UiO : **CEMO – Centre for Educational Measurement**
University of Oslo

Mild Introduction to Structural Equation Modeling

Workshop Oslo UseR! Group

Ronny Scherer
Centre for Educational Measurement at the University of Oslo (CEMO)

UseR! Oslo Workshop
28.05.2020


BAR
KAFE
10

Openness → Prejudice

Agreeableness → Prejudice

Research Questions

1. To what extent do agreeableness and openness explain variation in prejudice?
2. Do they explain variation to the same extent?

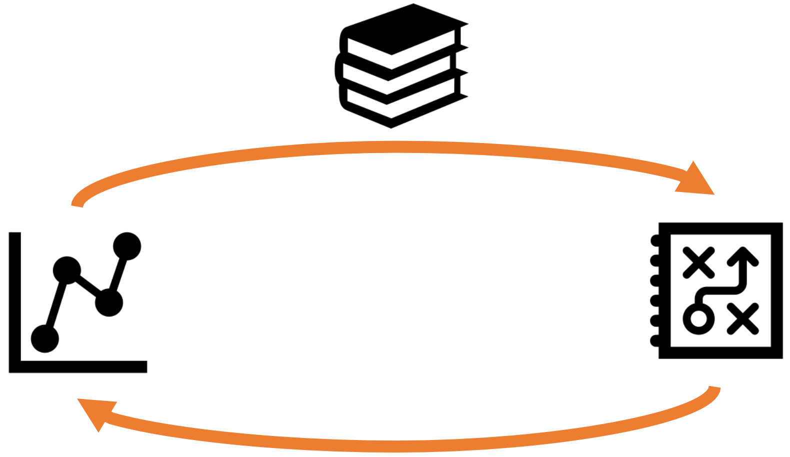
Some Background

Structural Equation Modeling (SEM)—Why?

Definition

“Structural equation modeling can be defined as a class of methodologies that seeks to **represent hypotheses** about the *means, variances, and covariances* of observed data in terms of a smaller number of ‘structural’ parameters defined by a **hypothesized underlying conceptual or theoretical model.**”

(Kaplan, 2001, p. 15215)



SEM is a multivariate statistical modelling technique.
We postulate a model, which may or may not fit the data.

(Kline, 2016)

Structural Equation Modeling (SEM)—Why?

What is so special about SEM?

- SEM can contain latent (unobserved) variables to represent constructs (i.e., measurement models) → Include random effects and control measurement error
- SEM can estimate derived effects, such as indirect effects in mediation models.
- SEM can be represented by path diagrams.
- SEM offers a broad range of models, including multi-group models, multilevel models, growth curve models, …
- SEM can handle full and summary data (i.e., covariance matrix and mean vector).

(Rosseel, 2020; Kline, 2016)

Structural Equation Modeling (SEM)—What?

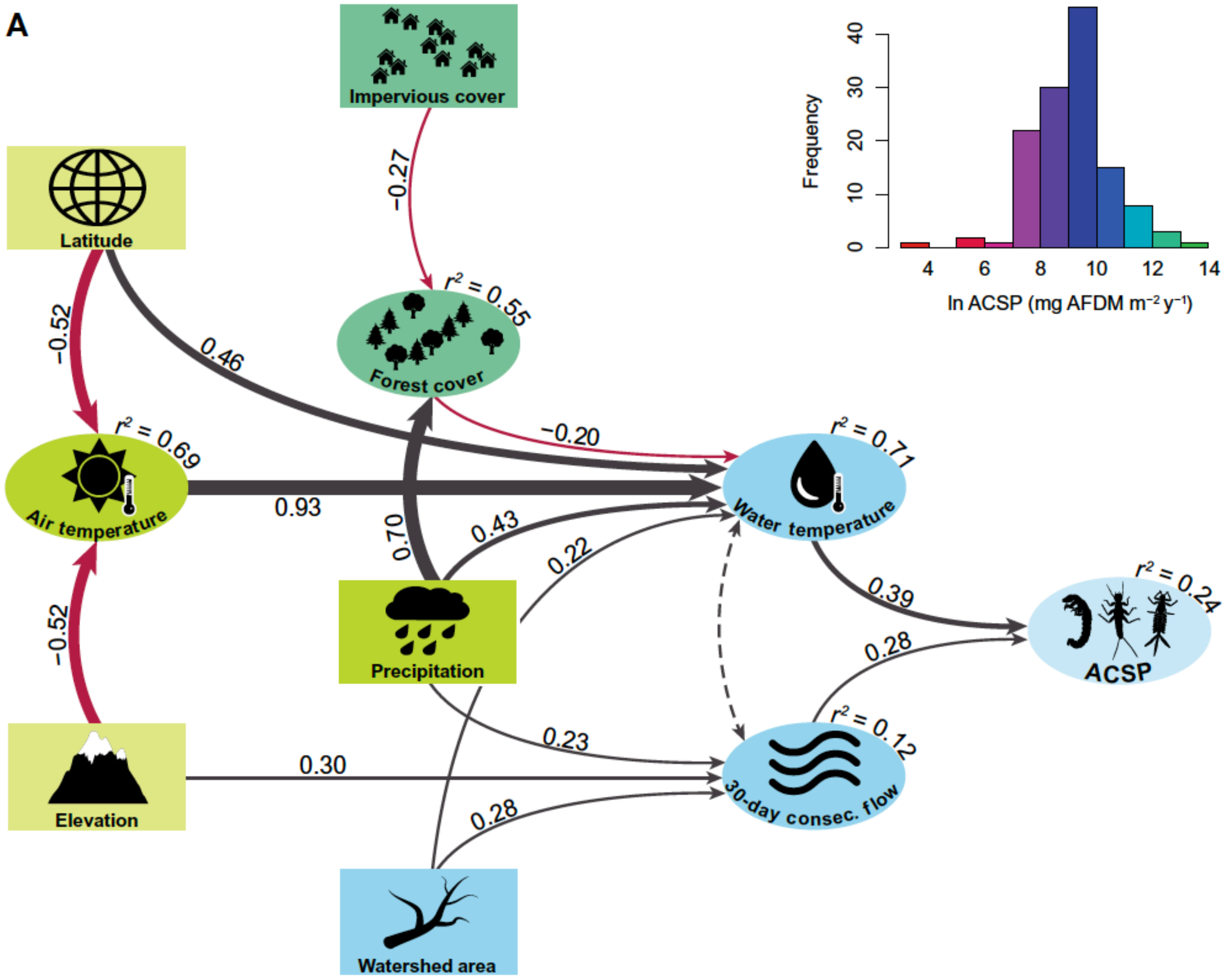
SCIENCE ADVANCES | RESEARCH ARTICLE

ECOLOGY

Precipitation and temperature drive continental-scale patterns in stream invertebrate production

C. J. Patrick[1]*, D. J. McGarvey[2], J. H. Larson[3], W. F. Cross[4], D. C. Allen[5], A. C. Benke[6], T. Brey[7], A. D. Huryn[6], J. Jones[8], C. A. Murphy[9], C. Ruffing[10], P. Saffarinia[11], M. R. Whiles[12], J. B. Wallace[13], G. Woodward[14]

Identifying drivers of annual community secondary production (ACSP), defined as the sum of annual production of all invertebrate populations within a community (*17*), is particularly challenging ...



Structural Equation Modeling (SEM)—What?

