Universidad Tecnológica de Pereira

MAESTRÍA EN INVESTIGACIÓN OPERATIVA Y ESTADÍSTICA

Multivariate statistics

Factor Analysis

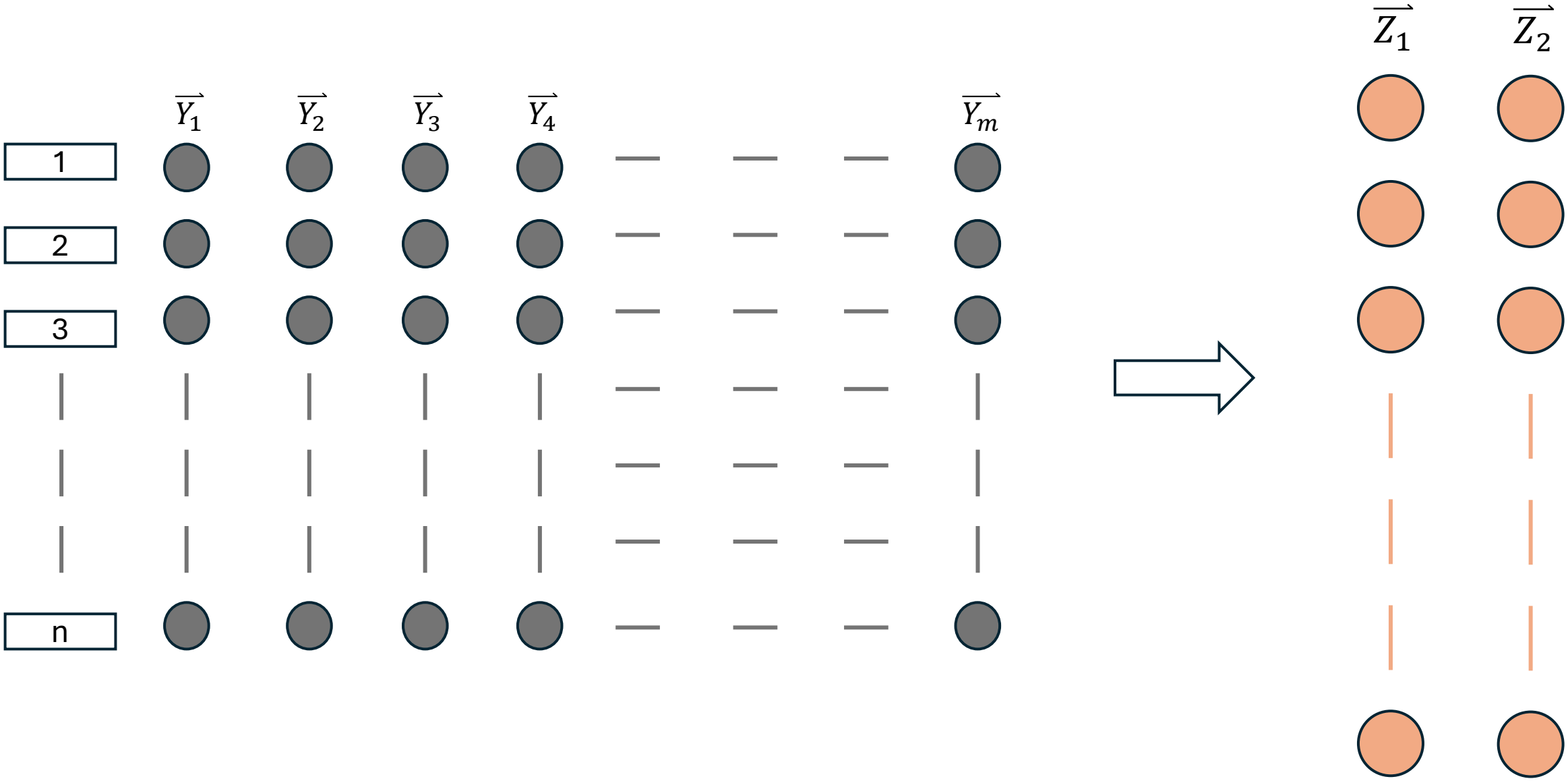
May 24, 2024

Orlando Sabogal-Cardona

@Antonio Sabogal

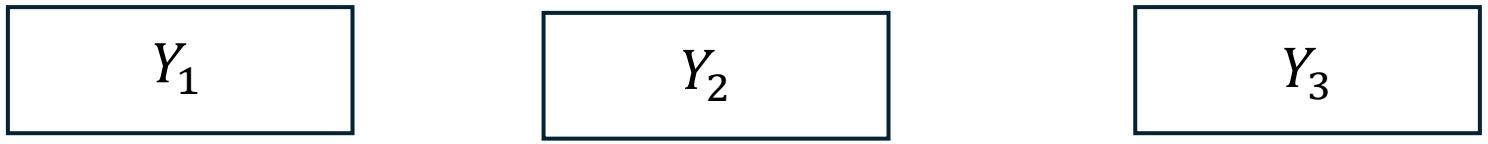
orlando.sabogal.20@ucl.ac.uk

General idea



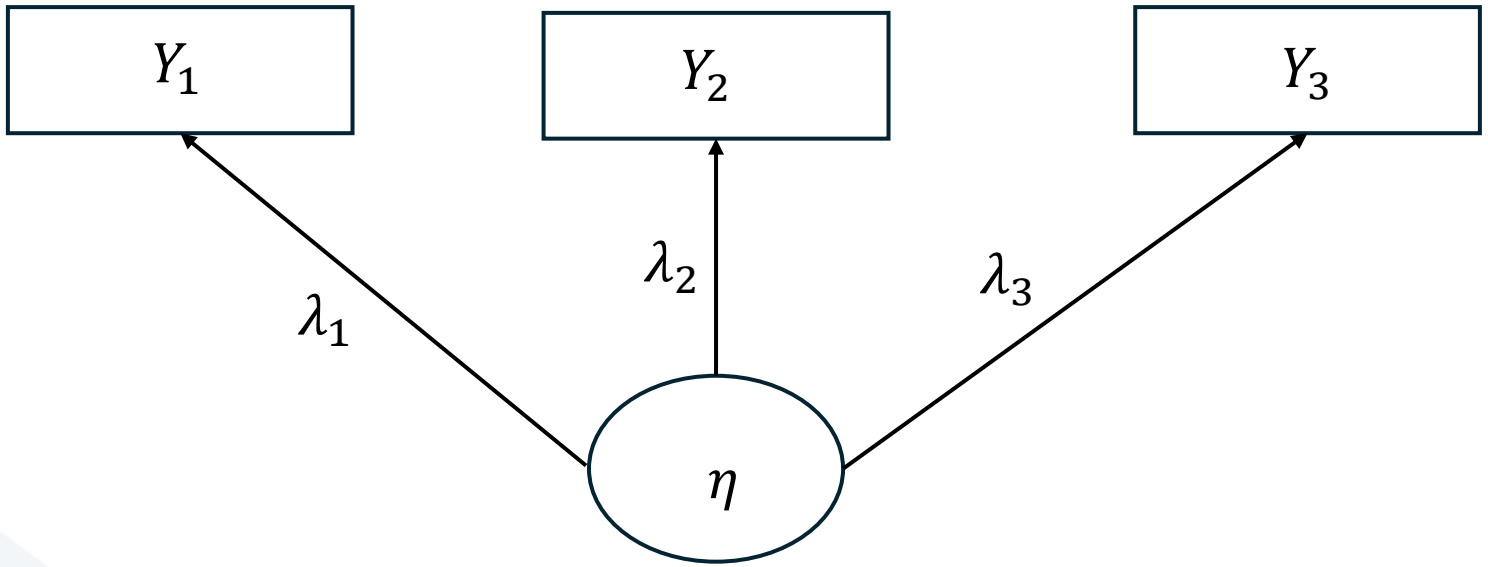
The common factor model

For the sake of simplicity, let's assume only three variables $\{Y_1, Y_2, Y_3\}$



