

# Isu-isu Penting dalam Penilaian

Dr. Memet Sudaryanto, M.Pd.

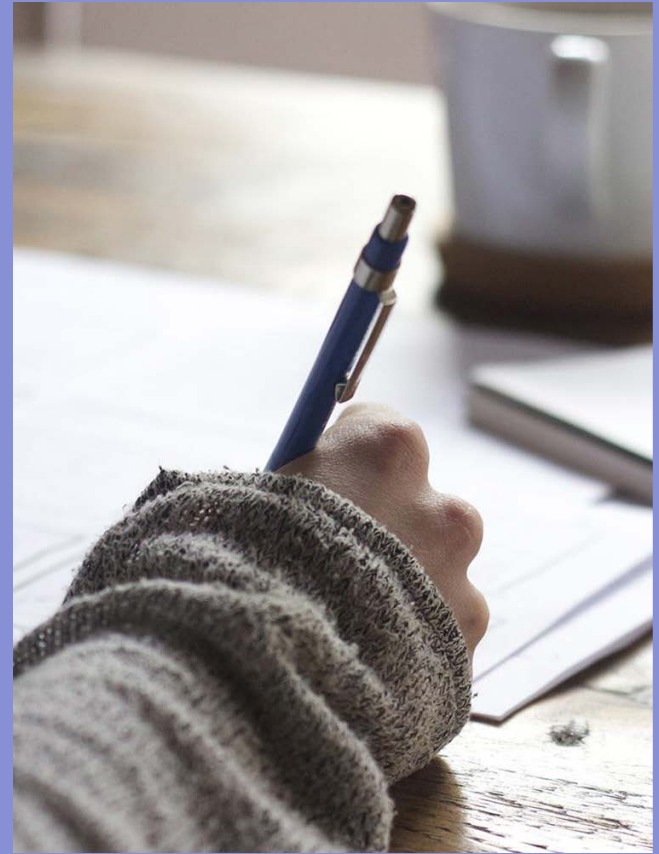
Pendidikan Bahasa Indonesia

Fakultas Keguruan dan Ilmu Pendidikan

Universitas Sebelas Maret

# Validitas Konstruk

Mengukur, apa yang diukur!







*Pembuktian validasi data  
melalui analisis faktor.*

- 1. Analisis Faktor  
Konfirmatori*
- 2. Analisis Faktor  
Eksploratori*



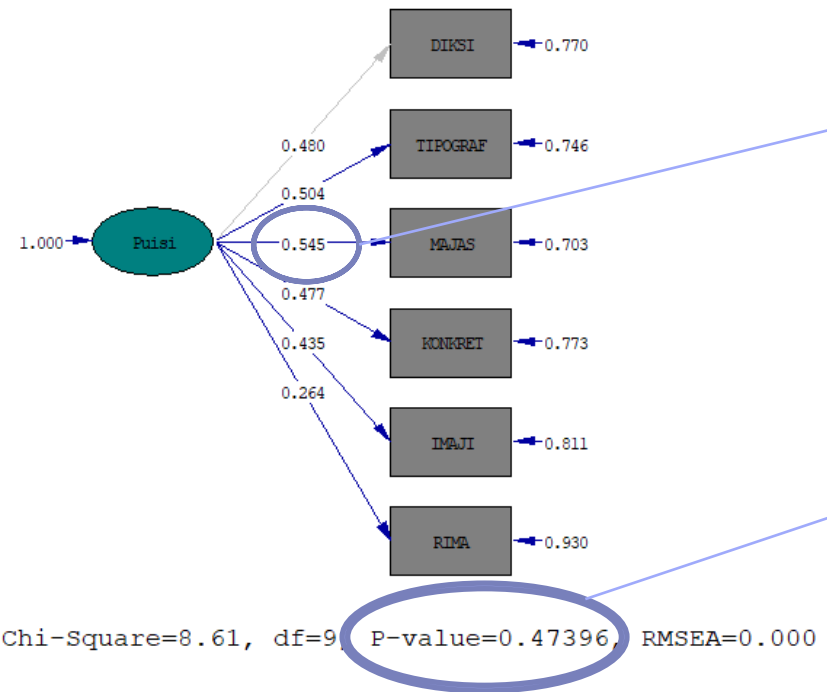
# CFA

## (Confirmatory Factor)

Mengonfirmasi variabel *latent* diukur oleh variabel *observed* dengan tepat.   


