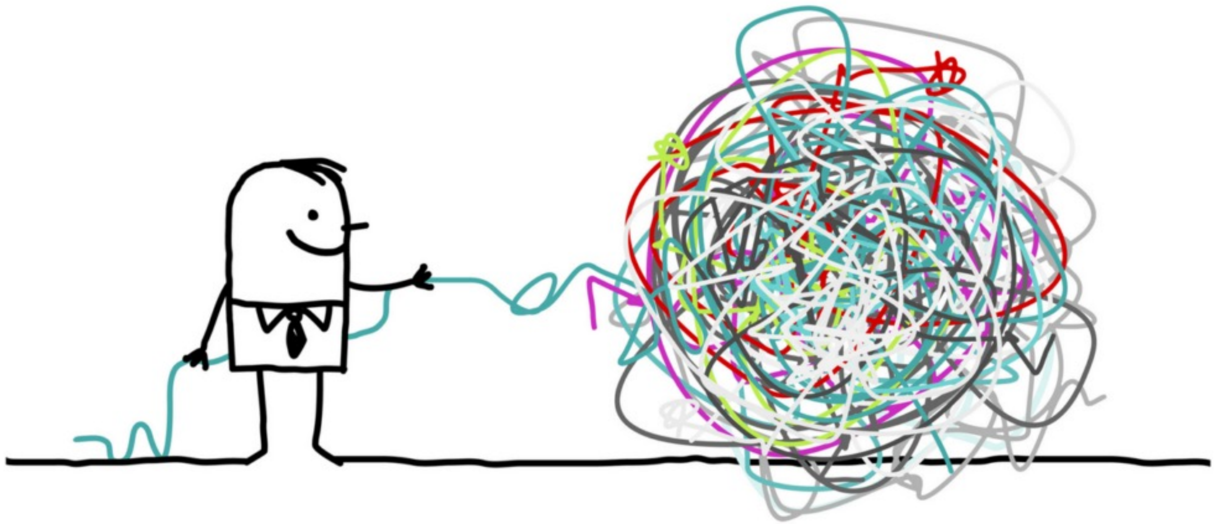
VENI
VIDI
VICI
NWO
Netherlands Organisation for Scientific Research



The mixture approach to finding measurement invariance across countries

Kim De Roover, KU Leuven/Tilburg University

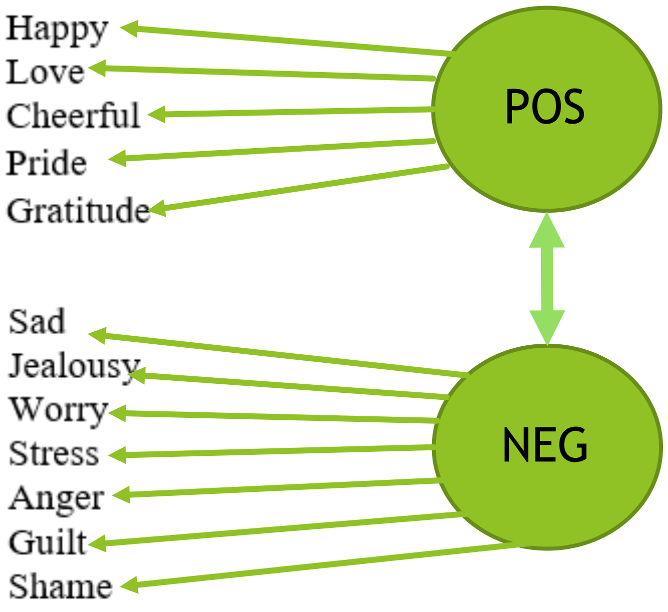
@Kim_De_Roover

Measurement invariance

- Social scientists often measure **latent constructs** (e.g., personality traits, attitudes, wellbeing, depression)
- To ensure valid conclusions about comparisons w.r.t. latent constructs, they should be **measured in exactly the same way** across the entire data set (e.g., across groups)

Empirical example: Social value of emotions in 47 countries

- Assumed measurement model (MM):



"How appropriate and valued is each of the following emotions in your society? Do people approve of this emotion?"

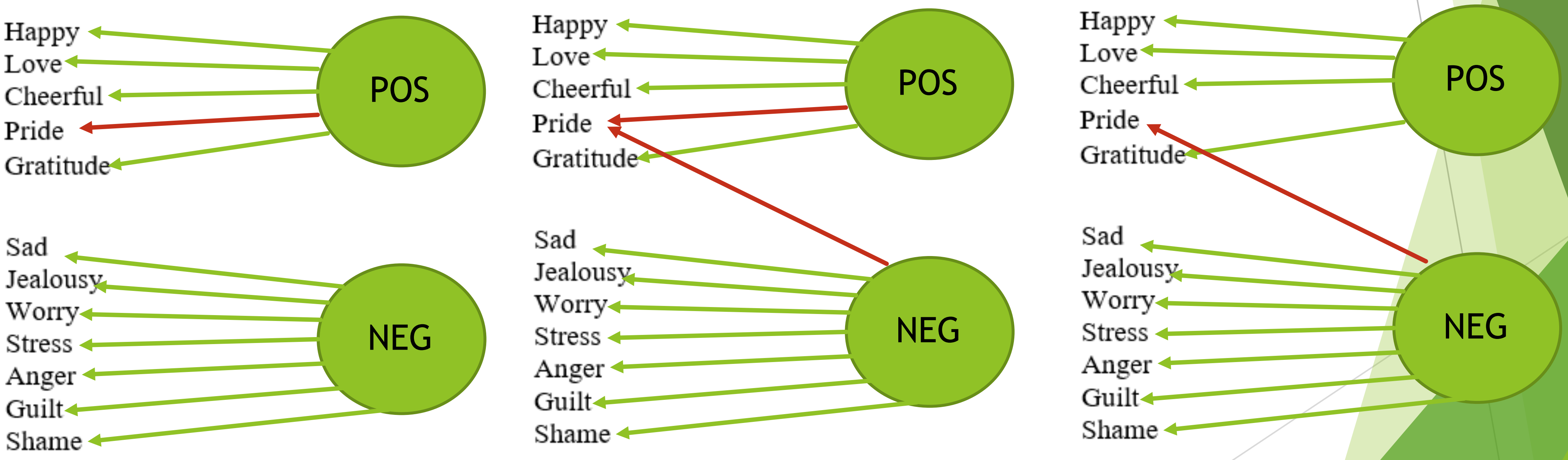
1-------2-------3-------4-------5-------6-------7-------8-------9
not at all very much

- Happy
- Love
- Sad
- Jealousy (in romantic situations)
- Cheerful
- Worry
- Stress
- Anger
- Pride
- Guilt
- Shame
- Gratitude

Bastian, B., Kuppens, P., De Roover, K., & Diener, E. (2014). Is valuing positive emotion associated with life satisfaction? *Emotion, 14*(4), 639–645.

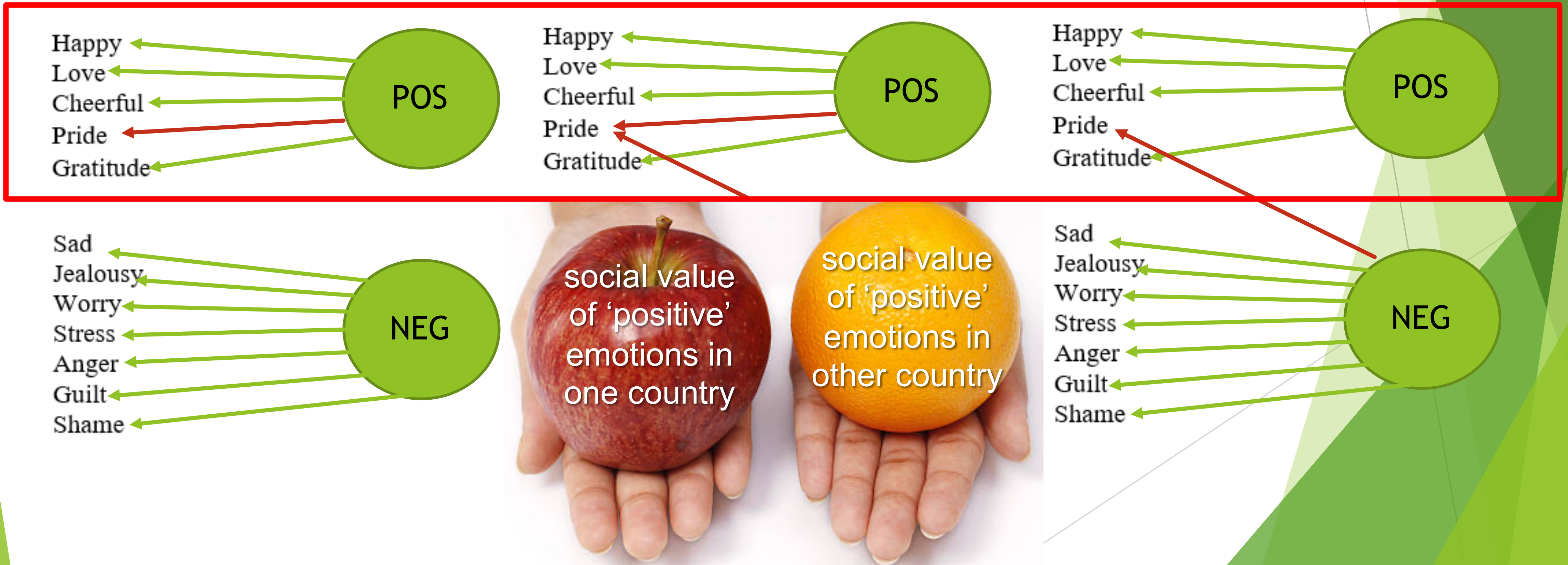
Empirical example: Social value of emotions in 47 countries

- Actual measurement model may differ across countries: For instance, pride is evoked when personal goals are achieved and is thus highly valued in individualistic cultures and less so in collectivistic ones (Eid & Diener, 2001).



Empirical example: Social value of emotions in 47 countries

- Actual measurement model may differ across countries: For instance, pride is evoked when personal goals are achieved and is thus highly valued in individualistic cultures and less so in collectivistic ones (Eid & Diener, 2001).



Levels of measurement (non-)invariance

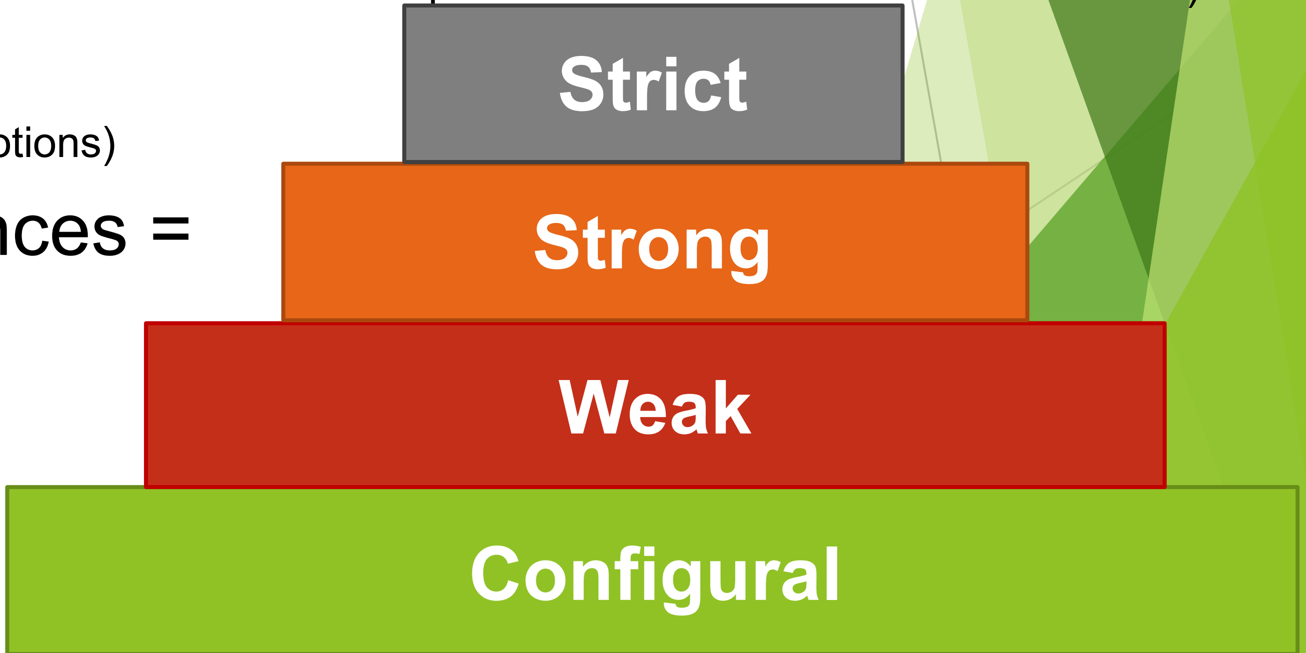
- Configural: number of factors & pattern of zero loadings =
- Weak/metric: size non-zero loadings =