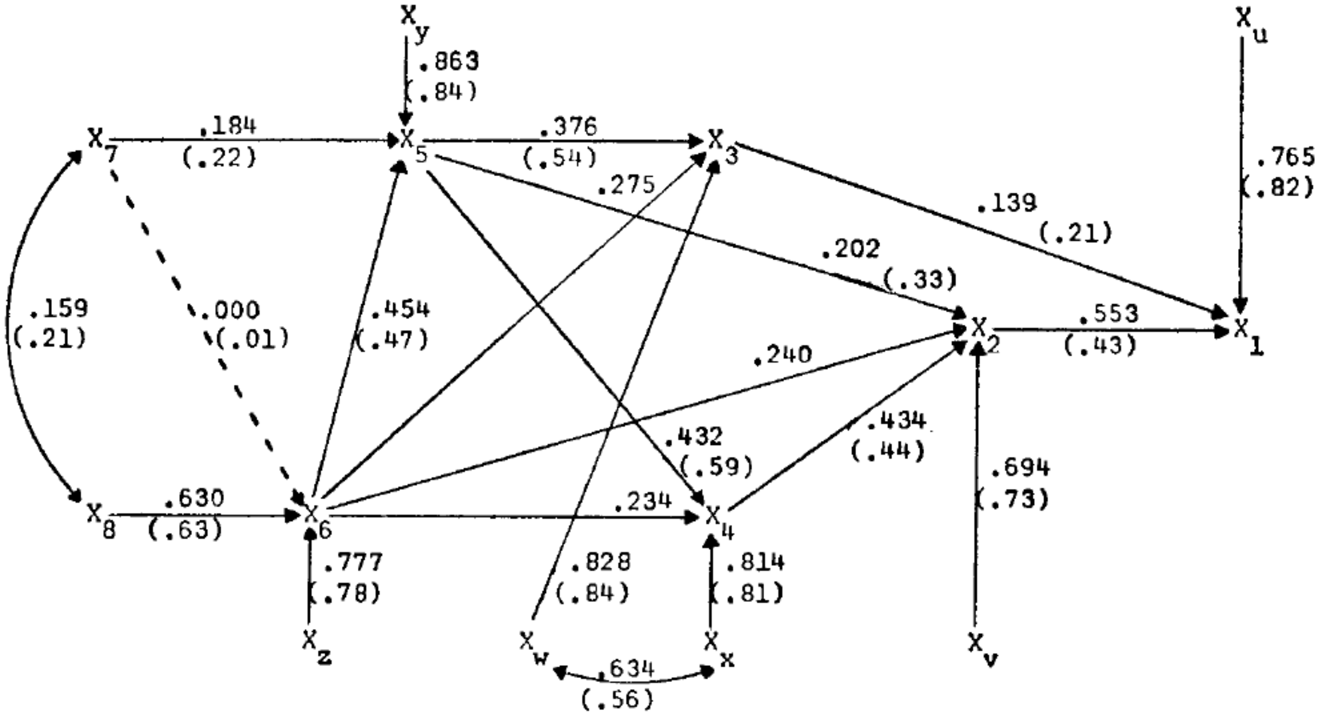
THE EDUCATIONAL AND EARLY OCCUPATIONAL STATUS ATTAINMENT PROCESS: REPLICATION AND REVISION[1]

WILLIAM H. SEWELL, ARCHIBALD O. HALLER AND GEORGE W. OHLENDORF
University of Wisconsin

American Sociological Review 1970, Vol. 35 (December):1014–1027

Relations between attainment, aspiration, socioeconomic status, mental ability, and academic performance

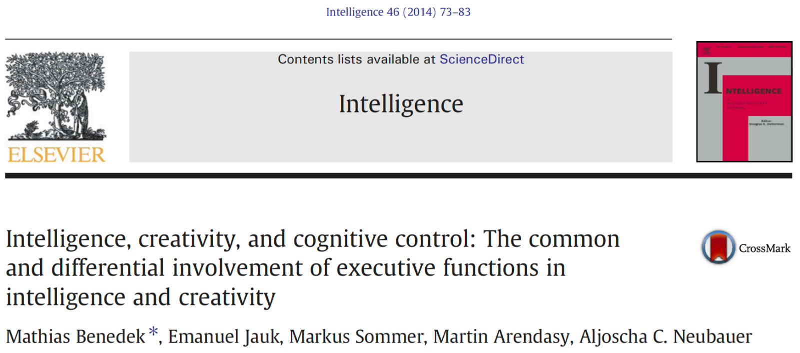
FIGURE 1. PATH COEFFICIENTS FOR ANTECEDENTS OF EDUCATIONAL AND OCCUPATIONAL ATTAINMENT WITH REVISED MODEL FOR FARM BOYS *



X_1 - Occupational Attainment
X_2 - Educational Attainment
X_3 - Level of Occupational Aspiration
X_4 - Level of Educational Aspiration
X_5 - Significant Others' Influence
X_6 - Academic Performance
X_7 - Socioeconomic Status
X_8 - Mental Ability

Structural Equation Modeling (SEM)—What?

Relations between creativity, intelligence, and executive functions


Intelligence 46 (2014) 73–83

Contents lists available at ScienceDirect

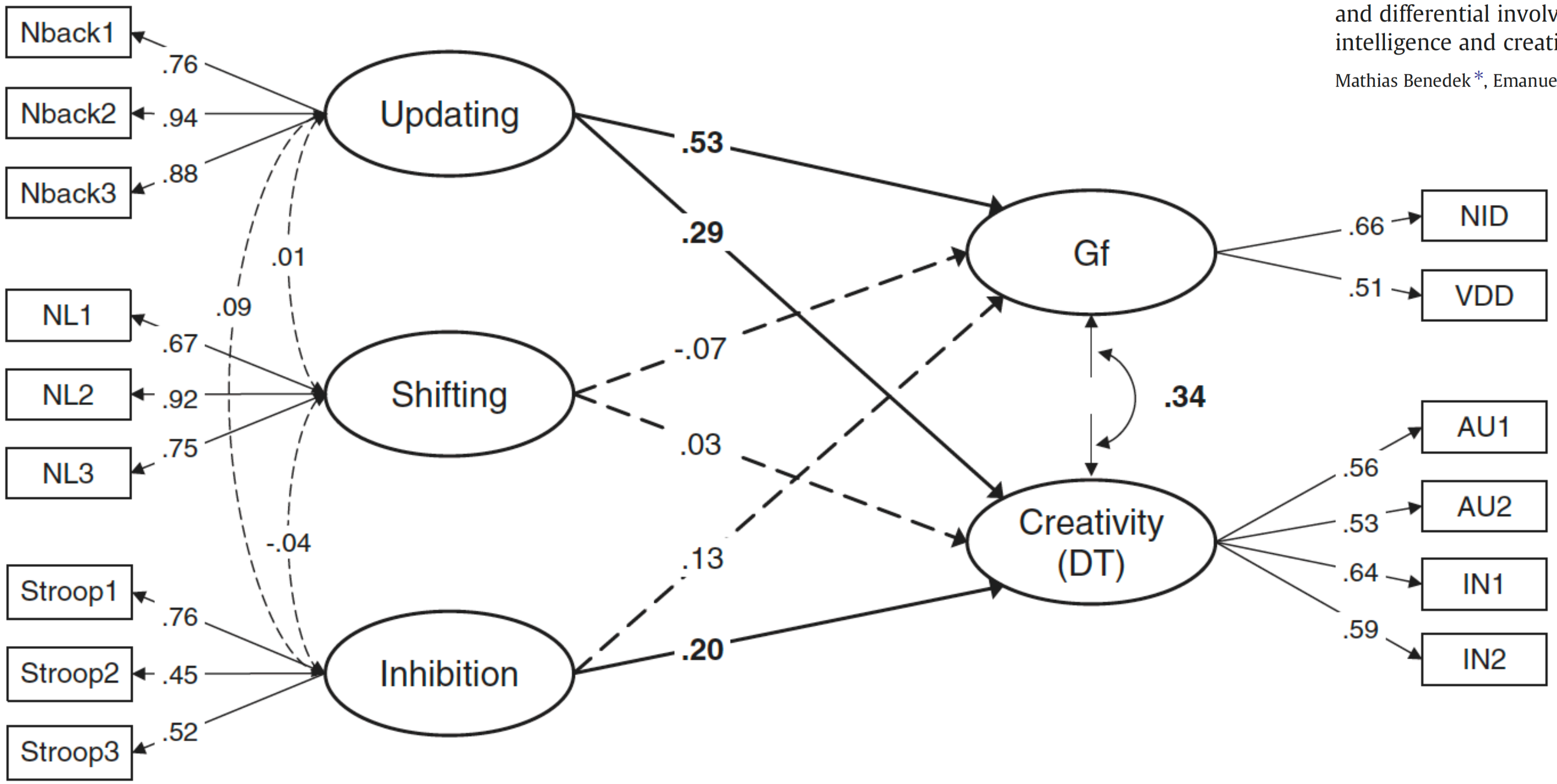
Intelligence

ELSEVIER

Intelligence, creativity, and cognitive control: The common and differential involvement of executive functions in intelligence and creativity