Pada analisis ini, data yang dikumpulkan mampu menunjukkan "kontribusi" variabel *observed* terhadap variabel *latent*.

# Contoh Output CFA (Lisrel)



Kontribusi Terbesar untuk Melihat *Loading* faktor

Di atas 0.05 dan RMSEA di bawah 0.05



# EFA

## Eksploratori Faktor Analisis

Menemukan ada berapa variabel dalam sebuah teori yang dikonstruksi oleh peneliti.



# EFA (Exploratory Factor Analysis)

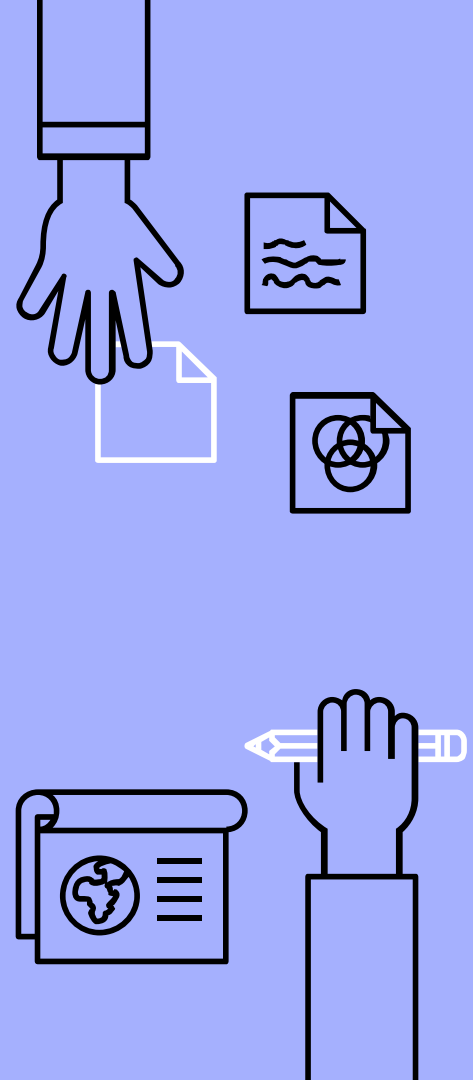
## Digunakan untuk

Menemukan ada berapa variabel dari sebuah teori yang dikonstruksi.

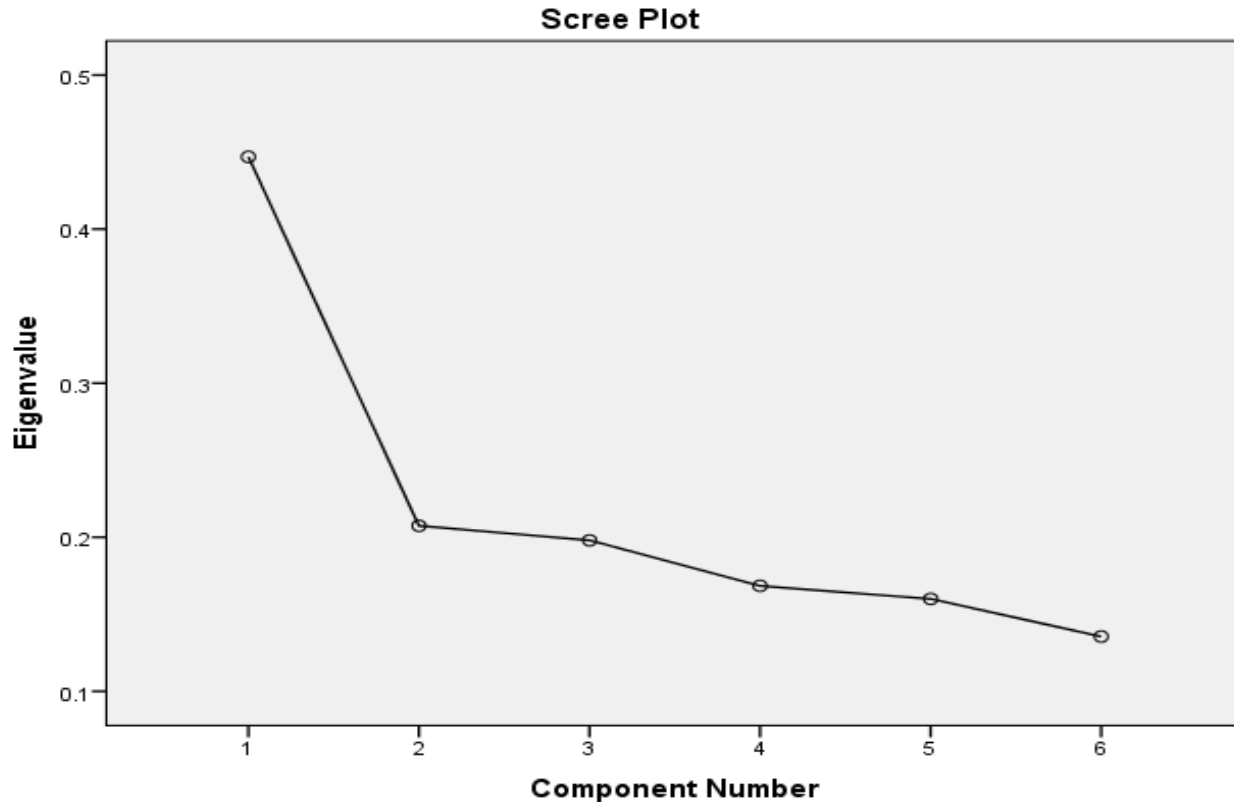
Biasanya, peneliti belum menemukan berapa faktor yang ada pada sebuah variabel.

