New Developments in Latent Variable Modeling Using Mplus

Bengt Muthén

Mplus
www.statmodel.com

bmuthen@statmodel.com

Workshop at the European Survey Research Association (ESRA) Conference, University of Ljubljana, July 15, 2013

Mplus

Several programs in one

- Path analysis
- Exploratory factor analysis
- Structural equation modeling
- Item response theory analysis
- Growth modeling
- Mixture modeling (latent class analysis)
- Longitudinal mixture modeling (Markov, LTA, LCGA, GMM)
- Survival analysis (continuous- and discrete-time)
- Multilevel analysis
- Complex survey data analysis
- Bayesian analysis
- Monte Carlo simulation

Fully integrated in a general latent variable framework

Latent Variable Modeling in Mplus: Integration of a Multitude of Analyses

- Exploratory factor analysis
- Structural equation modeling
- Item response theory analysis
- Growth modeling
- Latent class analysis
- Latent transition analysis (Hidden Markov modeling)
- Growth mixture modeling
- Survival analysis
- Missing data modeling
- Multilevel analysis
- Complex survey data analysis
- Bayesian analysis
- Causal inference

Latent Variable Modeling in Mplus: Integration of a Multitude of Analyses

- Exploratory factor analysis
- Structural equation modeling

Latent Variable Modeling in Mplus: Integration of a Multitude of Analyses

- Structural equation modeling
- Bayesian analysis

Latent Variable Modeling in Mplus: Integration of a Multitude of Analyses

- Survival analysis
- Latent class analysis

Bengt Muthén Mplus Version 7.11 3/ 134

Latent Variable Modeling in Mplus: Integration of a Multitude of Analyses

- Growth mixture modeling
- Survival analysis
- Missing data modeling

Bengt Muthén Mplus Version 7.11 3/ 134

Latent Variable Modeling in Mplus: Integration of a Multitude of Analyses

- Latent class analysis
- Causal inference

Bengt Muthén Mplus Version 7.11 3/ 134

What's New in Mplus Version 7? Released September, 2012

5 big new features:

1. Diagrammer
2. Factor analysis
   - Bi-factor EFA rotations, bi-factor ESEM, two-tier modeling
   - Bayesian EFA and CFA (BSEM), bi-factor BSEM
3. Analysis of several groups with approx. measurement invariance
   - using a Bayes approach (multiple-group BSEM)
   - using a two-level analysis with random intercepts and loadings
4. Analysis of individual differences SEM using measurement parameters that vary across subjects
5. Mixture analysis
   - Using a proper 3-step analyze-classify-analyze approach to investigate covariates and distal outcomes
   - Latent transition analysis with new output, covariates influencing transition probabilities, and probability parameterization
   - Exploratory LCA using Bayesian analysis

What's New in Mplus Version 7, Continued

5 more big features:

1. 3-level SEM analysis, complex survey data handling, and multiple imputation
2. Cross-classified SEM analysis including random subjects and contexts (2 random modes)
3. IRT analysis with random items
4. Longitudinal analysis with approx. measurement invariance
   - using a Bayes approach (multiple-time point BSEM)
   - using cross-classified analysis of time and subjects
5. Analysis of changing membership over time

What's New in Mplus Version 7, Continued

and 5 other new features:

1. Parallel analysis
2. LOOP plots (moderated mediation, cross-level interactions, etc)
3. Bayes plausible value factor score distribution plots for each subject
4. Two-tier algorithm
5. New convenience options: LOOP, DO, COV, DIFF, DO DIFF, MODEL=ALLFREE, auto-labeling, BY with random loadings, BITER = (minimum), TECH15, TECH16

What's New In Mplus Version 7.1?

Version 7.1 was released in May 2012 and has the following new features:

1. Multiple-group factor analysis: A new method
2. Multiple-group factor analysis: Convenience features
3. Exploratory factor analysis: Convenience features
4. Mixture modeling: 3-step modifications
5. Mixture modeling: A new distal outcome stepwise method
6. New TECH4 output
7. GROUPING and KNOWNCLASS convenience features
8. DO option for MODEL TEST

More on Mplus Version 7 and 7.1

For more information, see Version History at the Mplus web site www.statmodel.com, including 22 new User's Guide examples

Videos and pdfs from the Mplus Version 7 training at Utrecht University August 27-29 can be found via the Mplus home page

2. Bayesian Analysis

- Bayesian analysis firmly established and its use is growing in mainstream statistics
- Much less use of Bayes outside statistics
- Bayesian analysis not sufficiently accessible in other programs
- Bayesian analysis was introduced in Mplus Version 6 and greatly expanded in Version 7: Easy to use
- Bayes provides a broad platform for further Mplus development

Why Bayes?

Why do we have to learn about Bayes?

- More can be learned about parameter estimates and model fit
- Better small-sample performance, large-sample theory not needed
- Non-informative versus informative priors
- Frequentists can see Bayes with non-informative priors as a computing algorithm to get answers that would be the same as ML if ML could have been done
- Informative priors can better reflect substantive hypotheses
- Analyses can be made less computationally demanding
- New types of models can be analyzed

For a Bayes introduction with further references, see, e.g., Muthén (2010). Bayesian analysis in Mplus: A brief introduction. Technical Report. Version 3.

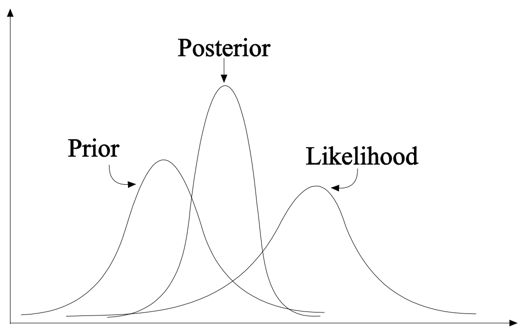
Writings On The Bayes Implementation In Mplus

- Muthén (2010). Bayesian analysis in Mplus: A brief introduction. Technical Report. Version 3.
- Asparouhov & Muthén (2010). Bayesian analysis using Mplus: Technical implementation. Technical Report. Version 3.
- Asparouhov & Muthén (2010). Bayesian analysis of latent variable models using Mplus. Technical Report. Version 4.
- Asparouhov & Muthén (2010). Multiple imputation with Mplus. Technical Report. Version 2.
- Asparouhov & Muthén (2010). Plausible values for latent variable using Mplus. Technical Report.
- Muthén & Asparouhov (2012). Bayesian SEM: A more flexible representation of substantive theory. Psychological Methods
- Asparouhov & Muthén (2011). Using Bayesian priors for more flexible latent class analysis.
- Asparouhov & Muthén (2012). General random effect latent variable modeling: Random subjects, items, contexts, and parameters.
- Asparouhov & Muthén (2012). Comparison of computational methods for high dimensional item factor analysis.

Posted under Papers, Bayesian Analysis and Latent Class Analysis

Prior, Likelihood, And Posterior

- Frequentist view: Parameters are fixed. ML estimates have an asymptotically-normal distribution
- Bayesian view: Parameters are variables that have a prior distribution. Estimates have a possibly non-normal posterior distribution. Does not depend on large-sample theory
  - Non-informative (diffuse) priors vs informative priors



Bayesian Estimation Obtained Iteratively Using Markov Chain Monte Carlo (MCMC) Algorithms

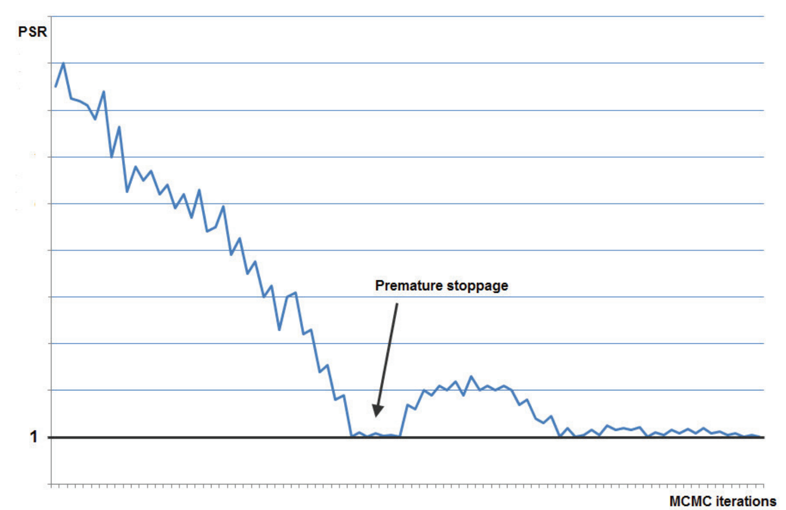
- θ_i: vector of parameters, latent variables, and missing observations at iteration i
- θ_i is divided into S sets:
  $\theta_i = (\theta_{1i}, ..., \theta_{Si})$
- Updated θ using Gibbs sampling over i = 1, 2, ..., n iterations:
  $\theta_{1i}|\theta_{2i-1}, ..., \theta_{Si-1}$, data, priors
  $\theta_{2i}|\theta_{3i-1}, ..., \theta_{Si-1}$, data, priors
  ...
  $\theta_{Si}|\theta_{1i}, ..., \theta_{S-1i-1}$, data, priors

Asparouhov & Muthén (2010). Bayesian analysis using Mplus. Technical implementation.Technical Report.