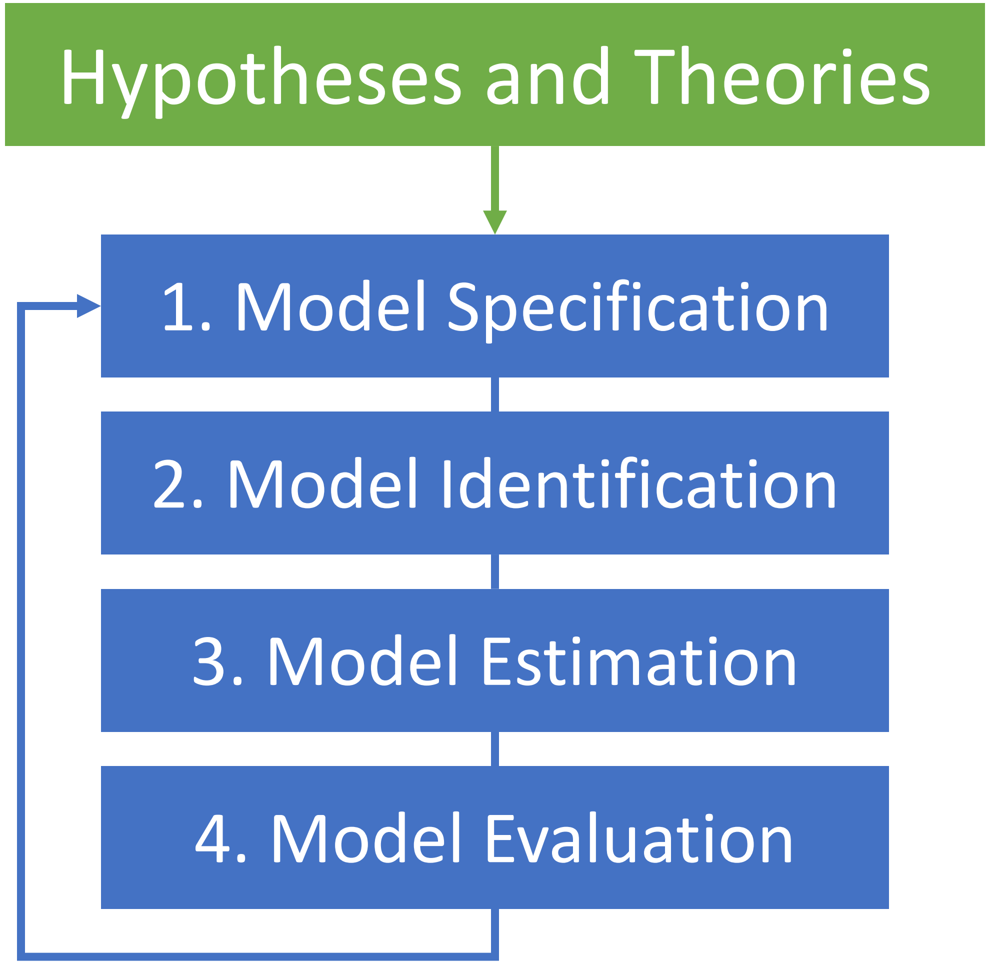
CrossMark

Mathias Benedek *, Emanuel Jauk, Markus Sommer, Martin Arendasy, Aljoscha C. Neubauer



Structural Equation Modeling (SEM)—How?

(Hoyle, 2012)



Structural Equation Modeling (SEM)—How?

(Hoyle, 2012)

Hypotheses and Theories

1. Model Specification
2. Model Identification
3. Model Estimation
4. Model Evaluation

The “Grammar” of a model:

Translate the hypotheses/theories into testable models.

- Specification equations
- Structural and measurement part
- Define regressions, (co-)variances, latent variables, …

Lavaan:

- `Regression: Y ~ X`
- `Covariance: Y ~~ X`
- `Latent variable: eta =~ x1+x2+x3`

Structural Equation Modeling (SEM)—How?

(Hoyle, 2012)

Hypotheses and Theories

1. Model Specification

2. Model Identification

3. Model Estimation

4. Model Evaluation

The "Feasibility Check" of a model:

Do we have enough information to estimate all model parameters?

- Observed information (from the data): variances, covariances, means
- Model parameters: variances, covariances, regression coefficients, intercepts, …

Rule:
Have at least the same amount of observed information than model parameters → degrees of freedom of the model ≥ 0

Structural Equation Modeling (SEM)—How?

(Hoyle, 2012)

Hypotheses and Theories

1. Model Specification
2. Model Identification
3. Model Estimation
4. Model Evaluation

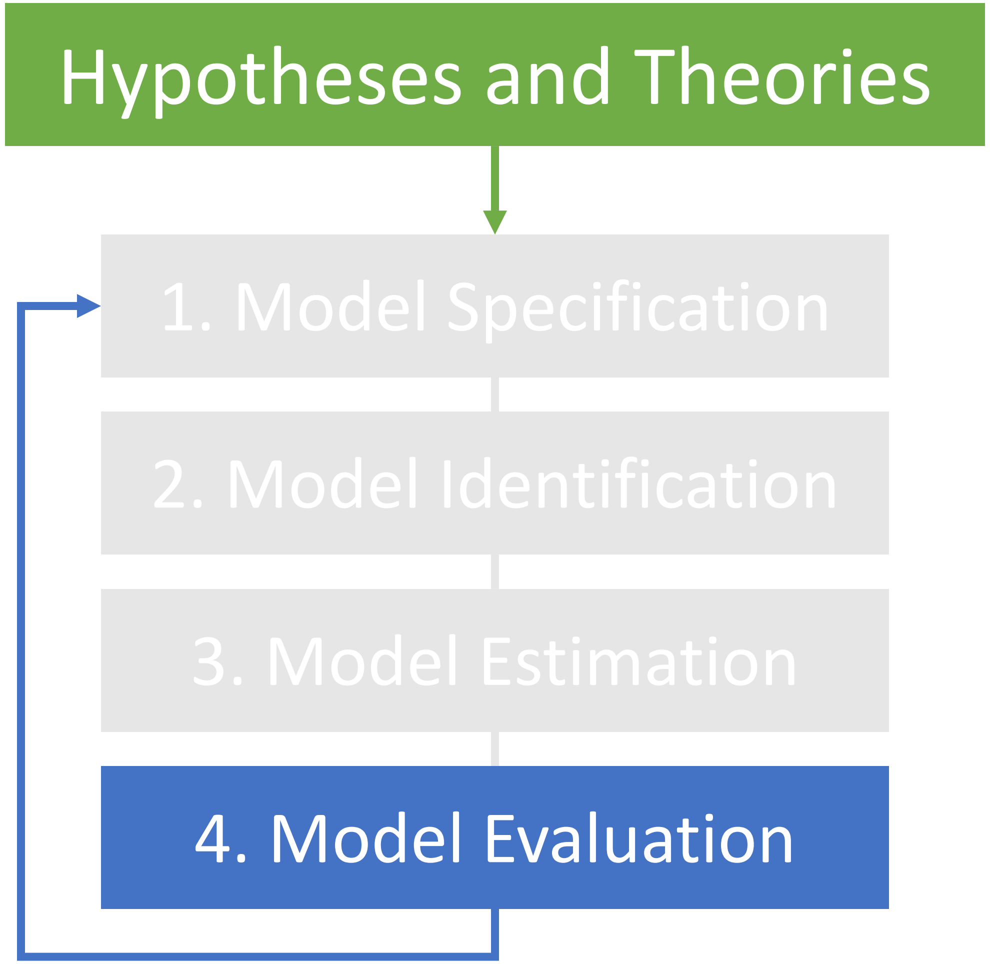
Estimating model parameters:

Reduce the discrepancy between the observed covariance matrix (S) and the covariance matrix implied by the model (Σ).