## Konstruksi

Peneliti harus mengumpulkan teori-teori tentang sebuah variabel dan mengujinya dengan SPSS.



# Contoh Output EFA (SPSS)



**Component Matrix<sup>a</sup>**

	Raw	Rescaled
	Component	Component
	1	1
Diksi	.255	.568
Tipografi	.237	.557
Majas	.318	.667
Konkret	.300	.623
Imaji	.313	.627
Rima	.191	.404

Extraction Method: Principal Component Analysis.  
a. 1 components extracted.





# ITEM FIT

## (Analyzed by Quest)

# Kecocokan Butir

Item yang disusun oleh peneliti “FIT” atau COCOK dengan

1. Peserta Tes
2. Model Tes
3. Kriteria Pengukuran



Item Fit

all on all (N = 120 L = 31 Probability Level= .50)

-----  
INFITMNSQ .63 .71 .83 1.00 1.20 1.40 1.60  
-----

Item	MNSQ	INFIT
1 item 1	.63	*
2 item 2	.71	*
3 item 3	.83	*
4 item 4	1.00	*
5 item 5	1.20	*
6 item 6	1.40	*
7 item 7	1.60	*
8 item 8	1.40	*
9 item 9	1.20	*
10 item 10	1.00	*
11 item 11	.83	*
12 item 12	.71	*
13 item 13	.63	*
14 item 14	.63	*
15 item 15	.71	*
16 item 16	.83	*
17 item 17	.95	*
18 item 18	1.00	*
19 item 19	1.00	*
20 item 20	1.00	*
21 item 21	1.20	*
22 item 22	1.20	*
23 item 23	.83	*
24 item 24	.95	*
25 item 25	.71	*
26 item 26	.83	*
27 item 27	1.00	*
28 item 28	.83	*
29 item 29	1.00	*
30 item 30	1.20	*
31 item 31	1.40	*

Batas bawah butir  
FIT

Batas atas butir FIT

# EQUATING (Kesetaraan)

Instrumen tes dibuat berjenjang untuk melihat peningkatan kemampuan peserta tes.

