One-stage meta-analytic structural equation modeling

AERA Systematic Review and Meta-Analysis SIG
December 17 2021

dr. Suzanne Jak (S.Jak@uva.nl)
Joint work with prof. dr. Mike W.-L. Cheung (National University of Singapore)

This work is part the project VI.Vidi.201.009 which is financed by the Dutch Research Council (NWO)
Content

1. What is (MA)SEM?
2. Univariate versus multivariate methods for MASEM
3. Simulation study results
4. Discussion
What is SEM?

Confirmatory technique to fit hypothesized models to data
Evaluated indirect effects, relations between latent and observed variables, use fit-indices to compare competing models
E.g. Factor models, path models, full SEM

No raw data needed (covariance matrix and sample size are sufficient)
What is SEM?

Example of a path model: Theory of planned behavior from Ajzen and Fishbein (1980)
What is SEM?

Example of a factor model
What is SEM?

The covariance matrix between the observed variables is modeled as a function of SEM parameters.

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>$\sigma_{11}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V2</td>
<td>$\sigma_{21}$</td>
<td>$\sigma_{22}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>$\sigma_{31}$</td>
<td>$\sigma_{32}$</td>
<td>$\sigma_{33}$</td>
<td></td>
</tr>
<tr>
<td>V4</td>
<td>$\sigma_{41}$</td>
<td>$\sigma_{42}$</td>
<td>$\sigma_{43}$</td>
<td>$\sigma_{44}$</td>
</tr>
</tbody>
</table>
Meta-analytic structural equation modeling

Combining meta-analyse (MA) and structural equation modeling (SEM)

Complete theoretical models

Mediating variables

Model fit

Latent variables
Example
Eltanamly et al. (2021)

War exposure ➔ Parenting behavior ➔ Child behavioral problems
Study 1

War exposure → Behavior control

Behavior control → Anxiety
Behavior control → Depression
Behavior control → PTSD

Harshness → Anxiety
Harshness → Depression
Harshness → PTSD
Study 2

War exposure

Behavior control

Harshness

Anxiety

Depression

PTSD
Study k

War exposure

Behavior control

Harshness

Anxiety

Depression

PTSD
Eltanamly et al. (2021)
38 studies
Factors affecting adoption of online banking: A meta-analytic structural equation modeling study

Ali Reza Montazemi a,1, Hamed Gahri-Saremi b,1

1Fetzer School of Business, McMaster University, 1280 Main Street West, Hamilton, ON, Canada L8S 4M8
2College of Business and Management, University of Illinois at Springfield, One University Plaza, Springfield, IL 62701-5407, USA

© 2017 American Psychological Association
0003-066X/2017/
https://dx.doi.org/10.1037/ama0000097

The Structure of Common Emotion Regulation Strategies: A Meta-Analytic Examination

Kristin Naragon-Gainey, Tienney P. McMahon, and Thomas P. Chacko
University at Buffalo, the State University of New York

https://doi.org/10.1007/s11065-019-09423-6

The Factor Structure of Cognitive Functioning in Cognitively Healthy Participants: A Meta-Analysis and Meta-Analysis of Individual Participant Data

Joost A. Agelink van Rentergem 1,2, Nathalie R. de Vent 1, Ben A. Schmand 2,3,4, Jaap M. J. Murre 1, Janneke P. C. Staaks 3, ANDi Consortium 3, Hilde M. Hulzenga 1,4,6
Methods for MASEM

Option 1: Pool correlations (or covariances), then fit SEM
- Univariate-r (Viswesvaran and Ones, 1995)
- GLS (Becker, 1992)
- Two Stage SEM (Cheung and Chan, 2005; Cheung, 2015)
- One-stage MASEM (Jak and Cheung, 2020)

Option 2: Fit SEM, then pool the SEM parameters
Becker and Wu (2007)
Gnambs and Staufenbiel (2016)
  Disadvantage: Need complete data, and model should fit equally well in all samples
Ke, Zhang and Tong (2019)
  Bayesian method, solves the issues with parameter-based MASEM
Methods for MASEM

Stage 1

Stage 2

Observed correlation matrices in \( k \) studies

Pooled correlation matrix

Pooled correlation matrix

Meta-analytic structural equation model
Univariate-r approach

Stage 1: Use several univariate meta-analyses to pool each correlation coefficient
Univariate-\( r \) approach