- Maximum likelihood estimation (ML)
- Weighted least squares (WLS)
- Bayesian estimation

Structural Equation Modeling (SEM)—How?

(Hoyle, 2012)



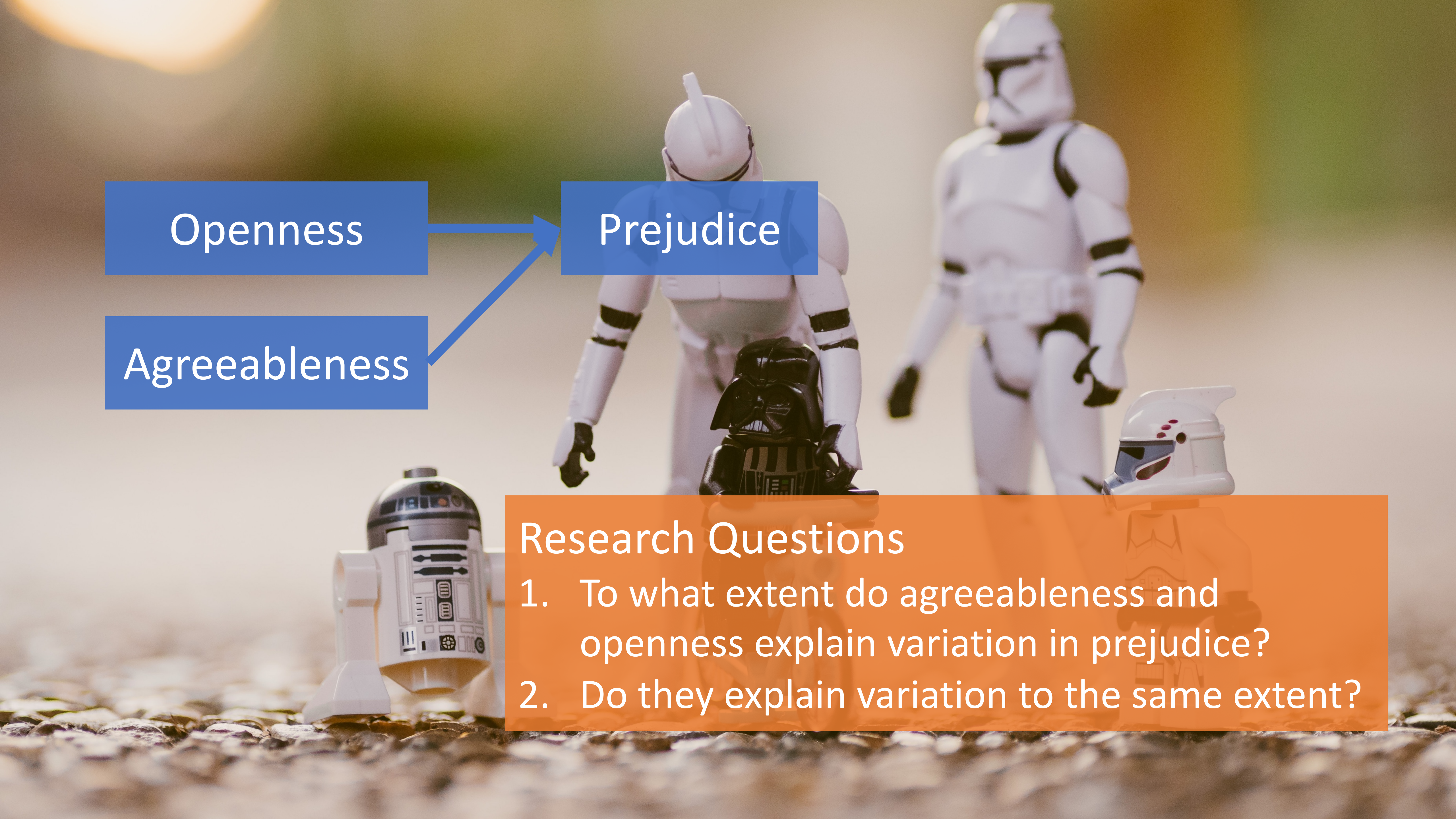
Evaluate the model fit:

Well-fitting models show a small discrepancy between the observed covariance matrix (S) and the covariance matrix implied by the model (Σ).

Global fit

- Chi-square statistic χ^2
- Comparative Fit Index (CFI)
- Root Mean Square Error of Approximation (RMSEA)
- Standardized Root Mean Square Residual (SRMR)

“Rules”: χ^2 ns, CFI$\geq$.95, RMSEA$\leq$.08, SRMR$\leq$.08


Openness
Agreeableness
Prejudice
Research Questions
1. To what extent do agreeableness and openness explain variation in prejudice?
2. Do they explain variation to the same extent?

UiO : CEMO – Centre for Educational Measurement
University of Oslo

SEM—Our Example for Today

Journal of Personality and Social Psychology
2016, Vol. 111, No. 3, 367–395

0022-3514/16/$12.00 http://dx.doi.org/10.1037/pspi0000064

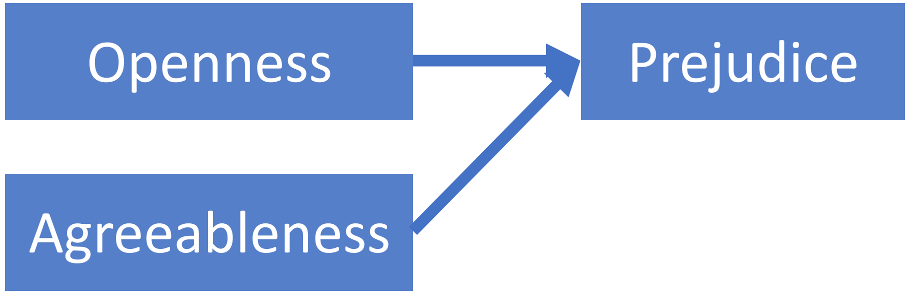
Is Group Membership Necessary for Understanding Generalized Prejudice? A Re-Evaluation of Why Prejudices Are Interrelated

Robin Bergh
Harvard University and Uppsala University

Nazar Akrami
Uppsala University

Jim Sidanius
Harvard University

Chris G. Sibley
University of Auckland



Format

A data frame with 861 individuals, 10 composite scores, and gender:

`EP` Ethnic prejudice

`SP` Sexism

`HP` Sexual prejudice against gays and lesbians

`DP` Prejudice toward mentally people with disabilities

`A1` Agreeableness indicator 1

`A2` Agreeableness indicator 2

`A3` Agreeableness indicator 3

`O1` Openness indicator 1

`O2` Openness indicator 2

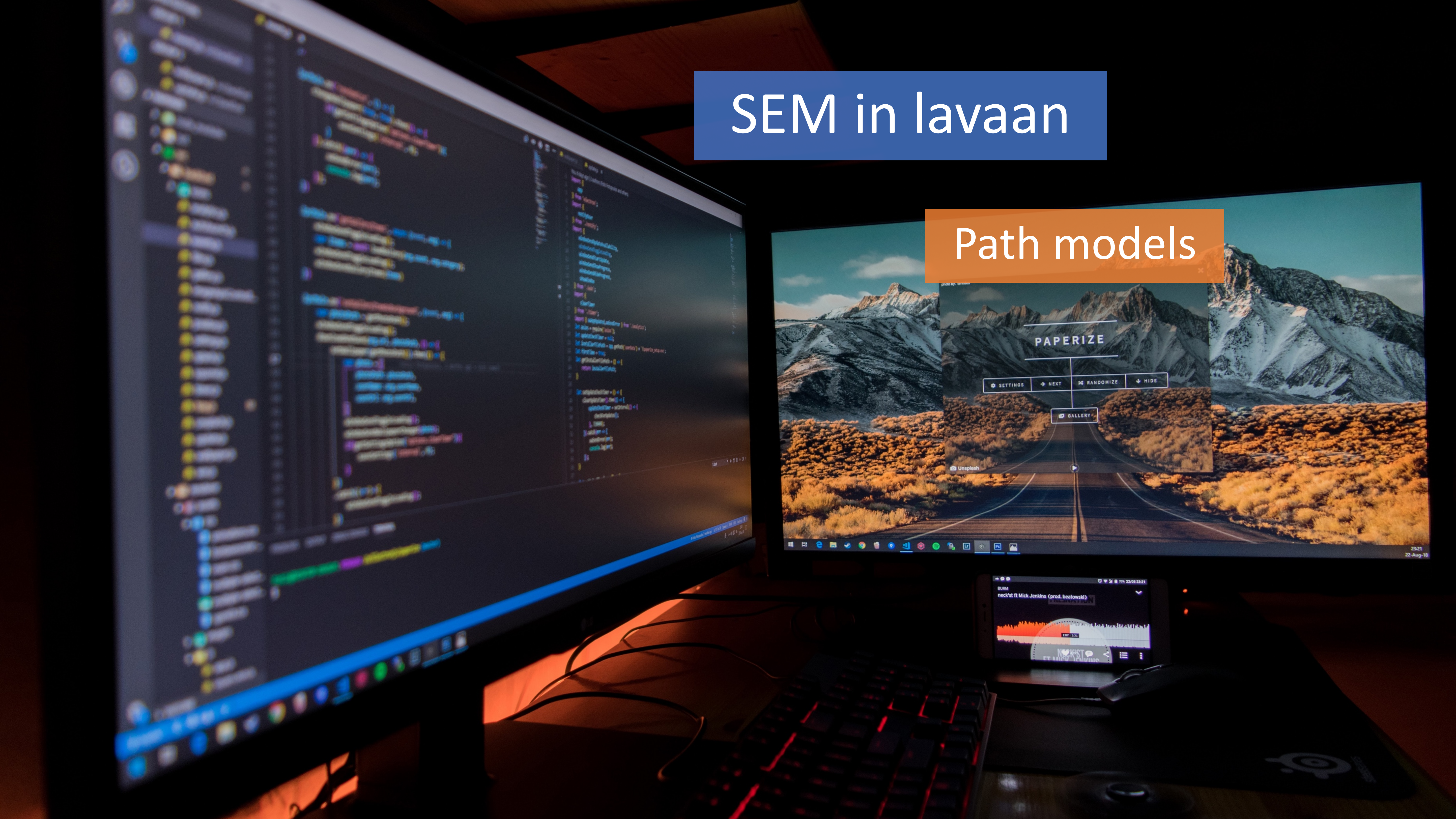
`O3` Openness indicator 3

`gender` gender

R package `MPsychoR`
`data(Bergh)`

SEM in lavaan

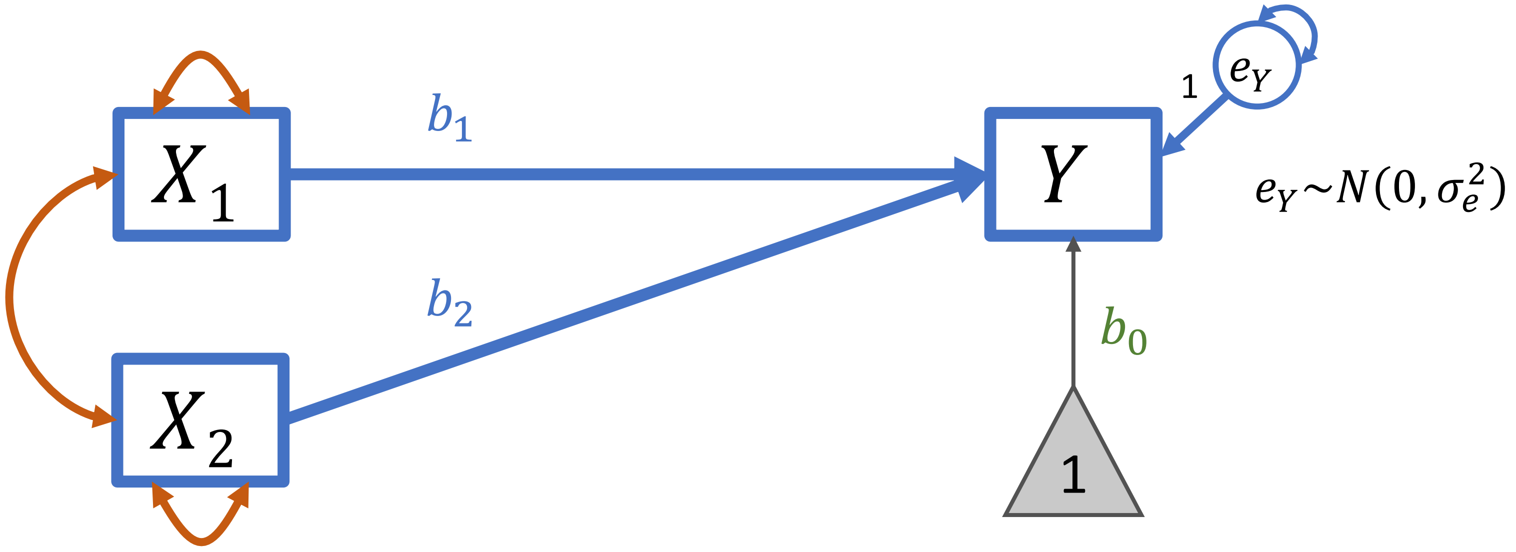
Path models



Model 1

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + e_{Yi}$$

Simple linear regression model with two predictors



Data
(y_i, x_{1i}, x_{2i}) for persons $i = 1, \dots, N$

Y	Prejudice
X_1	Openness to experience (sum score)
X_2	Agreeableness (sum score)

(Bergh et al., 2016; Mair, 2018)