➔ Latent *covariances/regression effects* can be compared (e.g., how social value of positive emotions affects life satisfaction)

- Strong/scalar: intercepts =

➔ Latent *means* can be compared as well
(e.g., group-means of social value of positive emotions)

- Strict: residual/unique variances =

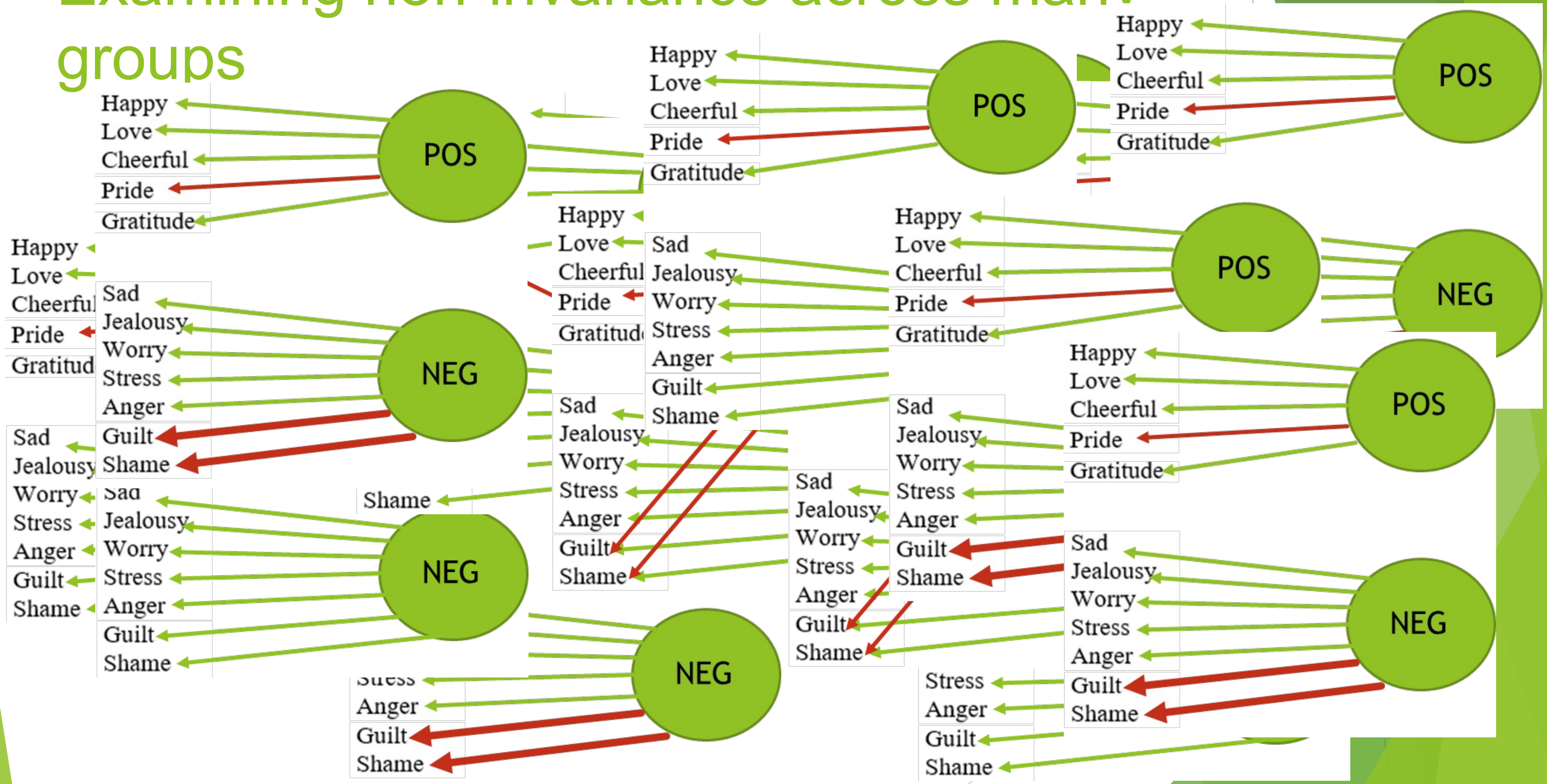


Measurement invariance across many groups

- Measurement invariance often does *not* hold across many groups (Boer, Hanke, & He, 2018)
- Methods for capturing non-invariance across many groups (Kim et al., 2017):
  - Multigroup CFA
  - Multilevel CFA
  - Approximate measurement invariance (Bayesian multigroup SEM)
  - Multigroup factor alignment
  - Multilevel factor mixture modeling

Kim, E. S., Cao, C., Wang, Y., & Nguyen, D. T. (2017). Measurement invariance testing with many groups: A comparison of five approaches. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(4), 524-544.

Examining non-invariance across many groups



Measurement invariance across many groups

- Measurement invariance often does *not* hold across many groups (Boer, Hanke, & He, 2018)
- Methods for *capturing* non-invariance across many groups (Kim et al., 2017):
  - ~~Multigroup CFA~~
  - ~~Multilevel CFA~~
  - ~~Approximate measurement invariance (Bayesian multigroup SEM)~~
  - ~~Multigroup factor alignment~~ ➔ 659 pages of output for the emotions data!
  - Multilevel factor mixture modeling

Not suitable for ***comparing*** MM parameters *across many groups*

Measurement invariance across many groups

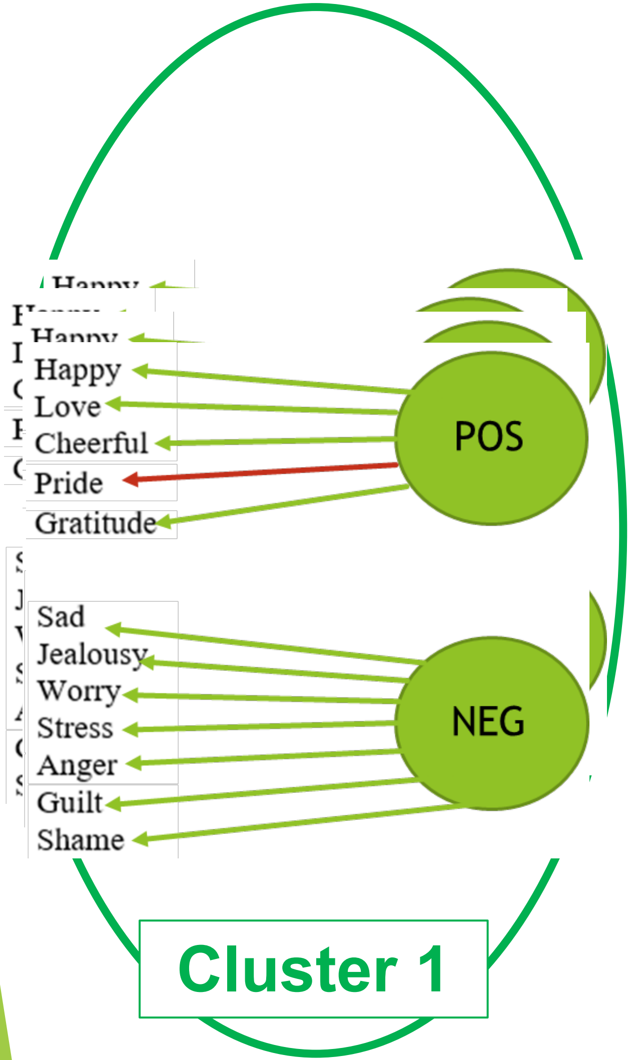
- With (Bayesian) multigroup CFA, multilevel CFA and multigroup factor alignment, we are left to wonder about the following:
  - Do (some) non-invariant groups share MM parameters?
  - What do the alternative measurement models look like?
- In case of many groups, it is likely that some groups have the same MM, so that latent classes emerge
- These latent classes can be captured by mixture approaches
- Like multilevel factor mixture modeling (Kim et al., 2016, 2017)

Kim, E. S., Joo, S. H., Lee, P., Wang, Y., & Stark, S. (2016). Measurement invariance testing across between-level latent classes using multilevel factor mixture modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *23*(6), 870-887.

