FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

SEM WITH EXPERIMENTAL DATA

CARMA, Bert Weijters, Oct 18, 2024

GHENT
UNIVERSITY

ABSTRACT

Structural Equation Modeling is a popular analytic approach but remains underutilized when working with experimental data. This talk discusses some of the advantages of using SEM for experimental data and provides some recommendations on when and how to apply SEM for between-subject, within-subject and mixed designs using dependent variables that are measured with one or many indicators, and including mediation and/or moderation.

GHENT UNIVERSITY

OVERVIEW

1. Between-subjects experiments
    a) Measurement and scaling
    b) Main effects
    c) Interaction effects
2. Mediation
    a) Common Method Variance (CMV)
    b) Accounting for CMV in measurement
3. Within-subject and mixed experiments
    a) Unconditional SEMWISE
    b) Conditional SEMWISE

GHENT UNIVERSITY

PART 1: BETWEEN-SUBJECTS EXPERIMENTS

GHENT UNIVERSITY

OVERVIEW

1. Between-subjects experiments
   a) Measurement and scaling
   b) Main effects
   c) Interaction effects
2. Mediation
   a) Common Method Variance (CMV)
   b) Accounting for CMV in measurement
3. Within-subject and mixed experiments
   a) Unconditional SEMWISE
   b) Conditional SEMWISE

GHENT UNIVERSITY

MODELING CONSIDERATIONS

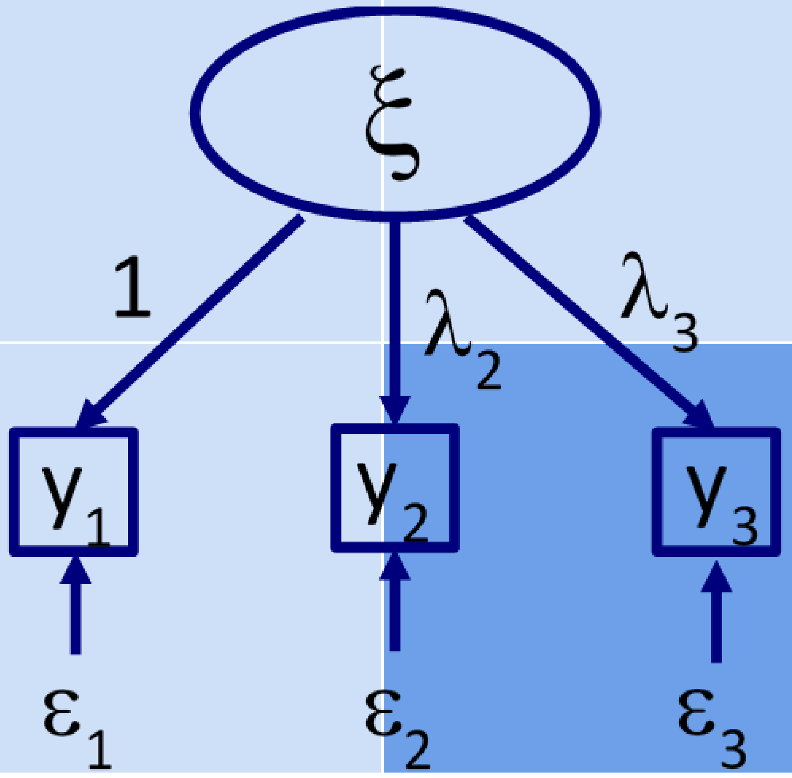
1. Type of indicators
2. Type of measurement model
3. Number of indicators

CLASSIFICATION OF OBSERVED VARIABLES

	Nominal scales	Ordinal scales	Metric scales
Discrete variables	Gender identity measured as 1 = male, 2 = female, 3 = transgender, and 4 = do not identify with a particular gender.	Extent of (dis)agreement measured on a 5-point scale (e.g., 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree).	Number of coupons redeemed during the last trip to the supermarket.
Continuous variables	n.a.	Degree of liking measured on a 0 to 100 slider scale	Response time

FOCUS : DV USING REFLECTIVE MODEL FOR METRIC CONTINUOUS VARIABLES (BUT APPLIED MORE BROADLY)

	Nominal scales	**Ordinal scales**	**Metric scales**
Discrete variables			
Continuous variables			



TYPES OF STANDARDIZED OUTPUT

1. StdYX (‘Completely standardized’)
   - Uses variances of continuous latent, background, and outcome variables.
   - Applicable in standard linear regression analysis.

2. StdY
   - Uses variances of continuous latent and outcome variables.
   - Best suited for binary covariates; interprets change in y for a unit change in x.

3. Std
   - Uses variances of continuous latent variables.
   - Standardizes covariances and residual covariances based on their respective variances.

GHENT UNIVERSITY

WHICH STANDARDIZATION APPROACH TO USE?

My personal decision tree:

- If the units of the variables have meaning (e.g., seconds, meters, EURO, percentage, 100-point scale,…) use unstandardized estimates
- If the Independent variable is a dummy (0/1 or -1/+1), but the dependent variables are latent, use STDY
- If all are latent, use STDYX

NUMBER OF INDICATORS

- Single vs. multi-indicator measurement
- Sum/mean score
- Parceling

SINGLE VS. MULTI-ITEM MEASURES

Single-item measures

- if “completely concrete construct”, i.e. one for which both the object of measurement and the attribute to be measured are concrete (easily and uniformly imagined)
- Example: attitude towards the ad / brand

Bergkvist and Rossiter (2007)

Multi-item measures

- allow to assess reliability/internal consistency
- allow to correct for measurement error
- better predictive validity (even for concrete constructs)

Kamakura (2015)

GHENT UNIVERSITY

NUMBER OF INDICATORS

- Single vs. multi-indicator measurement
- Sum/mean score
- Parceling

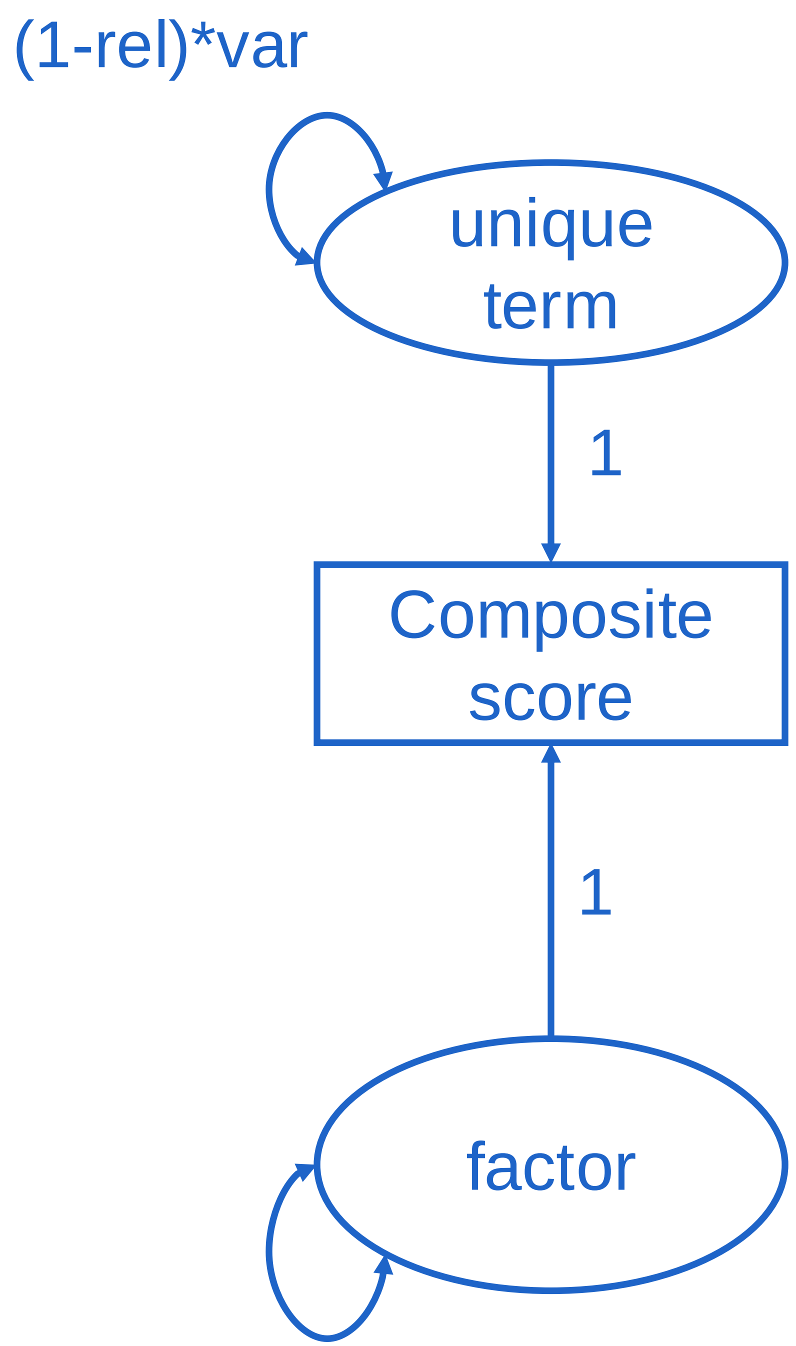
COLLAPSING MULTI-ITEM MEASURES INTO A SINGLE OVERALL COMPOSITE

- Averaging individual items will usually result in more reliable and valid assessments of the intended construct (compared to single-item measures)
- BUT
  - only for well-validated scales, or
  - after a careful measurement analysis
- Coefficient alpha is not a substitute for measurement analysis
- Best to account for unreliability by means of a measurement model

SINGLE-INDICATOR

Correct for measurement error by

1. using an average of the available items as a single indicator of the underlying construct (after establishing unidimensionality)
2. fixing the factor loading to one
3. setting the unique variance to one minus the reliability of the composite (e.g., based on coefficient alpha) multiplied by the variance of the composite
4. freely estimating the factor variance



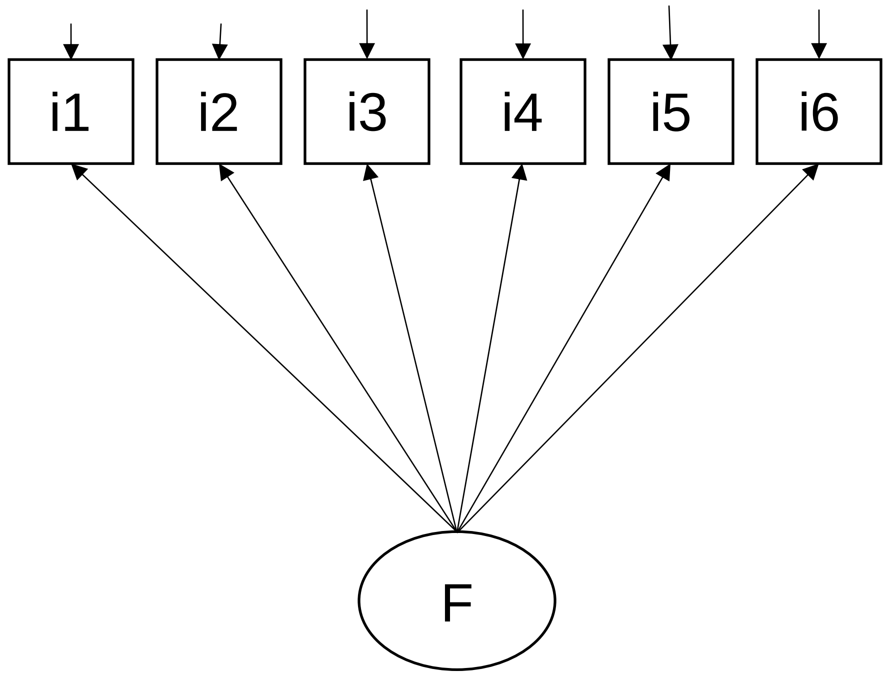
GHENT UNIVERSITY

NUMBER OF INDICATORS

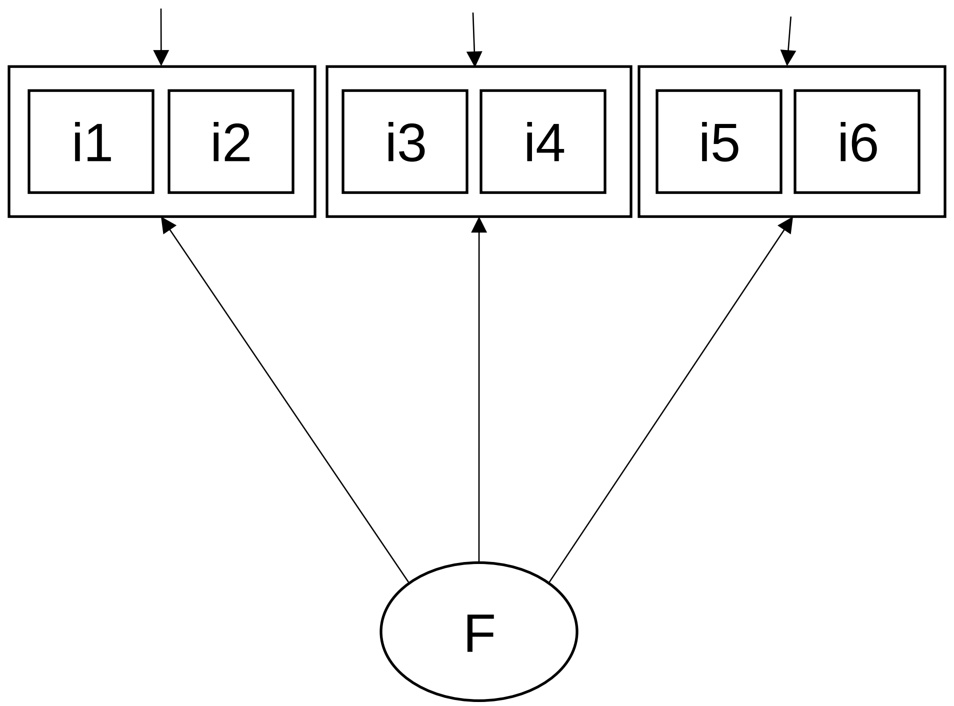
– Single vs. multi-indicator measurement
– Sum/mean score
– Parceling