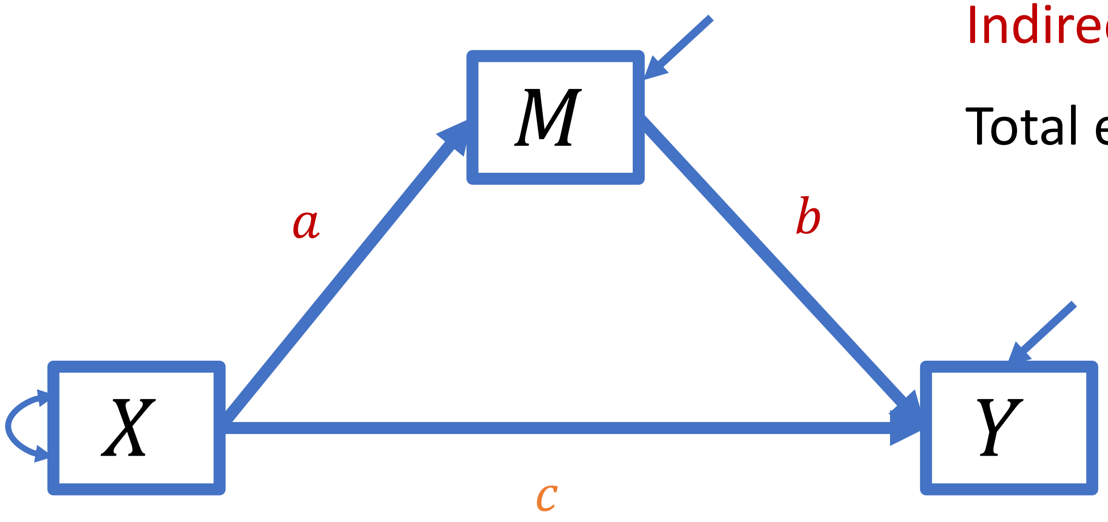
Model 2

$$M_i = i_M + aX_i + e_{Mi}$$
$$Y_i = i_Y + bM_i + cX_i + e_{Yi}$$

Mediation model with direct and indirect effects

Direct effect c
Indirect effect ab

Total effect $ab + c$



Your turn
Specify and estimate this model in lavaan.

Data
(y_i, m_i, x_i) for persons $i = 1, \dots, N$

Y	Prejudice
M	Openness to experience (sum score)
X	Agreeableness (sum score)

SEM in lavaan

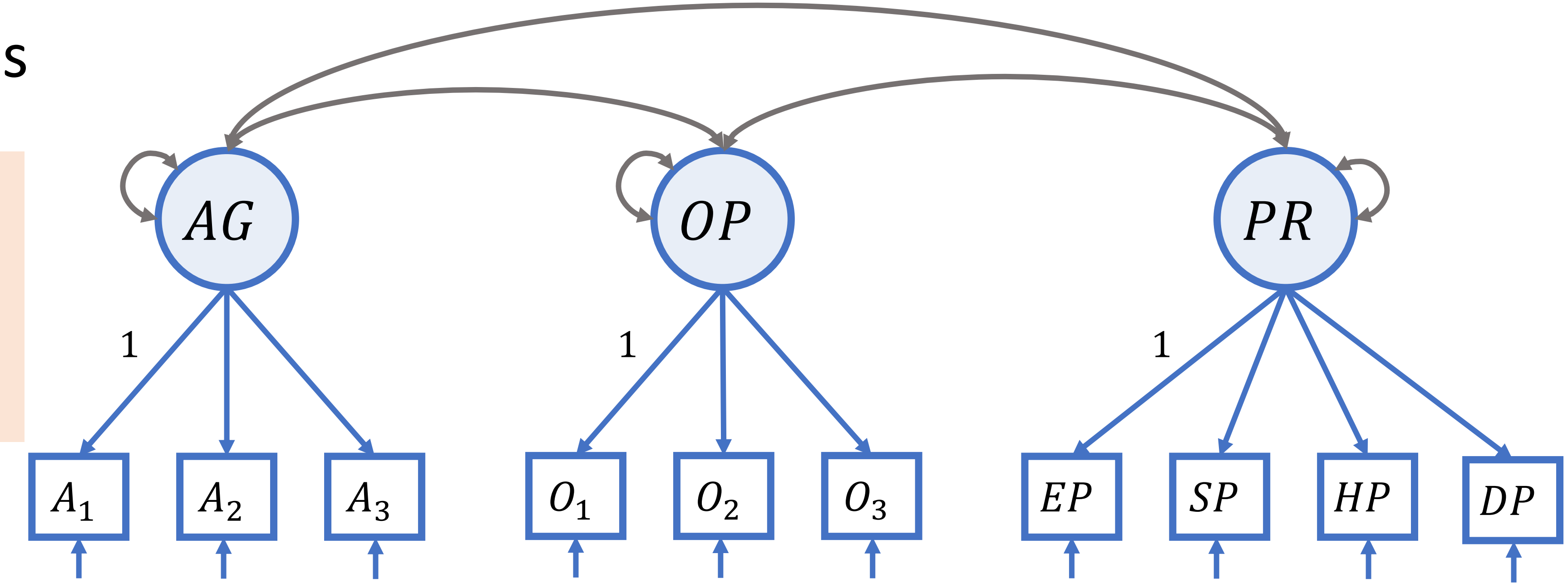
Models with latent variables

Model 3: Measurement Model

Confirmatory factor analysis with three latent (unobserved) variables

Represent constructs as latent variables and separate the “true score” and measurement error.

Data
$(y_i, a_{1i}, \ldots, dp_i)$ for persons $i = 1, \ldots, N$

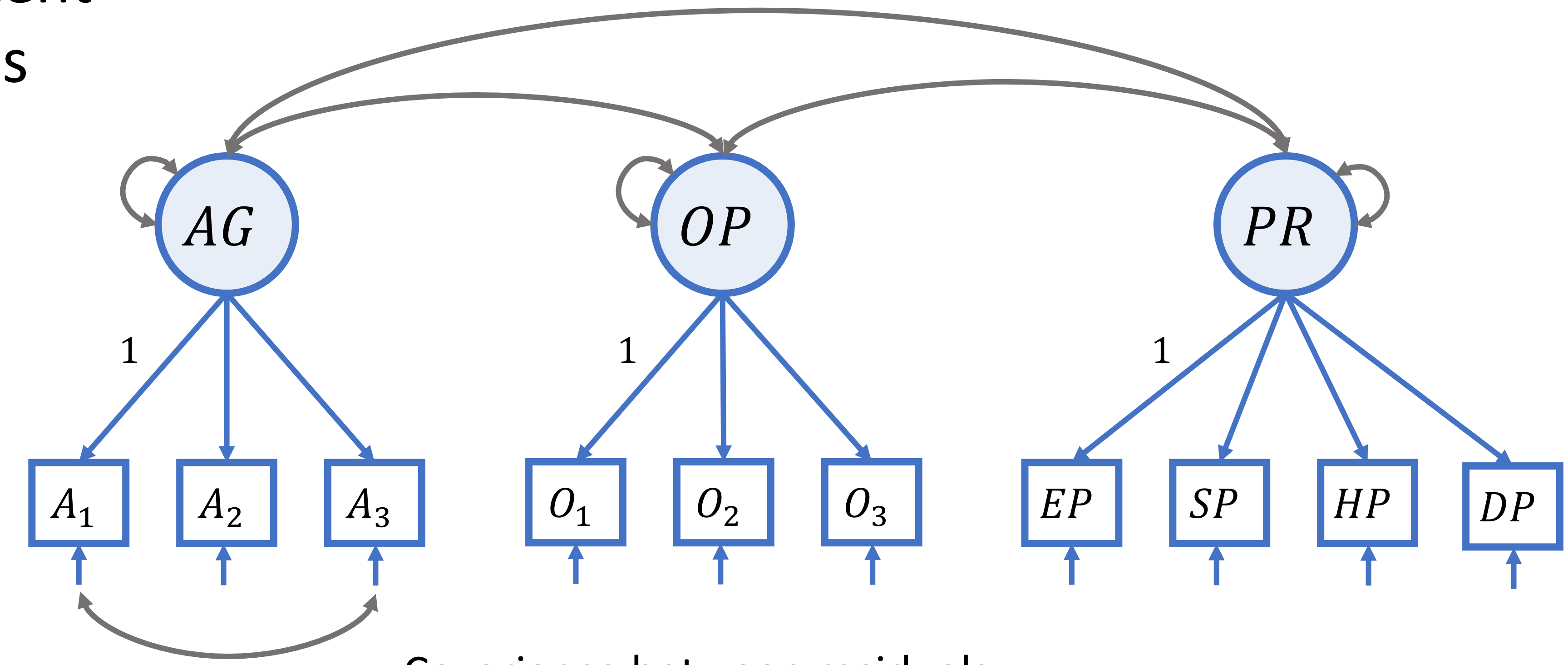


(Bergh et al., 2016; Mair, 2018)

Model 3b: Refined Measurement Model

Confirmatory factor analysis with three latent (unobserved) variables

Does this refined model fit better?



