Latent Variables: Reflective or Formative?

Advanced Master in Agricultural Economics and Policy

Giovanbattista Califano

University of Naples Federico II, giovanbattista.califano@unina.it

Attitudinal Scales for the Study
of Consumer Preferences
April 27, 2026

Roadmap

20/04	Hello, Psychometrics!	✓
23/04	The Questionnaire	✓
24/04	Reliability and Validity of a Measure	✓
27/04	Latent Variables: Reflective or Formative?	◯
30/04	A Bit of SEM	◯
07/05	Stata Stata Stata Stata Stata Stata Stata	◯

Recommended readings:

- Chapters 4 and 7 from Jhangiani et al.2019)
- Chapters 7, 10 and 14 from Olivero and Russo2022)
- Chapter 12 from Mehmetoglu and Jakobsen2022)

Jhangiani, R. S., Chiang, I. A., Cuttler, C., and Leighton, D. C. (2019). *Research Methods in Psychology*. Kwantlen Polytechnic University, 4th edition.

Mehmetoglu, M. and Jakobsen, T. G. (2022). *Applied Statistics Using Stata: A Guide for the Social Sciences*. SAGE Publications Ltd, 2nd edition.

Olivero, N. and Russo, V. (2022). *Psicologia dei Consumi*. McGraw-Hill Education, 3rd edition.

Moving on to...

Brief Recap

Latent Variables

Reflective Constructs

Formative Constructs

Brief recap

We have already seen how we often wish to study concepts that are *not directly observable and measurable*.

Through *scaling* techniques, however, we can measure the *reflections* (or *manifestations*) of the construct we are interested in studying.

Brief recap

For example

People who are more neophobic towards food technologies will tend to agree with these statements (precisely because they are more neophobic!)

FTNS 1: I don't need to eat food made using new technology because the food I already eat is fine

FTNS 2: New food technologies are something I am uncertain about

FTNS 3: New food products are not healthier than traditional foods

(...)

Strongly disagree	*Disagree*	*Neither agree nor disagree*	*Agree*	*Strongly agree*
1	2	3	4	5

Brief recap... and a problem?

The scores assigned to the responses for each item can therefore be aggregated, by sum or mean. The aggregated score represents, in this case, the participant's level of neophobia towards food technologies.
This classical approach is still widely used in the literature, but presents at least one problem...

Brief recap... and a problem?

The scores assigned to the responses for each item can therefore be aggregated, by sum or mean. The aggregated score represents, in this case, the participant's level of neophobia towards food technologies.
This classical approach is still widely used in the literature, but presents at least one problem...

It assumes that the items used for the aggregated score are a perfect measurement of the construct:

$$FTN_i = \frac{\sum_{j=1}^{k} FTNS_{ij}}{k}$$

In this average of items:

- No error is accounted for
- Items have equal weight in the aggregated score
- FTN is treated as a *consequence* of the responses to the scale items, but the causal direction should be reversed

Moving on to...