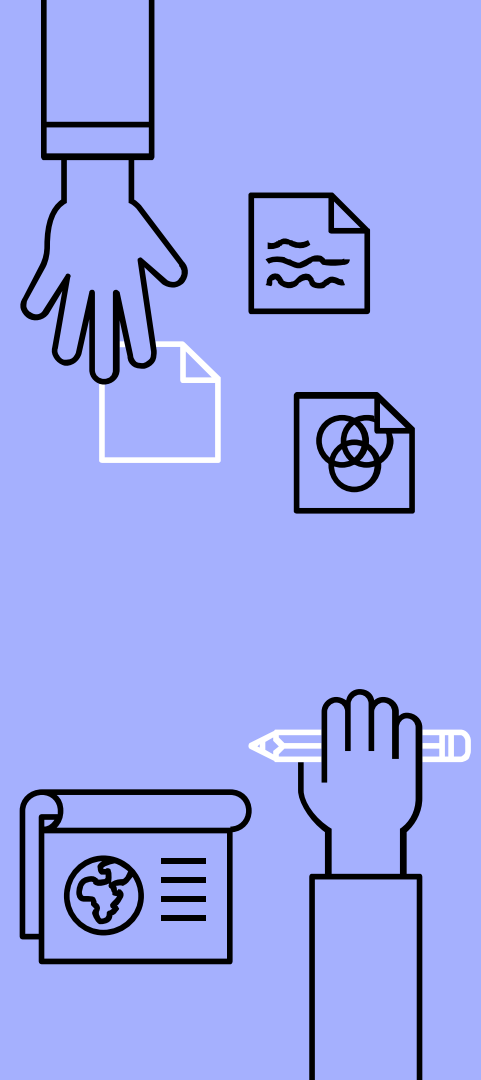
Equating dilakukan melalui butir *anchor* (butir bersama).

Ada butir yang indikatornya sama dan dikerjakan bersama-sama.



# Langkah Penyusunan Equating Test

1. Tentukan kelas tes misalnya kelas X, XI, dan XII.
2. Susun indikator tes yang sama dan buat butir tes.
3. Ukur kemampuan pada butir yang sama.
4. Apakah ada peningkatan atau malah penurunan?



# Standard Setting

## (Kriteria Ketuntasan)

Pemerintah dan hampir semua satuan pendidikan menggunakan

“alasan kebijakan”

untuk menentukan siswa lulus/tidak lulus.

Di sisi lain, kemampuan siswa antartahun tidak sama!





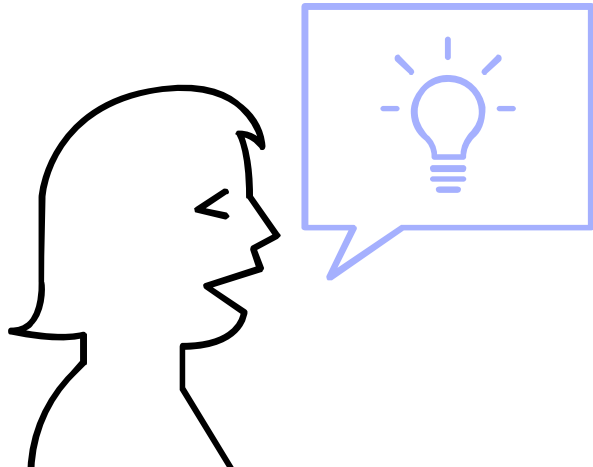
## Standard Setting dibagi menjadi 2 tujuan:

1. Lulus dan Tidak Lulus (Cut Score)
2. Placement (Penempatan Kemampuan)



# Standard Setting paling populer ditentukan dengan **BOOKMARK**

1. Tentukan tingkat kesukaran butir
2. Urutkan butir paling mudah ke sukar
3. *Focused Group Discussions* untuk menentukan kembali urutan butir soal
4. Tentukan skor batas kelulusan.