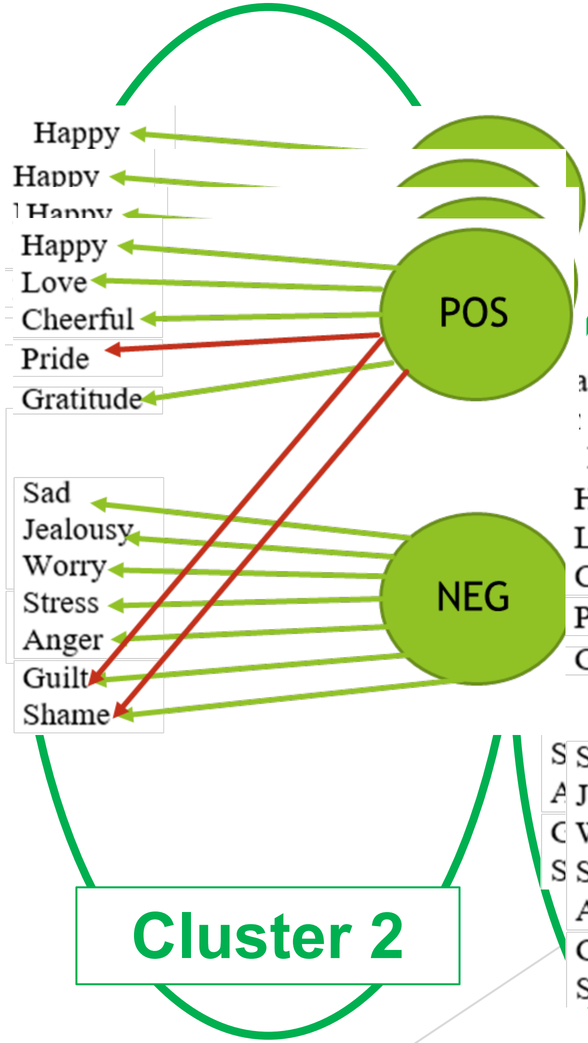
Multilevel factor mixture modeling

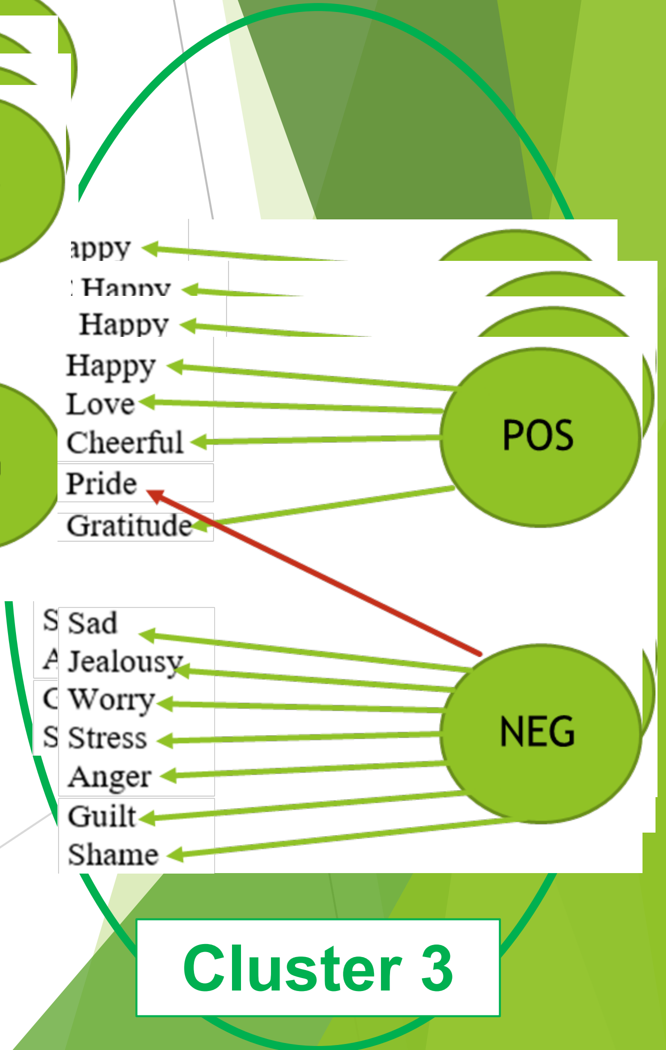
- Finds clusters of groups according to factor model parameters.
- Parameters can either be invariant or cluster-specific.
- ➔ Clusters groups on *all* measurement *and* structural parameters at the same time.
- ➔ Assumes that same clustering underlies all parameter differences, but this may not be the case. Some parameters may even be group-specific.
- ➔ Needs more clusters to capture all differences properly. Or a mix of differences is picked up by the clustering.
- ➔ Does not distinguish between different levels of measurement (non-)invariance.
- See simulation study in: De Roover, K. (2021). Finding clusters of groups with measurement invariance: Unraveling intercept non-invariance with mixture multigroup factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(5), 663-683.

Mixture multigroup factor analysis



- finds clusters of groups **focused on MM**
- a **specific level** of measurement invariance holds within each cluster = **clusterwise measurement invariance**
- cluster-specific models allow to find out *how* MMs differ by **comparing *less* models**
- **EFA- or CFA-based**





Mixture multigroup factor analysis for finding clusterwise **weak** invariance

For validly comparing latent *covariances* or *regression effects* within clusters

- Clusters groups based *on loadings only*!

$$f(\mathbf{X}_g;\boldsymbol{\theta}) = \sum_{k=1}^{K} \pi_k \prod_{n_g=1}^{N_g} MVN(\mathbf{x}_{n_g};\boldsymbol{\mu}_g,\boldsymbol{\Sigma}_{gk})$$

$$\text{with} \quad \boldsymbol{\Sigma}_{gk} = \boldsymbol{\Lambda}_k \boldsymbol{\Phi}_{gk} \boldsymbol{\Lambda}_k' + \boldsymbol{\Psi}_g \quad \text{and} \quad \boldsymbol{\mu}_g = \boldsymbol{\tau}_g \quad (\boldsymbol{\alpha}_g = 0)$$

structural parameters

De Roover, K., Vermunt, J.K., & Ceulemans, E. (2022). Mixture multigroup factor analysis for unraveling factor loading non-invariance across many groups. *Psychological Methods, 27*(3), 281–306.

Mixture multigroup factor analysis for finding clusterwise **strong** invariance (1)

For validly comparing latent *means* within clusters

- Building on overall weak invariance (invariant $\mathbf{\Lambda}$), clusters groups *based on intercepts only*

$$f\left(\mathbf{X}_g;\boldsymbol{\theta}\right)=\sum_{k=1}^{K}\pi_k\prod_{n_g=1}^{N_g}MVN(\mathbf{x}_{n_g};\boldsymbol{\mu}_{gk},\boldsymbol{\Sigma}_g)$$

$$with \quad \boldsymbol{\Sigma}_g=\mathbf{\Lambda}\mathbf{\Phi}_g\mathbf{\Lambda}'+\mathbf{\Psi}_g \quad and \quad \boldsymbol{\mu}_{gk}=\boldsymbol{\tau}_k+\mathbf{\Lambda}\boldsymbol{\alpha}_{gk}$$

De Roover, K. (2021). Finding clusters of groups with measurement invariance: Unraveling intercept non-invariance with mixture multigroup factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal, 28*(5), 663-683.

For validly comparing latent *means* within clusters

Mixture multigroup factor analysis for finding clusterwise **strong** invariance (2)

- Clusters groups *based on loadings AND intercepts*

$$f\left(\mathbf{X}_g;\boldsymbol{\theta}\right)=\sum_{k=1}^{K}\pi_k\prod_{n_g=1}^{N_g}MVN\left(\mathbf{x}_{n_g};\boldsymbol{\mu}_{gk},\boldsymbol{\Sigma}_{gk}\right)$$

$$with \quad \boldsymbol{\Sigma}_{gk}=\boldsymbol{\Lambda}_k\boldsymbol{\Phi}_{gk}\boldsymbol{\Lambda}_k'+\boldsymbol{\Psi}_g \quad and \quad \boldsymbol{\mu}_{gk}=\boldsymbol{\tau}_k+\boldsymbol{\Lambda}_k\boldsymbol{\alpha}_{gk}$$

Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., **De Roover, K.**, Jak, S., Meitinger, K., Menold, N., Muthén, B. Rudnev, M., Schmidt, P., & van de Schoot, R. (2022). Measurement Invariance in the Social Sciences: Historical Development, Methodological Challenges, State of the Art, and Future Perspectives. *Social Science Research*, 102805.

Mixture multigroup factor analysis for finding clusterwise **strict** invariance (1)

- Building on overall strong invariance, clusters groups *based on residual variances only*

$$f\left(\mathbf{X}_g;\boldsymbol{\theta}\right)=\sum_{k=1}^{K}\pi_k\prod_{n_g=1}^{N_g}MVN(\mathbf{x}_{n_g};\boldsymbol{\mu}_g,\boldsymbol{\Sigma}_{gk})$$

$$\textit{with}\quad \boldsymbol{\Sigma}_{gk}=\boldsymbol{\Lambda}\boldsymbol{\Phi}_g\boldsymbol{\Lambda}'+\boldsymbol{\Psi}_k\quad \textit{and}\quad \boldsymbol{\mu}_g=\boldsymbol{\tau}+\boldsymbol{\Lambda}\boldsymbol{\alpha}_g$$

Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., **De Roover, K.**, Jak, S., Meitinger, K., Menold, N., Muthén, B. Rudnev, M., Schmidt, P., & van de Schoot, R. (2022). Measurement Invariance in the Social Sciences: Historical Development, Methodological Challenges, State of the Art, and Future Perspectives. *Social Science Research*, 102805.

Mixture multigroup factor analysis for finding clusterwise **strict** invariance (2)

- Building on overall weak invariance (or not), clusters groups *based on (loadings,) intercepts and residual variances*

$$f\left(\mathbf{X}_g;\boldsymbol{\theta}\right)=\sum_{k=1}^{K}\pi_k\prod_{n_g=1}^{N_g}MVN(\mathbf{x}_{n_g};\boldsymbol{\mu}_{gk},\boldsymbol{\Sigma}_{gk})$$

$$with \quad \boldsymbol{\Sigma}_{gk}=\boldsymbol{\Lambda}\boldsymbol{\Phi}_g\boldsymbol{\Lambda}'+\boldsymbol{\Psi}_k \quad and \quad \boldsymbol{\mu}_{gk}=\boldsymbol{\tau}_k+\boldsymbol{\Lambda}\boldsymbol{\alpha}_{gk}$$

$$or \quad \boldsymbol{\Sigma}_{gk}=\boldsymbol{\Lambda}_k\boldsymbol{\Phi}_{gk}\boldsymbol{\Lambda}_k'+\boldsymbol{\Psi}_k \quad and \quad \boldsymbol{\mu}_{gk}=\boldsymbol{\tau}_k+\boldsymbol{\Lambda}_k\boldsymbol{\alpha}_{gk}$$

Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., **De Roover, K.**, Jak, S., Meitinger, K., Menold, N., Muthén, B. Rudnev, M., Schmidt, P., & van de Schoot, R. (2022). Measurement Invariance in the Social Sciences: Historical Development, Methodological Challenges, State of the Art, and Future Perspectives. *Social Science Research*, 102805.

‘mixmgfa’ package

- Estimated with tailor-made **EM algorithms**
- Available in LatentGOLD 6.0 (using ‘emfa’ option)
- and in ‘mixmgfa’ package (github.com/KimDeRoover/mixmgfa)

Mixture Multigroup Factor Analysis

Description

Perform mixture multigroup factor analyses (MMG-FA) with multiple numbers of clusters.

Usage

```
mixmgfa(
  data,
  N_gs = c(),
  nfactors = 1,
  cluster.spec = c("loadings", "intercepts", "residuals"),
  nsclust = c(1, 5),
  maxiter = 5000,
  nruns = 25,
  design = 0,
  rotation = 0,
  preselect = 10
)
```

Arguments

`data`	A list consisting of "$covariances" (a vertically concatenated matrix or list of group-specific (co)variance matrices) and "$means" (a matrix with rows = group-specific means); or a matrix containing the vertically concatenated raw data for all groups. Note: In case of raw data input without specifying N_gs, the first column of the data should contain group IDs. The remaining variables are then factor-analyzed.
`N_gs`	Vector with number of subjects (sample size) for each group (in the same order as they appear in the data). If left unspecified in case of raw data input, this vector is derived from the first column of the data matrix. If left unspecified in case of covariance matrix & means input, a warning is issued.
`nfactors`	Number of factors.
`cluster.spec`	Measurement parameters you want to cluster the groups on; "loadings", "intercepts", "residuals", c("loadings","intercepts"), c("intercepts","residuals"), or c("loadings","intercepts","residuals"). Note: cluster.spec = "intercepts" and cluster.spec = c("intercepts","residuals") impose invariant loadings across all groups, cluster.spec = "residuals" also imposes invariant intercepts across all groups.
`nsclust`	Vector of length two, indicating the minimal and maximal number of clusters (it is recommended to set the minimal number to one).
`maxiter`	Maximum number of iterations used in each MMG-FA analysis. Increase in case of non-convergence.
`nruns`	Number of (preselected) random starts (important for avoiding local maxima in case of few groups and/or small groups).
`design`	For confirmatory factor analysis, matrix (with ncol = nfactors) indicating position of zero loadings with '0' and non-zero loadings with '1'. Leave unspecified for exploratory factor analysis (EFA). (Using different design matrices for different clusters is currently not supported.)
`rotation`	Rotation criterion to use in case of EFA; currently either "oblimin" or "varimax" (0 = no rotation). (Note: For now, you need to install the GPArotation package for using rotation options.)
`preselect`	Percentage of best starts taken in pre-selection of initial partitions (for huge datasets, increase to speed up multistart procedure).