Brief Recap

Latent Variables

Reflective Constructs

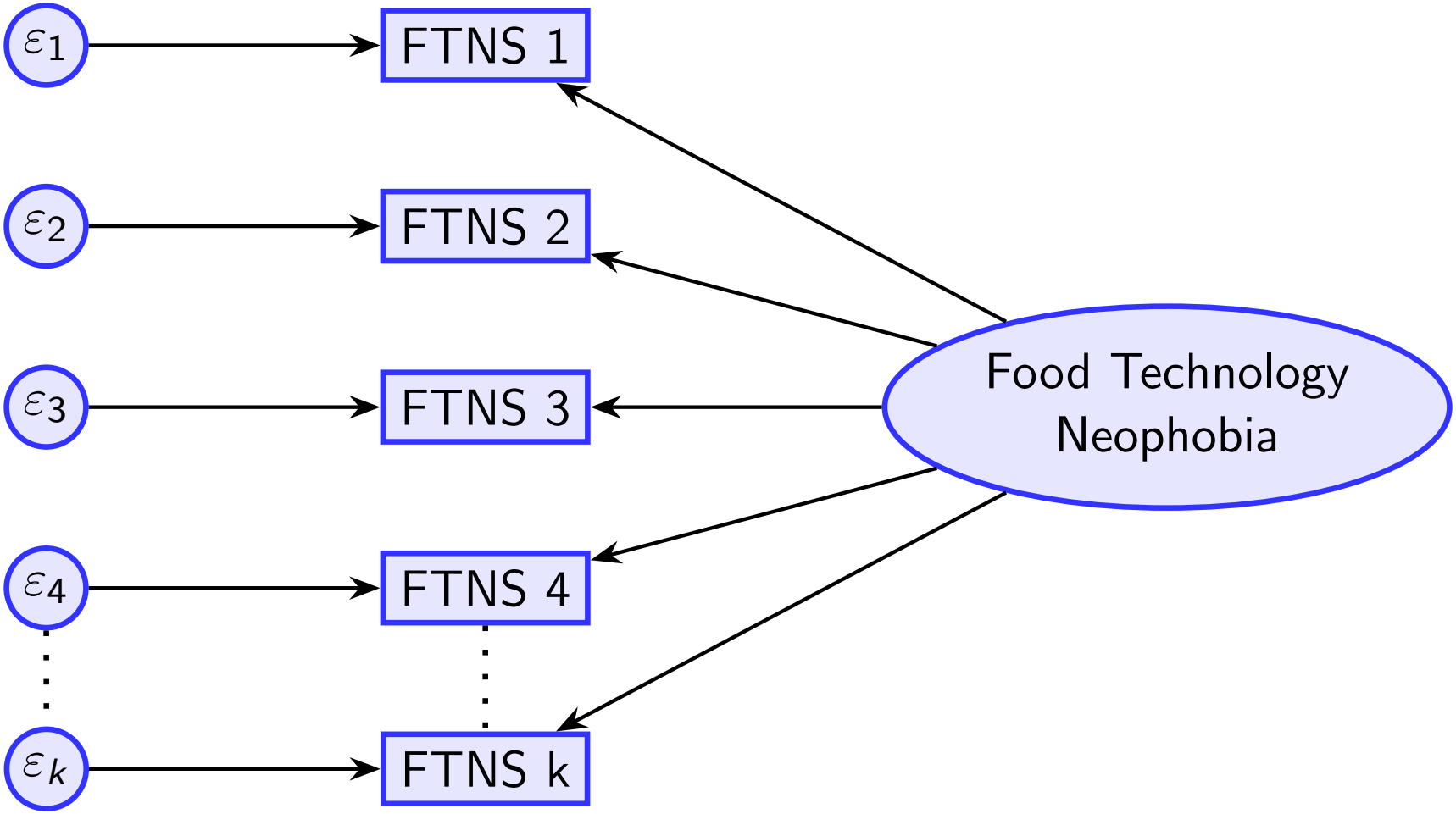
Formative Constructs

What we want

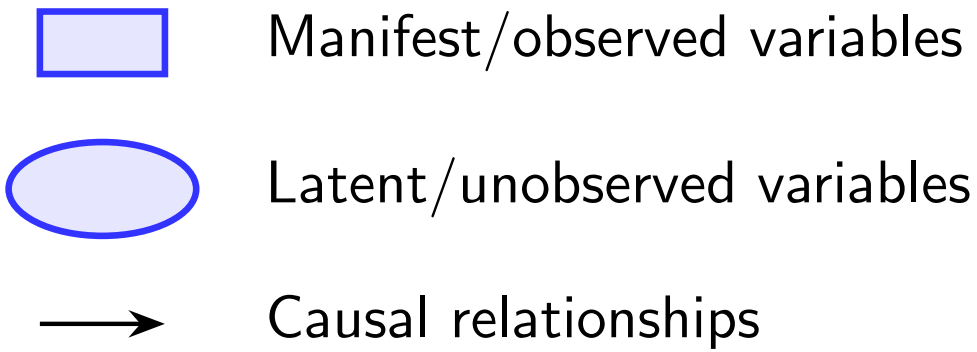
$$
\begin{aligned}
FTNS_1 &= \lambda_1 \cdot FTN + \varepsilon_1 \\
FTNS_2 &= \lambda_2 \cdot FTN + \varepsilon_2 \\
FTNS_3 &= \lambda_3 \cdot FTN + \varepsilon_3 \\
FTNS_4 &= \lambda_4 \cdot FTN + \varepsilon_4 \\
&\vdots \\
FTNS_k &= \lambda_k \cdot FTN + \varepsilon_k
\end{aligned}
$$

- The direction of causality is from the construct (FTN) to the individual items/indicators $FTNS_j$, $j = 1, \ldots, k$.
- The observed variables $FTNS_j$ are manifestations of the construct, each subject to measurement error ε_j, $j = 1, \ldots, k$.
- The coefficients λ_j are the factor loadings that measure the association between the latent factor FTN and item j.

In SEM language



In SEM language



Latent variables: Reflective vs Formative

Reflective constructs

Until now we have spoken of **reflective** latent variables: they refer to unobservable theoretical concepts (such as attitudes, emotions, etc.) that are assumed to exist independently of the scientific investigation.

But not all latent variables are reflective in nature...

Latent variables: Reflective vs Formative

Reflective constructs

Until now we have spoken of **reflective** latent variables: they refer to unobservable theoretical concepts (such as attitudes, emotions, etc.) that are assumed to exist independently of the scientific investigation.

But not all latent variables are reflective in nature...