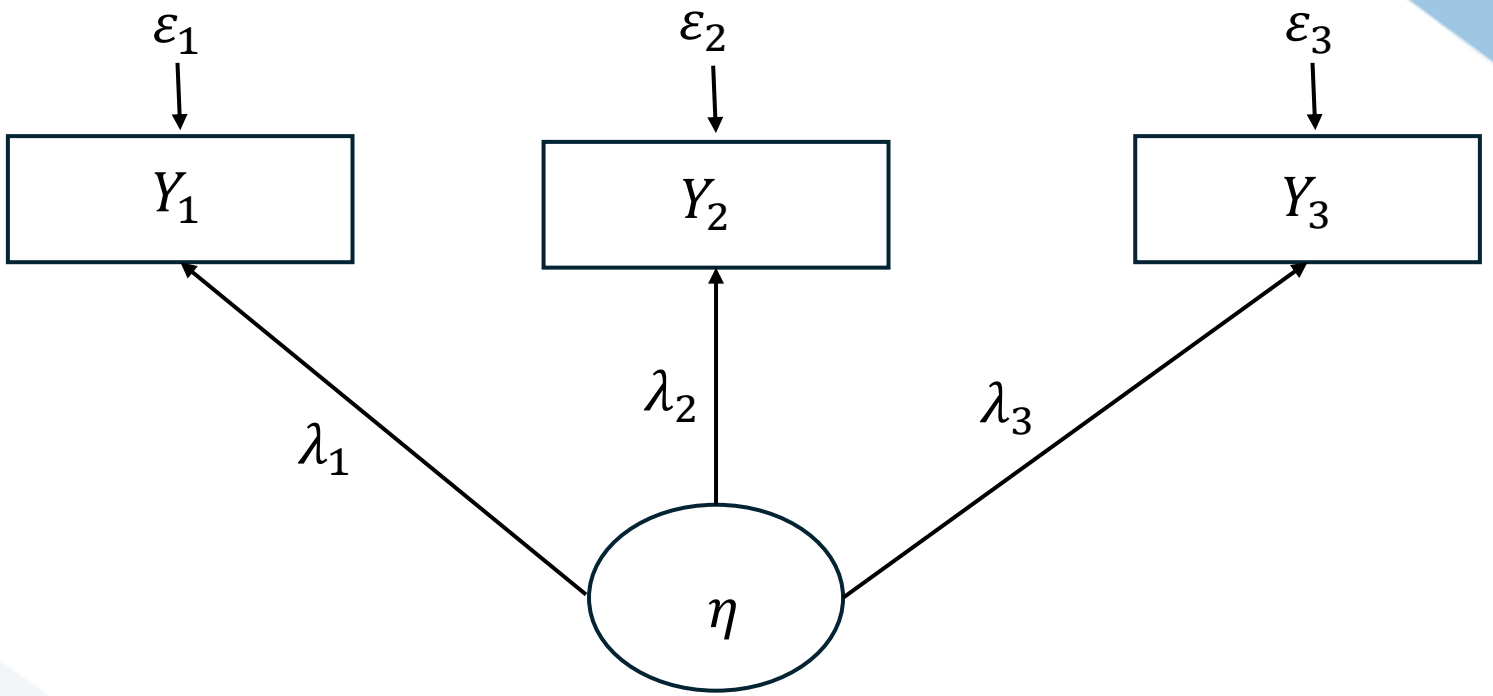
The common factor model

For the sake of simplicity, let's assume only three variables $\{Y_1, Y_2, Y_3\}$



The common factor model

For the sake of simplicity, let's assume only three variables $\{Y_1, Y_2, Y_3\}$



The common factor model

$$Y_1 = \lambda_1 * \eta + \varepsilon_1$$

$$Y_2 = \lambda_2 * \eta + \varepsilon_2$$

$$Y_3 = \lambda_3 * \eta + \varepsilon_3$$

Photo by Behzad Ghaffarian on Unsplash

Latent variables

- How would you define well-being? How would you measure it?
- What about anxiety or intelligence?
- Fear of crime?
- Satisfaction with travel?

Photo by Behzad Ghaffarian on Unsplash

Latent variables

- Unobserved or hidden variables that are inferred from observed data.
- They represent underlying concepts or constructs that cannot be directly measured.
- They may represent complex phenomena.