GHENT UNIVERSITY

ITEM PARCELING



CFA using items



CFA using item parcels

Parcels of items = indicators created by summing or averaging subsets of individual items within scales or subscales (Holt, 2004)

- simpler and better-fitting models
- improves the variable to sample size ratio
- better distributional properties of the aggregated items
- increased reliability of resulting indicators
- more stable parameter estimates (Bandalos & Finney, 2001)
- Not okay for scale construction/validation

GHENT UNIVERSITY

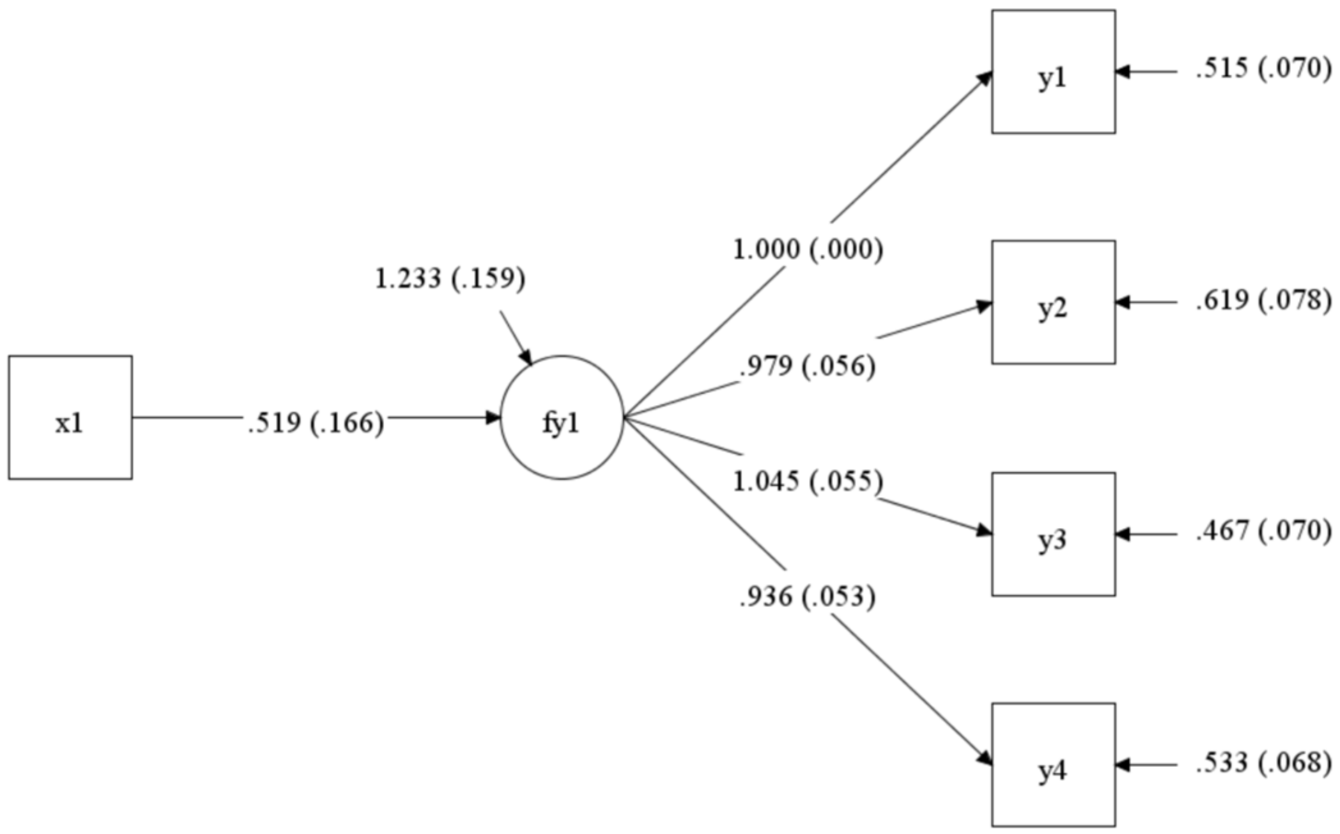
OVERVIEW

1. Between-subjects experiments
   a) Measurement and scaling
   b) Main effects
   c) Interaction effects
2. Mediation
   a) Common Method Variance (CMV)
   b) Accounting for CMV in measurement
3. Within-subject and mixed experiments
   a) Unconditional SEMWISE
   b) Conditional SEMWISE

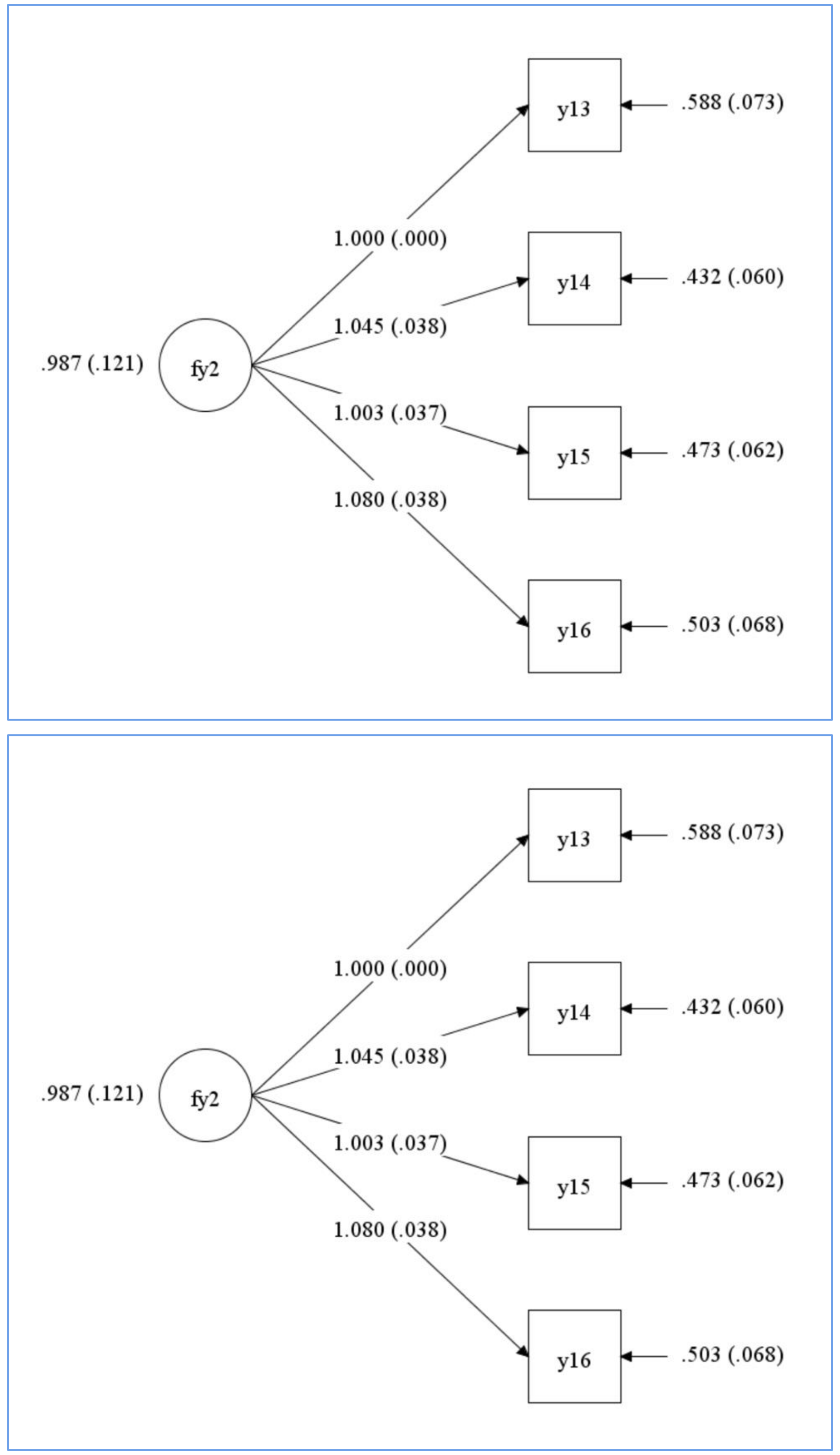
GHENT UNIVERSITY

HOW TO MODEL EXPERIMENTAL EFFECTS

Dummy covariate:



Multi-group model:



GHENT UNIVERSITY

HOW TO MODEL EXPERIMENTAL EFFECTS

Dummy covariate:

- Like linear regression
- Two conditions
- Parsimonious
- Measurement invariance assumed
- Homoscedasticity assumed
- Difference in means across conditions

Multi-group model:

- Like t-tests / ANOVA
- Any number of conditions
- Flexible
- Measurement invariance can be tested
- Heteroscedasticity okay
- Differences in any parameter: mean, intercept, loading, (residual) variance

GHENT UNIVERSITY

PLANNED MEAN COMPARISONS

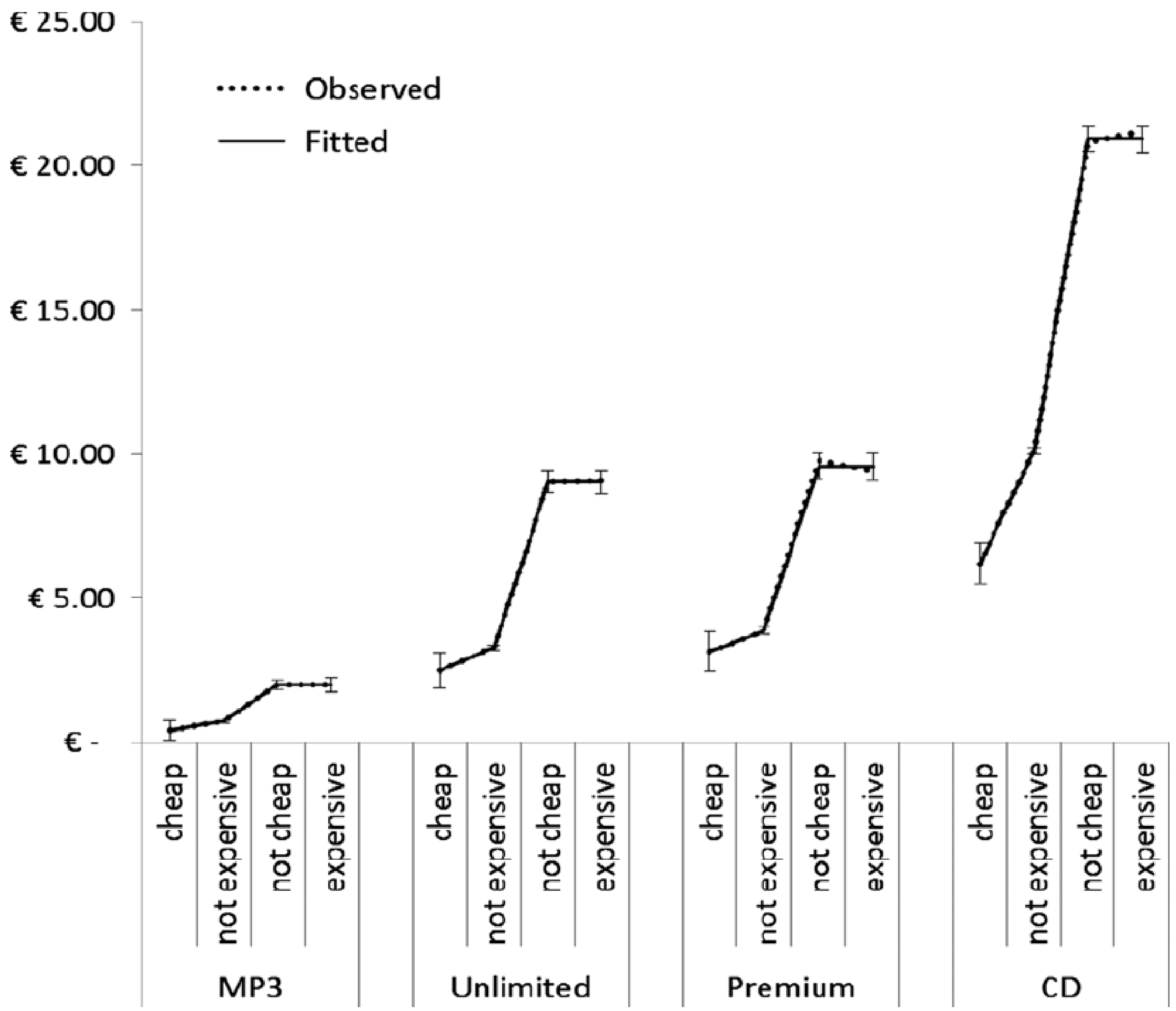


Figure 2. Study 1A: Mean price level by Generic Price term. *Note.* Error bars refer to standard errors (fitted). Mean-centered reference price is included as a control variable.