Stage 2: Fit the path model on the pooled correlation matrix

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
</tr>
</thead>
<tbody>
<tr>
<td>V2</td>
<td></td>
<td>( r_{21} )</td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>( r_{31} )</td>
<td>( r_{32} )</td>
<td></td>
</tr>
<tr>
<td>V4</td>
<td>( r_{41} )</td>
<td>( r_{42} )</td>
<td>( r_{34} )</td>
</tr>
</tbody>
</table>
Univariate-r approach

Stage 1: Use several univariate meta-analyses to pool each correlation coefficient
   - Ignores sampling covariance
   - Can lead to non-positive definite correlation matrices

Stage 2: Fit the path model on the pooled correlation matrix using standard SEM software
   - What is the sample size?
   - Differences in precision of estimated correlations ignored
   - Between-studies variance not taken into account
Multivariate approaches

GLS, TSSEM and one-stage MASEM involve multivariate meta-analysis of correlation coefficients

![Diagram showing correlations between variables V1, V2, V3, and V4]

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
</tr>
</thead>
<tbody>
<tr>
<td>V2</td>
<td></td>
<td>r_{21}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>r_{31}</td>
<td></td>
<td>r_{32}</td>
<td></td>
</tr>
<tr>
<td>V4</td>
<td>r_{41}</td>
<td>r_{42}</td>
<td></td>
<td>r_{34}</td>
</tr>
</tbody>
</table>

Estimates of average correlations: \( \hat{\rho}_R \)

And between-study (co)variances of correlations across studies: \( \hat{T}^2 \)
Between-studies heterogeneity

GLS, TSSEM and one-stage MASEM

4 variables
6 mean correlations:

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
</tr>
</thead>
<tbody>
<tr>
<td>V2</td>
<td></td>
<td>$r_{21}$</td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>$r_{31}$</td>
<td>$r_{32}$</td>
<td></td>
</tr>
<tr>
<td>V4</td>
<td>$r_{41}$</td>
<td>$r_{42}$</td>
<td>$r_{43}$</td>
</tr>
</tbody>
</table>
Between-studies heterogeneity

GLS, TSSEM and one-stage MASEM

With 6 mean correlations:

<table>
<thead>
<tr>
<th></th>
<th>$r_{21}$</th>
<th>$r_{31}$</th>
<th>$r_{41}$</th>
<th>$r_{32}$</th>
<th>$r_{42}$</th>
<th>$r_{43}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_{21}$</td>
<td>$\tau_{11}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{31}$</td>
<td>$\tau_{21}$</td>
<td>$\tau_{22}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{41}$</td>
<td>$\tau_{31}$</td>
<td>$\tau_{32}$</td>
<td>$\tau_{33}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{32}$</td>
<td>$\tau_{41}$</td>
<td>$\tau_{42}$</td>
<td>$\tau_{43}$</td>
<td>$\tau_{44}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$r_{42}$</td>
<td>$\tau_{51}$</td>
<td>$\tau_{52}$</td>
<td>$\tau_{53}$</td>
<td>$\tau_{54}$</td>
<td>$\tau_{55}$</td>
<td></td>
</tr>
<tr>
<td>$r_{43}$</td>
<td>$\tau_{61}$</td>
<td>$\tau_{62}$</td>
<td>$\tau_{63}$</td>
<td>$\tau_{64}$</td>
<td>$\tau_{65}$</td>
<td>$\tau_{66}$</td>
</tr>
</tbody>
</table>

In practice:
Use diagonal $T^2$
(Becker & Aloe, 2019)
TSSEMs

Stage 1: Random effects multivariate meta-analysis of correlation coefficients using ML estimation

\[ r_i = \rho_R + u_i + \varepsilon_i \]

\[ \text{cov}(\varepsilon_i) = V_i \quad \text{(Within studies covariances)} \]

\[ \text{cov}(u_i) = T^2 \quad \text{(Between studies (co)variances)} \]

Stage 2: Fit SEM with WLS

\[ \hat{F}_{WLS} = (\hat{\rho}_R - \hat{\rho}_{MODEL})^T \hat{V}_R^{-1} (\hat{\rho}_R - \hat{\rho}_{MODEL}) \]