https://quantitudepod.org/s2e21-yes-virginia-there-are-latent-variables/

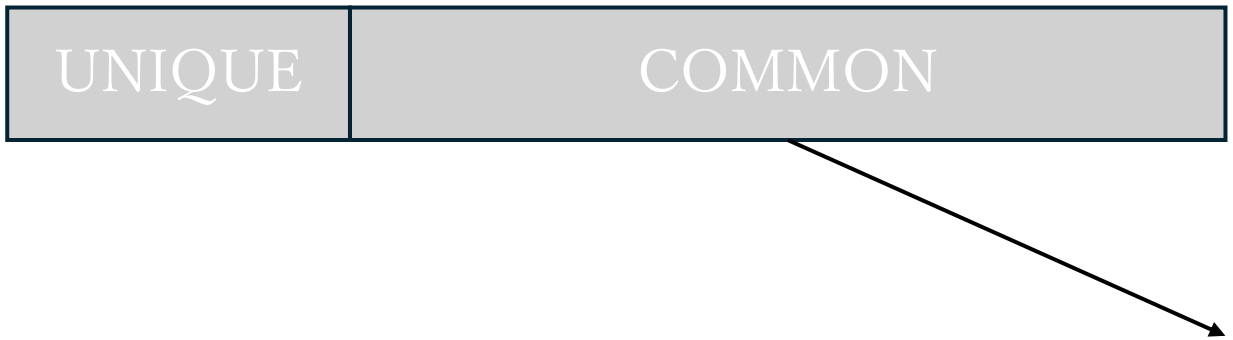
Variance

$$Y_j = \lambda_j * \eta + \varepsilon_j$$

UNIQUE	COMMON

Variance

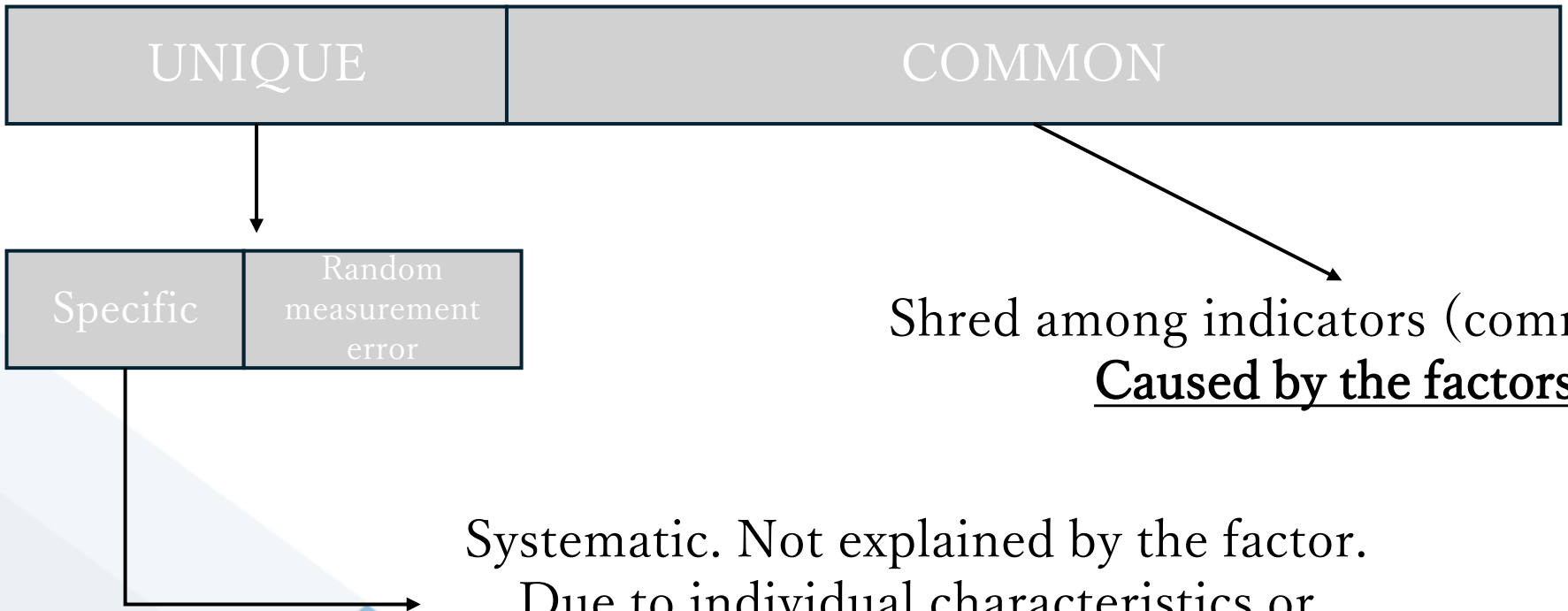