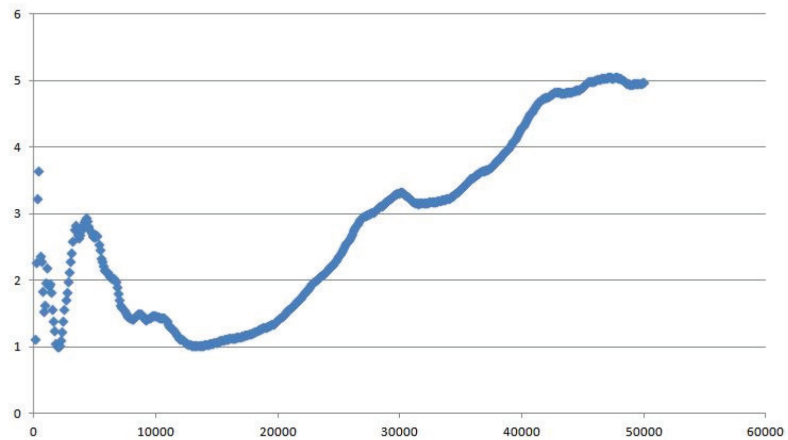
MCMC Iteration Issues

- Trace plot: Graph of the value of a parameter at different iterations
- Burnin phase: Discarding early iterations. Mplus discards first half
- Posterior distribution: Mplus uses the last half as a sample representing the posterior distribution
- Autocorrelation plot: Correlation between consecutive iterations for a parameter. Low correlation desired
- Mixing: The MCMC chain should visit the full range of parameter values, i.e. sample from all areas of the posterior density
- Convergence: Stationary process
- Potential Scale Reduction (PSR): Between-chain variation small relative to total variation. Convergence when PSR $\approx$ 1

PSR Convergence Issues: Premature Stoppage



PSR Convergence Issues: Premature Stoppages Due to Non-Identification



3. Factor Analysis

Types of factor analyses in Mplus:

- EFA: Exploratory Factor Analysis - Regular and bi-factor rotations
- CFA: Confirmatory Factor Analysis
- ESEM: Exploratory Structural Equation Modeling (Asparouhov & Muthén, 2009 in Structural Equation Modeling)
- BSEM: Bayesian Structural Equation Modeling (Muthén & Asparouhov, 2012 in Psychological Methods)

Multiple-Group ESEM: Factor Analysis of Aggressive Behavior of Males and Females in Grade 3

- 261 males and 248 females in third grade (Baltimore Cohort 3)
- Teacher-rated aggressive-disruptive behavior
- Outcomes treated as non-normal continuous variables
- Research question:
  - Does the measurement instrument function the same way for males and females?

Summary Of Separate Male/Female Exploratory Factor Analysis (Geomin Rotation)

	Loadings for Males			Loadings for Females		
Variables	Verbal	Person	Property	Verbal	Person	Property
Stubborn	**0.82***	-0.05	0.01	**0.88***	0.03	-0.22
Breaks Rules	0.47*	0.34*	0.01	**0.76***	0.06	-0.17
Harms Others and Property	-0.01	**0.63***	0.31*	0.45*	0.03	0.36
Breaks Things	-0.02	0.02	**0.66***	-0.02	0.19	**0.43***
Yells At Others	**0.66***	0.23	-0.03	**0.97***	-0.23	0.05
Takes Other's Property	0.27*	0.08	**0.52***	0.02	**0.79***	0.10
Fights	0.22*	**0.75***	-0.00	**0.81***	-0.01	0.18
Harms Property	0.03	-0.02	**0.93***	0.27	0.20	**0.57***
Lies	**0.58***	0.01	0.27*	0.42*	0.50*	-0.00
Talks Back to Adults	**0.61***	-0.02	0.30*	**0.69***	0.09	-0.02
Teases Classmates	0.46*	0.44*	-0.04	**0.71***	-0.01	0.10
Fights With Classmates	0.30*	**0.64***	0.08	**0.83***	0.03	0.21*
Loses Temper	**0.64***	0.16*	0.04	**1.05***	-0.29	-0.01

Are the Factor Loading Patterns Significantly Different in the Different Groups?

Measurement invariance can be tested by multiple-group analysis

- But this involves a move from EFA to CFA
- CFA often premature
- CFA often rejected

- Why should we have to switch from EFA to CFA to test measurement invariance?

Multiple-Group Exploratory Factor Analysis (ESEM)

Asparouhov & Muthén (2009). Exploratory structural equation modeling. **Structural Equation Modeling**, 16, 397-438.

- Estimate by ML using a group-invariant unrotated factor loading matrix with a reference group having uncorrelated unit variance factors (m^2 restrictions), allowing group-varying factor covariance matrices and residual variances
- Rotate the common factor loading matrix, e.g. by oblique Geomin
- Transform the factor covariance matrices by the rotation matrix
- Factor loading invariance across groups can be tested by LR chi-square test: Not rejected for gender invariance

Male And Female Estimates From Multiple-Group EFA Using Invariant Factor Loadings (Standardized)

	Males			Females		
Variables	Verbal	Person	Property	Verbal	Person	Property
Stubborn	**0.80***	-0.01	-0.02	**0.86***	-0.00	-0.01
Breaks Rules	0.53*	0.27*	0.01	**0.59***	0.20*	0.01
Harms Others & Property	0.00	**0.57***	0.35*	0.00	**0.56***	0.24*
Breaks Things	-0.01	-0.02	**0.67***	-0.03	-0.03	**0.63***
Yells At Others	**0.66***	0.25	-0.03	**0.69***	0.18	-0.01
Takes Others' Property	0.32*	0.04	0.53*	0.39*	0.03	0.31*
Fights	0.28*	**0.74***	-0.03	0.35*	**0.61***	-0.02
Harms Property	0.11	0.03	**0.83***	0.19	0.04	**0.68***
Lies	**0.58***	0.01	0.30*	**0.67***	0.00	0.16*
Talks Back To Adults	**0.64***	-0.03	0.29*	**0.71***	-0.02	0.15*
Teases Classmates	0.44*	0.40*	0.02	0.49*	0.30*	0.01
Fights With Classmates	0.33*	0.65*	0.05	0.41*	0.53*	0.03
Loses Temper	**0.64***	0.19	0.00	**0.74***	0.14	0.00

Further ESEM Possibilities

- Measurement intercept invariance testing and group differences in factor means
- Single-group invariance testing such as invariance across time with longitudinal factor analysis
- Exploratory SEM

Asparouhov & Muthén (2009). Exploratory structural equation modeling. **Structural Equation Modeling**, 16, 397-438.

3.2 Bayesian CFA (BSEM)

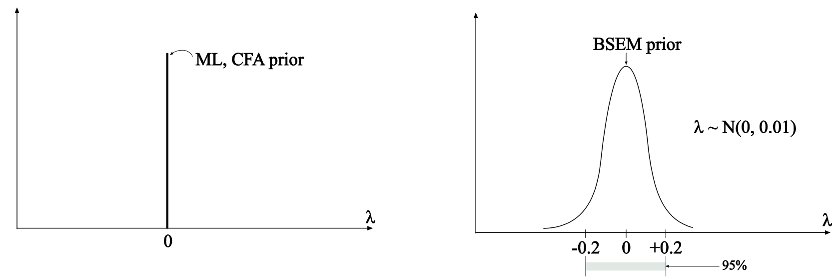
- Regular CFA is too strict, seldom fits well, and overestimates factor correlations
- Bayes CFA (BSEM) is more flexible, using zero-mean-small-variance informative priors to allow for cross-loadings, residual correlations, and direct effects which are not identified in ML

Muthén & Asparouhov (2012). Bayesian SEM: A more flexible representation of substantive theory. Psychological Methods, 17, 313-335. With commentaries and a rejoinder.

Golay, Reverte, Rossier, Favez & Lecerf (2012, November 12). Further insights on the French WISCIV factor structure through Bayesian structural equation modeling. Psychological Assessment. Advance online publication. DOI: 10.1037/a0030676

ML versus BSEM Priors

- ML CFA is characterized by many zero factor loadings
- ML CFA implicitly uses a strong prior with an exact zero loading
- BSEM uses an approximate zero loading using a zero-mean, small-variance prior for the loading:



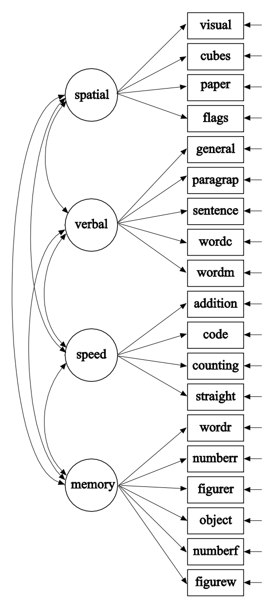
BSEM can be used to specify approximate zeros for

- Cross-loadings
- Residual correlations
- Direct effects from covariates
- Group and time differences in intercepts and loadings

3.3 Holzinger-Swineford Mental Abilities Data: BSEM CFA vs ML CFA

- Classic 1939 factor analysis study by Holzinger and Swineford (1939) in Illinois schools
- Twenty-six tests intended to measure a general factor and five specific factors
- Administered to seventh and eighth grade students in two schools
  - Grant-White school ($n = 145$). Students came from homes where the parents were mostly American-born
  - Pasteur school ($n = 156$). Students came largely from working-class parents of whom many were foreign-born and where their native language was used at home
- Source:
  - Holzinger, K. J. & Swineford, F. (1939). A study in factor analysis: The stability of a bi- factor solution. Supplementary Educational Monographs. Chicago, Ill.: The University of Chicago

Holzinger-Swineford 19 Vbles



CFA Factor Loading Pattern:

	Spatial	Verbal	Speed	Memory
visual	x	0	0	0
cubes	x	0	0	0
paper	x	0	0	0
flags	x	0	0	0
general	0	x	0	0
paragrap	0	x	0	0
sentence	0	x	0	0
wordc	0	x	0	0
wordm	0	x	0	0
addition	0	0	x	0
code	0	0	x	0
counting	0	0	x	0
straight	0	0	x	0
wordr	0	0	0	x
numberr	0	0	0	x
figurer	0	0	0	x
object	0	0	0	x
numberf	0	0	0	x
figurew	0	0	0	x