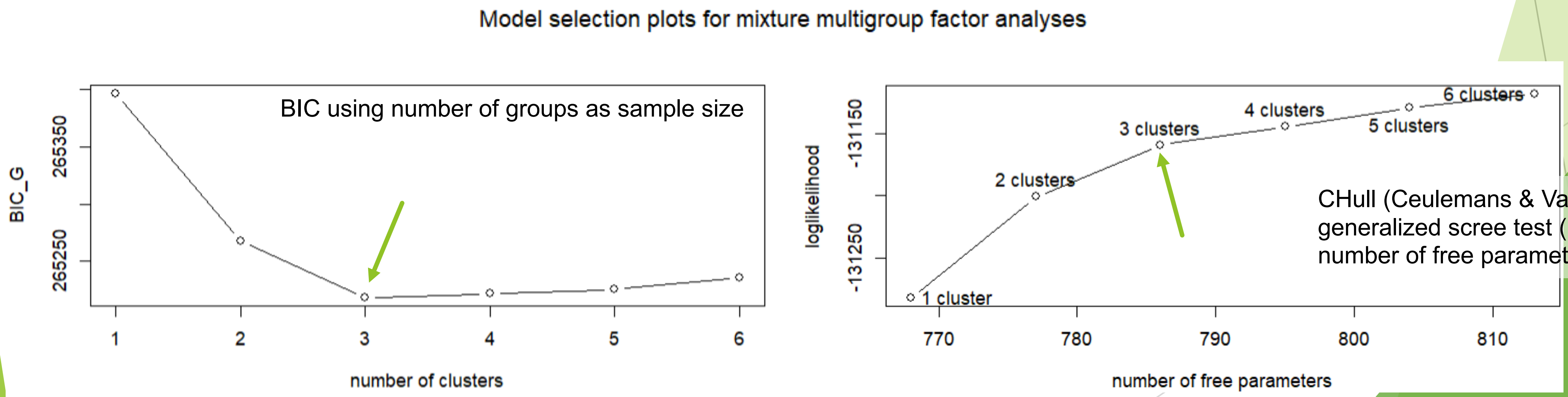
Model selection: How many clusters?

When in doubt: compare solutions with different # of clusters and/or test (full/partial/approximate) measurement invariance per cluster with 'lavaan' or 'blavaan'

```
> MSoutput<-mixmgfa(S_means_gs,N_gs, nfactors=1, cluster.spec=c("loadings"),nsclust=c(1,6),maxiter = 5000,nruns = 50,design=0)
Fitting MMG-FA with 1 cluster ...
Fitting MMG-FA with 2 clusters ...
Fitting MMG-FA with 3 clusters ...
Fitting MMG-FA with 4 clusters ...
Fitting MMG-FA with 5 clusters ...
Fitting MMG-FA with 6 clusters ...

     nr of clusters    loglik nrpars    BIC_N    BIC_G screeratios convergence nr.activated.constraints
[1,]              1 -131281.7    768 269393.8 265396.4          NA           1                        0
[2,]              2 -131200.9    777 269312.3 265268.0    1.947881           1                        0
[3,]              3 -131159.4    786 269309.4 265218.3    2.779855           1                        0
[4,]              4 -131144.5    795 269359.6 265221.6    1.008594           1                        0
[5,]              5 -131129.7    804 269410.1 265225.3    1.312947           1                        0
[6,]              6 -131118.4    813 269467.6 265235.9          NA           1                        0

Choose the best number of clusters ('K_best') based on the BIC_G and CHull scree ratios and the plots. For plots, use 'plot(OutputObject$overview)'.
```



CHull (Ceulemans & Van Mechelen, 2005): generalized scree test (model fit versus number of free parameters)

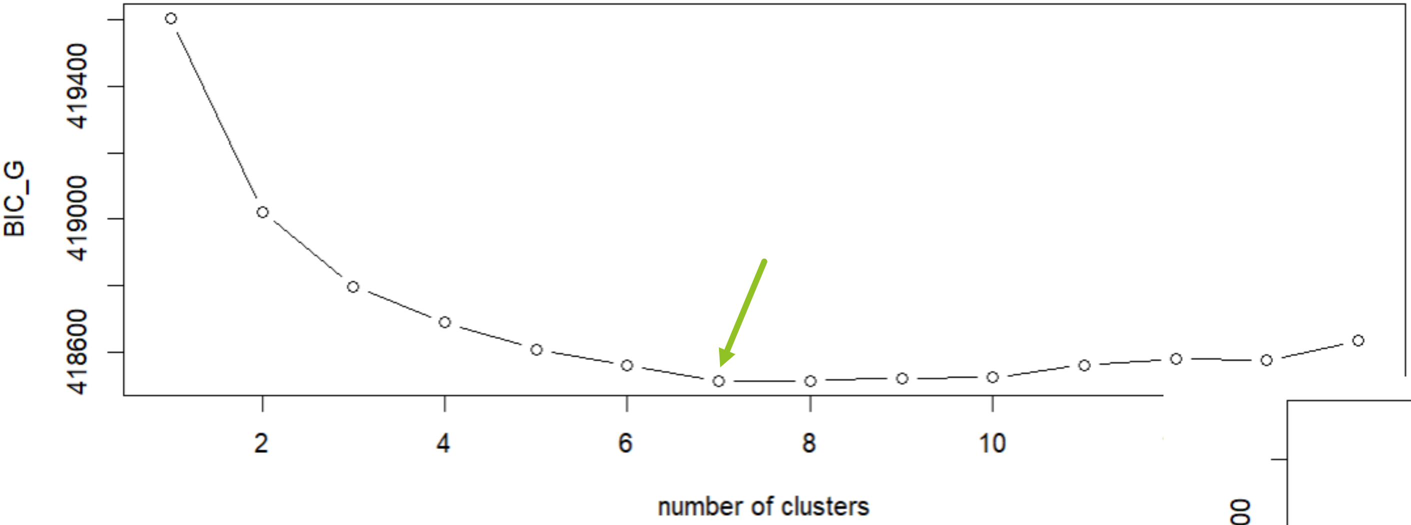
Empirical example on social value of emotions

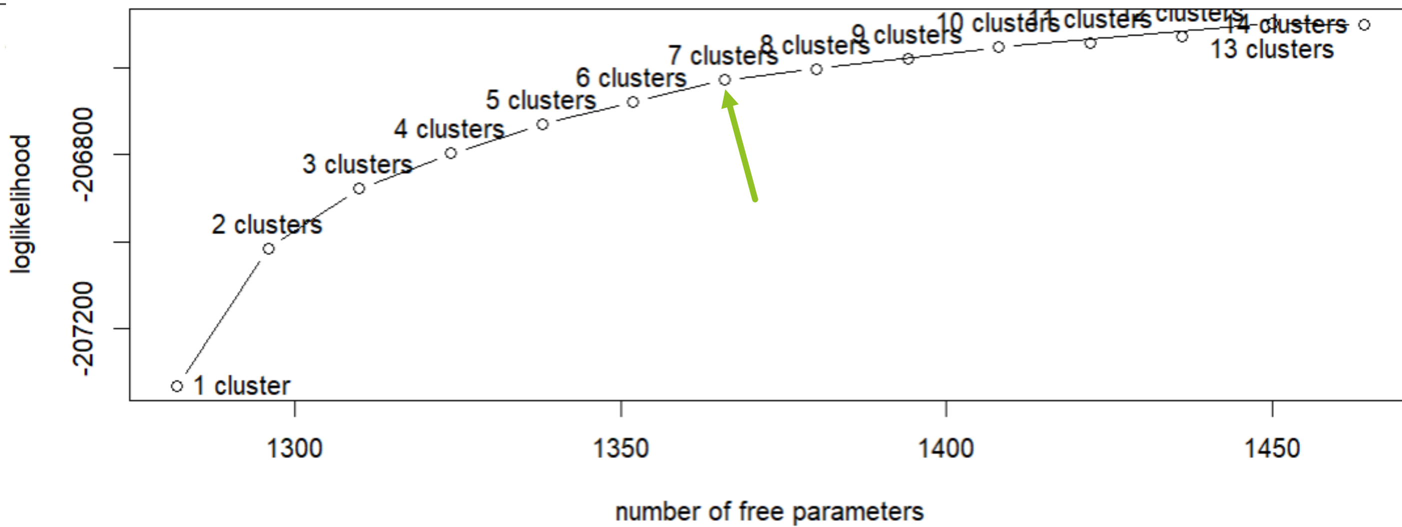
- Configural invariance model (imposing assumed MM, estimator = MLM): CFI = .819, RMSEA = .106
- EFA-based analyses ➔ added crossloadings for pride, guilt and shame in CFA-based model
  - This strategy is called ‘ECFA’ and is preferred over performing many model modifications (see Nájera, Abad, & Sorrel, 2023)
  - See De Roover et al. (2022, Psychological Methods) for the results of EFA-based MMG-FA
- Using MMG-FA to find clusterwise metric invariance→ how many clusters?

Nájera, P., Abad, F. J., & Sorrel, M. A. (2023). Is exploratory factor analysis always to be preferred? A systematic comparison of factor analytic techniques throughout the confirmatory–exploratory continuum. *Psychological Methods*. Advance online publication.

Model selection for empirical application

nr of clusters	loglik	nrpars	BIC_N	BIC_G	screeratios	convergence	nr.activated.constraints
1	-207333.7	1282	426307.3	419603.3	NA	1	0
2	-207015.9	1296	425798.7	419021.6	2.289808	1	0
3	-206877.1	1310	425648.2	418797.9	1.716673	1	0
4	-206796.2	1324	425613.6	418690.1	1.199854	1	0
5	-206728.8	1338	425606.0	418609.2	1.313638	1	0
6	-206677.6	1352	425630.5	418560.5	1.017998	1	0
7	-206627.2	1366	425656.8	418513.6	1.930388	1	0
8	-206601.1	1380	425731.7	418515.3	1.059951	1	0
9	-206577.5	1394	425811.8	418522.1	NA	1	0
10	-206551.8	1408	425887.4	418524.6	1.331920	1	0
11	-206543.2	1422	425997.4	418561.3	NA	1	0
12	-206526.6	1436	426091.2	418582.0	NA	1	0
13	-206496.3	1450	426157.8	418575.4	NA	1	0
14	-206499.2	1464	426290.6	418634.9	NA	1	0





Source

Console | Terminal | Background Jobs

R 4.2.2 · ~/mixmgfa/

```
Netherlands     0.9924   0.0004   0   0   0.0035   0.0001   0.0036
Malaysia        0.0000   0.0000   0   0   0.0000   1.0000   0.0000
Georgia         0.0000   0.9318   0   0   0.0063   0.0000   0.0619
Croatia         0.0000   0.0001   0   0   0.0000   0.0000   0.9999
Ghana           0.0000   0.0000   0   0   0.0000   1.0000   0.0000
Bulgaria        0.0000   0.0000   1   0   0.0000   0.0000   0.0000
Bangladesh      0.0000   0.9580   0   0   0.0371   0.0000   0.0049
Russia          0.0000   0.1447   0   0   0.0001   0.0124   0.8428
Slovakia        0.0000   1.0000   0   0   0.0000   0.0000   0.0000
Zimbabwe        0.9990   0.0000   0   0   0.0000   0.0000   0.0010
Germany         0.0020   0.0016   0   0   0.0000   0.0000   0.9964
Kuwait          0.1329   0.0032   0   0   0.0000   0.0000   0.8639
Colombia        1.0000   0.0000   0   0   0.0000   0.0000   0.0000
Brazil          0.0000   0.0000   0   0   0.0000   1.0000   0.0000
Cameroon        0.0000   0.0001   0   0   0.0000   0.9998   0.0001
Canada          0.0002   0.0000   0   0   0.0000   0.0000   0.9998
India           0.0000   0.9970   0   0   0.0003   0.0001   0.0026
South Africa    0.0000   0.0128   0   0   0.0000   0.9716   0.0156
Austria         0.0000   0.0001   0   0   0.0000   0.9999   0.0000


Groups modally assigned to cluster 1:
Chile Spain Mexico Venezuela Netherlands Zimbabwe Colombia
Groups modally assigned to cluster 2:
Hungary Poland Georgia Bangladesh Slovakia India
Groups modally assigned to cluster 3:
Thailand Bulgaria
Groups modally assigned to cluster 4:
Uganda
Groups modally assigned to cluster 5:
Korea Rep. Japan Indonesia
Groups modally assigned to cluster 6:
Turkey Nigeria China Hong Kong Iran Philippines Nepal Italy Belgium Portugal Malaysia Ghana Brazil Cameroon South Africa Austria
Groups modally assigned to cluster 7:
United States Slovenia Australia Greece Cyprus Switzerland Singapore Croatia Russia Germany Kuwait Canada


cluster proportions:

Cluster_1 Cluster_2 Cluster_3 Cluster_4 Cluster_5 Cluster_6 Cluster_7
   0.1521    0.1291    0.0426    0.0213    0.0648    0.3405    0.2496
```