$$Y_j = \lambda_j * \eta + \varepsilon_j$$



Shred among indicators (communality). **Caused by the factors.**

Variance

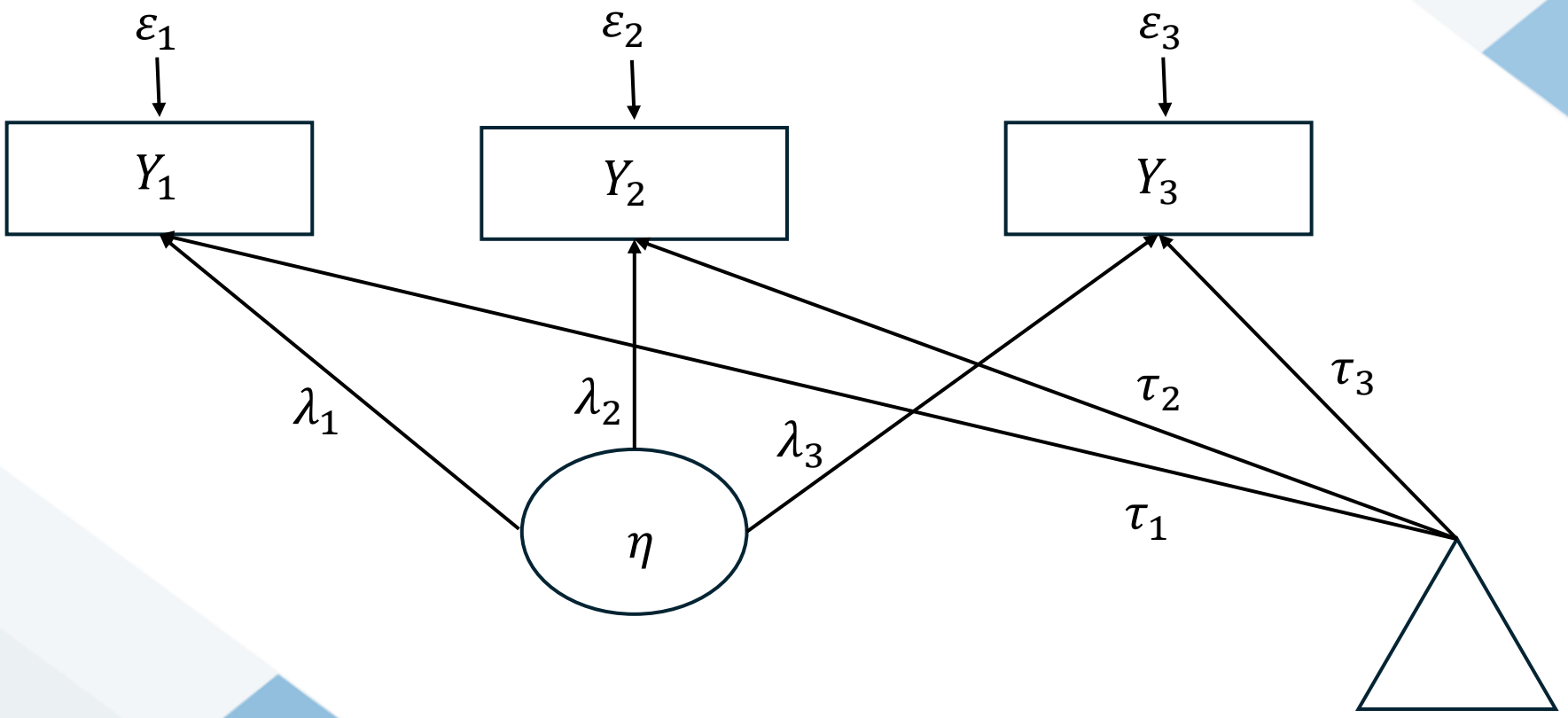
$$Y_j = \lambda_j * \eta + \varepsilon_j$$