ML CFA Testing Results For Holzinger-Swineford Data For Grant-White (n =145) And Pasteur (n=156)

Model	χ^2	df	P-value	RMSEA	CFI
Grant-White					
CFA	216	146	0.000	0.057	0.930
EFA	110	101	0.248	0.025	0.991
Pasteur					
CFA	261	146	0.000	0.071	0.882
EFA	128	101	0.036	0.041	0.972

EFA has 6 (Grant-White) and 9 (Pasteur) significant cross-loadings

	Grant-White Factor Loading Patterns For EFA				Pasteur Factor Loading Pattern For EFA			
	Spatial	Verbal	Speed	Memory	Spatial	Verbal	Speed	Memory
visual	**0.628***	0.065	0.091	0.085	**0.580***	0.307*	-0.001	0.053
cubes	**0.485***	0.050	0.007	-0.003	**0.521***	0.027	-0.078	-0.059
paper	**0.406***	0.107	0.084	0.083	**0.484***	0.101	-0.016	-0.229*
flags	**0.579***	0.160	0.013	0.026	**0.687***	-0.051	0.067	0.101
general	0.042	**0.752***	0.126	-0.051	-0.043	**0.838***	0.042	-0.118
paragrap	0.021	**0.804***	-0.056	0.098	0.026	**0.800***	-0.006	0.069
sentence	-0.039	**0.844***	0.085	-0.057	-0.045	**0.911***	-0.054	-0.029
wordc	0.094	**0.556***	0.197*	0.019	0.098	**0.695***	0.008	0.083
wordm	0.004	**0.852***	-0.074	0.069	0.143*	**0.793***	0.029	-0.023
addition	-0.302*	0.029	**0.824***	0.078	-0.247*	0.067	**0.664***	0.026
code	0.012	0.050	**0.479***	0.279*	0.004	0.262*	**0.552***	0.082
counting	0.045	-0.159	**0.826***	-0.014	0.073	-0.034	**0.656***	-0.166
straight	0.346*	0.043	**0.570***	-0.055	0.266*	-0.034	**0.526***	-0.056
wordr	-0.024	0.117	-0.020	**0.523***	-0.005	0.020	-0.039	**0.726***
numberr	0.069	0.021	-0.026	**0.515***	-0.026	-0.057	-0.057	**0.604***
figurer	0.354*	-0.033	-0.077	**0.515***	0.329*	0.042	0.168	**0.403***
object	-0.195	0.045	0.154	**0.685***	-0.123	-0.005	0.333*	**0.469***
numberf	0.225	-0.127	0.246*	**0.450***	-0.014	0.092	0.092	**0.427***
figurew	0.069	0.099	0.058	**0.365***	0.139	0.013	0.237*	**0.291***

BSEM CFA for Holzinger-Swineford

- CFA: Cross-loadings fixed at zero - the model is rejected
- A more realistic hypothesis: Small cross-loadings allowed
- Cross-loadings are not all identified in terms of ML
- Different alternative: Bayesian CFA with informative priors for cross-loadings: $\lambda \sim N(0, 0.01)$.

  This means that 95% of the prior is in the range -0.2 to 0.2

Input BSEM CFA 19 Items 4 Factors Crossloading Priors

```
VARIABLE:   NAMES = id female grade agey agem school
            ! grade = 7/8
            ! school = 0/1 for Grant-White/Pasteur
            visual cubes paper flags general paragrap sentence wordc
            wordm addition code counting straight wordr numberr figurer
            object numberf figurew deduct numeric problemr series arith-
            met;
            USEV = visual-figurew;
            USEOBS = school eq 0;
DEFINE:     STANDARDIZE visual-figurew;
ANALYSIS:   ESTIMATOR = BAYES;
            PROCESSORS = 2;
            FBITER = 10000;
```

Input BSEM CFA 19 Items 4 Factors Crossloading Priors (Continued)

```
MODEL:        spatial BY visual* cubes paper flags;
              verbal BY general* paragrap sentence wordc wordm;
              speed BY addition* code counting straight;
              memory BY wordr* numberr figurer object numberf figurew;
              spatial-memory@1;
              ! cross-loadings:
              spatial BY general-figurew*0 (a1-a15);
              verbal BY visual-flags*0 (b1-b4);
              verbal BY addition-figurew*0 (b5-b14);
              speed BY visual-wordm*0 (c1-c9);
              speed BY wordr-figurew*0 (c10-c15);
              memory BY visual-straight*0 (d1-d13);
MODEL PRIORS:
              a1-d13 ~ N(0,.01);
OUTPUT:       TECH1 TECH8 STDY;
PLOT:         TYPE = PLOT2;
```


ML analysis

Model	χ^2	Df	P-value	RMSEA	CFI
Grant-White					
CFA	216	146	0.000	0.057	0.930
EFA	110	101	0.248	0.025	0.991
Pasteur					
CFA	261	146	0.000	0.071	0.882
EFA	128	101	0.036	0.041	0.972

Bayesian analysis

Model	Sample LRT	2.5% PP limit	97.5% PP limit	PP p-value
Grant-White				
CFA	219	12	112	0.006
CFA w/ cross-loadings	142	-39	61	0.361
Pasteur				
CFA	264	56	156	0.000
CFA w/ cross-loadings	156	-28	76	0.162

	Grant-White Factor Loadings Using Informative Priors				Pasteur Factor Loadings Using Informative Priors			
	Spatial	Verbal	Speed	Memory	Spatial	Verbal	Speed	Memory
visual	**0.640***	0.012	0.050	0.047	**0.633***	0.145	0.027	0.039
cubes	**0.521***	-0.008	-0.010	-0.012	**0.504***	-0.027	-0.041	-0.030
paper	**0.456***	0.040	0.041	0.047	**0.515***	0.018	-0.024	-0.118
flags	**0.672***	0.046	-0.020	0.005	**0.677***	-0.095	0.026	0.093
general	0.037	**0.788***	0.049	-0.040	-0.056	**0.856***	0.027	-0.084
paragrap	-0.001	**0.837***	-0.053	0.030	0.015	**0.801***	-0.011	0.050
sentence	-0.045	**0.885***	0.021	-0.055	-0.063	**0.925***	-0.032	-0.036
wordc	0.053	**0.612***	0.096	0.029	0.055	**0.694***	0.013	0.063
wordm	-0.012	**0.886***	-0.086	0.020	0.092	**0.803***	0.001	0.012
addition	-0.172*	0.030	**0.795***	0.004	-0.147	-0.004	**0.655***	0.010
code	-0.002	0.054	**0.560***	0.130	-0.004	0.111	**0.655***	0.049
counting	0.013	-0.092	**0.828***	-0.049	0.025	-0.058	**0.616***	-0.057
straight	0.189*	0.043	**0.633***	-0.035	0.132	-0.067	**0.558***	0.001
wordr	-0.040	0.044	-0.031	**0.556***	-0.058	0.006	-0.090	**0.731***
numberr	0.003	-0.004	-0.038	**0.552***	0.006	-0.098	-0.106	**0.634***
figurer	0.132	-0.024	-0.049	**0.573***	0.156*	0.027	0.064	**0.517***
object	-0.139	0.014	0.029	**0.724***	-0.097	0.007	0.122	**0.545***
numberf	0.099	-0.071	0.095	**0.564***	-0.029	0.041	0.003	**0.474***
figurew	0.012	0.045	0.007	**0.445***	0.049	0.018	0.085	**0.397***

Number of significant cross-loadings: 2 for Grant-White and 1 for Pasteur

Summary of Analyses of Holzinger-Swineford 19-Variable Data

- Conventional, frequentist, CFA model rejected
- Bayesian CFA with informative cross-loadings not rejected
- The Bayesian approach uses an intermediate hypothesis:
  - Less strict than conventional CFA
  - Stricter than EFA, where the hypothesis only concerns the number of factors
  - Cross-loadings shrunken towards zero; acceptable degree of shrinkage monitored by PPP
- Bayes modification indices obtained by estimated cross-loadings
- Factor correlations: EFA < BSEM < CFA

Comparing BSEM And ESEM

- ESEM: Structural equation modeling with EFA measurement model (Asparouhov & Muthén (2009). Exploratory structural equation modeling. Structural Equation Modeling, 16, 397-438)
- Similarities: Both ESEM and BSEM can be used for measurement models in SEM
- Differences:
  - ESEM is EFA-oriented while BSEM is CFA-oriented
  - ESEM uses a mechanical rotation and the rotation is not based on information from other parts of the model
  - BSEM is applicable not only to measurement models

4. Analysis Choices for Multiple Groups/Clusters: Fixed vs Random Effect Factor Analysis (IRT)

- Fixed mode: Multiple-group factor analysis
  - Inference to the groups in the sample
  - Usually a relatively small number of groups
- Random mode: Two-level factor analysis
  - Inference to a population from which the groups/clusters have been sampled
  - Usually a relatively large number of groups/clusters

New Techniques

- Fixed mode:
  - ESEM: Asparouhov & Muthén (2009). Exploratory structural equation modeling. Structural Equation Modeling, 16, 397-438.
  - Alignment: Asparouhov & Muthén (2013). Multiple group factor analysis alignment. Web note 18.
  - BSEM:
    - Muthén & Asparouhov (2012). Bayesian SEM: A more flexible representation of substantive theory. Psychological Methods, 17, 313-335.
    - Muthén & Asparouhov (2013). BSEM measurement invariance analysis. Web note 17.
- Random mode:
  - Two-level (random intercepts and loadings):
    - Fox (2010). Bayesian IRT.
    - Asparouhov & Muthén (2012). General random effect latent variable modeling: Random subjects, items, contexts, and parameters.