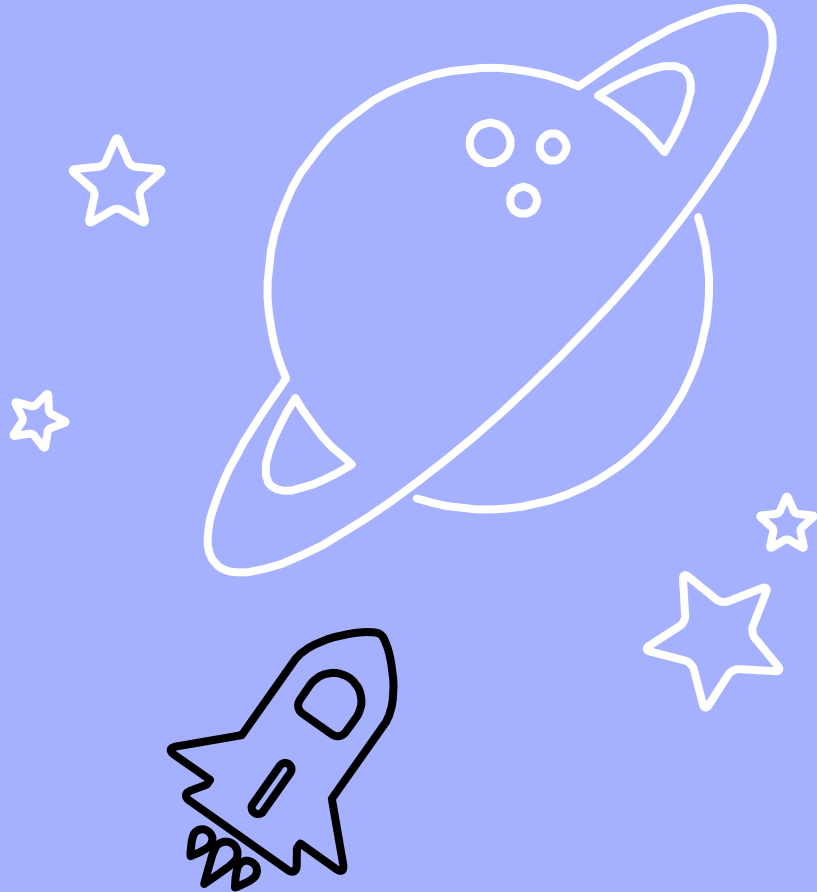




# CAT

## Computerized Adaptive Testing

Berbantuan teknologi komputer, penilaian tidak hanya memindahkan butir tes dari *paper and pencil test* menjadi *computer based test*.



# CAT

## Computerized Adaptive Testing

CAT merupakan aplikasi yang di dalamnya terdapat BANK SOAL dengan tingkat kesukaran butir yang sama.

Misalnya:

Indikator 1 terdapat 10 butir

Indikator 2 terdapat 15 butir, dst.



# Kelebihan CAT

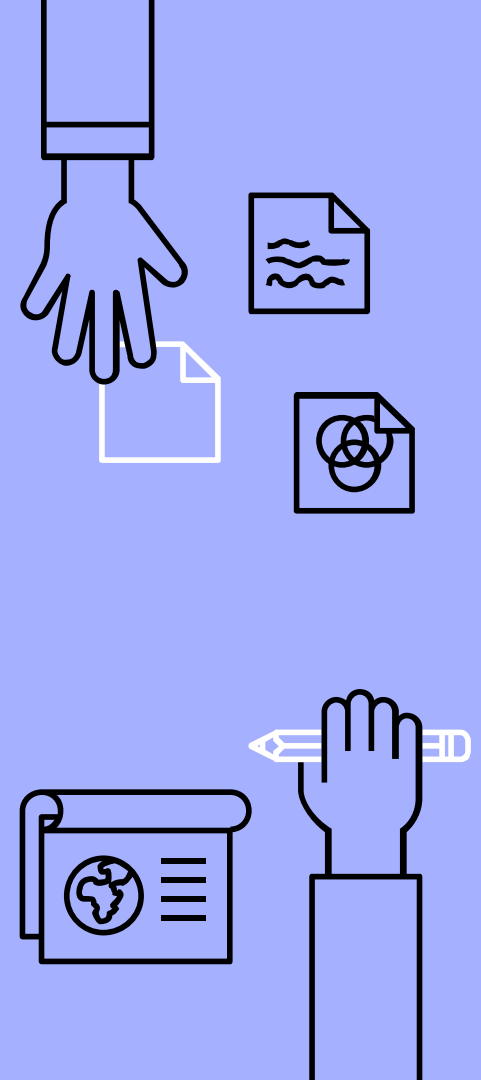
Tidak semua peserta tes mengerjakan jumlah butir yang sama: efektif dan efisien.