Modal cluster assignments
(classification probabilities < .99 between brackets)

- Cluster 1: Chile, Spain, Mexico, Venezuela, Colombia, Netherlands, Zimbabwe
- Cluster 2: Hungary, Poland, Georgia (.93), Slovakia, Bangladesh (.96), India
- Cluster 3: Thailand, Bulgaria
- Cluster 4: Uganda
- Cluster 5: Korea Rep., Japan, Indonesia
- Cluster 6: Turkey, Nigeria, China, Hong Kong, Iran, Philippines, Nepal, Italy, Belgium (.96), Portugal, Malaysia, Ghana, Brazil, Cameroon, South Africa, Austria
- Cluster 7: United States, Slovenia, Australia, Greece, Cyprus, Switzerland, Singapore (.92), Croatia, Russia (.84), Germany, Kuwait (.86), Canada

cluster-specif Cluster 1 s:

	Cluster 1		Cluster 2		Cluster 3		Cluster 4		Cluster 5		Cluster 6	
	NEG	POS	NEG	POS	NEG	POS	NEG	POS	NEG	POS	NEG	POS
Happy	0.0000	1.3869	0.0000	1.3408	0.0000	0.7165	0.0000	1.3210	0.0000	1.2389	0.0000	1.4856
Love	0.0000	1.3036	0.0000	1.3610	0.0000	0.6208	0.0000	1.2630	0.0000	1.1500	0.0000	1.4794
Sad	1.2906	0.0000	1.3671	0.0000	1.2772	0.0000	1.4776	0.0000	1.3171	0.0000	1.2282	0.0000
Jealousy	1.3518	0.0000	1.2417	0.0000	1.0316	0.0000	1.0410	0.0000	1.0386	0.0000	1.2573	0.0000
Cheerful	0.0000	1.2427	0.0000	1.1287	0.0000	0.7078	0.0000	1.6859	0.0000	1.0245	0.0000	1.1517
Worry	1.3937	0.0000	1.6689	0.0000	1.5705	0.0000	3.0439	0.0000	1.6309	0.0000	1.3780	0.0000
Stress	1.8052	0.0000	1.9194	0.0000	1.6384	0.0000	3.0012	0.0000	1.6745	0.0000	1.7002	0.0000
Anger	1.6660	0.0000	1.7081	0.0000	1.5718	0.0000	2.7378	0.0000	1.3327	0.0000	1.7595	0.0000
Pride	-0.1513	1.0678	0.5863	0.4291	0.9093	0.3626	1.4307	0.4445	0.1380	1.0648	1.1523	0.3501
Guilt	1.5011	0.0748	0.9173	0.4182	1.5416	2.0898	0.2162	-0.0154	0.6798	0.5500	1.3111	0.2556
Shame	1.4445	0.0396	0.9329	0.4448	1.5551	2.2140	0.2651	-0.4244	0.6254	0.5186	1.1858	0.2857
Gratitude	0.0000	1.1398	0.0000	0.8915	0.0000	1.0381	0.0000	-0.1392	0.0000	1.0681	0.0000	0.9213

	Cluster 7	
	NEG	POS
Happy	0.0000	1.2512
Love	0.0000	1.1873
Sad	1.2380	0.0000
Jealousy	1.3598	0.0000
Cheerful	0.0000	1.0502
Worry	1.5244	0.0000
Stress	1.7330	0.0000
Anger	1.5306	0.0000
Pride	0.4847	0.5213
Guilt	1.1019	0.0361
Shame	1.0774	-0.0141
Gratitude	0.0000	0.8975

To optimally compare loadings between clusters, an alignment (with clusters as groups) can be performed.

Does metric invariance hold per cluster?

Cluster 1: Chile, Spain, Mexico, Venezuela, Colombia, Netherlands, Zimbabwe	ΔCFI = 0.015
Cluster 2: Hungary, Poland, Georgia (.93), Slovakia, Bangladesh (.96), India	ΔCFI = 0.015
Cluster 3: Thailand, Bulgaria	ΔCFI = 0.014
Cluster 4: Uganda	
Cluster 5: Korea Rep., Japan, Indonesia	ΔCFI = 0.009
Cluster 6: Turkey, Nigeria, China, Hong Kong, Iran, Philippines, Nepal, Italy, Belgium (.96), Portugal, Malaysia, Ghana, Brazil, Cameroon, South Africa, Austria	ΔCFI = 0.020
Cluster 7: United States, Slovenia, Australia, Greece, Cyprus, Switzerland, Singapore (.92), Croatia, Russia (.84), Germany, Kuwait (.86), Canada	ΔCFI = 0.011

How to continue based on MMG-FA results?

A few ways to move forward:

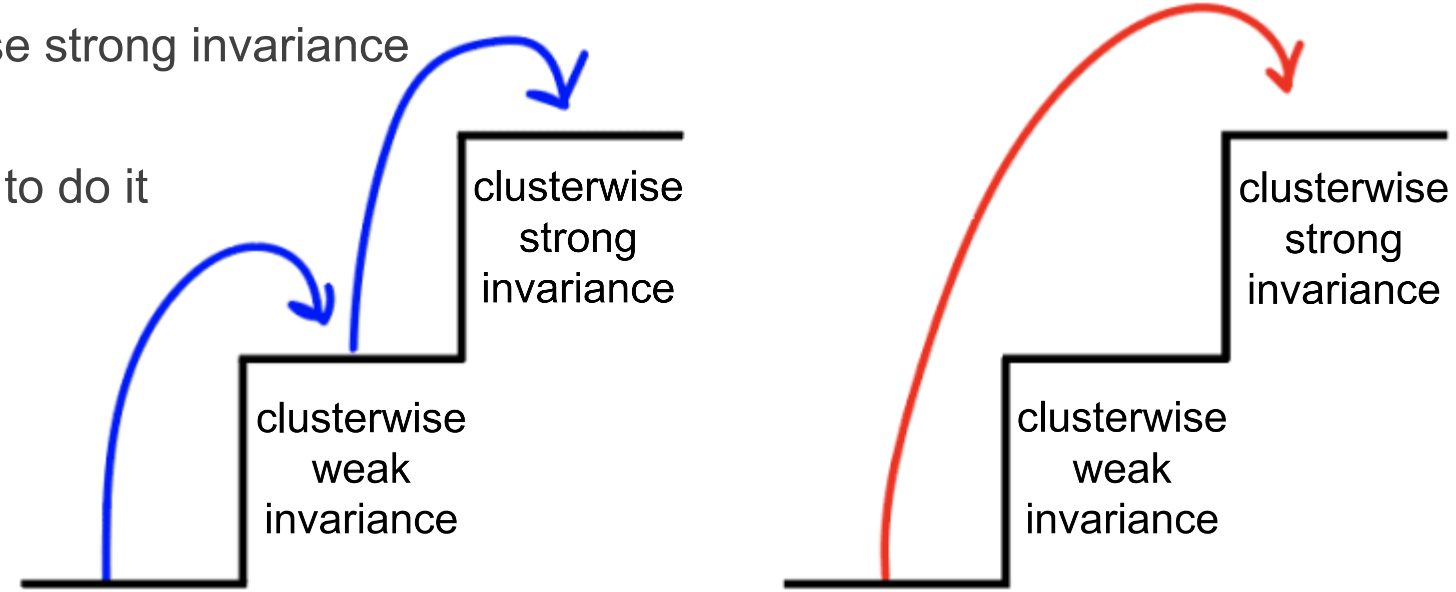
- Identify problematic items and delete them or continue with partial invariance
- And/or identify problematic groups and exclude them
- Or: continue comparison or invariance testing per cluster ✓

We have metric invariance per cluster:

- Between-group comparison of the predictive effect of ‘POS’ & ‘NEG’ on life satisfaction is allowed across groups within each cluster
- The original paper on the data (Bastian et al., 2014) used ***country-level*** social value indices (group means) – rather than individual indices – to predict the life satisfaction of a country’s inhabitants
  ➔ **scalar (intercept) invariance also required** ➔ What is the best way to pursue this? (Next slide)

How to take the steps to clusterwise measurement invariance?

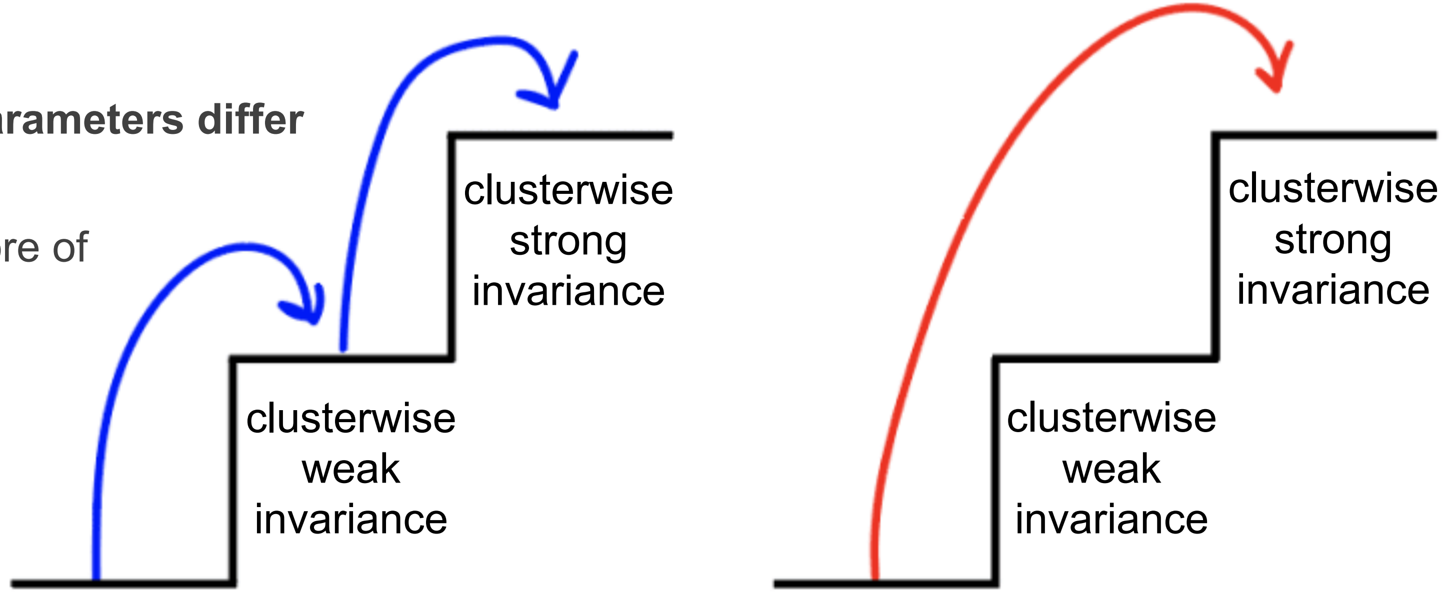
- MMG-FA operates in a **level-specific** way ➔ allows to investigate non-invariances in a **stepwise** manner, e.g., when going from overall *configural* invariance to clusterwise *strong* invariance
  - Step 1: Find clusters of groups with weak invariance
  - Step 2: Per 'loading-cluster' of groups, find clusters of groups with strong invariance
- But it is also possible to pursue clusterwise strong invariance in one step.
- Is it a good idea (or maybe the best idea) to do it in a stepwise way?
  (When) does it make a difference?



