan soal

- Jika 30 orang dari 50 orang menjawab dgn benar pada soal nomor 5 maka indeks kesulitan adalah 0,6 atau 60%
- Semakin tinggi indeks-nya maka semakin mudah soal yg diberikan



THANK YOU

Universitas Bhayangkara Jakarta Raya

PENGUKURAN TES DAN PSIKOMETRI

Univesitas Bhayangkara Jakarta Raya



Membuat Alat Tes Prestasi

Narasumber: Dr. Netty Merdiaty. MM., M. Si
Tanggal

PERTEMUAN 11



KUALIFIKASI PENULISAN AITEM

Bukan hanya blue print dan spesifikasi tes

Penulisan aitem yang berkualitas

Memahami bahan yang di ajarkan

Evaluasi

KOMPETENSI / KISI-KISI

Sebagai patokan prestasi belajar

Sebagai hasil dari mempelajari topik tertentu

Mengandung level kompetensi

DOMAIN UKUR/WILAYAH PENGUKURAN

BLUE PRINT / KISI-KISI

- Batas domain ukur berdasarkan subdomain
- Subdomain berdasarkan isi sub bab
- Memuat bobot % dan jumlah
- Semakin tidak terperinci semakin baik

CONTOH: uji matematika dalam persentase

Domain / bab	Subdomain / sub bab	Jumlah	Bobot
Hitungan sederhana	Penjumlahan	10	50 %
	Pengurangan		
Hitungan kompleks	Perkalian	10	50 %
	Pembagian		
Total		20 aitem	100%

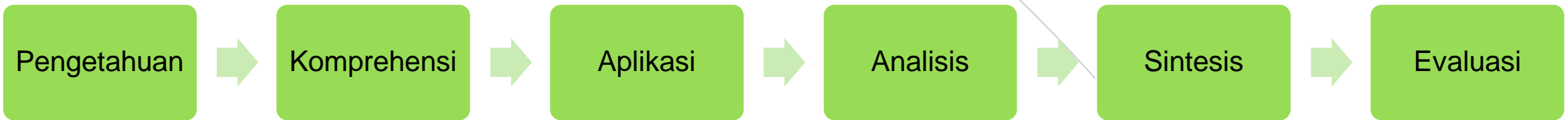
TAKSONOMI

Acuan dalam menentukan kompetensi

Terbagi menjadi beberapa level

Taksonomi yg digunakan dalam ranah kognitif : taksonomi BLOOM

LEVEL KOMPETENSI TAKSONOMI BLOOM



TAKSONOMI BLOOM

TAKSONOMI	COTOH KATA KERJA
Pengetahuan	Menyebutkan, memasang, memilih, menamakan, mengidentifikasi
Komprehensi/ Pemahaman	Menjelaskan, mendeskripsikan, membedakan, meramalkan
Aplikasi	Menghitung, mengoperasikan, menyusun, menerapkan
Analisis	Menganalisa, membandingkan, mengambil kesimpulan, mengkategorisasikan
Sintesis	Membuat rancangan, membuat hipotesa, merumuskan, merevisi
Evaluasi	Menggunakan, mengkritik, menilai, menyarankan, memberikan alasan

CONTOH TAKSONOMI

BAB	SUB BAB	KOMPETENSI
Hitungan sederhana	Penjumlahan	
	- Teori	Komprehensi
	- Perhitungan	Aplikasi
	Pengurangan	
Hitungan kompleks	- Perhitungan	Aplikasi
	Perkalian	
	- Teori	Komprehensi
	- Perhitungan	Aplikasi
	Pembagian	
	- Perhitungan	Aplikasi

CONTOH : uji matematika (dalam aitem)

Domain	Sub domain	Kompetensi			total
		pengetahuan	komprehensi	aplikasi	
Hitungan sederhana	Penjumlahan	-	1	4	5
	Pengurangan	-	-	5	5
Hitungan kompleks	Perkalian	-	2	3	5
	Pembagian	-	-	5	5
Total					20 aitem

MEMBUAT BLUE PRINT bobot soal masing masing 20%

DOMAIN	SUB DOMAIN	LEVEL KOMPETENSI			BOBOT
		Pengetahuan	Komprehensi	Aplikasi	
A	1	10			20 %
	2	10			
B	1	10			20 %
	2			10	
C	1			10	20 %
	2	10			
D	1			10	20 %
	2			10	
E	1			5	20 %
	2			5	
	3			5	
	4		5		
TOTAL		40	5	55	100 %