Table 1
Chi-square and BIC values by Interpretational Pattern (Studies 1A, 1B, and 5)

			BIC	χ^2	*df*	*p*	
Study 1	Study 1A	Null model (all means equal)	7,069.7	402.9	12	.000	
		Dual fusion	6,725.8	35.1	8	.000	
		Dual mitigation	6,738.4	0.0	0	N.A.	
		Lower fusion	6,748.4	33.9	4	.000	
		Upper fusion	**6,715.7**	1.2	4	.876	NS

GHENT UNIVERSITY

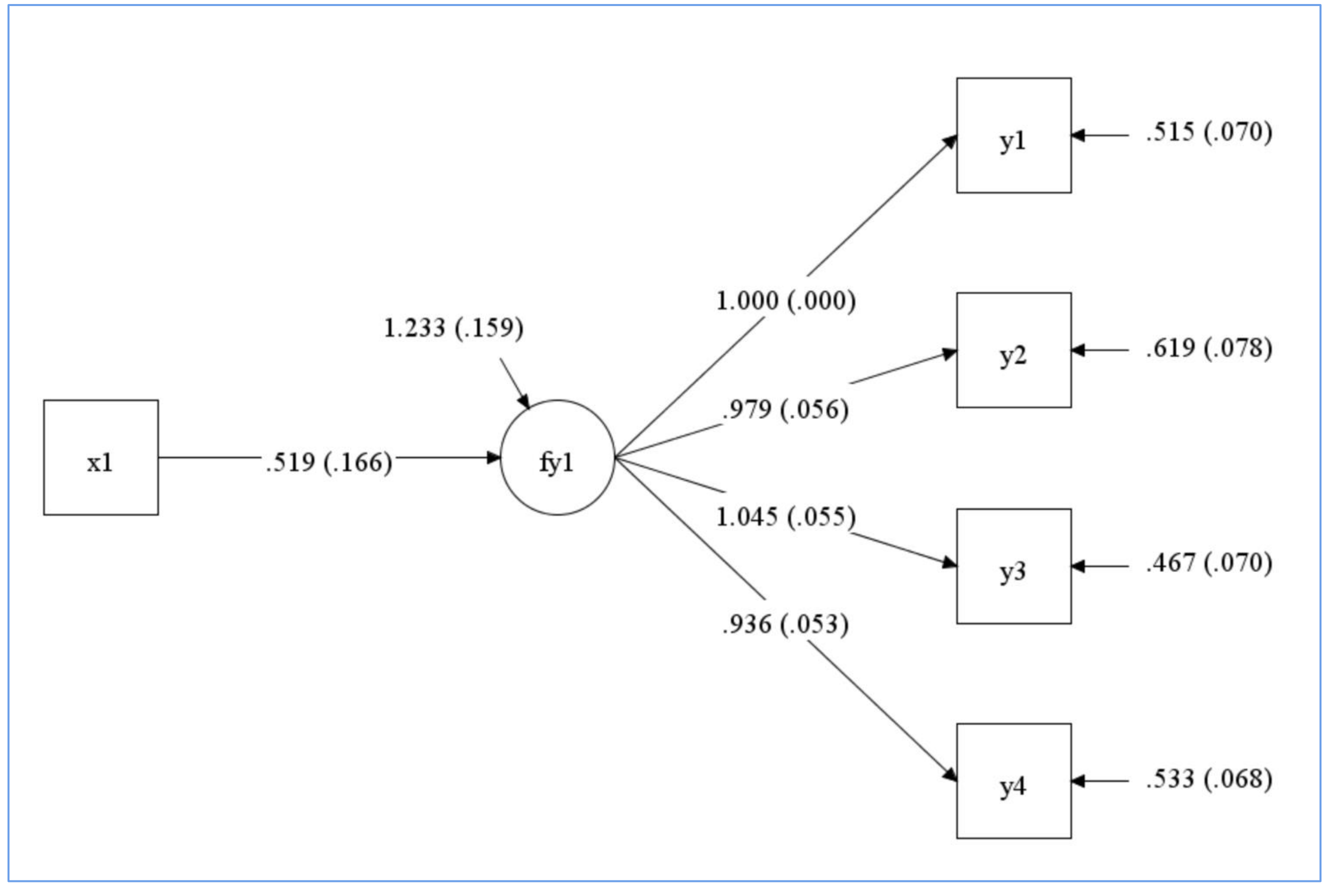
OVERVIEW

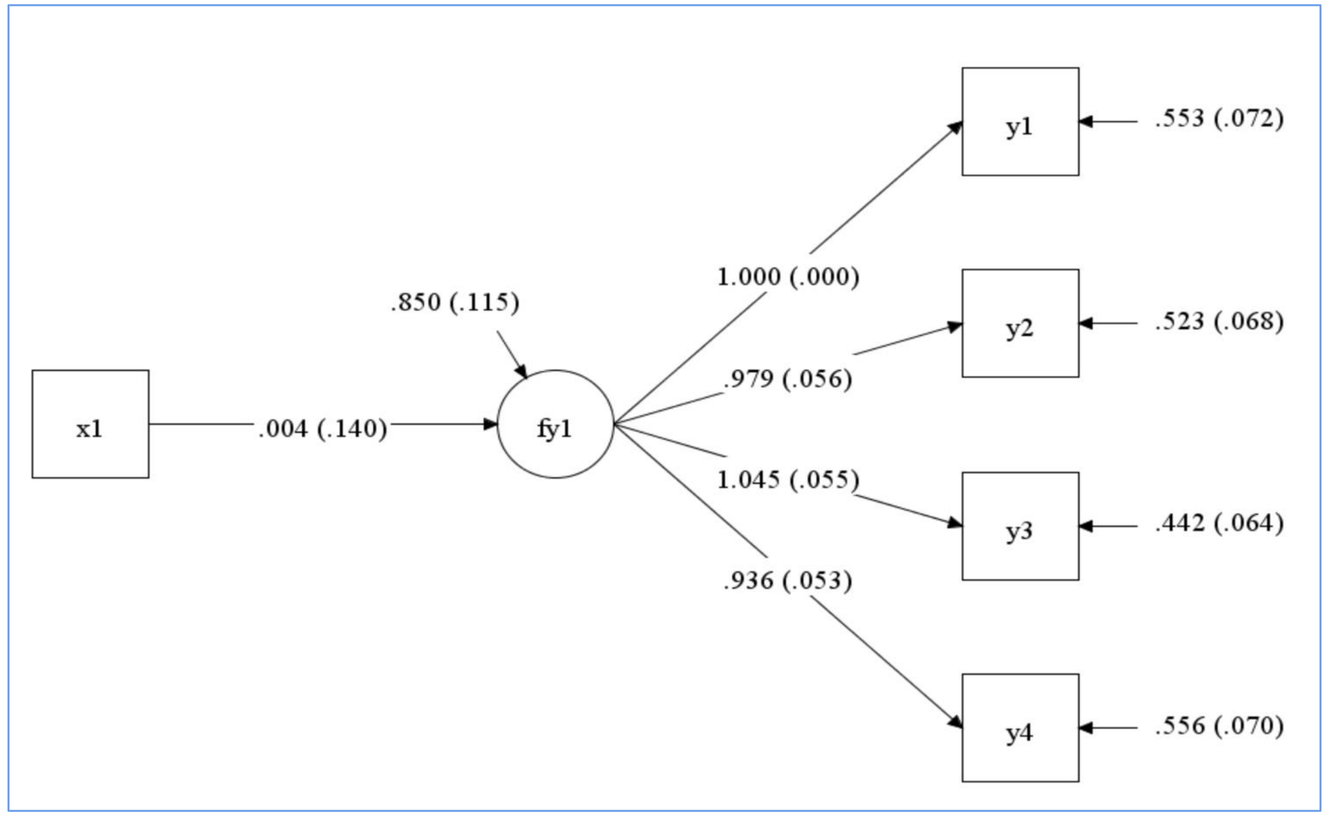
1. Between-subjects experiments
   a) Measurement and scaling
   b) Main effects
   c) Interaction effects
2. Mediation
   a) Common Method Variance (CMV)
   b) Accounting for CMV in measurement
3. Within-subject and mixed experiments
   a) Unconditional SEMWISE
   b) Conditional SEMWISE

GHENT UNIVERSITY

INTERACTION: MULTI-GROUP SEM

- Intuitive interpretation: effect of x in group 1 v group 2
- Flexible in terms of assumptions
- Invariance testing





GHENT UNIVERSITY

PART 2: MEDIATION

GHENT UNIVERSITY

OVERVIEW

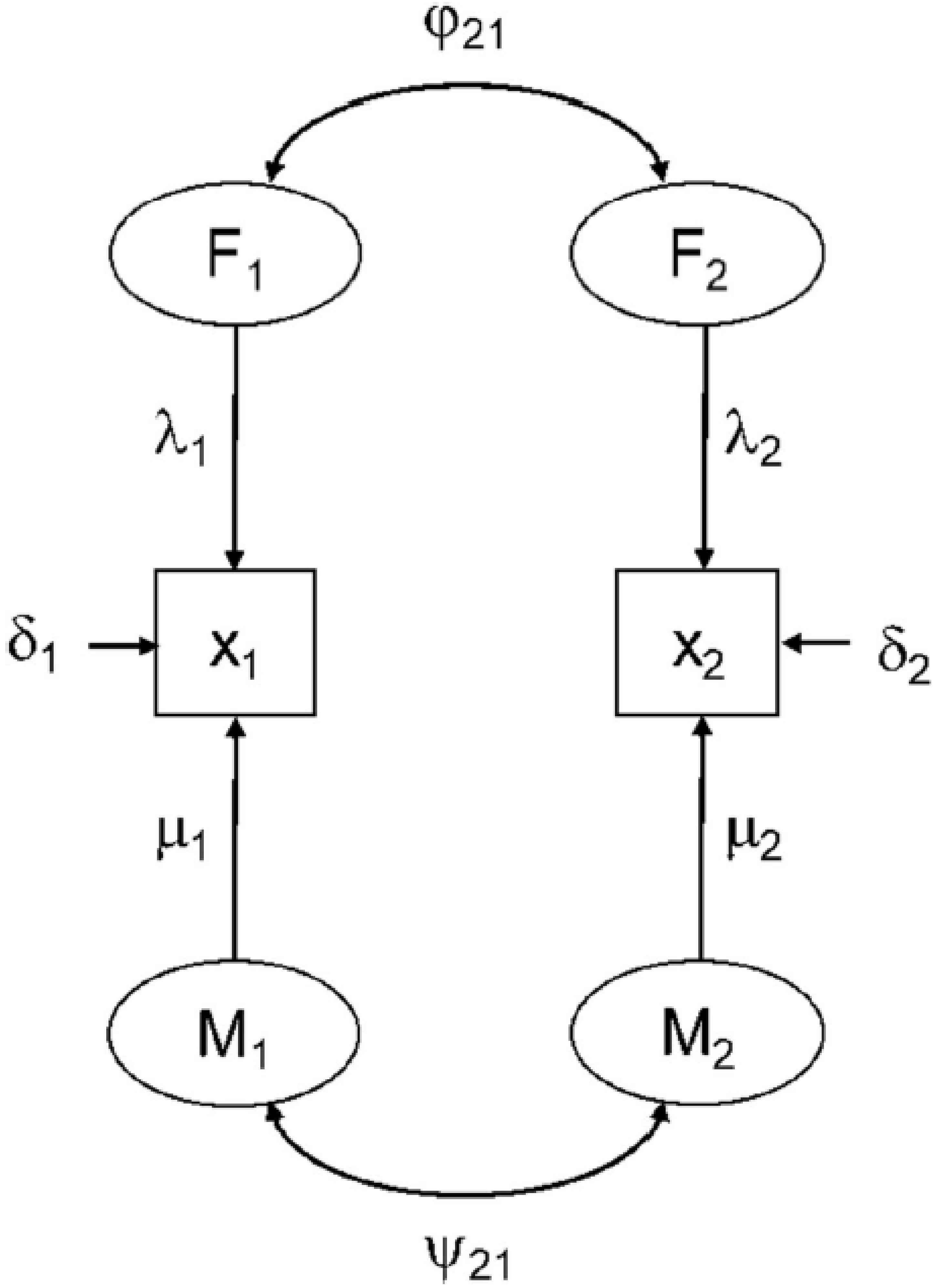
1. Between-subjects experiments
   a) Measurement and scaling
   b) Main effects
   c) Interaction effects
2. Mediation
   a) Common Method Variance (CMV)
   b) Accounting for CMV in measurement
3. Within-subject and mixed experiments
   a) Unconditional SEMWISE
   b) Conditional SEMWISE

COMMON METHOD VARIANCE

- Definition: CMV refers to systematic errors in data due to the measurement method used, rather than true variance in the construct.
- Context: CMV is a well-known concern in marketing, psychology, and management research, particularly in survey-based studies.
- Key Issue: Misconceptions persist about CMV's biasing effects, including the belief that it is either negligible or easy to detect.

GHENT UNIVERSITY

Baumgartner, H., Weijters, B., & Pieters, R. (2021). The biasing effect of common method variance: Some clarifications. *Journal of the Academy of Marketing Science*, *49*, 221-235.

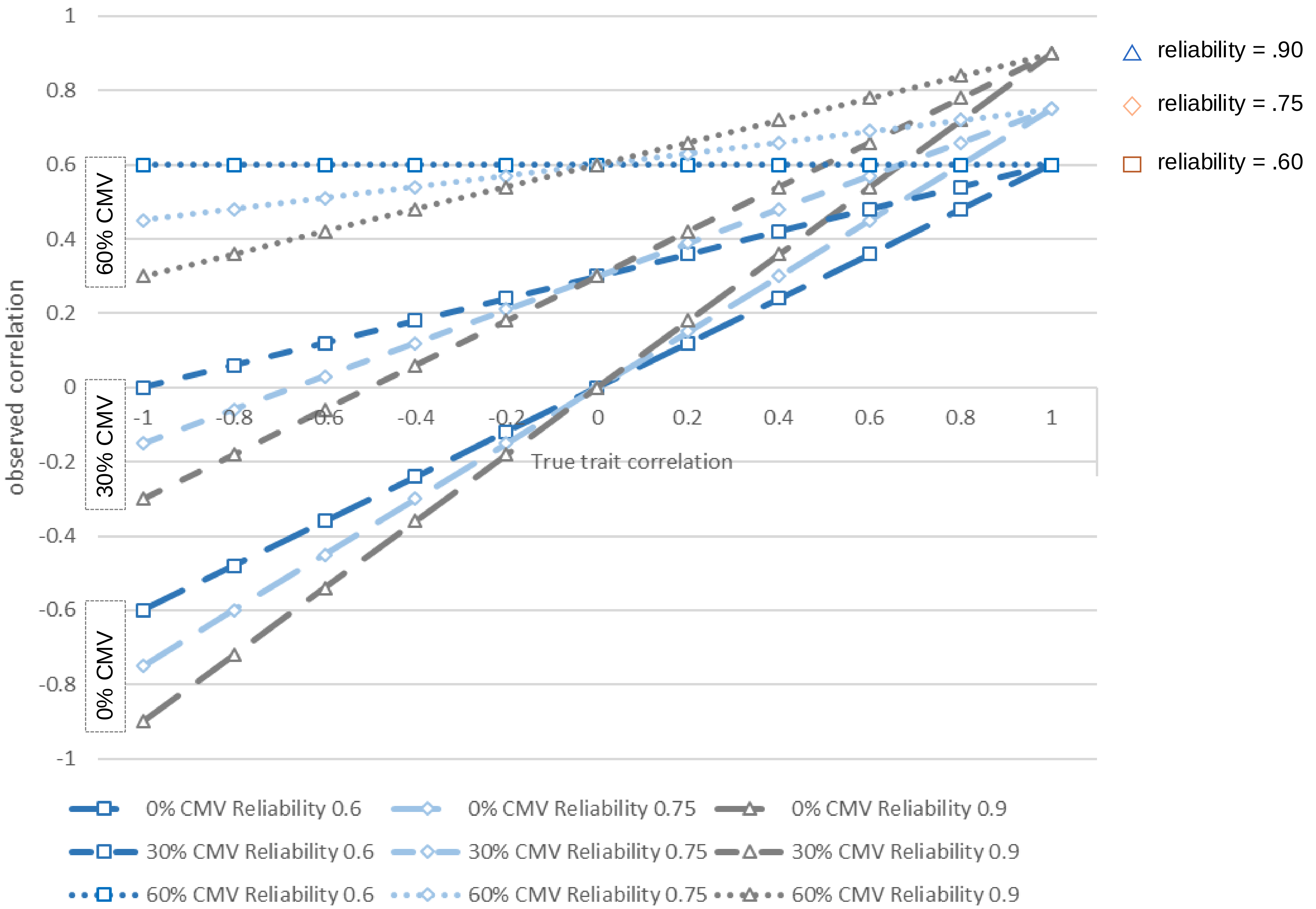


GHENT
UNIVERSITY

IMPACT OF COMMON METHOD VARIANCE

– Effect on Correlations: CMV can either inflate or deflate observed correlations between variables, depending on the sign and magnitude of the trait and method correlations.
– Negative Trait Correlations: CMV tends to make negative correlations less negative or even positive, which can lead to significant distortion.
– Positive Trait Correlations: CMV may amplify positive correlations but is counterbalanced by measurement unreliability (to a usually unknown extent).