4.1 Multiple-Group Factor Analysis: A New Method - Alignment Optimization

There is a need for a new approach to multiple-group factor analysis for many groups such as with country comparisons of achievement (PISA, TIMSS, PIRL) or cross-cultural studies (ISSP, ESS etc):

- Goal is to study measurement invariance and also group differences in factor means and variances
- Standard approach is confirmatory factor analysis with equality constraints, followed by model modifications
- The standard approach is too cumbersome to be practical for analysis of many groups where there can be a large number of non-invariant measurement parameters
- A radically different method is introduced in Mplus Version 7.1: Alignment optimization

Multiple-Group CFA Alignment Optimization

1. Estimate the configural model (loadings and intercepts free across groups, factor means fixed @0, factor variances fixed @1)
2. Alignment optimization:
   - Free the factor means and variances and choose their values to minimize the total amount of non-invariance using a simplicity function

$$F = \sum_p \sum_{j_1<j_2} w_{j_1,j_2} f(\lambda_{pj_1} - \lambda_{pj_2}) + \sum_p \sum_{j_1<j_2} w_{j_1,j_2} f(\nu_{pj_1} - \nu_{pj_2}),$$

   for every pair of groups and every intercept and loading using a component loss function (CLF) f from EFA rotations (Jennrich, 2006)
   - The simplicity function F is optimized at a few large non-invariant parameters and many approximately invariant parameters rather than many medium-sized non-invariant parameters (compare with EFA rotations using functions that aim for either large or small loadings, not mid-sized loadings)

Alignment Optimization, Continued

- In this way, a non-identified model where factor means and factor variances are added to the configural model is made identified by adding a simplicity requirement
- This model has the same fit as the configural model:
  - Free the factor means α_j and variances ψ_j, noting that for every set of factor means and variances the same fit as the configural model is obtained with loadings λ_j and intercepts ν_j changed as:

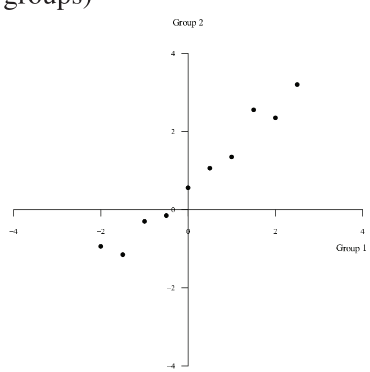
$$\lambda_j = \lambda_{j,configural} / \sqrt{\psi_j},$$
$$\nu_j = \nu_{j,configural} - \alpha_j \, \lambda_{j,configural} / \sqrt{\psi_j}.$$

- Simulation studies show that the alignment method works very well unless there is a majority of significant non-invariant parameters or small group sizes
- For well-known examples with few groups and few non-invariances, the results agree with the alignment method

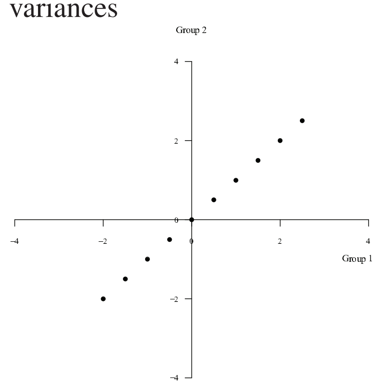
A Visual Answer to Why it is Called Alignment

Consider group-invariant intercepts for 10 items and 2 groups with factor means = 0, -1 and factor variances = 1, 2

Unaligned: Configural model (mean=0, variance=1 in both groups)



Aligned: Taking into account the group differences in means and variances



How Do We use the Alignment Results?

In addition to the estimated aligned model, the alignment procedure gives

- Measurement invariance test results produced by an algorithm that determines the largest set of parameters that has no significant difference between the parameters
- Factor mean ordering among groups and significant differences produced by z-tests

4.2 Alignment Example: Cross-Cultural Data on Nationalism and Patriotism

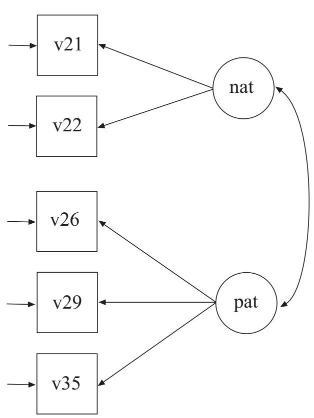
Davidov (2009). Measurement equivalence of nationalism and constructive patriotism in the ISSP: 34 countries in a comparative perspective. Political Analysis,17, 64-82.

- Data from the International Social Survey Program (ISSP) 2003 National Identity Module
- 34 countries, n=45,546
- 5 measurements of nationalism and patriotism
- Expected 2-factor structure

Nationalism and Patriotism Data: Item Wording

- Nationalism factor:
  - V21: The world would be a better place if people from other countries were more like in [own country]
  - V22: Generally speaking, [own country] is better than most other countries
- Constructive Patriotism factor:
  - V26: How proud are you of [respondent's country] in the way democracy works?
  - V29: How proud are you of [respondent's country] in its social security system?
  - V35: How proud are you of [respondent's country] in its fair and equal treatment of all groups in society?

Nationalism and Patriotism Data: Confirmatory Factor Analysis (CFA) Model



	nat	pat
v21	x	0
v22	x	0
v26	0	x
v27	0	x
v35	0	x

Nationalism and Patriotism Data: Multiple-Group CFA with ML ($n = 45,546$)

Two-factor CFA with scalar measurement invariance across all 34 countries: χ^2 (334) = 9669, p = 0, RMSEA $=$ 0.144, CFI $=$ 0.721

Group-specific misfit evenly spread over the countries

Modification indices show a multitude of similarly large values

The usual multiple-group CFA approach fails

4.3 Input for Nationalism & Patriotism Alignment in 34 Countries

```
DATA:       FILE = issp.txt;
VARIABLE:   NAMES = country v21 v22 v26 v29 v35;
            USEVARIABLES = v21-v35;
            MISSING = v21-v35 (0 8 9);
            CLASSES = c(34);
            !KNOWNCLASS = c(country = 1 2 4 6-8 10-22 24-28 30-33 36
            !40-43);
            KNOWNCLASS = c(country);
ANALYSIS:   TYPE = MIXTURE;
            ESTIMATOR = ML;
            ALIGNMENT = FREE;
MODEL:      %OVERALL%
            nat BY v21-v22;
            pat BY v26v35;
OUTPUT:     TECH1 TECH8 ALIGN;
```


Nationalism & Patriotism

- STANDARD ERROR COMPARISON INDICATES THAT THE FREE ALIGNMENT MODEL MAY BE POORLY IDENTIFIED. USING THE FIXED ALIGNMENT OPTION MAY RESOLVE THIS PROBLEM.

Choosing group with smallest factor mean to be the reference groups, this leads to the fixed alignment run:

```
ANALYSIS:
            TYPE = MIXTURE;
            ESTIMATOR = ML;
            ALIGNMENT = FIXED(28);
```


Nationalism and Patriotism Example: Alignment Results

Approximate Measurement (**Non**-) Invariance by Group

Intercepts for Nationalism indicators (V21, V22) and Patriotism indicators (V26, V29, V35)

V21	1	2	3	4	5	6	7	8	9	10	11	12
	13	14	15	16	17	18	19	20	21	22	23	24
	25	26	27	28	29	30	31	32	33	34		
V22	**(1)**	2	3	**(4)**	5	**(6)**	7	8	**(9)**	10	11	12
	13	14	**(15)**	**(16)**	17	18	**(19)**	**(20)**	21	**(22)**	**(23)**	24
	(25)	26	27	28	**(29)**	30	31	**(32)**	33	34		
V26	1	2	3	4	5	6	7	8	9	10	11	12
	13	14	15	16	17	18	19	20	21	22	23	24
	25	26	27	28	29	30	31	32	33	34		
V29	**(1)**	2	3	**(4)**	**(5)**	6	7	**(8)**	**(9)**	10	11	12
	(13)	14	15	16	**(17)**	18	**(19)**	**(20)**	**(21)**	**(22)**	**(23)**	**(24)**
	(25)	26	27	28	29	**(30)**	31	32	33	**(34)**		
V35	**(1)**	**(2)**	3	**(4)**	5	6	7	**(8)**	**(9)**	**(10)**	11	12
	13	14	15	16	17	18	**(19)**	**(20)**	21	**(22)**	23	**(24)**
	25	26	**(27)**	**(28)**	**(29)**	**(30)**	31	32	**(33)**	34		

Nationalism and Patriotism Example: Alignment Results

Loadings for NATIONALISM factor

V21	1	**(2)**	**(3)**	4	5	6	7	**(8)**	**(9)**	**(10)**	11	12
	13	14	15	16	17	18	19	20	21	22	**(23)**	**(24)**
	(25)	26	27	28	29	**(30)**	31	32	33	34		
V22	1	2	3	4	5	6	7	8	9	10	11	12
	13	14	15	16	17	18	19	20	21	22	23	24
	25	26	27	28	29	30	31	32	33	34		

Loadings for PATRIOTISM factor

V26	1	2	3	4	5	6	7	8	9	10	11	12
	13	14	15	16	17	18	19	20	**(21)**	**(22)**	23	24
	25	26	27	**(28)**	29	30	31	32	33	34		
V29	1	2	3	4	5	6	7	8	9	10	11	12
	13	14	15	16	17	18	**(19)**	20	21	22	23	**(24)**
	25	26	27	28	29	30	31	32	33	34		
V35	1	2	3	4	5	6	7	8	9	10	11	12
	13	14	15	16	17	18	19	20	21	22	23	24
	25	26	27	28	29	30	31	32	33	34		