Formative constructs

Formative constructs *emerge* from reality as composite variables obtained from a linear combination of multiple observed variables. They are never natural phenomena, but artificial constructions created to achieve specific purposes (such as quality indices, satisfaction, etc.)

Moving on to...

Brief Recap

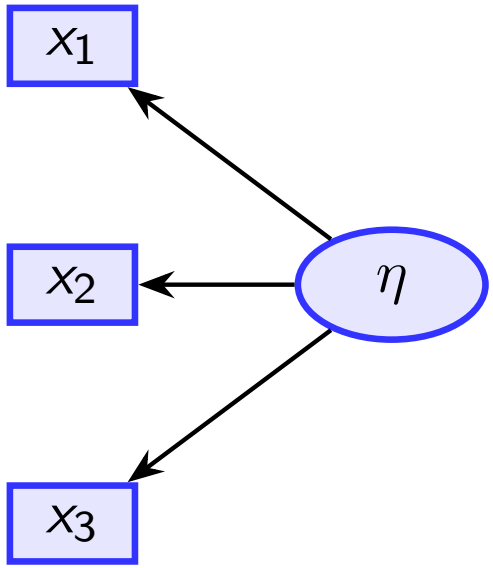
Latent Variables

Reflective Constructs

Formative Constructs

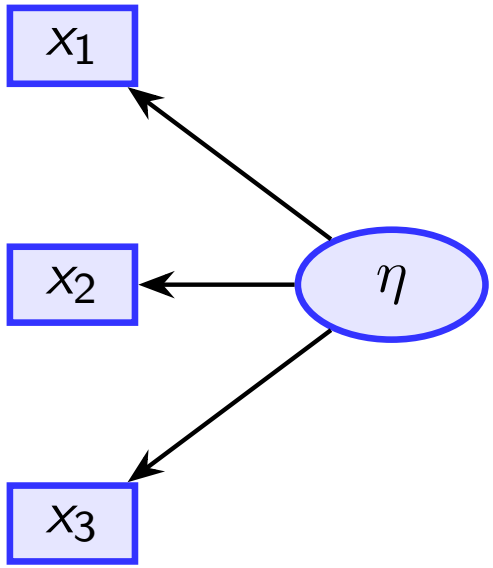
Reflective latent variables

$$x_1 = \lambda_1 \cdot \eta + \varepsilon_1$$
$$x_2 = \lambda_2 \cdot \eta + \varepsilon_2$$
$$x_3 = \lambda_3 \cdot \eta + \varepsilon_3$$



Reflective latent variables

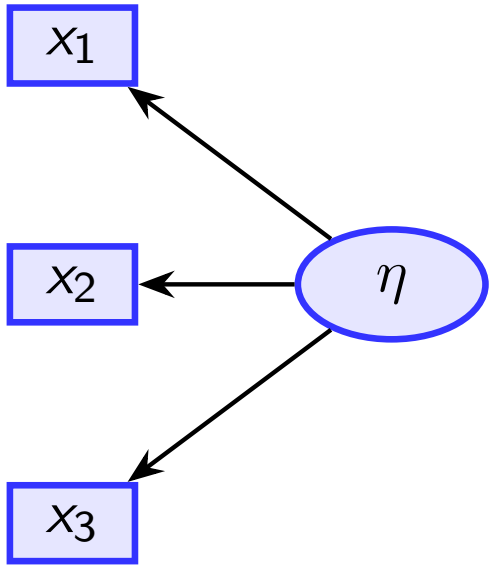
$$x_1 = \lambda_1 \cdot \eta + \varepsilon_1$$
$$x_2 = \lambda_2 \cdot \eta + \varepsilon_2$$
$$x_3 = \lambda_3 \cdot \eta + \varepsilon_3$$



- Variations in the latent variable η manifest as changes in **every** observed variable x.
- The direction of causality is from the construct to the indicators.
- Each observed variable is treated as a manifestation, affected by a specific error, of the latent variable.

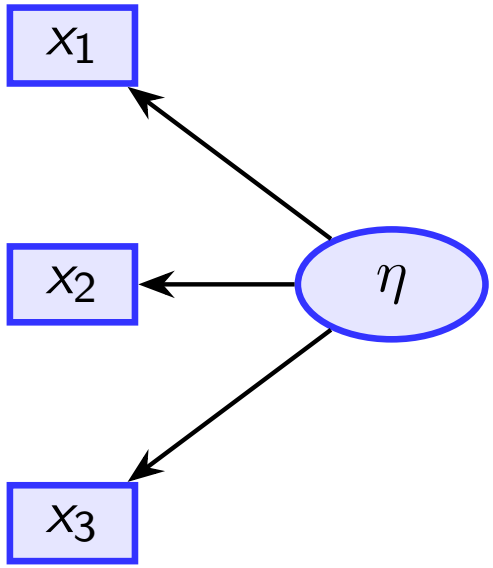
Reflective latent variables

$$x_1 = \lambda_1 \cdot \eta + \varepsilon_1$$
$$x_2 = \lambda_2 \cdot \eta + \varepsilon_2$$
$$x_3 = \lambda_3 \cdot \eta + \varepsilon_3$$