Baumgartner, H., Weijters, B., & Pieters, R. (2021). The biasing effect of common method variance: Some clarifications. *Journal of the Academy of Marketing Science*, *49*, 221-235.



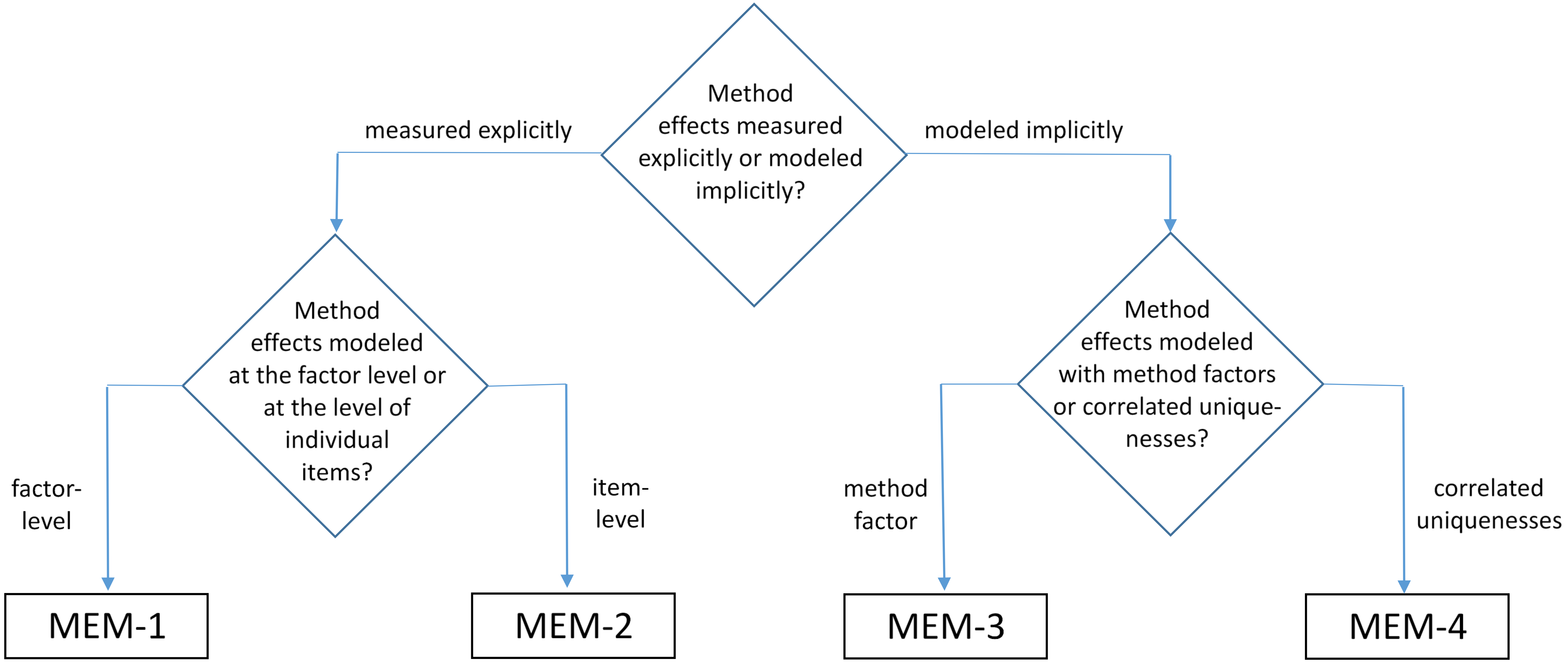
observed correlation
True trait correlation
60% CMV
30% CMV
0% CMV
reliability = .90
reliability = .75
reliability = .60
0% CMV Reliability 0.6
0% CMV Reliability 0.75
0% CMV Reliability 0.9
30% CMV Reliability 0.6
30% CMV Reliability 0.75
30% CMV Reliability 0.9
60% CMV Reliability 0.6
60% CMV Reliability 0.75
60% CMV Reliability 0.9

OVERVIEW

1. Between-subjects experiments
   a) Measurement and scaling
   b) Main effects
   c) Interaction effects
2. Mediation
   a) Common Method Variance (CMV)
   b) Accounting for CMV in measurement
3. Within-subject and mixed experiments
   a) Unconditional SEMWISE
   b) Conditional SEMWISE

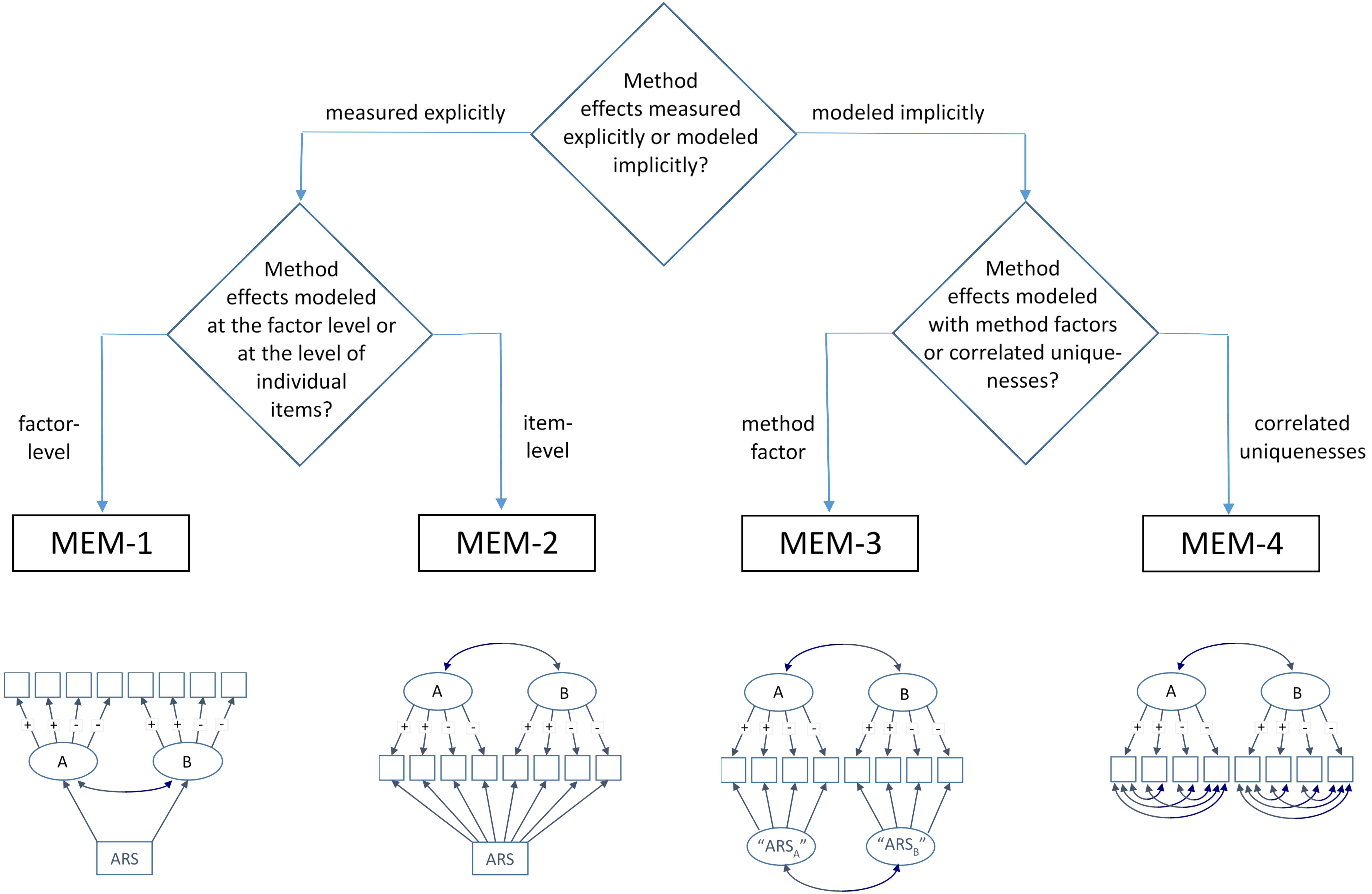
GHENT UNIVERSITY

CLASSIFICATION OF METHOD EFFECT MODELS



GHENT UNIVERSITY

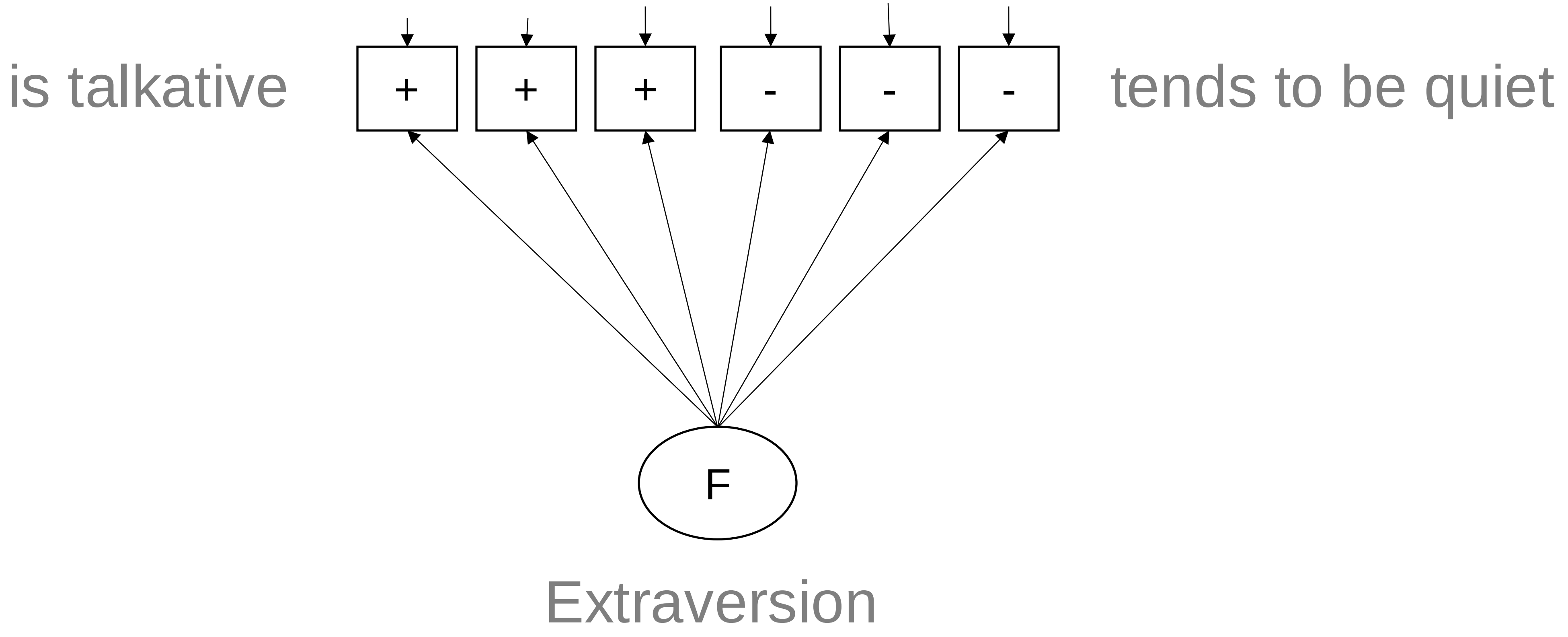
Baumgartner, H., & Weijters, B. (2019). Measurement in marketing. *Foundations and Trends® in Marketing*, *12*(4), 278-400.



Baumgartner, H., & Weijters, B. (2019). Measurement in marketing. *Foundations and Trends® in Marketing*, *12*(4), 278-400.