MEMBUAT BLUE PRINT menggunakan aitem



DOMAIN	SUB DOMAIN	LEVEL KOMPETENSI			BOBOT
		Pengetahuan	Komprehensi	Aplikasi	
A	1	(3)			(6)
	2	(3)			
B	1	(3)			20 % 6
	2			(3)	
C	1			(3)	20 % 6
	2	(3)			
D	1			(3)	20 % 6
	2			(3)	
E	1			(1)	20 % 6
	2			(1)	
	3			(2)	
	4		(2)		
TOTAL		40 12 aitem	10 3 aitem	50 9 aitem	100 % 30



THANK YOU

Universitas Bhayangkara Jakarta Raya

PENGUKURAN TES DAN PSIKOMETRI

Univesitas Bhayangkara Jakarta Raya



Format Aitem

Narasumber: Dr. Netty Merdiaty. MM., M. Si
Tanggal

PERTEMUAN 12



Penulisan Aitem

Kualitas aitem tergantung pada Penulisan aitem

Aitem dibuat sesuai indikator perilaku

Ditulisa dengan pertimbangan taraf kesukaran aitem

KAIDAH PENULISAN AITEM

Ditulis sesuai
EYD

Efisien, jelas,
dan mudah
difahami

Menanyakan
yang sesuai
indikator

Aitem harus
dinyatakan
secara jelas

FORMAT AITEM

FORMAT OBJEKTIF

- Hanya ada satu jawaban yang benar

Format essay

- Pertanyaan terbuka, jawaban panjang

KEUNGGULAN

Format objektif

- Waktu lebih cepat dalam mengerjakan
- Isi tes lebih komprehensif

Format essay

- Untuk level kompetensi yang tinggi
- Untuk mengungkapkan kemampuan individu

AITEM FORMAT OBJEKTIF

Tipe pilihan ganda

- Pertanyaan tidak lengkap
- pertanyaan

Tipe jawaban pendek

- Berupa pertanyaan fakta dan jawaban hanya 1-2 kata

Tipe benar-salah

Kaidah penulisan ganda

Jawaban pasti dan tidak ambigu

Kalimat pernyataan bukan negatif

Jawaban berdasarkan rumpun yang sama

Jawaban terlihat tampak benar/salahnya

Jangan membuat alternatif yang bertentangan

Alternatif "semuanya benar" hanya diberikan jika jawaban MUTLAK benar

Alternatif kuantitatif diurutkan dari kecil hingga besar

Contoh aitem format objektif tipe pilihan ganda

1. Pernyataan tidak lengkap

- Konsep reliabilitas menurut teori skor klasik adalah..
- A B C D

2. Pertanyaan

- Apa yang disebut dengan reliabilitas?
- a b c d

KAIDAH PENULISAN BENAR-SALAH

Untuk menguji pemahaman/komprehensi dan analisis

Benar/salah sifatnya MUTLAK

Menguji hal yang spesifik

Contoh

1. Tipe benar salah

- Perilaku merokok termasuk perilaku kenakalan remaja
- Benar salah

Contoh aitem format objektif tipe jawaban pendek

1. Pertanyaan fakta

- Tokoh utaa pendiri psikologi adalah...

2. Melengkapi

- “memahami” berada dalam level... menurut taksonomi bloom

RULES

BENAR / SALAH

Berisi tentang ANALISIS

Berisi tentang KOMPREHENS I

PILIHAN GANDA

Berisi tentang pengetahuan

Berisi tentang komprehensi

JAWBAN PENDEK

Berisi tentang PENGETAHUAN

Berisi tentang KOMPREHENS I

JUMLAH SOAL

Disesuaikan dengan level kompetensi yg dibuat pd blueprint

Aitem format esay

Lebih mudah dibuat

Berupa pertanyaan

Dapat juga berupa perintah

Kaidah penulisan aitem esay





THANK YOU

Universitas Bhayangkara Jakarta Raya

PENGUKURAN TES DAN PSIKOMETRI

Univesitas Bhayangkara Jakarta Raya



Menghitung Validitas

Narasumber: Dr.Netty Merdiaty.MM.,M.Si
Tanggal

PERTEMUAN 13

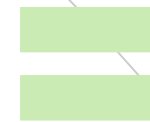


LANGKAH UJI KUALITAS AITEM

REVIEW
MATERI



REVIEW
PENGGUNAAN
KATA/UJI
KETERBACAAN



VALIDITAS
ISI

VALIDITAS ISI

Secara
Teoritik

- Aitem baik dan layak
- Disimpulkan melalui analisis statistic pd kelayakan

Bentuk
penilaian

- Dilakukan oleh experts, bukan penulis aitem
- Semakin tinggi kesepakatan antar experts, maka aitem semakin layak