Reflective latent variables

$$x_1 = \lambda_1 \cdot \eta + \varepsilon_1$$
$$x_2 = \lambda_2 \cdot \eta + \varepsilon_2$$
$$x_3 = \lambda_3 \cdot \eta + \varepsilon_3$$



- The indicators of a reflective latent variable are *interchangeable* and we expect a high correlation among them, as they should all be good approximations of the latent construct.
- We cannot assign a value to the latent variable, but we can study its relationships with other latent or manifest variables.
- The measurement errors ε allow us to account for the portion of variability in the indicators not explained by the latent factor.

Confirmatory Factor Analysis (CFA)

Confirmatory factor analysis allows us to verify whether the hypothesised *reflective* factor structure is supported by the observed data, thus enabling researchers to validate hypotheses about the relationships between observed variables and one or more latent factors.

Confirmatory Factor Analysis (CFA)

Confirmatory factor analysis allows us to verify whether the hypothesised *reflective* factor structure is supported by the observed data, thus enabling researchers to validate hypotheses about the relationships between observed variables and one or more latent factors.

Various indices are used to evaluate the goodness of fit of a CFA, among the most common:

Fit Index	**Desired Value**
χ^2 (Chi-square)	Close to 0
CFI (Comparative Fit Index)	> 0.95 (good), > 0.90 (acceptable)
TLI (Tucker-Lewis Index)	> 0.95 (good), > 0.90 (acceptable)
RMSEA (Root Mean Square Error of Approximation)	< 0.05 (good), < 0.08 (acceptable)
SRMR (Standardized Root Mean Square Residual)	< 0.08

Moving on to...

Brief Recap

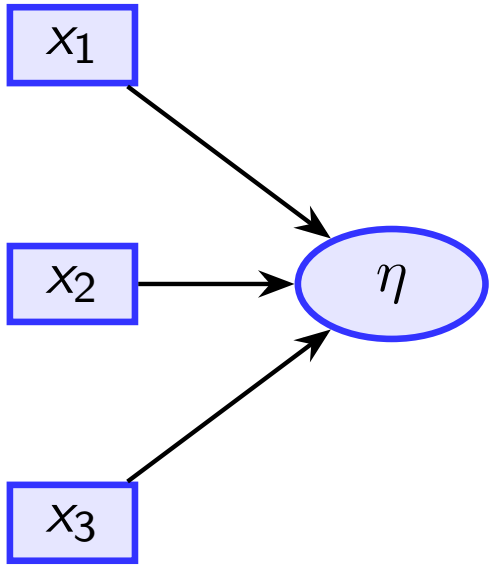
Latent Variables

Reflective Constructs

Formative Constructs

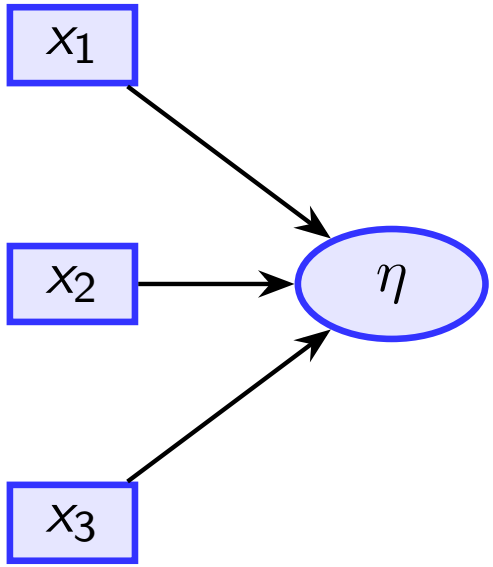
Formative latent variables

$$\eta = \sum_{j=1}^{J} w_j \cdot x_j$$



Formative latent variables

$$\eta = \sum_{j=1}^{J} w_j \cdot x_j$$



- The relationship between manifest and latent variables is not to be understood as causal, but rather as prescriptive.
- These latent variables are **formed** by the manifest variables, so their definition depends on the indicators used.
- Indicators are specific ingredients: they are not interchangeable and we do not expect a high correlation among them: if we change indicators, the very meaning of the formative latent variable changes.

Examples of formative constructs

Socioeconomic Status (SES)

1. Income
2. Education level
3. Occupation

Hotel service quality

1. Staff courtesy
2. Room cleanliness
3. Breakfast quality
4. Check-in efficiency

- If the indicators change, the definition of the construct changes.
- Indicators are not necessarily intercorrelated.