GHENT UNIVERSITY

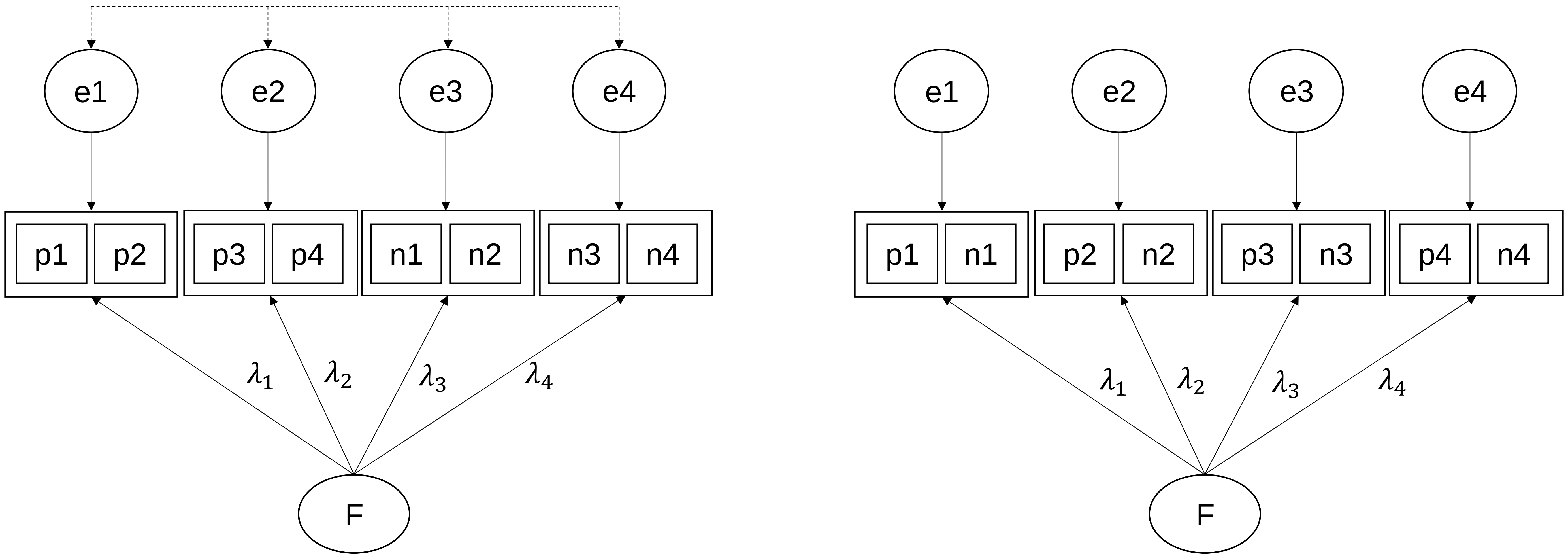
BUT WHAT IF THE SCALE IS BALANCED?



- Many scales in organizational research are balanced = consist of equal numbers of regular vs. reversed items
- Balancing
  - counters inattention and acquiescence
  - but messes up factor structures
- Strategic parceling to the rescue!

GHENT UNIVERSITY

Weijters, B., & Baumgartner, H. (2022). On the use of balanced item parceling to counter acquiescence bias in structural equation models. Organizational Research Methods, 25(1), 170-180.



A. Isolated parceling

B. Balanced parceling

Simulation with 1000 replications, substantive factor accounting for 81 percent of the total variance in the individual items and the method and unique factors for 3 and 16%, resp.

N	ParcelAlloc	chisq_m	df	tli_m	cfi_m	rmsea_m	srmr_m
200	iso	60.30	2	0.828	0.943	0.379	0.0248
200	bal	1.99	2	1.00	0.999	0.0228	0.0025

GHENT UNIVERSITY

Weijters, B., & Baumgartner, H. (2022). On the use of balanced item parceling to counter acquiescence bias in structural equation models. Organizational Research Methods, 25(1), 170-180.

SEMWISE

PART 3:

STRUCTURAL **E**QUATION **M**ODELING

FOR **WI**THIN-**S**UBJECT **E**XPERIMENTS

GHENT UNIVERSITY

Weijters, B., & Baumgartner, H. (2019). Analyzing Policy Capturing Data Using Structural Equation Modeling for Within-Subject Experiments (SEMWISE). *Organizational Research Methods*, *22*(3), 623-648.

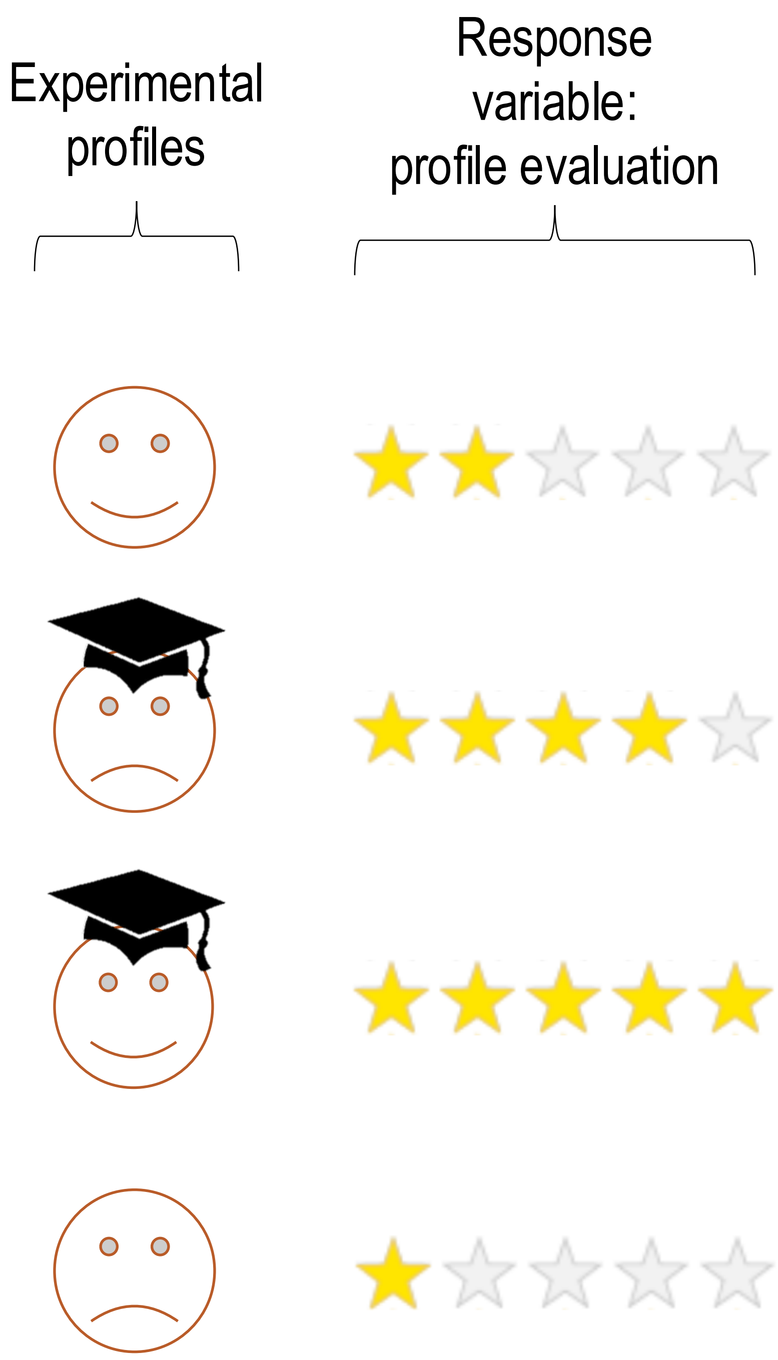
OVERVIEW

1. Between-subjects experiments
   a) Measurement and scaling
   b) Main effects
   c) Interaction effects
2. Mediation
   a) Common Method Variance (CMV)
   b) Accounting for CMV in measurement
3. Within-subject and mixed experiments
   a) Unconditional SEMWISE
   b) Conditional SEMWISE

WHO HAS USED...

- Conjoint analysis
- Policy capturing
- Factorial survey designs

STYLIZED EXAMPLE:
WARM (LO/HI) X COMPETENT (LO/HI)



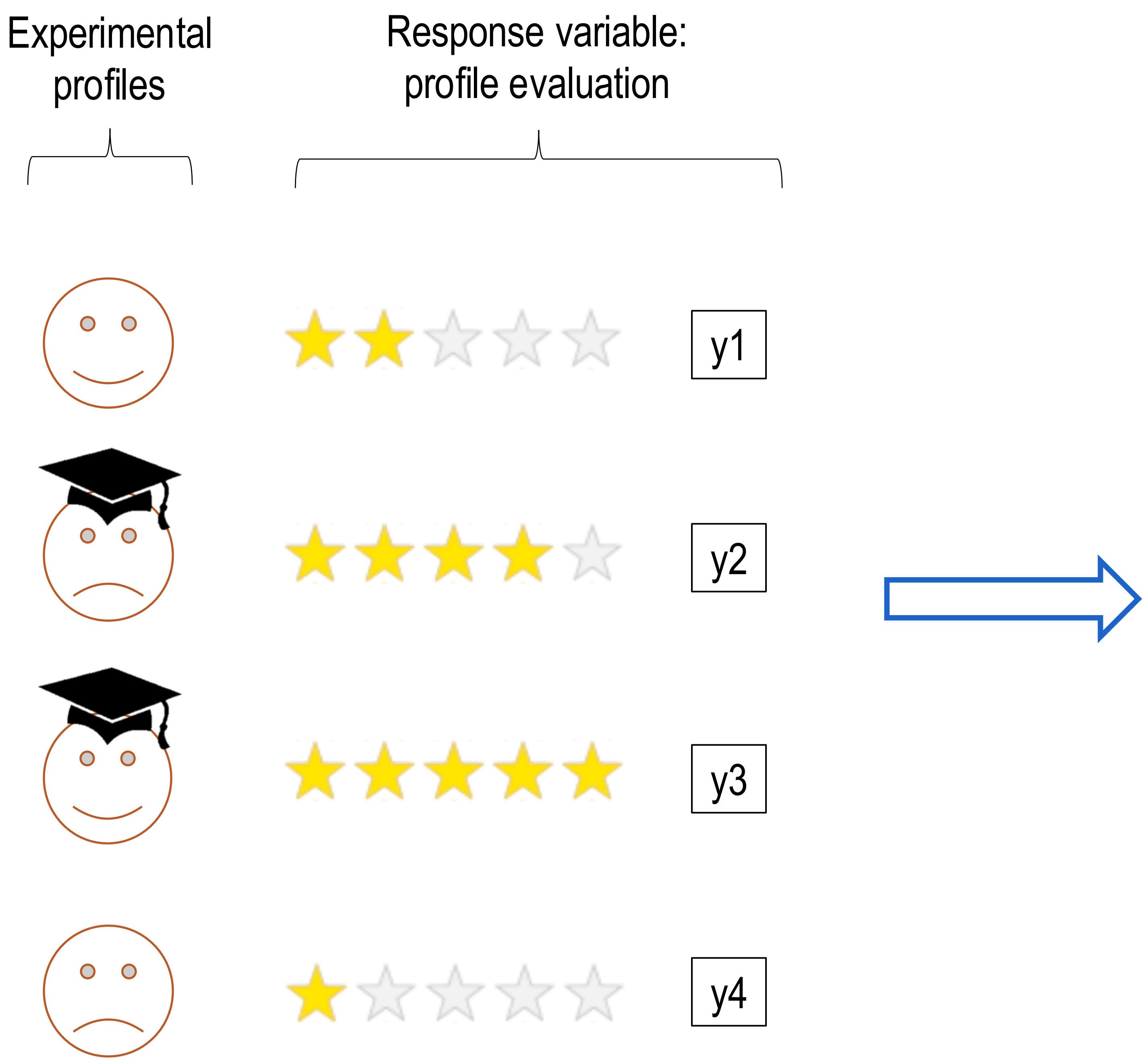
"long format"

	Profile	Warm	Competent	Rating
Respondent 1	1	1	-1	2
	2	-1	1	4
	3	1	1	5
	4	-1	-1	1
Respondent 2	1	1	-1	...
	2	-1	1	...
	3	1	1	...
	4	-1	-1	...

Respondent 1

	Coefficients
Intercept	3
Warm	0.5
Competent	1.5

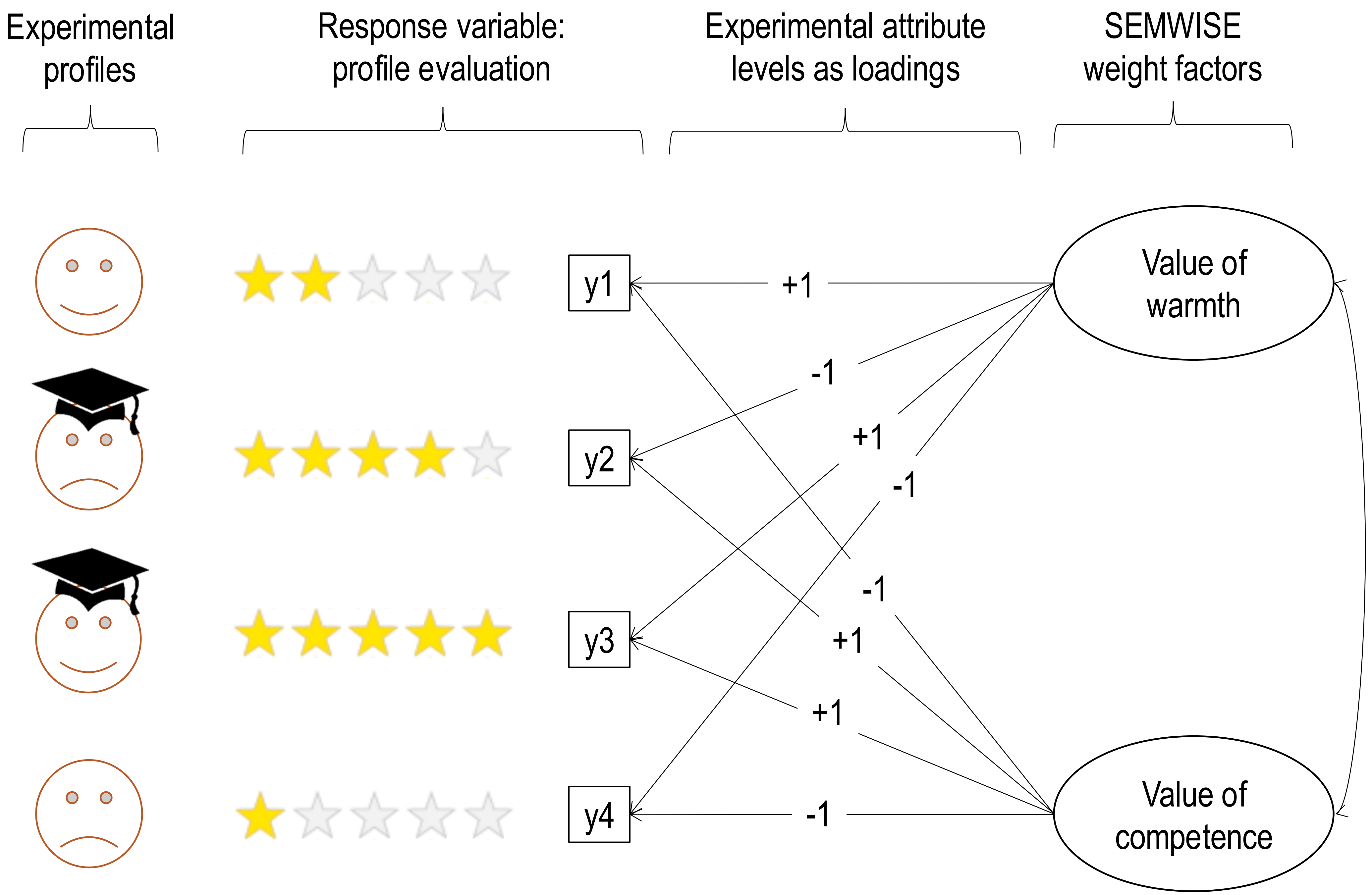
STYLIZED EXAMPLE:
WARM (LO/HI) X COMPETENT (LO/HI)



"wide format"

	y1	y2	y3	y4
Respondent 1	2	4	5	1
Respondent 2	...			...