Stepwise versus simultaneous approach

- Simultaneous approach:
  - **Many clusters may be required** to capture loading AND intercept differences (e.g., if intercepts require a very different clustering or are group-specific), meaning you still have a lot of MM parameters to compare
  - May **mix up differences** or **pick up most dominant differences only** (e.g., the intercept differences)
- Stepwise approach:
  - Allows to gain **more insight in *which* parameters differ for *which* groups**
  - But performing **separate analyses** is more of a hassle

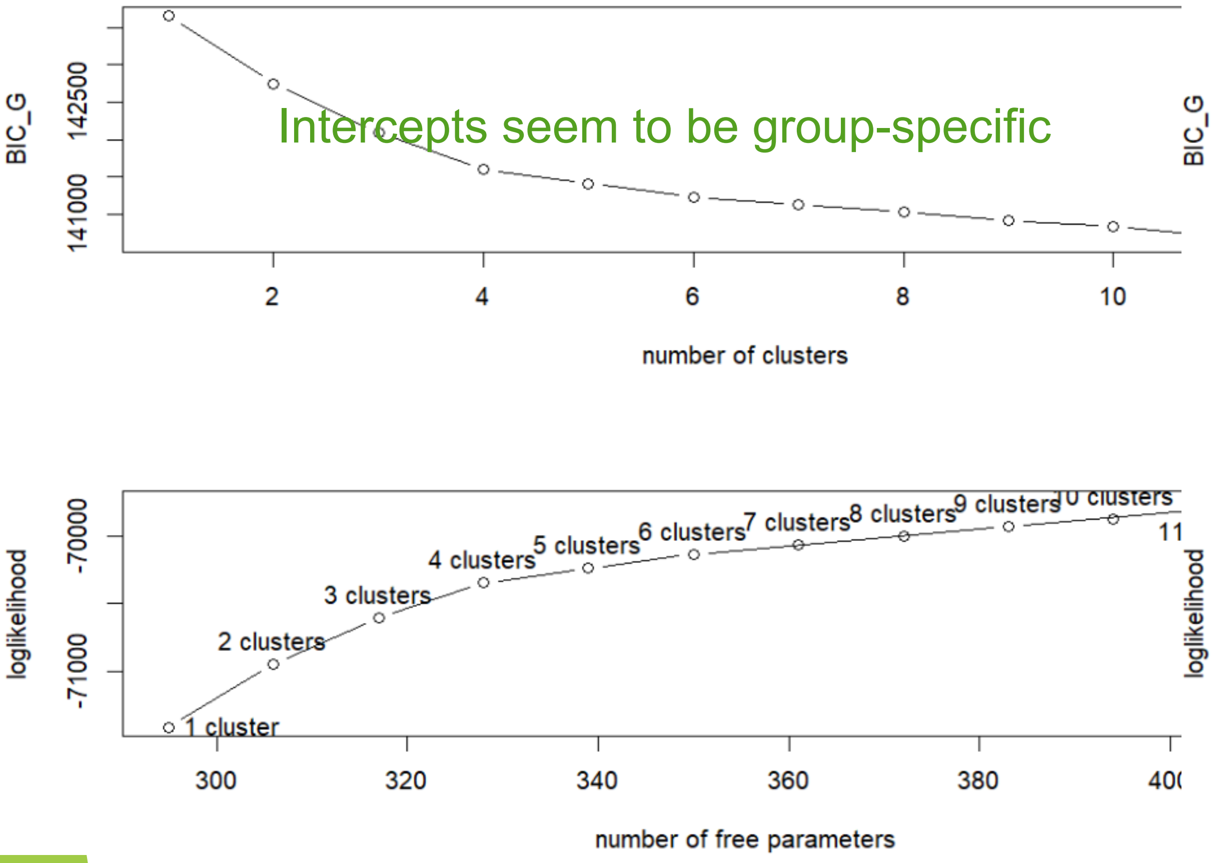
Simulation study (presented at IMPS 2022) ➔ stepwise disentanglement of measurement (non-)invariances (i.e., per level) is possible & recommended!



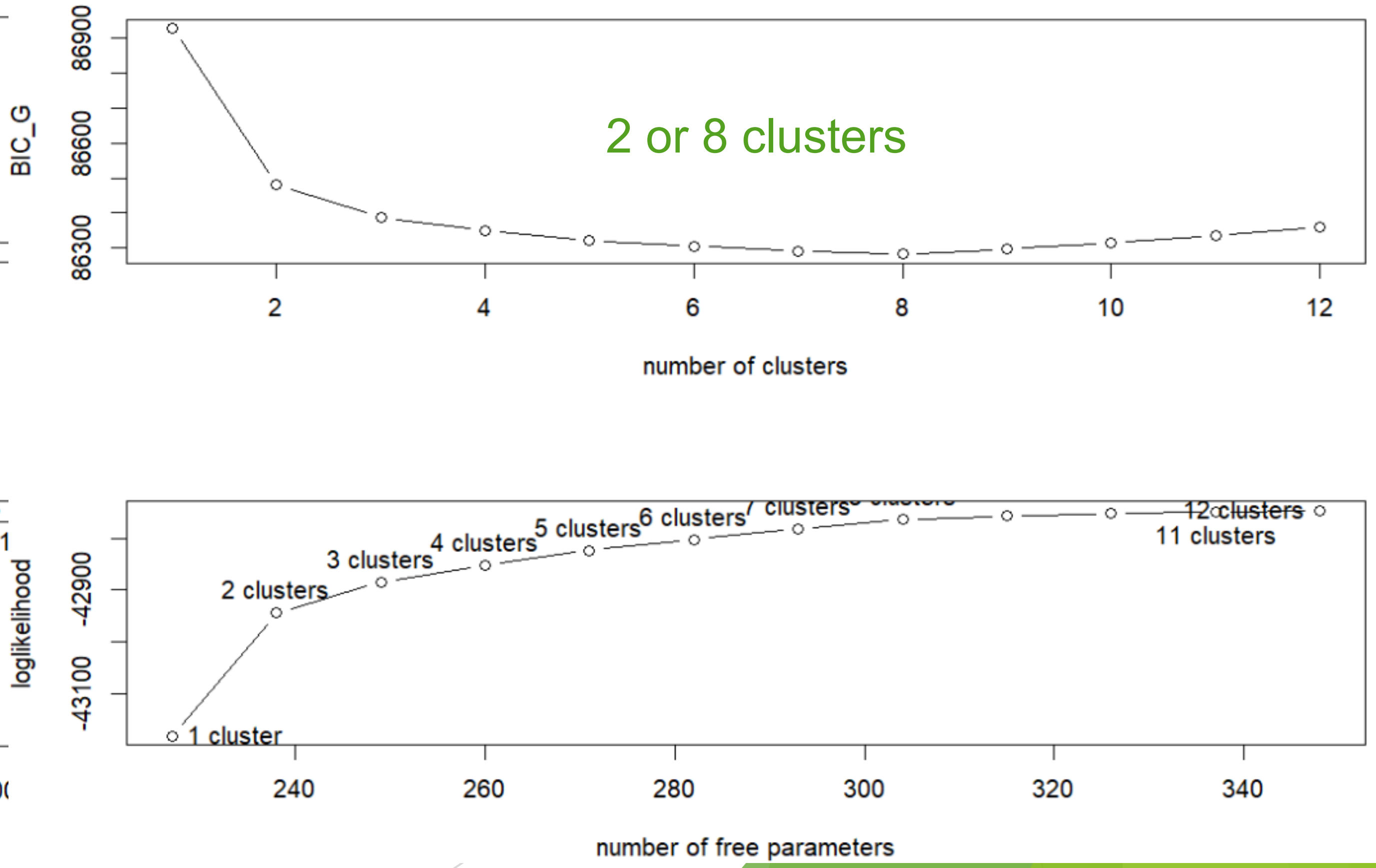
Stepwise approach for empirical application

- Let's try to find intercept-clusters for the largest loading-clusters: Cluster 6 and 7

Loading-cluster 6



Loading-cluster 7



Modal cluster assignments (all classification probabilities are equal to 1.000)

- Cluster 1: Russia
- Cluster 2: Cyprus
- Cluster 3: United States, Slovenia
- Cluster 4: Switzerland
- Cluster 5: Australia, Germany, Kuwait
- Cluster 6: Greece, Croatia
- Cluster 7: Canada
- Cluster 8: Singapore

	Happy	Love	Sad	Jealous	Cheerful	Worry	Stress	Anger	Pride	Guilt	Shame	Gratitude
Cluster_1	5,86	5,87	6,01	5,99	6,35	6,16	6,58	5,77	5,62	6,10	5,84	6,60
Cluster_2	7,11	7,26	7,30	7,20	7,08	7,45	7,13	6,18	6,33	6,97	7,58	7,44
Cluster_3	7,76	7,67	7,73	6,87	7,00	7,70	7,30	6,84	6,42	7,31	7,59	7,66
Cluster_4	5,44	4,91	4,92	5,44	5,33	5,46	5,72	5,31	4,37	5,04	4,68	4,81
Cluster_5	4,50	4,38	4,50	4,61	4,73	4,58	4,53	4,62	4,13	4,40	4,53	4,19
Cluster_6	4,40	4,39	4,10	4,71	4,78	4,51	5,36	4,99	4,71	4,34	4,27	4,23
Cluster_7	7,08	7,21	7,07	6,31	6,26	6,66	5,36	6,02	5,72	6,54	7,44	6,80
Cluster_8	5,75	5,50	4,99	4,68	4,99	5,75	5,63	5,27	4,69	5,46	4,09	4,85

Overall identification restriction on factor means across groups within a cluster

➔ Factor means can be compared WITHIN a cluster

➔ For optimal comparison of factor means and intercepts BETWEEN clusters, re-alignment is necessary

➔ with multigroup factor alignment using clusters as groups, but ideally with equal loadings across groups (which is currently not possible)

➔ indicates significant intercept differences for jealous, stress & anger

Modal cluster assignments (all classification probabilities are equal to 1.000)

- Cluster 1: Russia
- Cluster 2: Cyprus
- Cluster 3: United States, Slovenia
- Cluster 4: Switzerland
- Cluster 5: Australia, Germany, Kuwait
- Cluster 6: Greece, Croatia
- Cluster 7: Canada
- Cluster 8: Singapore

APPROXIMATE MEASUREMENT INVARIANCE (NONINVARIANCE) FOR GROUPS

Intercepts/Thresholds

HAPPY	1 2 3 4 5 6 7 8
LOVE	1 2 3 4 5 6 7 8
SAD	1 2 3 4 5 6 7 8
JEALOUS	1 (2) 3 4 5 6 7 8
CHEERFUL	1 2 3 4 5 6 7 8
WORRY	1 2 3 4 5 6 7 8
STRESS	1 2 3 (4) 5 6 (7) (8)
ANGER	1 (2) 3 4 5 6 7 8
PRIDE	1 2 3 4 5 6 7 8
GUILT	1 2 3 4 5 6 7 8
SHAME	1 2 3 4 5 6 7 8
GRATITUD	1 2 3 4 5 6 7 8

Loadings for PA

HAPPY	1 2 3 4 5 6 7 8
LOVE	1 2 3 4 5 6 7 8
CHEERFUL	1 2 3 4 5 6 7 8
PRIDE	1 2 3 4 5 6 7 8
GUILT	1 2 3 4 5 6 7 8
SHAME	1 2 3 4 5 6 7 8
GRATITUD	1 2 3 4 5 6 7 8

Loadings for NA

SAD	1 2 3 4 5 6 7 8
JEALOUS	1 2 3 4 5 6 7 8
WORRY	1 2 3 4 5 6 7 8
STRESS	1 2 3 4 5 6 7 8
ANGER	1 2 3 4 5 6 7 8
PRIDE	1 2 3 4 5 6 7 8
GUILT	1 2 3 4 5 6 7 8
SHAME	1 2 3 4 5 6 7 8

Alignment finds significant intercept differences for ‘jealous’, ‘stress’ and ‘anger’