Shred among indicators (communality).
Caused by the factors.

Systematic. Not explained by the factor.
Due to individual characteristics or absent factor.

Actually…


ε_1
ε_2
ε_3
Y_1
Y_2
Y_3
λ_1
λ_2
λ_3
τ_1
τ_2
τ_3
η

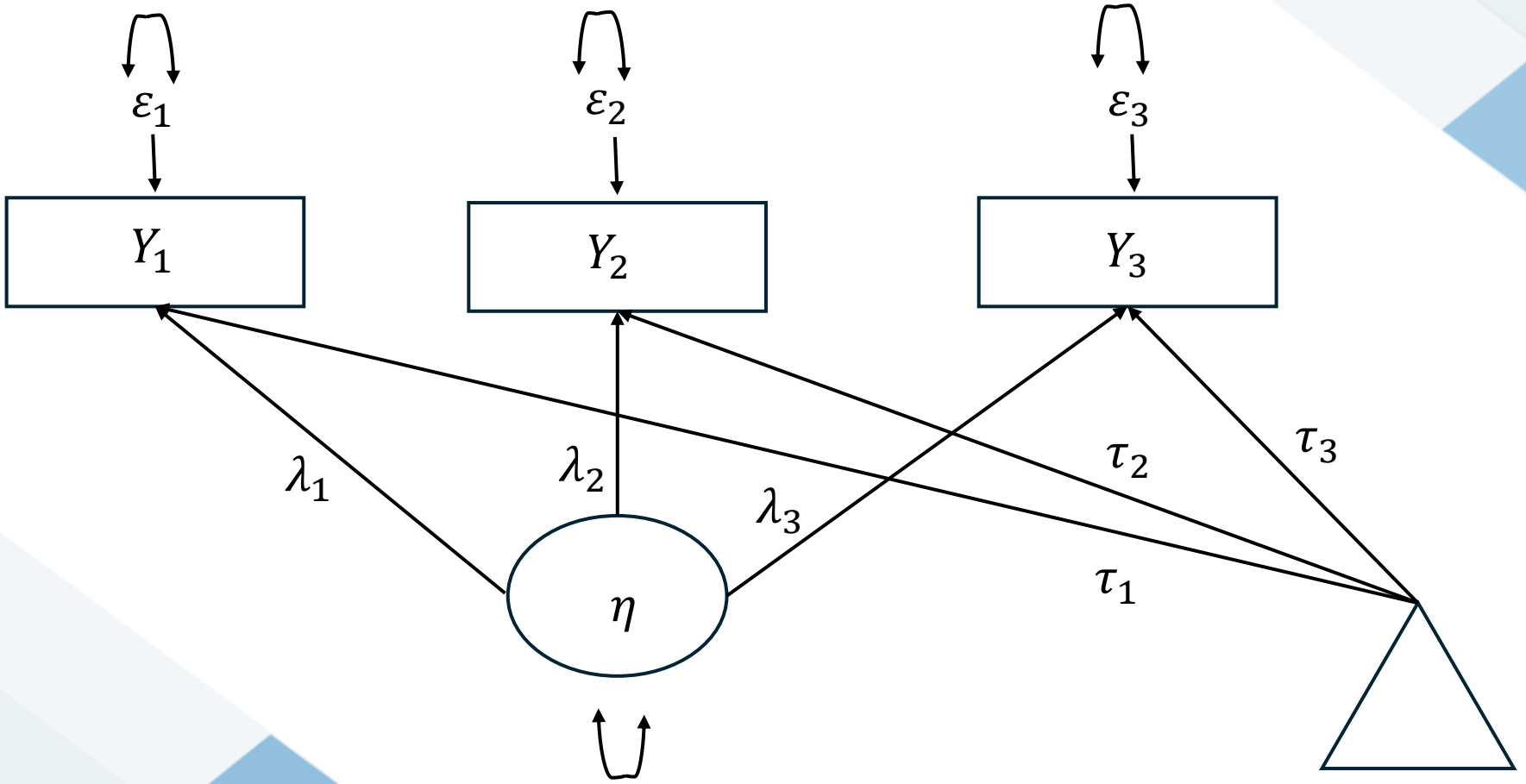
Actually…

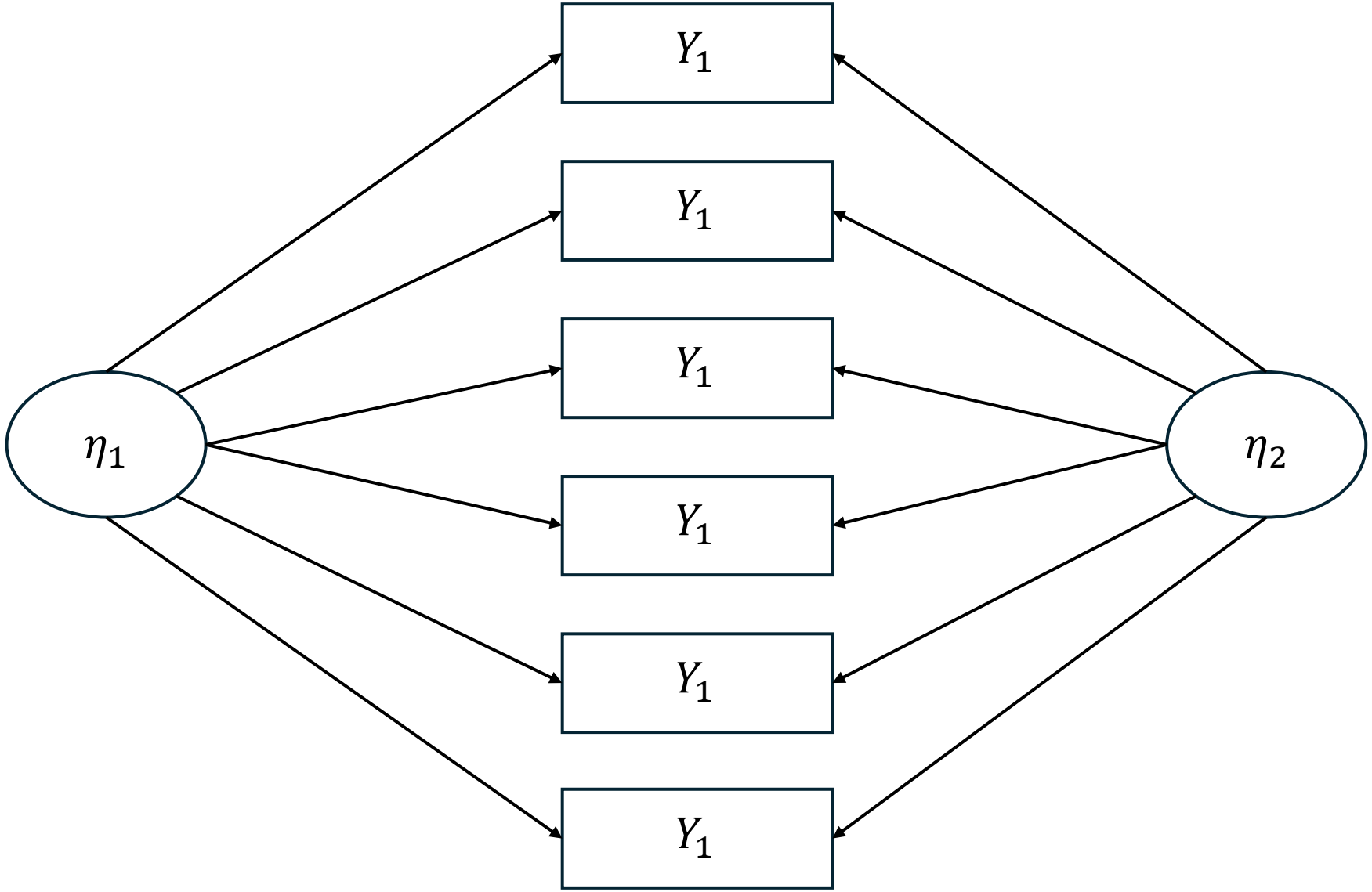
$$Y_1 = \tau_1 + \lambda_1 * \eta + \varepsilon_1$$