STYLIZED EXAMPLE: WARM (LO/HI) X COMPETENT (LO/HI)



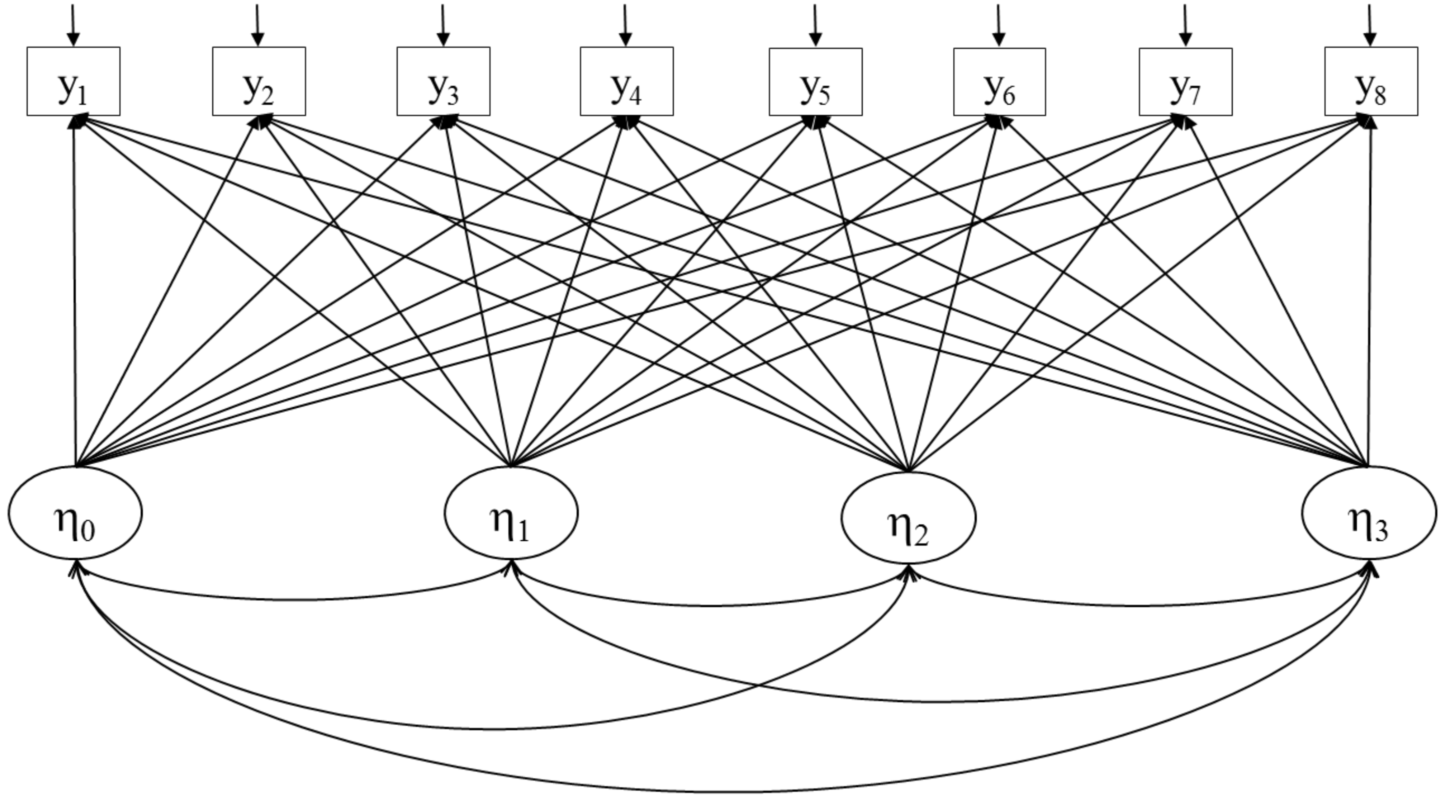
BUT WHY WOULD YOU WANT TO DO ALL THIS? ADVANTAGES OF USING SEMWISE

- Incorporate the weight factors into a broader nomological network
- Detailed model fit information
- Parameter restrictions can be evaluated based on indices of local misfit
- Measurement error in the dependent variable accounted for
- Multi-group modeling
  - testing cross-group parameter differences
  - measurement invariance
- Account for method effects

SEMWISE CORE MODEL

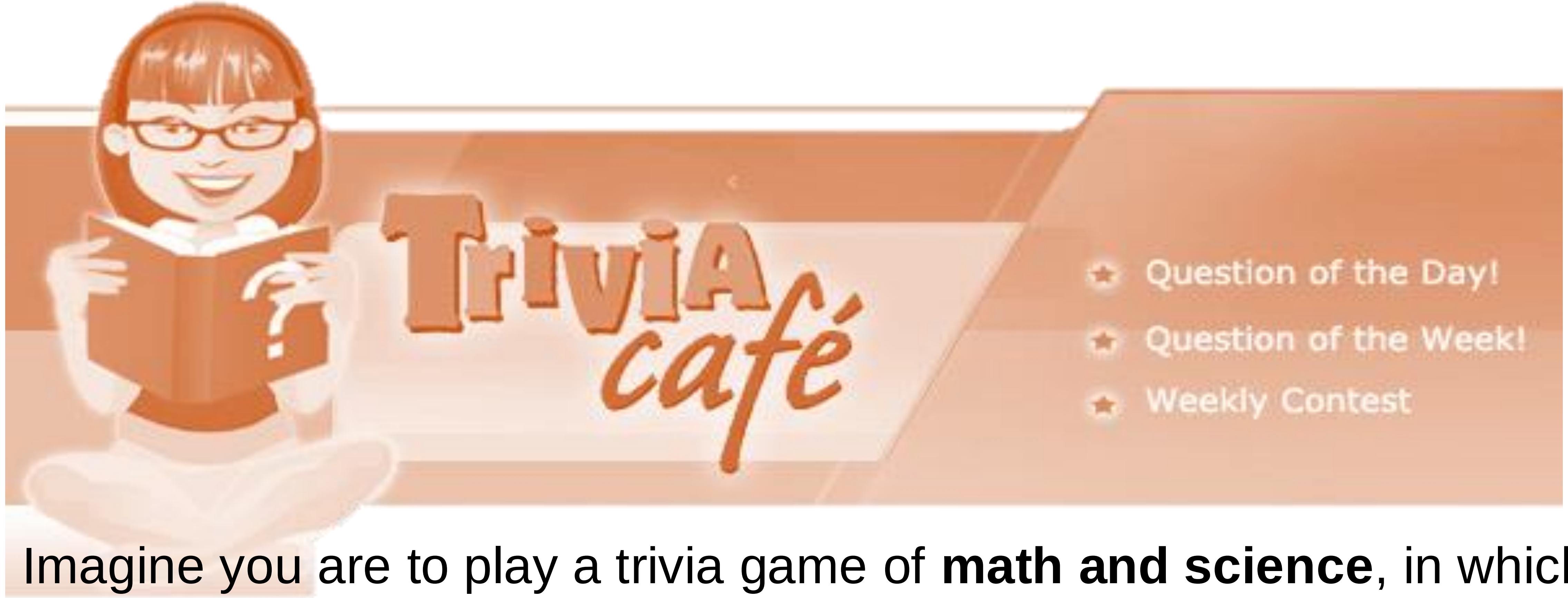
Model for three binary experimental attributes: $2^3 = 8$ profiles (full factorial design)

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \alpha_6 \\ \alpha_7 \\ \alpha_8 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{bmatrix} * \begin{bmatrix} \eta_0 \\ \eta_1 \\ \eta_2 \\ \eta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \end{bmatrix}$$




```
MODEL:
eta0 by y1-y8@1;!intercept factor;
eta1 by y1@1 y2@-1 y3@1 y4@-1 y5@1 y6@-1 y7@1 y8@-1;
eta2 by y1@1 y2@1 y3@-1 y4@-1 y5@1 y6@1 y7@-1 y8@-1;
eta3 by y1@1 y2@1 y3@1 y4@1 y5@-1 y6@-1 y7@-1 y8@-1;
```

EMPIRICAL APPLICATION 1: TEAM MATE PREFERENCE



Imagine you are to play a trivia game of **math and science**, in which teams of two compete to win a cash prize. You are given short problems involving math and the natural sciences, and you and your teammate have to solve the problems and answer some questions. The team that answers the most questions correctly wins a cash prize.

For each of the women described on the following pages, please indicate how likely you would be to select them as your teammate.

Please indicate to what extent you would like to have this person as your teammate for a trivia game of math and science:

Jackelyn is described by her friends as trustworthy. She has an IQ of 104.

>>

Please indicate to what extent you would like to have this person as your teammate for a trivia game of math and science:

Dottie is described by her friends as good-natured. She has an IQ of 101.

>>

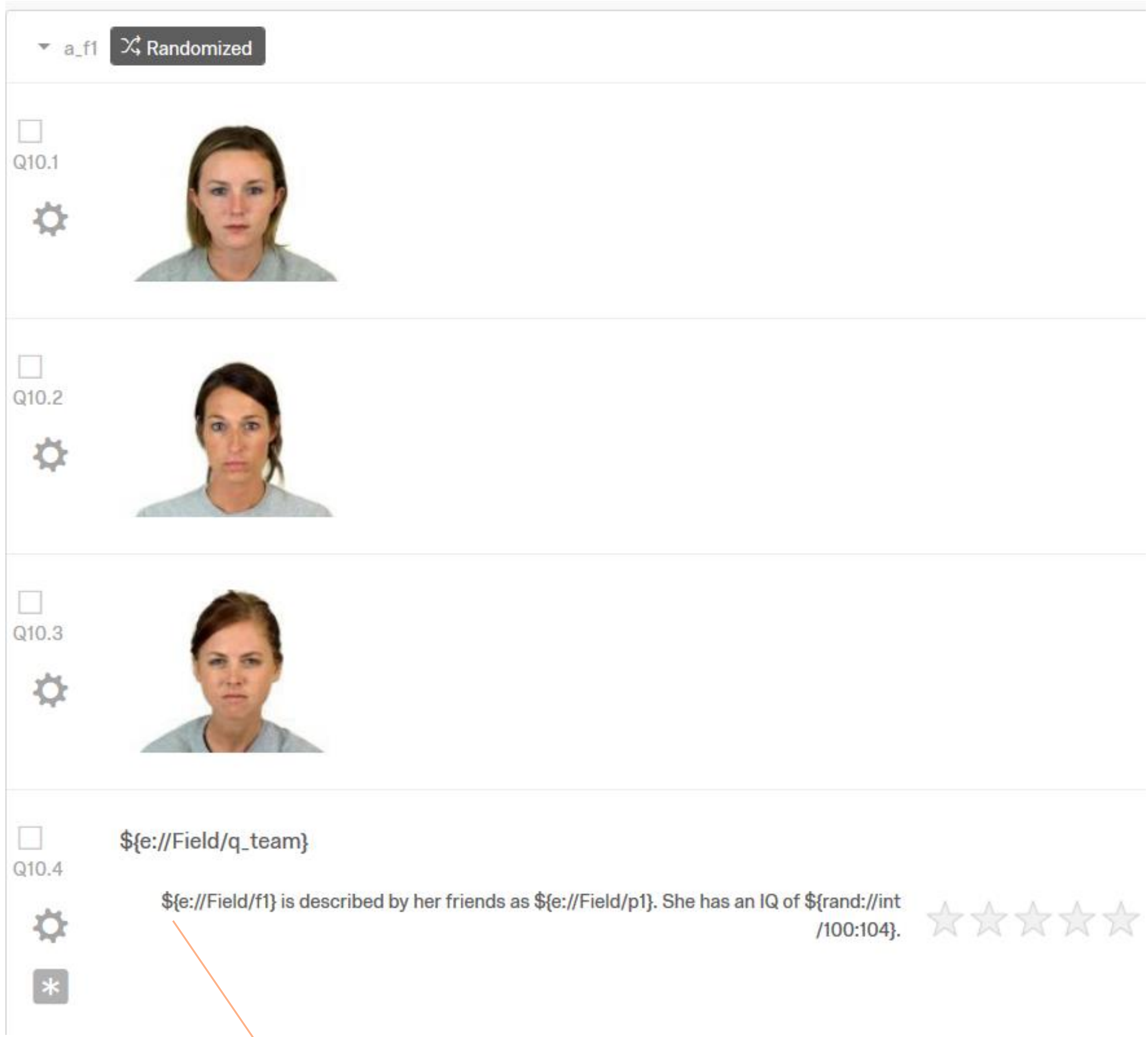
Please indicate to what extent you would like to have this person as your teammate for a trivia game of math and science:

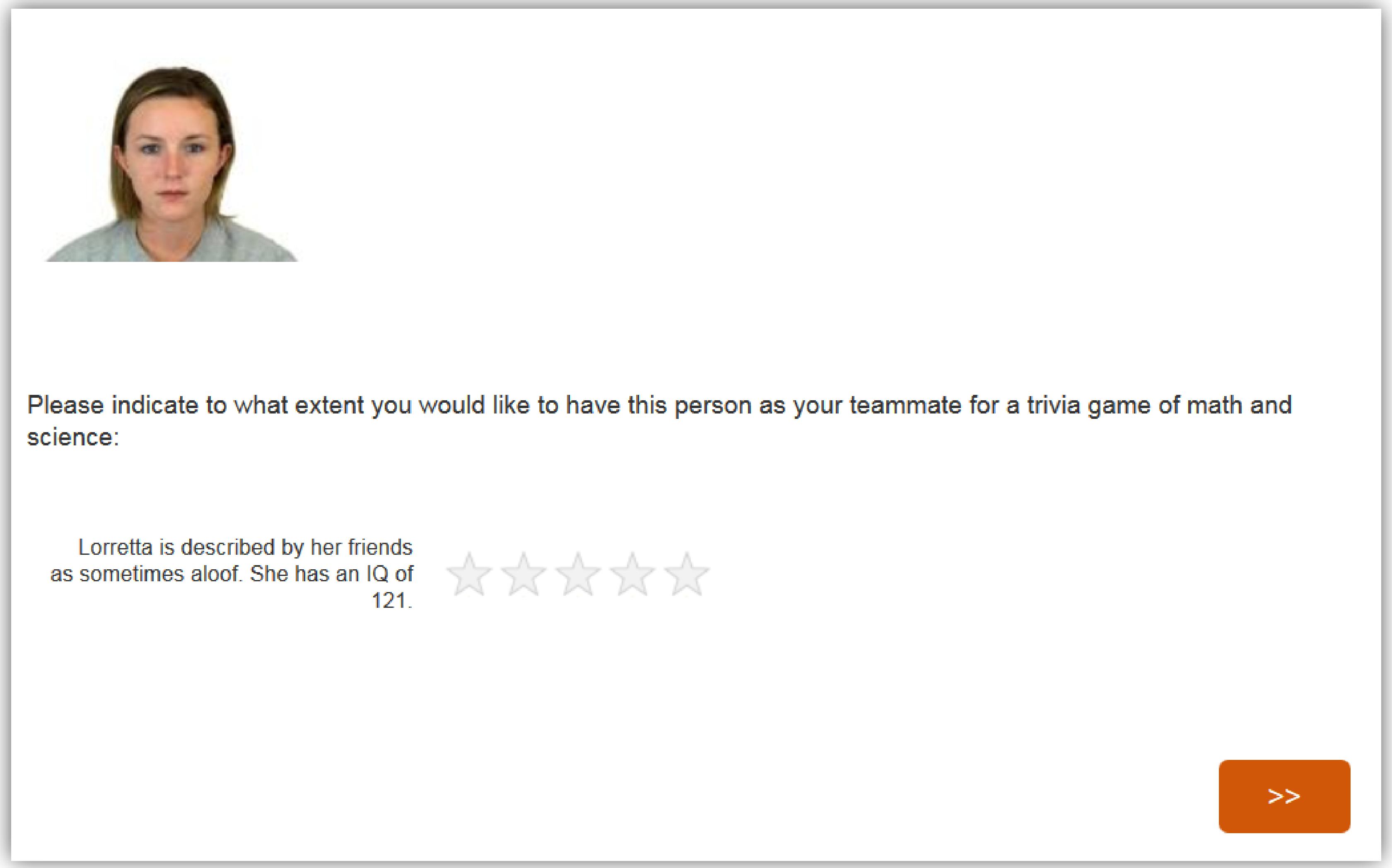
Lorretta is described by her friends as sometimes aloof. She has an IQ of 121.

☆☆☆☆☆

>>

QUALTRICS IMPLEMENTATION





${e://Field/f1} is described by her friends as ${e://Field/p1}. She has an IQ of ${rand://int/100:104}.

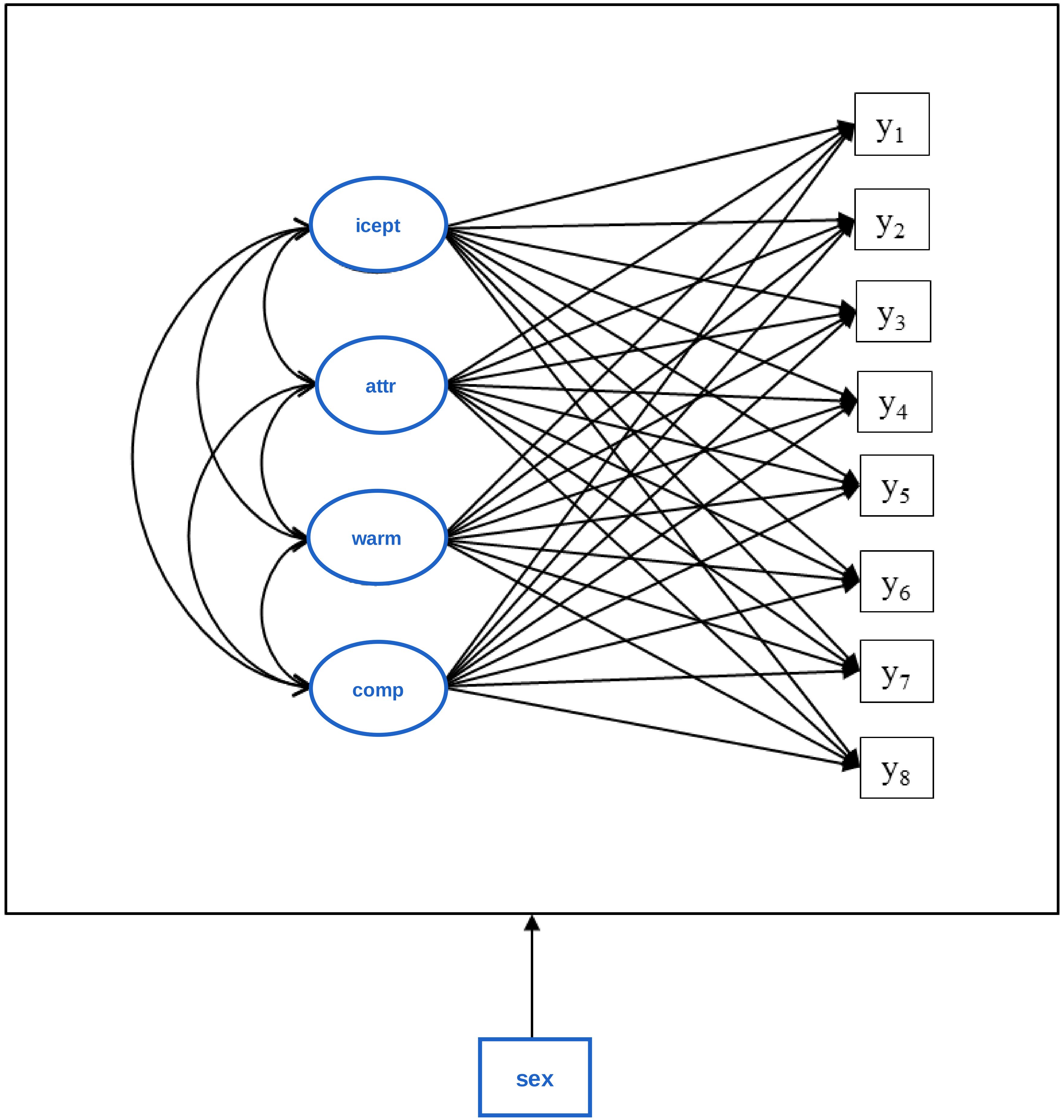
GHENT UNIVERSITY

EMPIRICAL APPLICATION 1: TEAM MATE PREFERENCE

Element	Attribute	Levels	Set	Source/reference
[Name]	/	n.a.	Jackelyn, Alison, Gabriele, Lorretta, Dottie, Marybeth, Sarina, Hester, Janet, Rosemary	listofrandomnames.com
[Picture]	Attractiveness	below average	12 pictures of white women close to attractiveness score of 2.57 (i.e., M - 1 SD)	Chicago Face Database (Ma, Correll, & Wittenbrink, 2015), adapted to a 160 x 112 format
		above average	12 pictures of white women close to attractiveness score of 4.34 (i.e., M + 1 SD)	
[Warmth]	Warmth	high	friendly, trustworthy, warm, good-natured	Fiske, Cuddy, Glick, and Xu (2002)
		low	not always kind, occasionally unfriendly, sometimes aloof, a bit standoffish	
[IQ]	Competence	average	randomly sampled number from the range [100-104] (inclusive)	Caruso et al. (2009)
		high	randomly sampled number from the range [118-122] (inclusive)	

GHENT UNIVERSITY

ANALYSES



Please indicate to what extent you would like to have this person as your teammate for a trivia game of math and science:

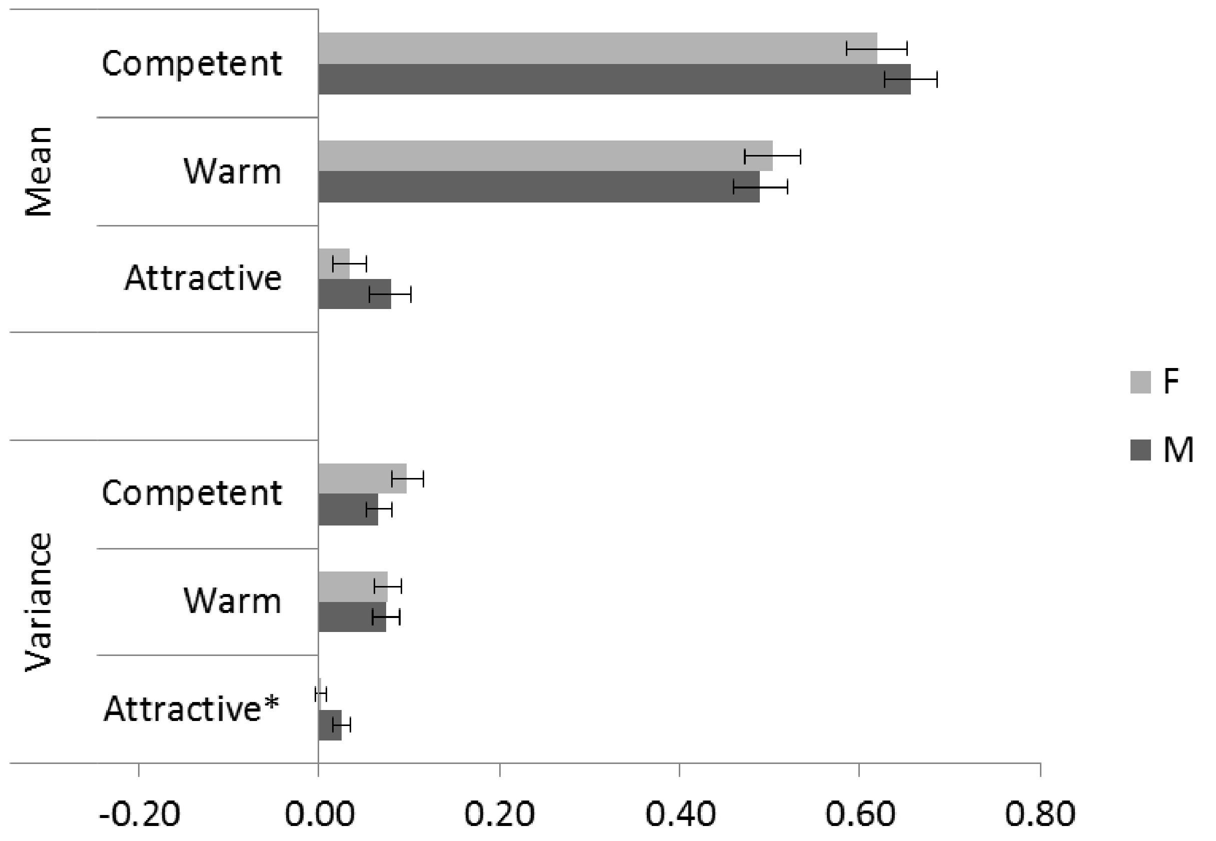
Jackelyn is described by her friends as trustworthy. She has an IQ of 104.