Andi mengerjakan 8 butir: constant

Suci mengerjakan 12 butir: constant

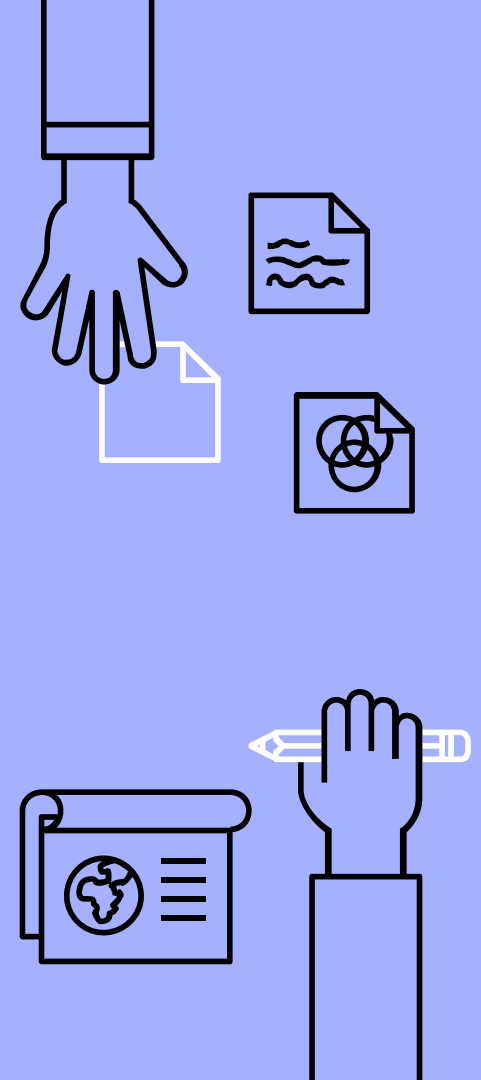
Emil mengerjakan 20 butir: constant

Kemampuan Peserta Test Berpengaruh pada jumlah butir yang akan dikerjakan.



## 4 Hal penting dalam CAT

1. Bank Soal
2. Starting Rules
3. Pemilihan Item Berikutnya
4. Stopping Rules



# DIF

*differential item functioning*

Jangan sampai tes yang dibuat mengandung DIF! DIF adalah keberpihakan suatu tes pada golongan tertentu.

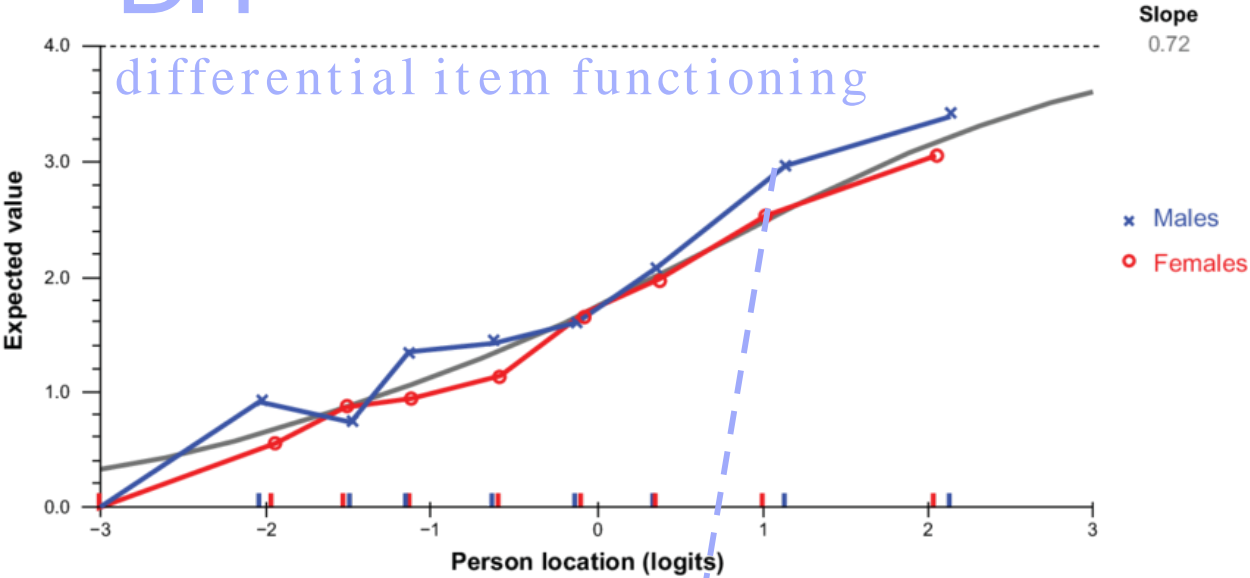
Misalnya:

Teks tentang jenis kosmetik akan menguntungkan perempuan.



# DIF

differential item functioning



Lelaki cenderung di atas, probabilitas menjawab benar lebih tinggi daripada perempuan.



# TERIMA KASIH

Memet Sudaryanto