Did not impose equal loadings, but (approx.) metric invariance is tenable

Applying MMG-FA to educational measurement: TIMMS 2015 data

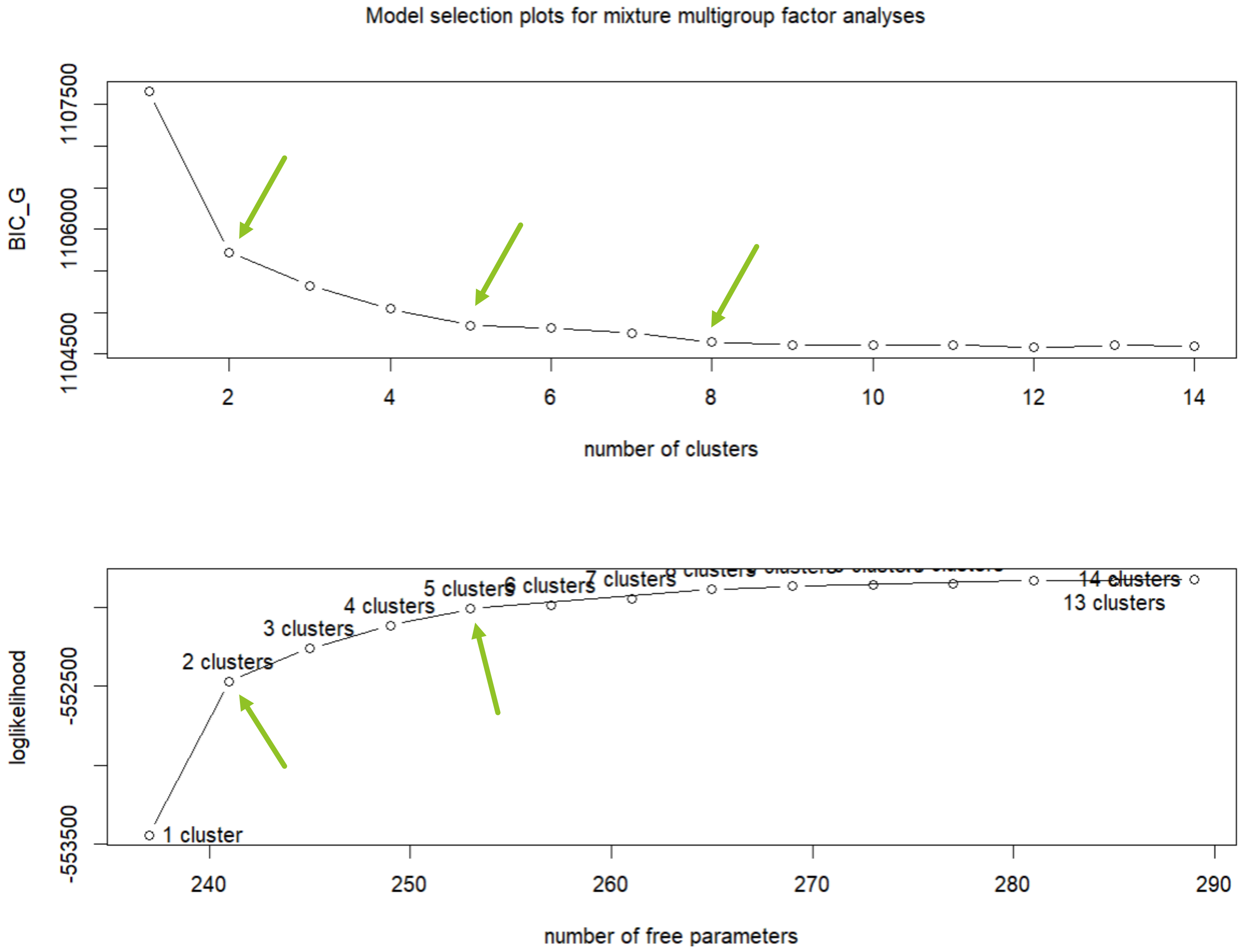
- In international assessment programs like PISA and TIMMS – student's achievements are often related to **non-cognitive constructs** such as academic self-concept and self-efficacy, enjoyment of science, sense of belonging, wellbeing, etc.
- The comparability (measurement invariance) of non-cognitive constructs is questionable and rarely evaluated (Wurster, 2022)
- The constructs are often measured with four response categories, so data are ordinal rather than continuous
- MMG-FA uses maximum likelihood (ML) that assumes data to be continuous and normally distributed
- Dolan (1994) ➔ From five response categories, ML can be used in case of non-severe non-normality. (Four was not evaluated.)
- When measurement invariance is evaluated for these data, ML is very often used.

Wurster, S. (2022). Measurement invariance of non-cognitive measures in TIMSS across countries and across time. An application and comparison of Multigroup Confirmatory Factor Analysis, Bayesian approximate measurement invariance and alignment optimization approach. *Studies in Educational Evaluation*, *73*, 101143.

Applying MMG-FA to educational measurement: TIMMS 2015 data

- TIMMS 2015 data (fourth grade) on **students' self-confidence in science** (SCS)
- **Four items**: (1) "I usually do well in science", (2) "Science is harder for me than for many of my classmates", (3) "I am just not good at science", (4) "I learn things quickly in science"
- Same selection of **26 countries** as in Wurster (2022)
- *Configural invariance* model: CFI = 0.848
  - Note: a 2-factor model with a separate factor for items 1 & 4, and one for items 2 & 3, and equality restrictions on the loadings (equal loadings for items 1 & 4, and for items 2 & 3) has a CFI of 0.966
- *Metric invariance* model: CFI = 0.815 ➔ does not hold
  - Partial metric invariance model with free loadings for *either* items 1 & 4 (positively keyed items) *or* items 2 & 3 (negatively keyed items): CFI = 0.846/0.847
- *Scalar invariance* model (building on partial metric invariance): CFI = 0.797
  - Partial scalar invariance with free intercepts for the items with free loadings holds (CFI = 0.838)

Applying MMG-FA to educational measurement: TIMMS 2015 data



- 2 clusters: Most important differences are captured. Metric invariance holds exactly in one cluster & approximately in the other one
- 5 clusters: Captures additional subtle differences. Metric invariance holds in all clusters.
- 8 clusters: Additional differences captured are negligible

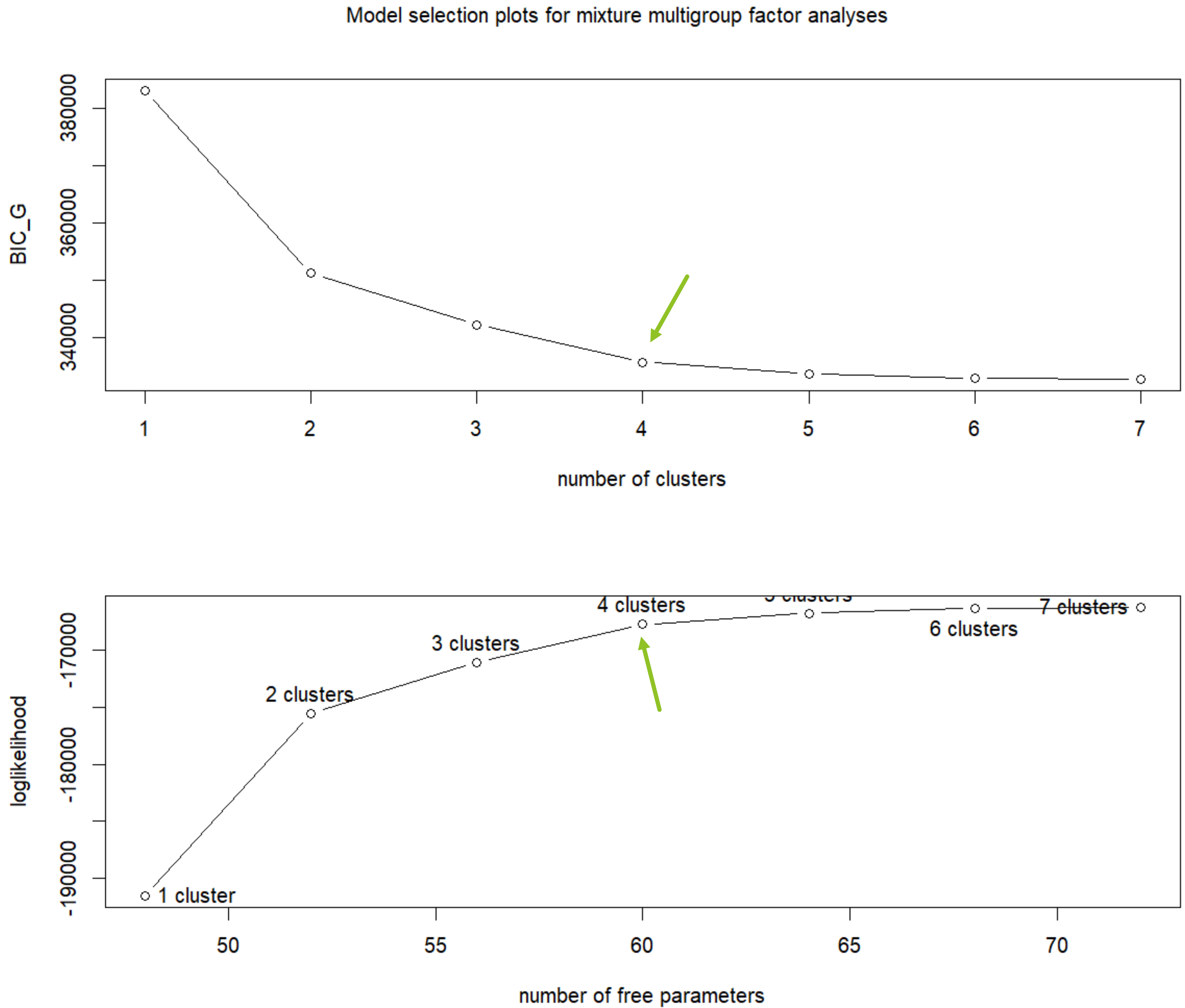
Applying MMG-FA to educational measurement: TIMMS 2015 data

- Groups modally assigned to cluster 1: **Hong Kong, Singapore, United States, New Zealand, Netherlands, Hungary, Slovak Republic, Russian Federation**
- Groups modally assigned to cluster 2: **Georgia, Iran, Kuwait, Morocco, Qatar**
- Groups modally assigned to cluster 3: **Kazakhstan, Lithuania**
- Groups modally assigned to cluster 4: **Australia, Czech Republic, Germany, Denmark England, Japan, Norway, Slovenia**
- Groups modally assigned to cluster 5: **Italy, Sweden, Chinese Taipei**

} merged in case of 4 clusters

Cluster-specific loadings:

	Fact1Cl1	Fact1Cl2	Fact1Cl3	Fact1Cl4	Fact1Cl5
ASBS06A (harder for me)	0.5255	0.1548	0.3072	0.5848	0.5596
ASBS06B (usually do well)	-0.7141	-0.8000	-0.7321	-0.5989	-0.5190
ASBS06C (just not good at science)	-0.7466	-0.8455	-0.7282	-0.6653	-0.4722
ASBS06D (learn things quickly)	0.5532	0.2186	0.3353	0.6331	0.6354



Model selection plots for mixture multigroup factor analyses
BIC_G
380000
360000
340000
1
2
3
4
5
6
7
number of clusters
loglikelihood
-170000
-180000
-190000
1 cluster
2 clusters
3 clusters
4 clusters
6 clusters
7 clusters
50
55
60
65
70
number of free parameters