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>V4</th>
</tr>
</thead>
<tbody>
<tr>
<td>V1</td>
<td>σ_{11}</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V2</td>
<td>σ_{21}</td>
<td>σ_{22}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V3</td>
<td>σ_{31}</td>
<td>σ_{32}</td>
<td>σ_{33}</td>
<td></td>
</tr>
<tr>
<td>V4</td>
<td>σ_{41}</td>
<td>σ_{42}</td>
<td>σ_{43}</td>
<td>σ_{44}</td>
</tr>
</tbody>
</table>
GLS approach

Stage 1: Random effects multivariate meta-analysis of correlation coefficients

\[ r = \rho_R + u + \varepsilon \]

\[ \hat{\rho}_R = (X^T \Sigma^{-1} X)^{-1} X^T \Sigma^{-1} r \]

\[ \hat{\mathbf{V}}_R = (X^T \Sigma^{-1} X)^{-1} \]

\[ r = \text{vector of observed correlations} \]
\[ X = \text{stacked selection matrices} \]
\[ \Sigma = \text{block-diagonal matrix with } V_i + T^2 \]

Stage 2: \[ \hat{\mathbf{B}} = \hat{\rho}_{XX}^{-1} \hat{\rho}_{YX} \]

Alternative: use WLS with \( \hat{\rho}_R \) and \( \hat{\mathbf{V}}_R \) like in TSSEM
One-stage MASEM
One-stage MASEM

Multivariate random-effects analysis of correlation coefficients

\[ r_i = \rho_R + u_i + \varepsilon_i \]

One-stage MASEM restricts the pooled correlation matrix to a SEM model:

\[ \rho_R = \text{vechs}(\Lambda \Phi \Lambda^{-1t} + \Theta) \]

The SEM parameters may be regressed on study-level moderator variables
Comparison of four methods

Data generating model
Based on Nohe et al. (2015)

Fit correct model using univariate-r, GLS, TSSEM and one-stage MASEM
Comparison of four methods

Fit correct path model using

**Univariate-r**

- Univariate meta-analyses of raw correlations, weighted by $N$
- Using harmonic mean as sample size in Stage 2

**Multivariate methods** (GLS, TSSEM, one-stage MASEM)

- With diagonal $T^2$ (although generated full $T^2$)
Conditions

Varying:
- Number of studies
  \( k = 16, 32 \text{ or } 64 \)
- Missing data
  0, 50 or 75\% of the studies missed 2 of the 4 variables
- Heterogeneity \((T^2)\)
  Between-studies SD .05 or .10
  Between studies correlations .10 or .30

Fixed:
- Sample size per study
Evaluation

Rejection percentages of test statistic
Relative bias in parameter estimates
Relative bias in standard errors
Test statistic (rejection percentages)

Rejection rates of the test statistics (expected counts=5 for alpha=.05)
Bias in parameter estimates

K = 32, SD = 0.1, cor = 0.3, missing studies = 0.5
Bias in parameter estimates

Relative percentage bias of parameter estimates

Method:
- gls
- osmasem
- tssem
- unir
Bias in standard errors

K = 32, SD = 0.1, cor = 0.3, missing studies = 0.5
Bias in standard errors
Conclusions

Univariate-r approach leads to unbiased parameter estimates, but:

- Extremely inflated test-statistics and associated Type 1 errors
- Extremely biased standard errors

The three multivariate methods generally lead to unbiased parameter estimates, well behaved test-statistics, and correct standard errors.
Discussion

Within the Journal of Applied Psychology, all-but-one of the MASEM applications since 2020 used the univariate-r method.
Discussion

Multivariate methods also need more research
   Minimum sample size conditions (e.g., how many studies are needed in which conditions?)
   Handling dependent effect sizes

However, it is worrying how often researchers apply the only method that clearly leads to wrong results in all situations
Discussion

Why is the univariate method (still) so popular?

Because it is easier to apply?

Not anymore!

 Tutorial and Shiny app for one-stage MASEM (Jak et al. 2021)

https://sjak.shinyapps.io/webMASEM/
Discussion

Why is the univariate method (still) so popular?

Because researchers copy the current practice in their field.

Harder to change, but we are hopeful.

Thanks for listening!