JENIS VALIDITAS



JENIS VALIDITAS ISI “CVR”

Dipopulerkan oleh Lawshe

Sejauhmana kelayakan suatu tes sbg sampel

Penting dalam mengukur prestasi belajar

Para ahli disebut sbg Subject Matter Expret → menguasai materi

Tingkatan: essensial, berguna tapi tdk essensial, tdk essensial

Dihitung
per
aitem

BENTUK TABEL CVR

DOMAIN	SUBDOMAIN / KOMPETENSI	AITEM	PENILAIAN		
			E	G	T
BAB I	Penjumlahan / Aplikasi	1. (Aplikasi)	√		
		2. (Aplikasi)		X	
BAB II	Pengurangan / Pengetahuan, Komprehensi	3. (Pengetahuan)		√	
		4. (Komprehensi)			X
... dst					

SKOR PADA CVR

Angka CVR bergerak dari 1-1.0 hingga + 1.0

Jika $CVR > 0.00$ artinya 50% SME menyatakan aitem essensial

Jika CVR berada dalam rentang POSITIF artinya aitem layak digunakan

RUMUS CVR

$$CVR = (2n_e : n) - 1$$

n_e = banyaknya SME yang menilai "ESSENSIAL"

N = jumlah SME yang melakukan penilaian

Buatlah tabulasi data sehingga dapat diketahui berapa banyak SME yang menilai essensial. Contoh :

Aitem I

$n_e = 15$ dan total SME adalah 20 orang, maka

$$CVR = (2n_e : n) - 1 = 2.15 : 20 - 1 = 1,5 - 1 = 0,5 \text{ validitas aitem baik}$$

Coba hitung

Jika ada 30 penilai, kemudian dinilai pada aitem yang kedua di peroleh skor essential sebanyak 18 penilai, skor tidak essential 7 penilai, dan skor berguna tapi tidak essential sebanyak 5 penilai. Kemudian pada aitem ketiga, diperoleh skor essential sebanyak 10 penilai, dan aitem berguna tapi tidak essential sebanyak 20 penilai. Hitunglah skor CVR pada aitem kedua dan ketiga! Apakah aitem tersebut berada dalam kategori aitem yang layak?

JENIS VALIDITAS ISI "AIKEN'S V"

Dipopulerkan oleh Haynes, Richard dan Kubany

Validitas isi dipandang sbg sejauhmana instrument alat ukur dan mempresentasikan konstruk

Berkaitan dengan tes potensi dan skala non-kognitif

Menggunakan statistic V sbg indicator = 1 hingga 5

Ditabulasikan per aitem seperti CVR

BENTUK TABEL AIKEN'S V

DOMAIN	SUBDOMAIN / KOMPETENSI	AITEM	Nilai Relevansi				
			1	2	3	4	5
BAB I	Penjumlahan / Aplikasi	1..... (A)			×		
		2..... (A)				×	
BAB II	Pengurangan / P, K	3..... (P)					×
		4..... (K)					×
... dst							

RUMUS AIKEN'S V

$$V = \sum S/[n(c - 1)]$$

S = r – lo (angla yang diberikan penilai- angka terendah)

R = angla yang diberikan penilai

lo = angka terenda (l)

C = angla tertinggi (5)

Buatlah tabulasi data sehingga dapat diketahui berapa banyak penilai yang menilai relevan

Contoh :