χ^2 (44) = 67.156, p = .014,
RMSEA = .064, CFI = .968,
TLI = .960, SRMR = .067;
n = 260

GHENT UNIVERSITY

RESULTS

		M			F			Diff.	
		Est.	SE		Est.	SE		p	
Mean	Intercept	2.82	0.05	*	2.85	0.06	*	0.710	
	Attractive	0.08	0.02	*	0.03	0.02		0.135	
	Warm	0.49	0.03	*	0.50	0.03	*	0.741	
	Competent	0.66	0.03	*	0.62	0.03	*	0.398	
Variance	Intercept	0.32	0.04	*	0.39	0.06	*	0.288	
	Attractive	0.03	0.01	*	0.00	0.01		0.040	*
	Warm	0.08	0.02	*	0.08	0.02	*	0.979	
	Competent	0.07	0.01	*	0.10	0.02	*	0.168	

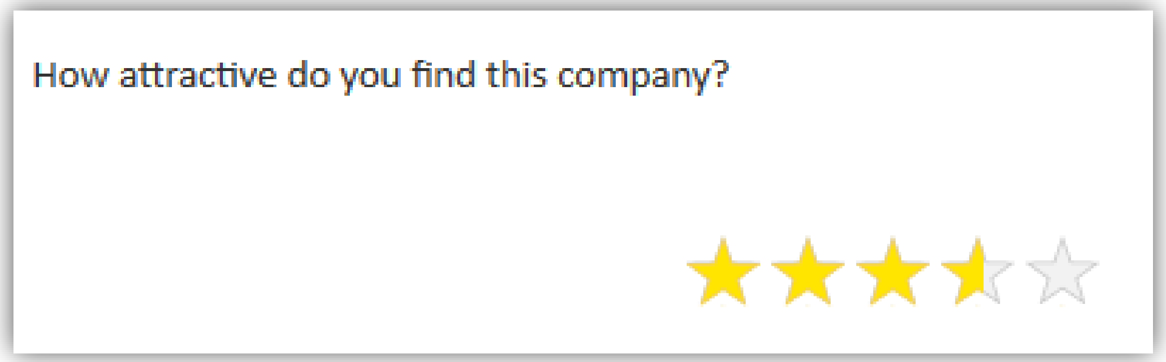


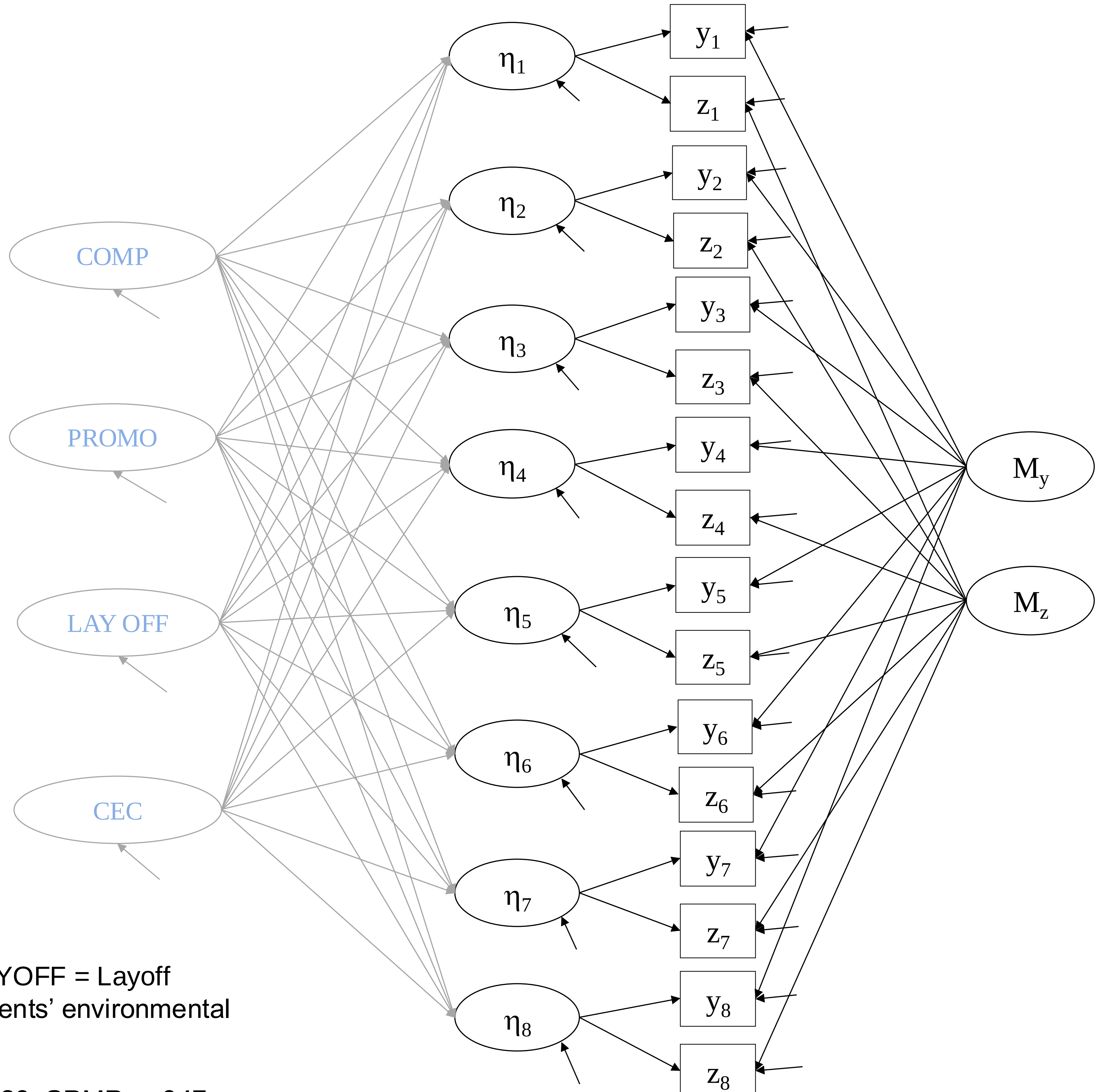
SEMWISE
BERT.WEIJTERS@UGENT.BE

EMPIRICAL APPLICATION 2: JOB/ORG CHARACTERISTICS AND ORGANIZATIONAL ATTRACTIVENESS

Attribute	Levels	
	-1	1
1. COMP: The firm's compensation package is somewhat [x] average for the industry	below	above
2. PROMO: the typical career path for the average graduate includes [x] promotion in five years	no	one
3. LAYOFF: The company's policy is that employees are [x] laid off.	sometimes	rarely
4. CEC: Concern for the environment is [x] priority in the company	not a	a

GHENT UNIVERSITY



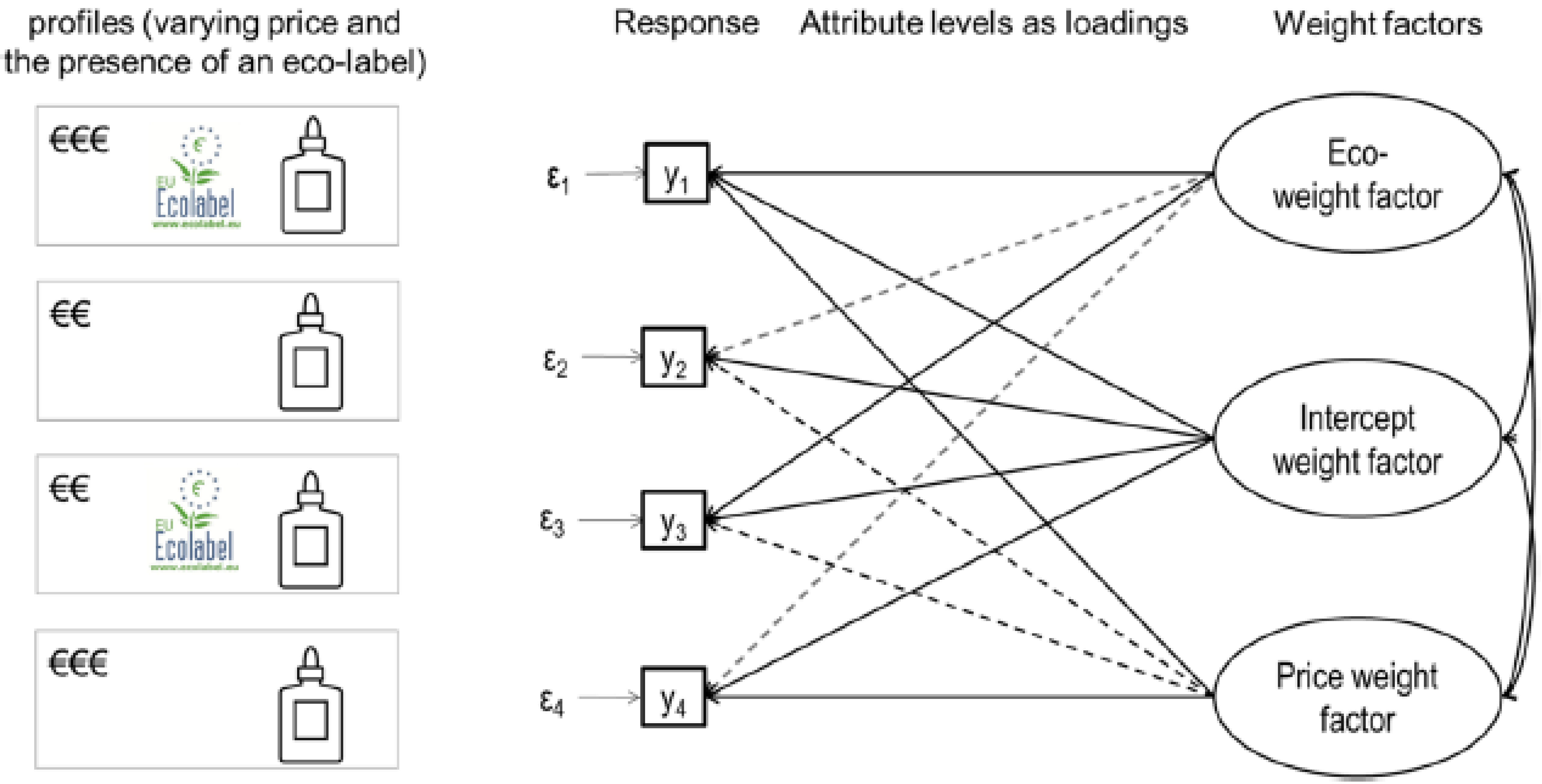


COMP = Compensation; PROMO = Promotion opportunities; LAYOFF = Layoff policy; CEC = Company's environmental concern; EC = respondents' environmental concern; IM = Impression mgt.

$\chi^2(100) = 144.174$, $p = .003$, RMSEA = .040, CFI = .983, TLI = .980, SRMR = .047

$n = 276$

APPLICATION 3: MEASURING ECO-UTILITY

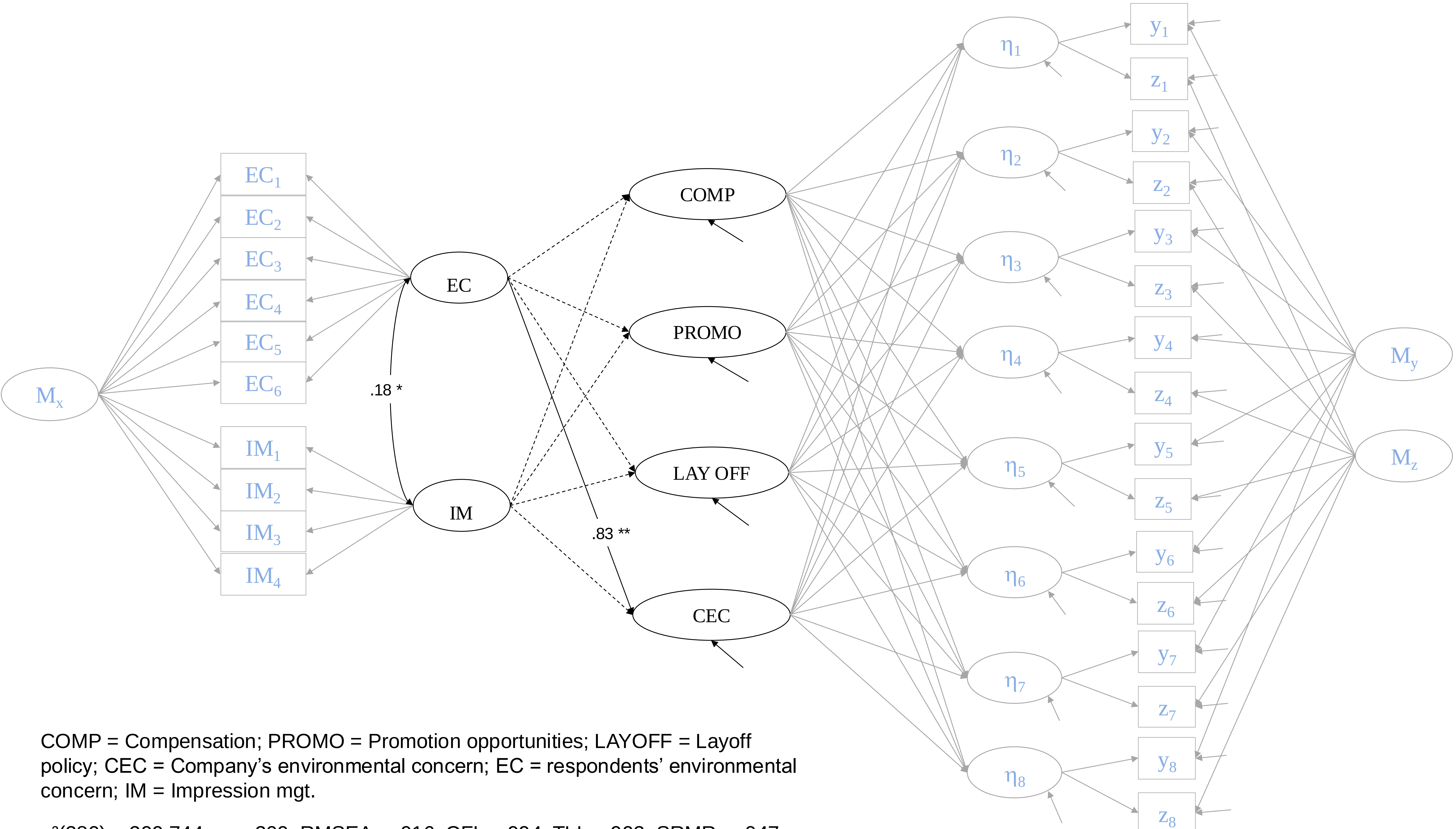


GHENT UNIVERSITY

OVERVIEW

1. Between-subjects experiments
   a) Measurement and scaling
   b) Main effects
   c) Interaction effects
2. Mediation
   a) Common Method Variance (CMV)
   b) Accounting for CMV in measurement
3. Within-subject and mixed experiments
   a) Unconditional SEMWISE
   b) Conditional SEMWISE

GHENT UNIVERSITY



COMP = Compensation; PROMO = Promotion opportunities; LAYOFF = Layoff policy; CEC = Company's environmental concern; EC = respondents' environmental concern; IM = Impression mgt.

$\chi^2(280) = 299.744$, $p = .200$; RMSEA = .016, CFI = .994, TLI = .993, SRMR = .047
n = 276

SEMWISE: DISCUSSION

SEMWISE is not intended to replace multi-level analysis for policy capturing, but probably preferable if

- the dependent variable is latent (measured with error)
- there is method variance in the data that needs modeling
- the variance-covariance structure needs to be compared across multiple samples
- alternative model specifications need to be statistically compared in terms of fit
- **a researcher is primarily interested in the broader nomological network in which the decision variables assessed through policy capturing analysis are embedded**


Questions?

FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

Bert Weijters

Ph.D., Associate Professor Market Research

DEPT. WORK, ORGANIZATION & SOCIETY

E bert.weijters@ugent.be

https://www.linkedin.com/in/bertweijters/

www.be4life.eu
http://archive.ugent.be/person/802000878995

GHENT
UNIVERSITY