Nationalism and Patriotism Example: Factor Mean Comparisons (5% Significance Level)

Results for NATIONALISM factor

Ranking	Group	Value	Groups with significantly smaller factor mean
1	22	0.067	2 19 11 12 9 24 23 10 15 20 33 14 32 29 13 7 6 8 16 4 21 1 26 27 34 30 31 3 25 5
2	28	0.000	19 11 12 9 24 23 15 20 33 14 32 29 13 7 6 8 16 4 21 1 26 27 34 30 31 3 25 5 18 17
3	2	-0.284	6 16 4 21 1 26 27 34 31 3 25 5 18 17
4	19	-0.333	32 13 7 6 16 4 21 1 26 27 34 31 3 25 5 18 17
5	11	-0.344	33 32 13 7 6 16 4 21 1 26 27 34 31 3 25 5 18 17
6	12	-0.352	13 7 6 16 4 21 1 26 27 34 31 3 25 5 18 17
7	9	-0.357	7 6 16 4 21 1 26 27 34 31 3 25 5 18 17
8	24	-0.379	6 16 4 21 1 26 27 34 31 3 25 5 18 17
9	23	-0.388	13 7 6 16 4 21 1 26 27 34 31 3 25 5 18 17
10	10	-0.395	16 4 21 1 26 27 34 31 3 25 5 18 17
11	15	-0.396	13 7 6 16 4 21 1 26 27 34 31 3 25 5 18 17
12	20	-0.413	13 7 6 16 4 21 1 26 27 34 31 3 25 5 18 17

4.4 Alignment Monte Carlo Studies: How Do We Know That We Can Trust The Alignment Results?

- Simulations in Asparouhov-Muthén Web Note 18
- Simulations based on the estimated model:
  - Request SVALUES for real-data alignment run (parameter estimates arranged as starting values)
  - Do a Monte Carlo run with these parameter values as population values, choosing the sample size and check parameter bias, SE bias, and the coverage
  - Do a ”real-data” run on Monte-Carlo generated data from one or more replications to study the measurement invariance assessment - does it look like the real-data run?

Input for Alignment Monte Carlo Study

- Copy SVALUES results from real-data run into Monte Carlo run
- Do a global change of the class label ”c” to ”g” (reverse unwanted changes: Montegarlo, Progessors, etc)
- Change f BY in OVERALL to give starting values

```
MONTECARLO:    NAMES = ipfrule ipmodst ipbhprp imptrad;
               NGROUPS= 26;
               NOBSERVATIONS = 26(2000);
               NREPS = 100;
               REPSAVE = ALL;
               SAVE = n2000f-22rep*.dat;
ANALYSIS:      TYPE = MIXTURE;
               ESTIMATOR = ML;
               ALIGNMENT = FIXED(22);
               PROCESSORS = 8;
MODEL POPULATION:
               %OVERALL%
               traco BY ipfrule-imptrad*1;
               [ g#1*-0.10053 ];
               etc
```


Nationalism and Patriotism Example: Monte Carlo Simulations

A cautionary tale:

Monte Carlo simulations based on these data show failure in recovering the population values. Reasons include:

- Too large degree of non-invariance
- Factor model weak by having only 2 indicators for one of the factors
- Several groups have Heywood cases

Simulations show that using a much larger sample size does not resolve the problem.

4.5 Alignment Optimization: Binary Math Items in 40 Countries (PISA)

- Items from the PISA (Program for International Student Assessment) survey of 2003
- A total of 9796 students from 40 countries
- Analyzed by Fox (2010). Bayesian Item Response Modeling
- A 40-group, one-factor model for eight mathematics test items
- 2-parameter probit IRT model that accommodates country measurement non-invariance for all difficulty (threshold) and discrimination (loading) parameters as well as country-specific factor means and variances

Input for PISA Alignment

```
DATA:       FILE = pisa2003.dat;
VARIABLE:   NAMES = cn y1-y8;
            CATEGORICAL = y1-y8; ! Requires Bayesian analysis
            USEVARIABLES = y1-y8;
            MISSING = y1-y8(9);
            CLASSES = c(40);
            KNOWNCLASS = c(cn = 1-40);
ANALYSIS:   TYPE = MIXTURE;
            ESTIMATOR = BAYES;
            PROCESSORS = 2;
            ALIGNMENT = FREE;
            THIN = 10; ! record only every 10th iter; saves alignment time
            BITERATIONS = (5000); ! do a minimum of 5000 iterations
MODEL:      %OVERALL%
            f BY y1-y8;
OUTPUT:     TECH1 TECH8 ALIGN;
PLOT:       TYPE = PLOT2;
```


4.6 Multiple-Group Analysis using Bayes and BSEM Alignment

The several uses of BSEM with zero-mean, small-variance priors:

- Single group analysis (2012 Psych Methods article):
  - Cross-loadings
  - Residual covariances
  - Direct effects in MIMIC
- Multiple-group analysis:
  - Configural and scalar analysis with cross-loadings and/or residual covariances
  - Approximate measurement invariance (Web Note 17)
  - BSEM-based alignment optimization (Web Note 18):
    - Residual covariances
    - Approximate measurement invariance

Bayes and BSEM Alignment

What does Bayes contribute?

1. Bayes with informative, zero-mean, small-variance priors for residual covariances can allow better configural fit - configural misfit in some groups is a common problem
2. Bayes with informative, zero-mean, small-variance priors for measurement parameter differences across groups (multiple-group BSEM) can allow better scalar fit
   - MG-BSEM as an alternative to alignment (finds non-invariance)
   - MG-BSEM-based alignment (advantageous for small samples?)
3. Bayes alignment can produce plausible values for the subjects' factor score values to be used in further analyses

Types of Alignment

- ML estimation:
  - ALIGNMENT = FREE
  - ALIGNMENT = FIXED(value)
- Bayes estimation:
  - ALIGNMENT = FREE
  - ALIGNMENT = FIXED(group)
  - ALIGNMENT = FREE(BSEM) - ”BSEM-based alignment”
  - ALIGNMENT = FIXED(group BSEM)
  - Adding Inverse Wishart (IW) priors for Theta to allow residual covariances

4.7 Multiple-Group BSEM

Muthén & Asparouhov (2013). BSEM measurement invariance analysis. Web Note 17.

- Approximate measurement invariance across groups using zero-mean, small-variance informative priors for the group differences
- Produces ”modification indices” by flagging non-invariant items as significantly deviating from average (ML-based MIs not available for categorical items)
- Freeing the non-invariant parameters gives proper ”alignment”, otherwise an alignment run is needed (BSEM-based alignment: ALIGNMENT = FREE(BSEM);)

Multiple-Group BSEM: Math Items in 40 PISA Countries

```
DATA:        FILE = pisa2003.dat;
VARIABLE:    NAMES = cn y1-y8;
             CATEGORICAL = y1-y8;
             USEVARIABLES = y1-y8;
             MISSING = y1-y8(9);
             CLASSES = c(40);
             KNOWNCLASS = c(cn = 1-40);
ANALYSIS:    TYPE = MIXTURE;
             ESTIMATOR = BAYES;
             PROCESSORS = 2;
             MODEL = ALLFREE ;
             BITERATIONS = (10000);
MODEL:       %OVERALL%
             f BY y1-y8* (lam#_1-lam#_8);
             [y1$1-y8$1] (tau#_1-tau#_8);
             %c#40%
             [f@0];
             f@1;
```


Multiple-Group BSEM: Math Items in 40 PISA Countries, Continued

```
MODEL PRIORS:
             DO(1,8) DIFF(tau1_#-tau40_#)~N(0,0.10);
             DO(1,8) DIFF(lam1_#-lam40_#)~N(0,0.10);
OUTPUT:      TECH1 TECH2;
PLOT:        TYPE = PLOT2;
```

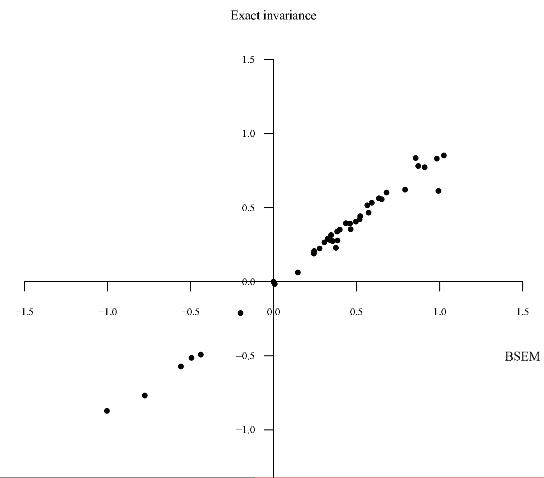
Multiple-Group BSEM: Non-Invariance Findings for PISA Items

Table : PISA countries with significant differences relative to the average across countries (prior variance = 0.10)

Item	Loading	Threshold
1	-	2, 12, 18, 22, 28, 39
2	15, 35	29, 38
3	15	23, 34, 35
4	-	12, 27, 40
5	3	7, 37
6	3, 33	5, 18, 25, 27, 37
7	-	9, 24, 27
8	24	-

Estimated Factor Means for 40 PISA Countries

Figure : Estimated factor means for 40 countries: Comparing BSEM analysis (X axis) with analysis imposing exact invariance (Y axis)