Covariance between residuals (perhaps due to item wording)

Your turn
Specify and estimate this model in lavaan.

Data
$(y_i, a_{1i}, \dots, dp_i)$ for persons $i = 1, \dots, N$

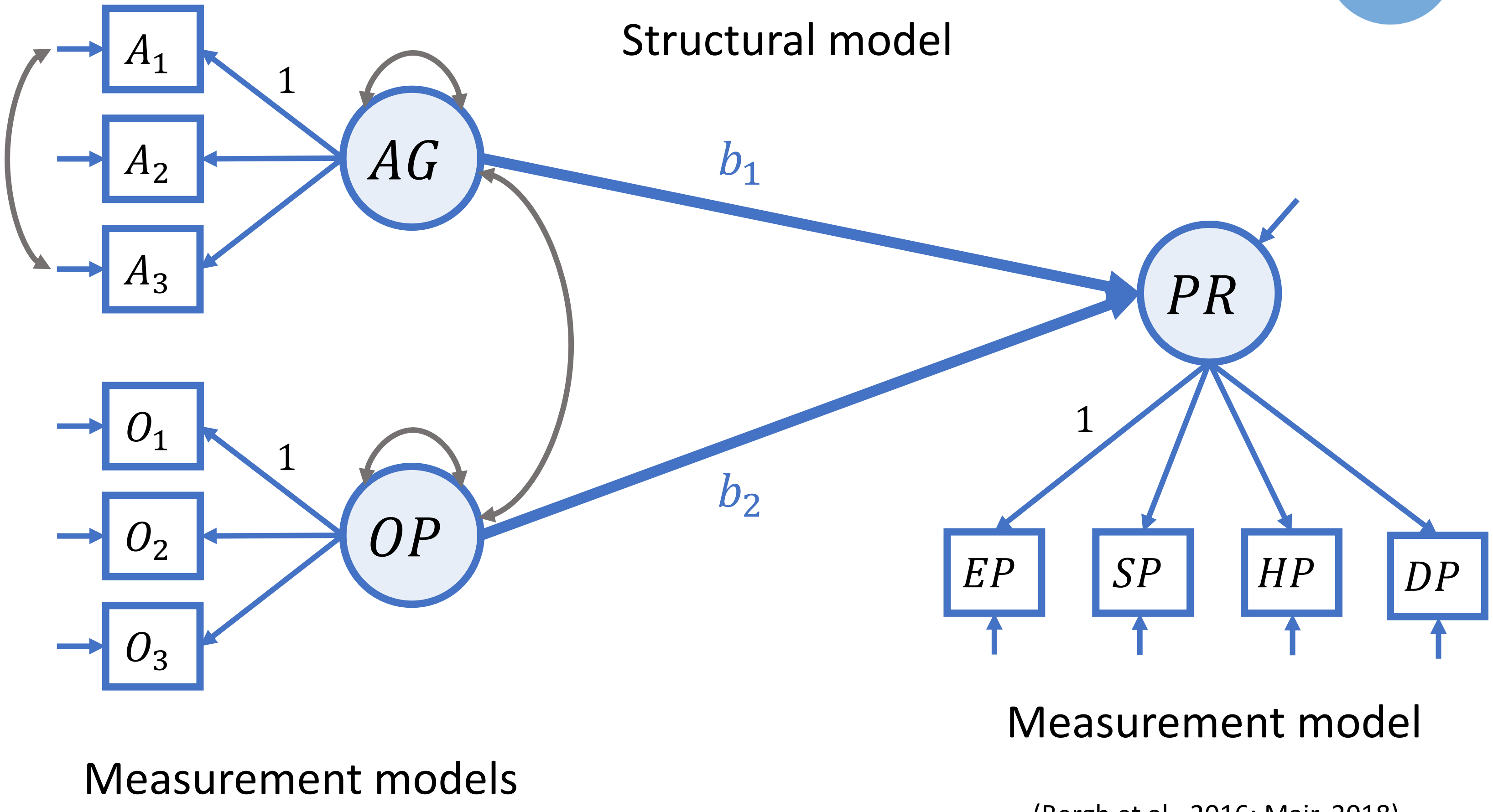
(Bergh et al., 2016; Mair, 2018)

Model 4: Structural Equation Model

Structural equation model with directional paths

Note
This model has a structural and a measurement part.

Data
$(y_i, a_{1i}, \dots, dp_i)$ for persons $i = 1, \dots, N$



(Bergh et al., 2016; Mair, 2018)