Ada 10 penilai, 3 orang menilai 5, 2 orang menilai 3, 5 orang menilai 4

$\Sigma s = 3 (5-1) + 2 (3-1) + 5 (4-1) = 12 + 6 + 15 = 33$, maka

$V = \Sigma s : n (c-1) = 33 : 10 (5-1) = 0,825$, validitas isi aitem baik

SKOR PADA AIKEN'S

Angka pd skor Aiken's V adalah 0 hingga 1

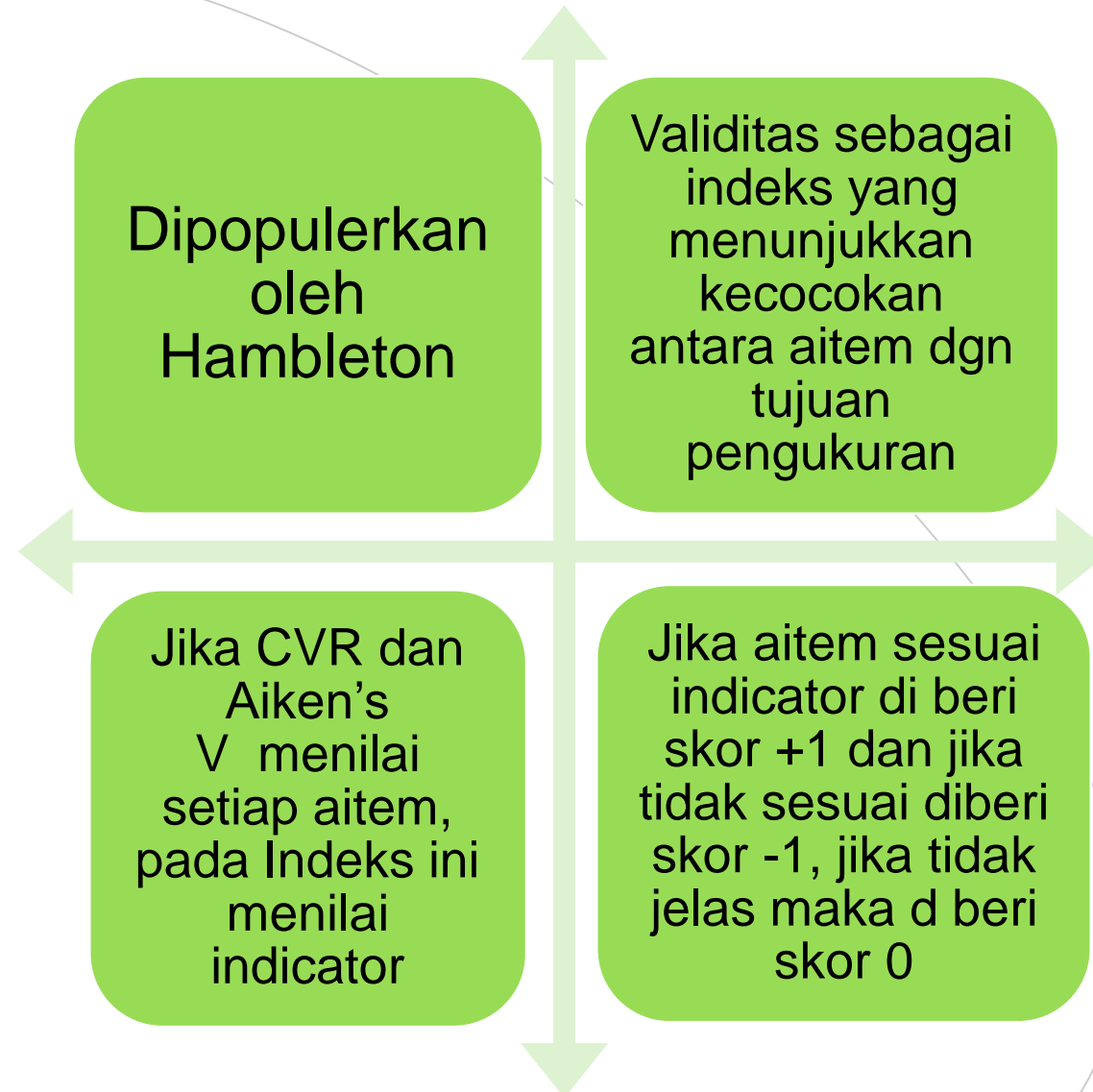
Jika skor diatas 0,5 maka dapat dikatakan maka aitem memiliki validitas isi yang baik

Semakin mendektai angka 1 maka validitasnya semakin baik

Coba hitung

Pada aitem ke-5, terdapat 10 penilai yang memberikan skor 4, 3 penilai memberikan skor 5, dan 7 penilai memberikan skor 3. sedangkan pada aitem ke-10, terdapat 7 penilai yang meberikan skor 3, 9 penilai memberikan skor 2, dan 4 penilai memberikan skor 4. bagaimana kelayakan aitem pada aitem ke-5 dan ke-10 berdasarkan validitas Aiken's V?