5. Two-Level Analysis with Random Item Parameters

- De Jong, Steenkamp & Fox (2007). Relaxing measurement invariance in cross-national consumer research using a hierarchical IRT model. Journal of Consumer Research, 34, 260-278.
- Fox (2010). Bayesian Item Response Modeling. Springer
- Fox & Verhagen (2011). Random item effects modeling for cross-national survey data. In E. Davidov & P. Schmidt, and J. Billiet (Eds.), Cross-cultural Analysis: Methods and Applications
- Asparouhov & Muthén (2012). General random effect latent variable modeling: Random subjects, items, contexts, and parameters
- Bayesian estimation needed because random loadings with ML give rise to numerical integration with many dimensions

Random Item Parameters In IRT

- Y_{ijk} - outcome for student i, in country j and item k

$$P(Y_{ijk}=1)=\Phi(a_{jk}\theta_{ij}+b_{jk})$$

$$a_{jk}\sim N(a_k,\sigma_{a,k}), b_{jk}\sim N(b_k,\sigma_{b,k})$$

  This is a 2-parameter probit IRT model where both discrimination (a) and difficulty (b) vary across country
- The θ ability factor is decomposed as

$$\theta_{ij}=\theta_j+\varepsilon_{ij}$$

- The mean and variance of the ability vary across country
- Model preserves common measurement scale while accommodating measurement non-invariance
- The ability for each country obtained by factor score estimation

5.1 Random Loadings: UG Ex9.19

Part 1: Random factor loadings (decomposition of the factor into within- and between-level parts)

```
TITLE:     this is an example of a two-level MIMIC
           model with continuous factor indicators,
           random factor loadings, two covariates on
           within, and one covariate on between
           with equal loadings across levels
DATA:      FILE = ex9.19.dat;
VARIABLE:  NAMES = y1-y4 x1 x2 w clus;
           WITHIN = x1 x2;
           BETWEEN = w;
           CLUSTER = clus;
ANALYSIS:  TYPE = TWOLEVEL RANDOM;
           ESTIMATOR = BAYES;
           PROCESSORS = 2;
           BITER = (1000);
MODEL:     %WITHIN%
           s1-s4 | f BY y1-y4;
           f@1;
           f ON x1 x2;
           %BETWEEN%
           f ON w;
           f; ! defaults: s1-s4; [s1-s4];
PLOT:      TYPE = PLOT2;
OUTPUT:    TECH1 TECH8;
```

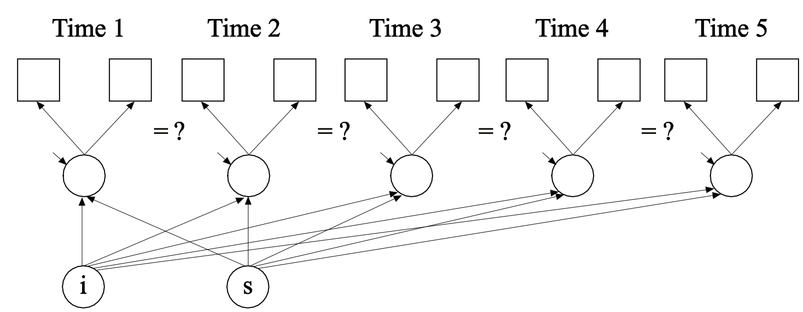

6. Longitudinal Analysis

- BSEM longitudinal approximate measurement invariance
  - Muthén & Asparouhov (2013). BSEM measurement invariance analysis. Web Note 17
- Intensive longitudinal data (many time points)
  - Individual differences factor analysis (TYPE=TWOLEVEL)
  - Cross-classified longitudinal analysis (TYPE=CROSSCLASSIFIED)

6.1 Advances In Multiple Indicator Growth Modeling

- An old dilemma
- Two new solutions

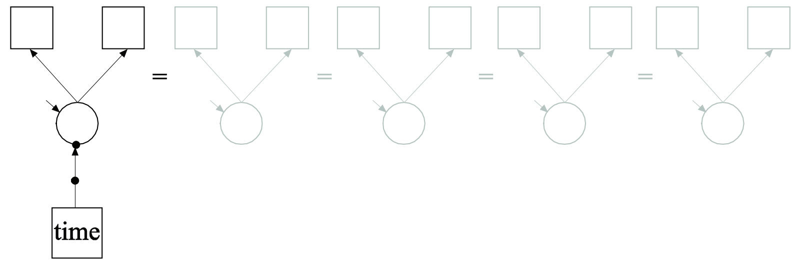
Categorical Items, Wide Format, Single-Level Approach



Single-level analysis with $p \times T = 2 \times 5 = 10$ variables, $T = 5$ factors.

- ML hard and impossible as T increases (numerical integration)
- WLSMV possible but hard when $p \times T$ increases and biased unless attrition is MCAR or multiple imputation is done first
- Bayes possible
- Searching for partial measurement invariance is cumbersome

Categorical Items, Long Format, Two-Level Approach



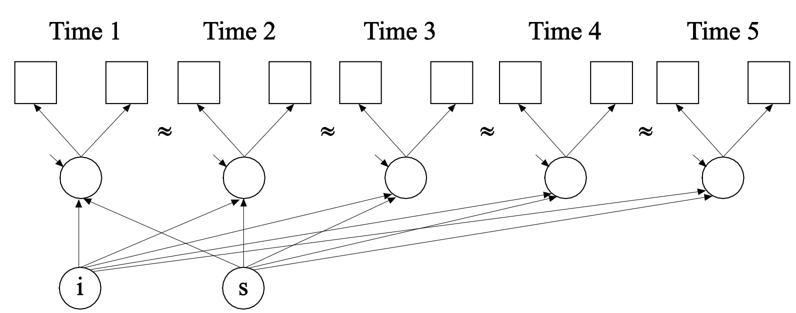
Two-level analysis with $p = 2$ variables, 1 within-factor, 2-between factors, **assuming full measurement invariance across time**.

- ML feasible
- WLSMV feasible (2-level WLSMV)
- Bayes feasible

Measurement Invariance Across Time

- Both old approaches have problems
  - Wide, single-level approach easily gets significant non-invariance and needs many modifications
  - Long, two-level approach has to assume invariance
- New solution no. 1, suitable for small to medium number of time points
  - A new wide, single-level approach where time is a fixed mode
- New solution no. 2, suitable for medium to large number of time points
  - A new long, two-level approach where time is a random mode
  - No limit on the number of time points

New Solution No. 1: Wide Format, Single-Level Approach



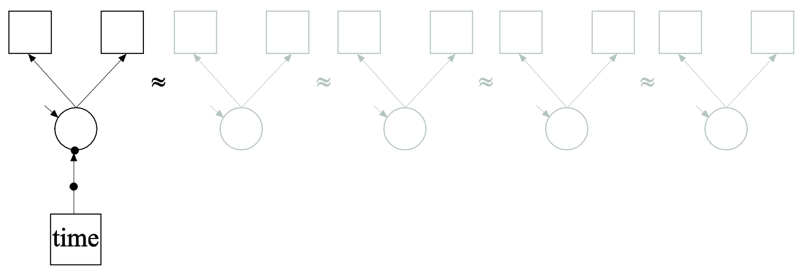
Single-level analysis with $p \times T = 2 \times 5 = 10$ variables, $T = 5$ factors.

- Bayes (”BSEM”) using approximate measurement invariance, still identifying factor mean and variance differences across time

Measurement Invariance Across Time

- New solution no. 2, time is a random mode
- A new long, two-level approach
  - Best of both worlds: Keeping the limited number of variables of the two-level approach without having to assume invariance

New Solution No. 2: Long Format, Two-Level (Cross-Classified) Approach



Two-level analysis with $p = 2$ variables.

- Bayes twolevel random approach with random measurement parameters and random factor means and variances using Type=Crossclassified: Clusters are time and person

6.2 BSEM for Aggressive-Disruptive Behavior in the Classroom

Randomized field experiment in Baltimore public schools with a classroom-based intervention aimed at reducing aggressive-disruptive behavior among elementary school students (Ialongo et al., 1999).

This analysis:

- Cohort 1
- 9 binary items at 8 time points, Grade 1 - Grade 7
- $n = 1174$

Aggressive-Disruptive Behavior in the Classroom: ML Versus BSEM For Binary Items

- Traditional ML analysis
  - 8 dimensions of integration
  - Computing time: 25:44 with INTEGRATION=MONTECARLO(5000)
  - Increasing the number of time points makes ML impossible
- BSEM analysis
  - 156 parameters
  - Computing time: 4:01
  - Increasing the number of time points has relatively less impact

BSEM Input Excerpts For Aggressive-Disruptive Behavior

```
VARIABLE:    USEVARIABLES = stub1f-tease7s;
             CATEGORICAL = stub1f-tease7s;
             MISSING = ALL (999);
DEFINE:      CUT stub1f-tease7s (1.5);
ANALYSIS:    ESTIMATOR = BAYES;
             PROCESSORS = 2;
MODEL:       f1f by stub1f-tease1f* (lam11-lam19);
             f1s by stub1s-tease1s* (lam21-lam29);
             f2s by stub2s-tease2s* (lam31-lam39);
             f3s by stub3s-tease3s* (lam41-lam49);
             f4s by stub4s-tease4s* (lam51-lam59);
             f5s by stub5s-tease5s* (lam61-lam69);
             f6s by stub6s-tease6s* (lam71-lam79);
             f7s by stub7s-tease7s* (lam81-lam89);
             f1f@1;
```