SEM in lavaan

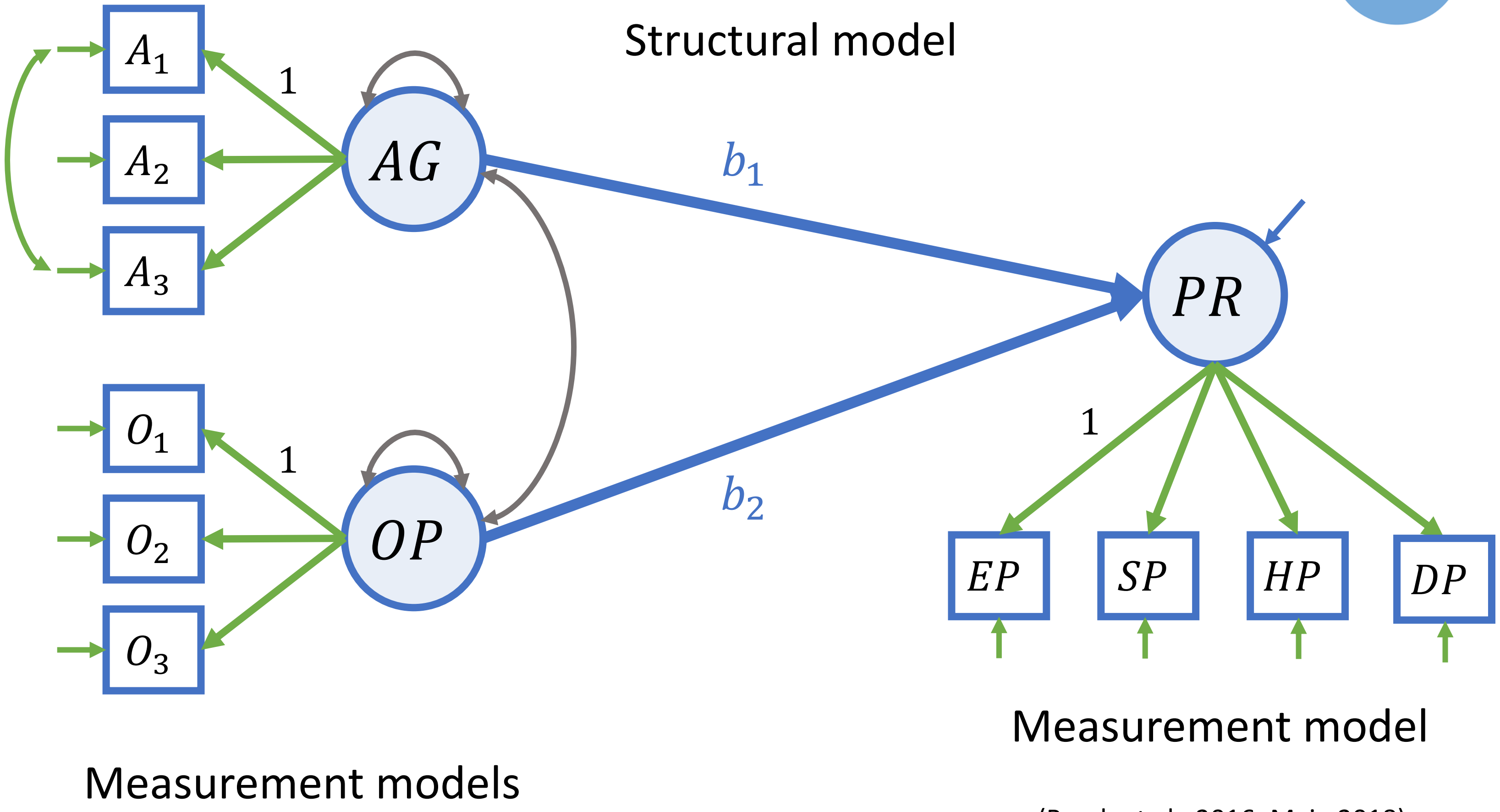
Add-on: Multi-group SEM

Model 5: Multi-Group Structural Equation Model

Structural equation model across gender groups

Note
This model assumes measurement invariance.

Data
$(y_i, a_{1i}, \dots, dp_i)$ for persons $i = 1, \dots, N$



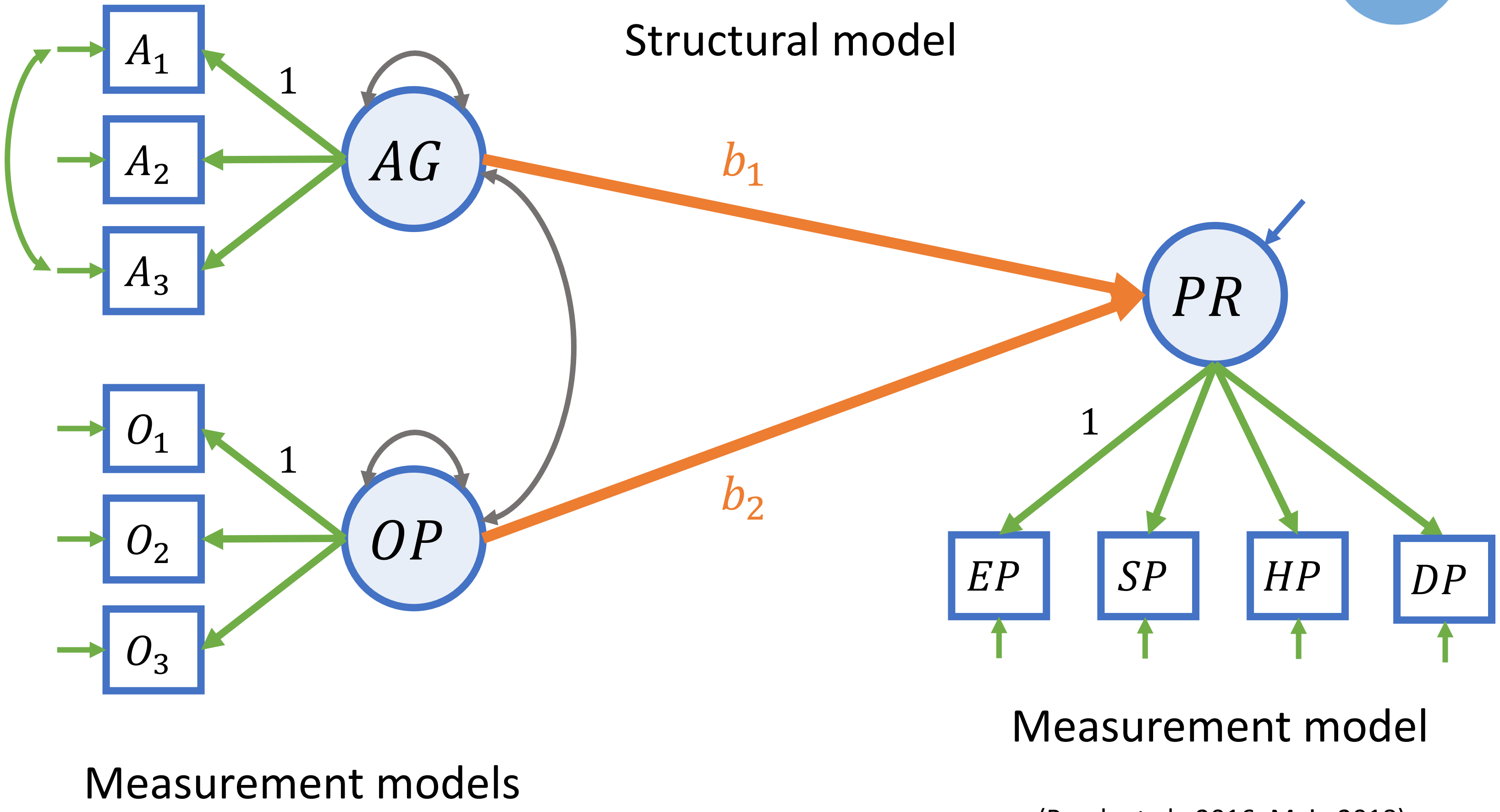
(Bergh et al., 2016; Mair, 2018)

Model 6: Multi-Group Structural Equation Model

Structural equation model across gender groups

Note
This model assumes measurement and structural invariance.

Data
$(y_i, a_{1i}, \dots, dp_i)$ for persons $i = 1, \dots, N$



(Bergh et al., 2016; Mair, 2018)

Wait a minute...

Structural Equation Modeling—What if?

(Rosseel, 2020)

Data are not normally distributed.

- Asymptotically distribution-free estimation
- ML estimation with robust standard errors and/or test statistics
- Bootstrapping

Lavaan:

- `estimator="WLS", "WLSMV",…; blavaan`
- `estimator="MLM", "MLR"; se="robust"; test="Satorra-Bentler"`
- `se="bootstrap"; test="bootstrap"`

Data are missing.

- Full-information-maximum-likelihood
- Two-stage approach (*step 1*: mean vector and covariance matrix, *step 2*: SEM)
- Multiple imputation

Lavaan:

- `missing="FIML"`
- `missing="two.stage"`
- Use `mice` to create multiply imputed datasets and `runMI()` in `semTools`

Data are dependent (hierarchical).

- Multilevel SEM

Lavaan:

- `cluster="…"` + level-specific model specification
- `lavaan.survey`

Thank you!

References

- Hoyle, R. H. (2012). Introduction and overview. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 3-16). New York, NY: The Guilford Press.
- Jöreskog, K., Olsson, K. G., Wallentin, F. T. (2016). *Multivariate Analysis with LISREL*. Cham: Springer International Publishing. DOI:10.1007/978-3-319-33153-9
- Kaplan, D. (2001). Structural equation modeling. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (Vol. 11, pp. 15215-15222). Amsterdam: Elsevier
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4 ed.). New York, London: Guilford Press.
- Mair, P. (2018). *Modern Psychometrics with* R. Cham: Springer International Publishing. DOI: 10.1007/978-3-319-93177-7
- Rosseel, Y. (2020, January). *Structural equation modeling with lavaan*. Geneva: CISA. Retrieved from https://users.ugent.be/~yrosseel/lavaan/geneve2020/lavaan_oneday_geneve2020.pdf, latest access: 27 May 2020