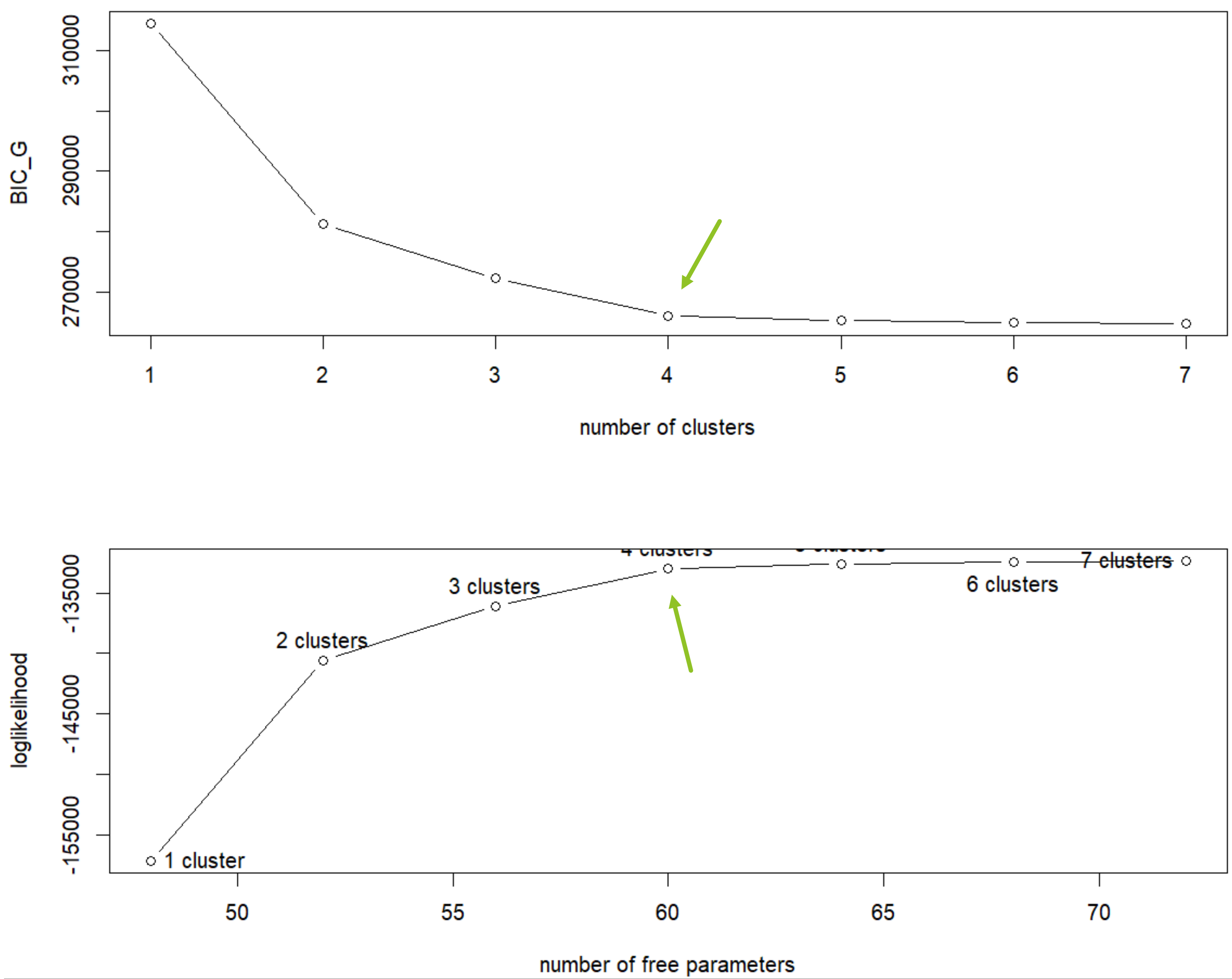
Scalar invariance clusters within metric invariance cluster 1

- Groups modally assigned to cluster 1: Hong Kong, Slovak Republic
- Groups modally assigned to cluster 2: Hungary
- Groups modally assigned to cluster 3: Netherlands, New Zealand, Russian Federation
- Groups modally assigned to cluster 4: Singapore

```
cluster-specific intercepts:

           ASBS06A ASBS06B ASBS06C ASBS06D
Cluster_1   1.9252  1.7008  1.8699  1.8002
Cluster_2   1.6388  1.9884  1.6886  2.9051
Cluster_3   3.1244  3.0639  3.0076  3.1005
Cluster_4   3.0457  1.9413  1.7996  1.7319
```

Model selection plots for mixture multigroup factor analyses



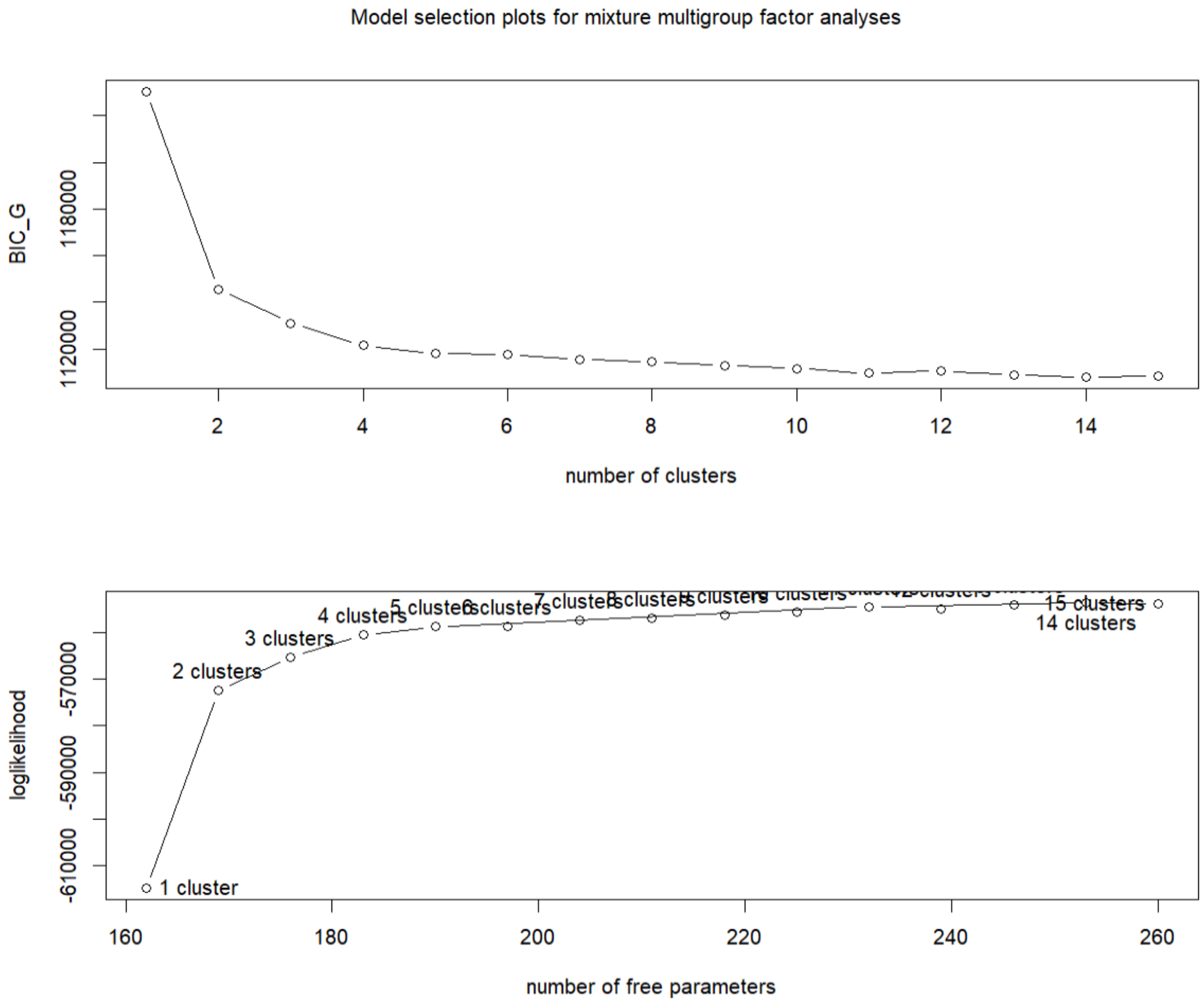
Scalar invariance clusters within metric invariance cluster 4

- Groups modally assigned to cluster 1: Germany, Denmark, England
- Groups modally assigned to cluster 2: Czech Republic
- Groups modally assigned to cluster 3: Australia, Slovenia
- Groups modally assigned to cluster 4: Norway

cluster-specific intercepts:

	ASBS06A	ASBS06B	ASBS06C	ASBS06D
Cluster_1	3.2261	3.1757	3.1464	3.0969
Cluster_2	1.8268	1.5909	1.7519	3.0101
Cluster_3	1.7911	1.8329	1.6692	1.7996
Cluster_4	3.1395	1.8519	1.9778	1.7096

What if we use the simultaneous approach?



- 4 clusters are selected (or 2)
- In case of four clusters, the loading differences (observed in stepwise approach) are not captured
- Even for 4 clusters **metric and scalar invariance fail within each cluster**

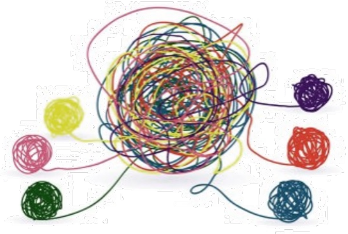
cluster-specific loadings:

	Fact1Cl1	Fact1Cl2	Fact1Cl3	Fact1Cl4
ASBS06A	0.5172	0.5618	0.5083	0.4767
ASBS06B	-0.7193	-0.6270	-0.4979	-0.7176
ASBS06C	-0.7497	-0.7008	-0.5199	-0.8477
ASBS06D	0.6131	0.5670	0.5477	0.5133

cluster-specific intercepts:

	ASBS06A	ASBS06B	ASBS06C	ASBS06D
Cluster_1	3.3551	3.3266	1.8519	1.9890
Cluster_2	1.7003	1.7229	1.7430	1.6936
Cluster_3	3.1333	3.1301	3.0809	3.1722
Cluster_4	1.6442	1.5571	3.0101	3.0386

Conclusion



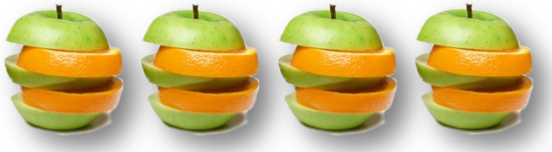
- MMG-FA finds **clusters of groups with a specific level of measurement invariance** (**clusterwise measurement invariance**) (
- And it disentangles *measurement* non-invariances from *structural* differences <−> existing mixture methods
- Estimated with tailor-made EM algorithms (LatentGOLD, mixmgfa package)
- Stepwise disentanglement of measurement (non-)invariances (i.e., per level) is possible & recommended
  - Or compare/combine results of stepwise and simultaneous approach
  - How to take classification uncertainty into account? Weighting or nested mixtures?
    - Note: classification uncertainty is limited

Future research

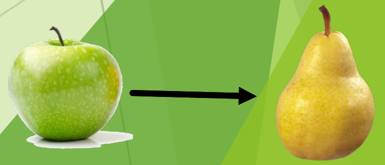
- Continue to extend **R-package** (other estimators, more rotation and scaling options, make it easier to use stepwise approach, residual covariances, SE's, etc.)

- Extension to build on **partial invariance** of loadings to find scalar invariance clusters (and/or to find partial invariance within clusters)

- Extension to accommodate **approximate invariance** within clusters



- Trace measurement non-invariance **within groups**



- Mixture multigroup SEM focused on structural relations (~ research question) rather than on measurement model

Funded by the European Union

NWO VIDI

Thanks to CEMO for the invitation!

References

- **De Roover, K.**, Vermunt, J.K., & Ceulemans, E. (2022). Mixture multigroup factor analysis for unraveling factor loading non-invariance across many groups. *Psychological Methods, 27*(3)*,* 281–306*.*
- **De Roover, K.** (2021). Finding clusters of groups with measurement invariance: Unraveling intercept non-invariance with mixture multigroup factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(5), 663–683.
- **De Roover, K.** (2022). Taking the steps toward clusterwise measurement invariance with mixture multigroup factor analysis. Manuscript in preparation.
- Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., **De Roover, K.**, Jak, S., Meitinger, K., Menold, N., Muthén, B. Rudnev, M., Schmidt, P., & van de Schoot, R. (2022). Measurement Invariance in the Social Sciences: Historical Development, Methodological Challenges, State of the Art, and Future Perspectives. *Social Science Research*, 102805.
- Perez Alonso, A. F., Rosseel, Y., Vermunt, J., & De Roover, K. (2023, July 5). Mixture Multigroup Structural Equation Modeling: A Novel Method for Comparing Structural Relations Across Many Groups. https://doi.org/10.31234/osf.io/mvd96

Thanks! Comments, suggestions, questions?

Want to know more? @Kim_De_Roover github.com/KimDeRoover/mixmgfa