BSEM Input For Aggressive-Disruptive Behavior, Continued

```
            [stub1f$1-tease1f$1] (tau11-tau19);
            [stub1s$1-tease1s$1] (tau21-tau29);
            [stub2s$1-tease2s$1] (tau31-tau39);
            [stub3s$1-tease3s$1] (tau41-tau49);
            [stub4s$1-tease4s$1] (tau51-tau59);
            [stub5s$1-tease5s$1] (tau61-tau69);
            [stub6s$1-tease6s$1] (tau71-tau79);
            [stub7s$1-tease7s$1] (tau81-tau89);
            [f1f-f7s@0];
            i s q | f1f@0 f1s@0.5 f2s@1.5 f3s@2.5 f4s@3.5
            f5s@4.5 f6s@5.5 f7s@6.5;
            q@0;
MODEL
PRIORS:     DO(1,9) DIFF(lam1#-lam8#) ~ N(0,.01);
            DO(1,9) DIFF(tau1#-tau8#) ~ N(0,.01);
OUTPUT:     TECH1 TECH8;
```


Estimates For Aggressive-Disruptive Behavior

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	
Means						
I	0.000	0.000	1.000	0.000	0.000	
S	0.238	0.068	0.000	0.108	0.366	*
Q	-0.022	0.011	0.023	-0.043	0.000	*
Variances						
I	9.258	2.076	0.000	6.766	14.259	*
S	0.258	0.068	0.000	0.169	0.411	*
Q	0.001	0.000	0.000	0.001	0.001	

Estimates For Aggressive-Disruptive Behavior, Continued

	Estimate	Posterior S.D.	One-Tailed P-Value	95% C.I. Lower 2.5%	95% C.I. Upper 2.5%	
F1F BY						
STUB1F	0.428	0.048	0.000	0.338	0.522	*
BKRULE1F	0.587	0.068	0.000	0.463	0.716	*
HARMO1F	0.832	0.082	0.000	0.677	0.985	*
BKTHIN1F	0.671	0.067	0.000	0.546	0.795	*
YELL1F	0.508	0.055	0.000	0.405	0.609	*
TAKEP1F	0.717	0.072	0.000	0.570	0.839	*
FIGHT1F	0.480	0.052	0.000	0.385	0.579	*
LIES1F	0.488	0.054	0.000	0.386	0.589	*
TEASE1F	0.503	0.055	0.000	0.404	0.608	*
...						
F7S BY						
STUB7S	0.360	0.049	0.000	0.273	0.458	*
BKRULE7S	0.512	0.068	0.000	0.392	0.654	*
HARMO7S	0.555	0.074	0.000	0.425	0.716	*
BKTHIN7S	0.459	0.063	0.000	0.344	0.581	*
YELL7S	0.525	0.062	0.000	0.409	0.643	*
TAKEP7S	0.500	0.069	0.000	0.372	0.634	*
FIGHT7S	0.515	0.067	0.000	0.404	0.652	*
LIES7S	0.520	0.070	0.000	0.392	0.653	*
TEASE7S	0.495	0.064	0.000	0.378	0.626	*

Displaying Non-Invariant Items: Time Points With Significant Differences Compared To The Mean (V = 0.01)

Item	Loading	Threshold
stub	3	1, 2, 3, 6, 8
bkrule	-	5, 8
harmo	1, 8	2, 8
bkthin	1, 2, 3, 7, 8	2, 8
yell	2, 3, 6	-
takep	1, 2, 5	1, 2, 5
fight	1, 5	1, 4
lies	-	-
tease	-	1, 4, 8

6.3 Cross-Classified Longitudinal Analysis

- Both subject and time are random modes of variation (2 cluster variables)
- Observations nested within time and subject
- A large number of time points can be handled via Bayesian analysis
- A relatively small number of subjects is needed
- Mplus TYPE = CROSSCLASSIFIED
- Allows multiple indicator growth modeling with item parameters varying across time and subject (see UG ex9.27)

6.4 Two-Level Analysis with Random Loadings: Intensive Longitudinal Data

- Intensive longitudinal data (ILD): More and more longitudinal data are collected with very frequent observations using new tools for data collection such as palm pilots, smartphones etc.
- Ecological Momentary Assessment (EMA) involves repeated sampling of subjects' current behaviors and experiences in real time, in subjects' natural environments
- Experience Sampling Methods (ESM)
- Many time points, small number of subjects

Some Intensive Longitudinal Data Methods References

- Walls & Schafer (2006). Intensive Longitudinal Data. New York: Oxford University Press
- Jahng, Wood & Trull (2008). Analysis of Affective Instability in Ecological Momentary Assessment: Indices Using Successive Difference and Group Comparison via Multilevel Modeling. Psychological Methods, 13, 354-375 (MSSD measure)
- Bolger & Laurenceau (2012). Intensive Longitudinal Methods: An Introduction to Diary and Experience Sampling Research. New York: Guilford Press
- Brose & Ram (2012). Within-Person Factor Analysis. In the new Handbook of Research Methods for Studying Daily Life

6.5 Individual Differences Factor Analysis: Two-Level Analysis with Random Factor Loadings

- Jahng S., Wood, P. K.,& Trull, T. J., (2008). Analysis of Affective Instability in Ecological Momentary Assessment: Indices Using Successive Difference and Group Comparison via Multilevel Modeling. Psychological Methods, 13, 354-375
- An example of the growing amount of EMA data
- 84 outpatient subjects: 46 meeting borderline personality disorder (BPD) and 38 meeting MDD or DYS
- Each individual is measured several times a day for 4 weeks for total of about 100 assessments
- A mood factor for each individual is measured with 21 self-rated continuous items
- The research question is if the BPD group demonstrates more temporal negative mood instability than the MDD/DYS group

Individual Differences Factor Analysis (IDFA)

- This data set is suitable for checking if a measurement instrument is interpreted the same way by different individuals. Some individuals responses may be more correlated for some items, i.e., the correlation matrix could be different for different individuals
- Suppose that one individual always answers item 1 and 2 the same way and a second individual doesn't. We need separate factor analysis models for the two individuals, that is, individual-specific factor loadings
- If the within-level correlation matrix varies across individuals that means that the loadings are individual-specific
- Should factor loadings be individually specific in general? This cannot be determined in cross-sectional studies, only in longitudinal studies with multiple assessments
- IDFA uses TYPE=TWOLEVEL where cluster = individual with many assessments per cluster

Individual Differences Factor Analysis (IDFA) Continued

- Large across-time variance of the mood factor is considered a core feature of BPD that distinguishes this disorder from other disorders like depressive disorders.
- The individual-specific factor variance is the most important feature in this study
- The individual-specific factor variance is confounded with individual-specific factor loadings
- How to separate the two? Answer: **Using IDFA with a factor model for the random factor loadings**
- Asparouhov & Muthén, B. (2012). General Random Effect Latent Variable Modeling: Random Subjects, Items, Contexts, and Parameters

7. Mixture Modeling

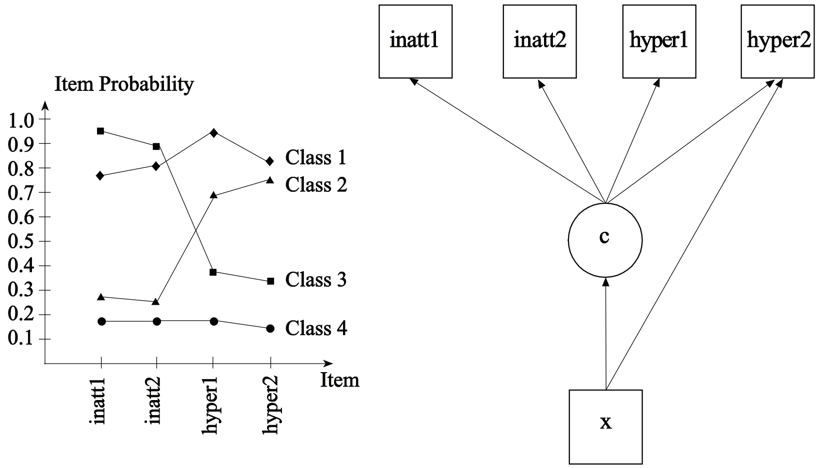
Analysis Methods

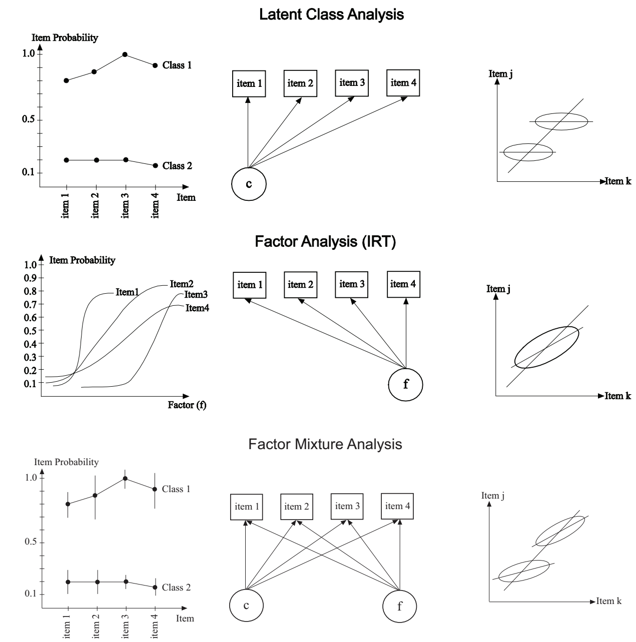
- Regression mixture models - Modeling of counts, randomized interventions with non-compliance
- Latent class analysis with and without covariates
- Latent transition analysis
- Latent class growth analysis
- Growth mixture modeling
- Survival mixture modeling

Mixture Modeling: Overview of Version 7 Developments

- 3-step mixture modeling: Analyze-classify-analyze approaches to investigate covariates and distal outcomes
  - LCA
  - Regression mixture analysis
  - GMM
  - LTA
- Latent transition analysis (LTA)
  - Introductory examples
  - New Mplus output
  - Covariates influencing transition probabilities
  - Probability parameterization useful for Mover-Stayer LTA
  - LTA extensions