$$Y_2 = \tau_2 + \lambda_2 * \eta + \varepsilon_2$$

$$Y_3 = \tau_3 + \lambda_3 * \eta + \varepsilon_3$$

Actually…


ε_1
ε_2
ε_3
Y_1
Y_2
Y_3
λ_1
λ_2
λ_3
τ_1
τ_2
τ_3
η


Y_1
Y_1
Y_1
Y_1
Y_1
Y_1
η_1
η_2

Does that make sense to you?
How likely do you think it is?

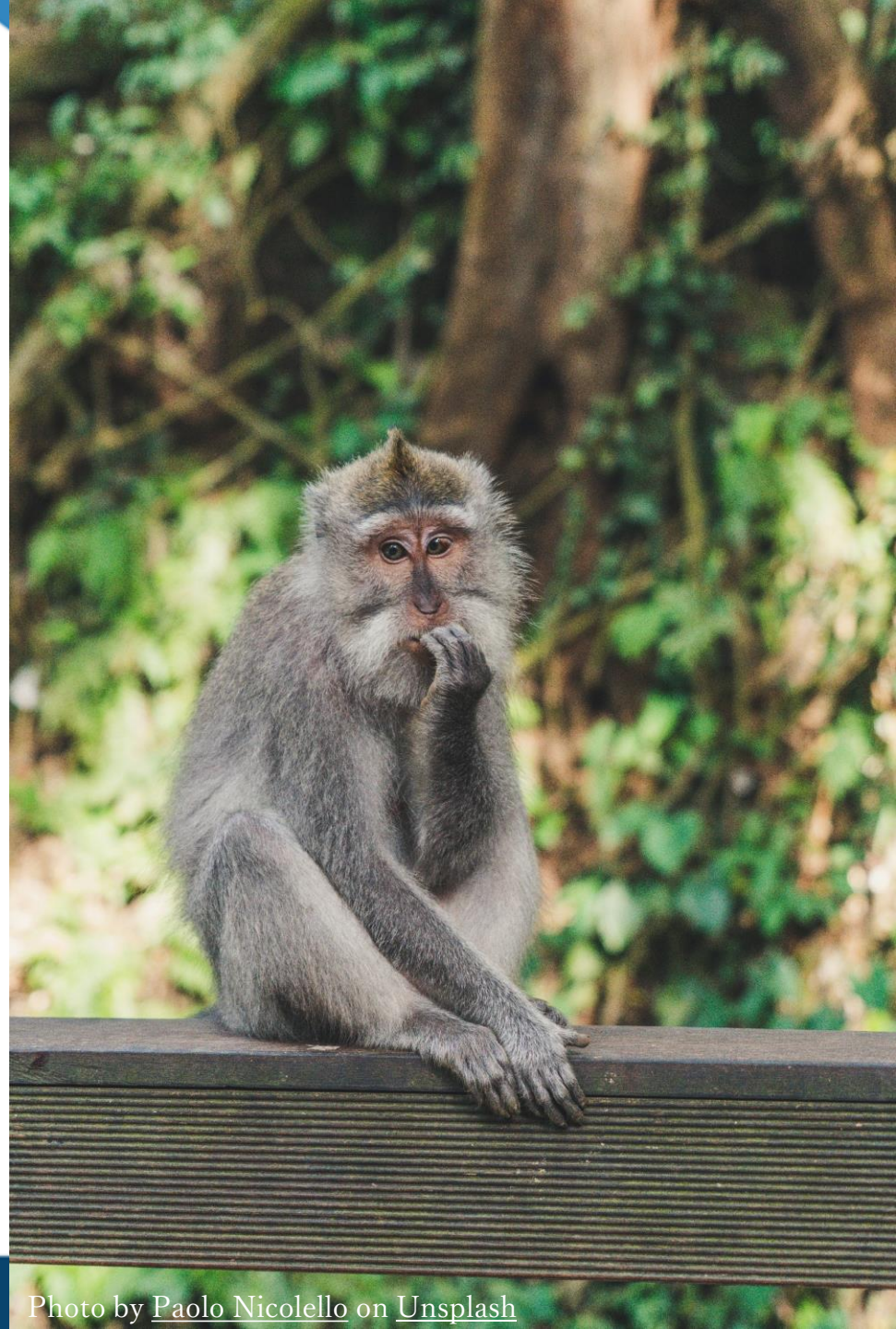
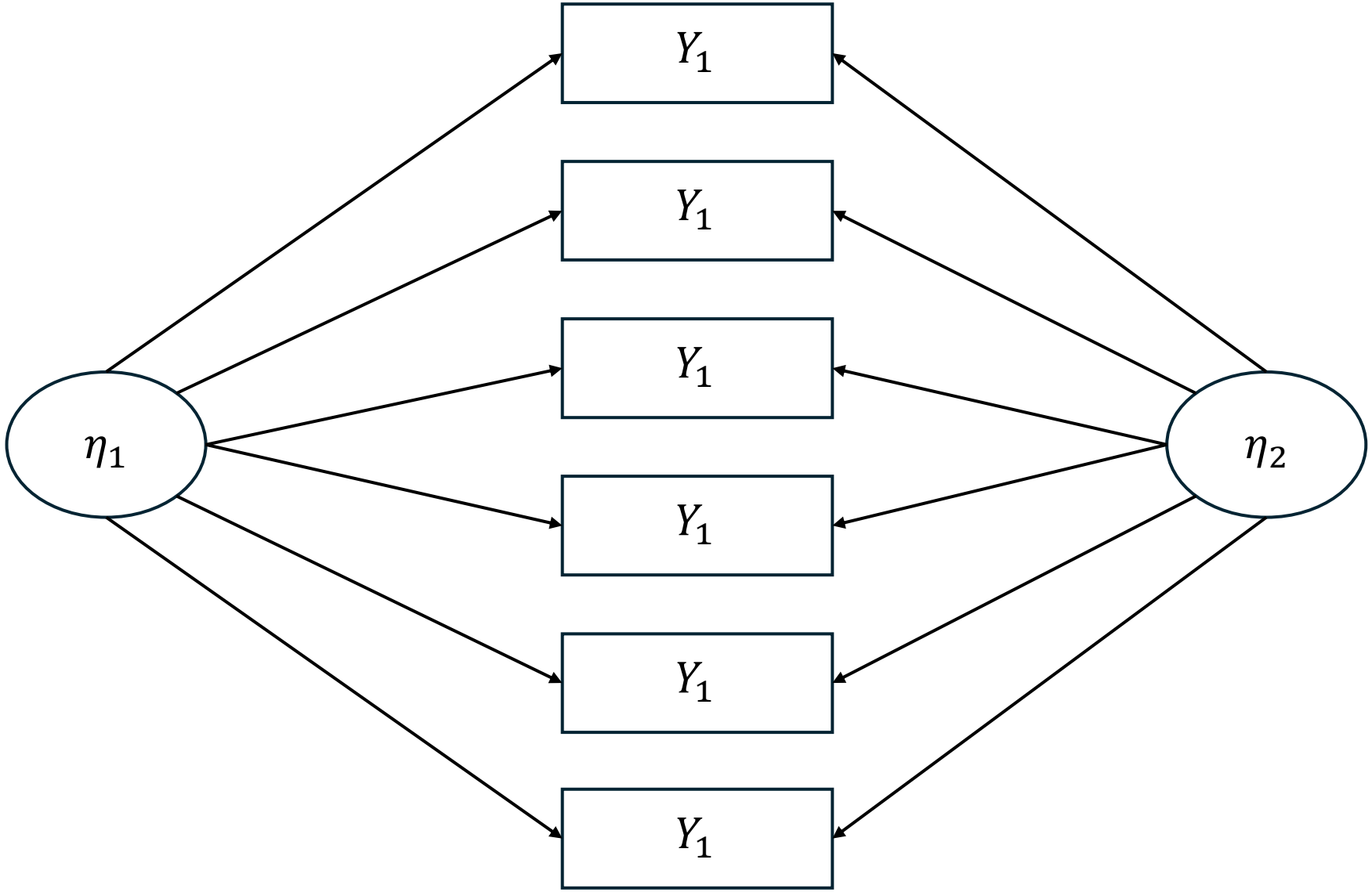


Photo by Paolo Nicolello on Unsplash


Y_1
Y_1
Y_1
Y_1
Y_1
Y_1
η_1
η_2

Factor Analysis

- Confirmatory Factor Analysis
- Exploratory Factor Analysis
- Principal Component Analysis (?)

Purpose

- Data reduction
- Find consistent **correlations** among groups of variables
- Confirm hypothesis of factor structures

EFA vs PCA

- Factors vs components. Factor cause variables
- PCA all variance is analyzed; in FA, only shared variance
- PCA is computationally simpler and does not produce improper solutions.
- They often produce similar results
- EFA generalizes better to CFA (and SEM) than EFA

Universidad Tecnológica de Pereira

MAESTRÍA EN INVESTIGACIÓN OPERATIVA Y ESTADÍSTICA

Thank you!

Orlando Sabogal-Cardona

@Antonio Sabogal

orlando.sabogal.20@ucl.ac.uk