JENIS VALIDITAS ISI “ INDEKS IIK”



BENTUK TABEL INDEKS IIK

Penilai/ raters	Sub bab I			
	1	2	3	4
A	1	-1	1	1
B	1	0	1	0
C	1	0	0	0
D	0	0	0	-1
E	0	-1	1	-1
F	1	-1	1	0
Total	4	-3	4	-1

RUMUS INDEKS IIK

I_{ik} = indeks kesesuaian aitem I dengan tujuan k

N = banyaknya tujuan ukur

\bar{r}_{ik} = rata-rata rating terhadap aitem I pada tujuan k

\bar{r}_i = rata-rata rating terhadap aitem I pada semua tujuan (total aitem x jumlah rater)





THANK YOU
Universitas Bhayangkara Jakarta Raya

PENGUKURAN TES DAN PSIKOMETRI

Univesitas Bhayangkara Jakarta Raya



Uji Coba dan Analisis Aitem

Narasumber: Dr. Netty Merdiaty. MM., M. Si
Tanggal

PERTEMUAN 14



Uji Daya Diskriminasi Aitem

Kemampuan aitem untuk membedakan antara individu atau kelompok kelompok individu yang memiliki dan tidak memiliki atribut yg diukur

Menghitung koefisien korelasi antara distrusi skor aitem dengan distribusi skor skala itu sendiri

Akan menghasilkan koefisien korelasiasitem-total

Uji Daya Diskriminasi aitem

Misalnya pada suatu skala yang disusun untuk mengungkap kecerdasan emosi, maka aitem yang berdaya tinggi adalah aitem yang mampu menunjukkan mana individu kelompok atau kelompok individu yang memiliki kecerdasan emosi tinggi dan mana yang tidak

Uji Taraf Kesukaran Aitem

Dilakukan pada tes performansi maksimal, dengan aitem skor benar / salah

Perhitungan uji taraf kesukaran umum digunakan dengan melihat proporsi peserta tes yang menjawab benar

Individu diharapkan lebih byk menjawab benar pada aitem yang tidak dapat dijawab individu yang kurang mampu

Uji Taraf Kesukaran Aitem

Tes perlu memiliki *item difficulty* yang bervariasi untuk dapat membedakan kemampuan individu

Dibagian awal tes perlu diberikan aitem mudah lalu semakin sulit agar pengerjaan tes memiliki kepercayaan diri mengerjakan tes dan tdk membuang waktu pd aitem sulit

Sifatnya ordinal hanya untuk mengurutkan aitem berdasarkan kesukarannya

Nilai p hanya mengurutkan aitem berdasarkan kesukaran, tapi tdk diketahui perbedaan kesukaran

UJI EFEKTIVITAS DISTRAKTOR

seberapa baik pilihan yang salah tersebut dapat mengecoh peserta tes yang memang tidak mengetahui kunci jawaban yang tersedia.

Uji Distraktor pada pilihan ganda

Tujuan dari penggunaan distractor pada butir aitem pilihan ganda agar dari sekian banyak testee yang mengikuti tes hasil belajar ada yang tertarik atau terangsang untuk memiliki, karena mereka mengira bahwa distraktor yang mereka pilih itu merupakan jawaban betul

Efektifitas Distraktor

Distraktor dipilih oleh siswa kelompok rendah

Pemilih distractor tersebar relative proporsional pada masing-masing distractor yang ada

Fungsi Distraktor

Distraktor : Alternatif jawaban salah pada tes PG yang berfungsi sebagai pengecoh

Analisis fungsi distractor dimaksudkan untuk melihat seberapa efektif suatu distractor dapat berfungsi

Suatu distractor dikatakan berfungsi secara efektif bila dipilih minimal oleh 5% dari peserta tes

Daya Distraktor (diambil dari Azwar, 1996)

Nomor Aitem	Kelompok	N	Alternatif Jawaban				
			A	B	C	D	0
1	T	20	0	4	15*	1	-
	R	20	8	6	2	20	
2	T	20	2	11*	6	0	
	R	20	16	4	0	0	
3	T	20	2	6	4	8*	
	R	20	0	6	6	8	
4	T	20	2	0	2	16*	
	R	20	15	3	1	1	
5	T	20	17*	2	-	1	
	R	20	2	7	-	9	



THANK YOU
Universitas Bhayangkara Jakarta Raya