7.1 Latent Class Analysis





Latent Class, Factor, And Factor Mixture Analysis Alcohol Dependence Criteria, NLSY 1989 (n = 8313)

Source: Muthén & Muthén (1995)

	Latent Classes				
	Two-class solution[1]		Three-class solution[2]		
	I	II	I	II	III
Prevalence	0.78	0.22	0.75	0.21	0.03
DSM-III-R criterion	conditional probability of fulfilling a criterion				
Withdrawal	0.00	0.14	0.00	0.07	0.49
Tolerance	0.01	0.45	0.01	0.35	0.81
Larger	0.15	0.96	0.12	0.94	0.99
Cut down	0.00	0.14	0.01	0.05	0.60
Time spent	0.00	0.19	0.00	0.09	0.65
Major role-hazard	0.03	0.83	0.02	0.73	0.96
Give up	0.00	0.10	0.00	0.03	0.43
Relief	0.00	0.08	0.00	0.02	0.40
Continue	0.00	0.24	0.02	0.11	0.83

[1]Likelihood ratio chi-square fit = 1779 with 492 degrees of freedom

[2]Likelihood ratio chi-square fit = 448 with 482 degrees of freedom

LCA, FA, And FMA For NLSY 1989

- LCA, 3 classes: logL = -14,139, 29 parameters, BIC = 28,539
- FA, 2 factors: logL = -14,083, 26 parameters, BIC = 28,401
- FMA 2 classes, 1 factor, loadings invariant:

  logL = -14,054, 29 parameters, BIC = 28,370

Models can be compared with respect to fit to the data:

- Standardized bivariate residuals
- Standardized residuals for most frequent response patterns

Estimated Frequencies And Standardized Residuals

Obs. Freq.	LCA 3c		FA 2f		FMA 1f, 2c	
	Est. Freq.	Res.	Est. Freq.	Res.	Est. Freq.	Res.
5335	5332	-0.07	5307	-0.64	5331	-0.08
941	945	0.12	985	1.48	946	0.18
601	551	**-2.22**	596	-0.22	606	0.21
217	284	**4.04**	211	-0.42	228	0.75
155	111	**-4.16**	118	**-3.48**	134	1.87
149	151	0.15	168	1.45	147	0.17
65	68	0.41	46	**-2.79**	53	1.60
49	52	0.42	84	**3.80**	59	1.27
48	54	0.81	44	-0.61	46	0.32
47	40	-1.09	45	-0.37	45	0.33

Bolded entries are significant at the 5% level.

Input For FMA Of 9 Alcohol Items In The NLSY 1989

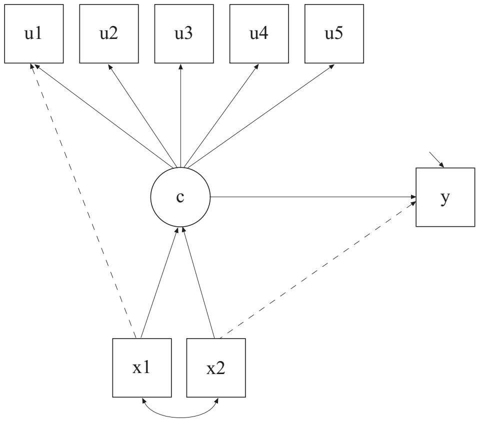
```
TITLE:      Alcohol LCA M & M (1995)
DATA:       FILE = bengt05_spread.dat;
VARIABLE:   NAMES = u1-u9;
            CATEGORICAL = u1-u9;
            CLASSES = c(2);
ANALYSIS:   TYPE = MIXTURE;
            ALGORITHM = INTEGRATION;
            STARTS = 200 10; STITER = 20;
            ADAPTIVE = OFF;
            PROCESSORS = 4(STARTS);
```


Input For FMA Of 9 Alcohol Items In The NLSY 1989 (Continued)

```
MODEL:      %OVERALL%
            f BY u1-u9;
            f*1; [f@0];
            %c#1%
            [u1$1-u9$1];
            f*1;
            %c#2%
            [u1$1-u9$1];
            f*1;
OUTPUT:     TECH1 TECH8 TECH10;
PLOT:       TYPE = PLOT3;
            SERIES = u1-u9(*);
```


7.2 3-Step Mixture Modeling

1-step analysis versus 3-step (analyze-classify-analyze) latent class analysis



1-Step vs 3-Step: A Hypothetical Genetic Example

Substantive question: Should the latent classes be defined by the indicators alone or also by covariates and distal outcomes (antecedents and consequences)?

- Example: Study of genotypes (x variables) influencing phenotypes (y variables)
- Phenotypes may be observed indicators of mental illness such as DSM criteria. The interest is in finding latent classes of subjects and then trying to see if certain genotype variables influence class membership
- Possible objection to 1-step: If the genotypes are part of deciding the latent classes, the assessment of the strength of relationship is compromised
- 3-step: Determine the latent classes based on only phenotype information. Then classify subjects. Then relate the classification to the genotypes

Substantive Checking of Latent Class Models

- Latent class models should be subjected to both statistical and substantive checking (Muthén, 2003 in Psychological Methods)
- Substantive checking can be done by relating latent classes to antecedents and consequences (covariates and distal outcomes)
- The 3-step approach is a useful tool for this

The Old 3-Step Approach

1. Estimate the LCA model
2. Determine each subject's most likely class membership
3. Relate the most likely class variable to other variables

The old 3-step approach is problematic: Unless the classification is very good (high entropy), this gives biased estimates and biased standard errors for the relationships with other variables.

The LCA Provides Information About the Classification Quality

Average Latent Class Probabilities for Most Likely Class Membership (Row) by Latent Class (Column)

	1	2	3
1	0.839	0.066	0.095
2	0.053	0.845	0.102
3	0.125	0.107	0.768

The New 3-Step Approach

- New Method in Mplus Version 7: 3-Step approach correcting for classification error
  1. Estimate the LCA model
  2. Create a nominal most likely class variable N
  3. Use a mixture model for N, C and X, where N is a C indicator with measurement error rates prefixed at the misclassification rate of N estimated in the step 1 LCA analysis
- Bolck, Croon, & Hagenaars (2004) Estimating latent structure models with categorical variables: One-step versus three-step estimators. Political Analysis, 12, 3-27.
- Vermunt (2010). Latent Class Modeling with Covariates: Two improved three-step approaches. Political Analysis, 18, 450-469
- Asparouhov & Muthén (2012). Auxiliary variables in mixture modeling: A 3-step approach using Mplus. Mplus Web Note 15.

Classification Information from Step 1 LCA

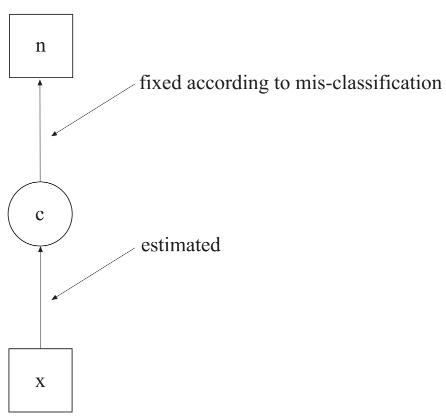
Average Latent Class Probabilities for Most Likely
Class Membership (Row) by Latent Class (Column)

	1	2	3
1	0.839	0.066	0.095
2	0.053	0.845	0.102
3	0.125	0.107	0.768

log(0.839/0.095) = 2.178
log(0.066/0.095) = -0.364
log(0.053/0.102) = -0.654
log(0.845/0.102) = 2.114
log(0.125/0.768) = -1.815
log(0.107/0.768) = -1.970

Step 3 Regression on a Covariate

- n: Most likely class membership from Step 2 (nominal variable)
- c: Latent class variable
- x: Covariate



Input File for Step 3 in the 3-Step Estimation

```
VARIABLE:   NAMES = u1-u5 x p1-p3 n;
            USEVARIABLES = x n;
            CLASSES = c(3);
            NOMINAL = n;
DATA:       FILE = man3step2.dat;
ANALYSIS:   TYPE = MIXTURE; STARTS = 0;
MODEL:      %OVERALL%
            c ON x;
            %c#1%
            [n#1@2.178];
            [n#2@-0.364];
            %c#2%
            [n#1@-0.654];
            [n#2@2.114];
            %c#3%
            [n#1@-1.815];
```

Auxiliary Variables In Mixtures: Covariate x and Distal y

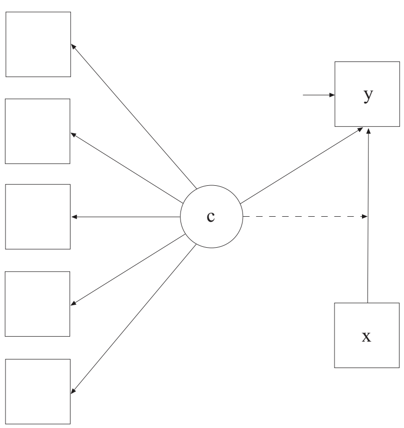
```
VARIABLE:   NAMES = u1-u5 x;
            CATEGORICAL = u1-u5;
            CLASSES = c(3);
            AUXILIARY = x(R3STEP);
DATA:       FILE = 3step.dat;
ANALYSIS:   TYPE = MIXTURE;
MODEL:      !no model is needed, LCA is default

VARIABLE:   NAMES = u1-u5 y;
            CATEGORICAL = u1-u5;
            CLASSES = c(3);
            AUXILIARY = y(DU3STEP);
DATA:       FILE = 3step.dat;
ANALYSIS:   TYPE = MIXTURE;
MODEL:      !no model is needed, LCA is default
```

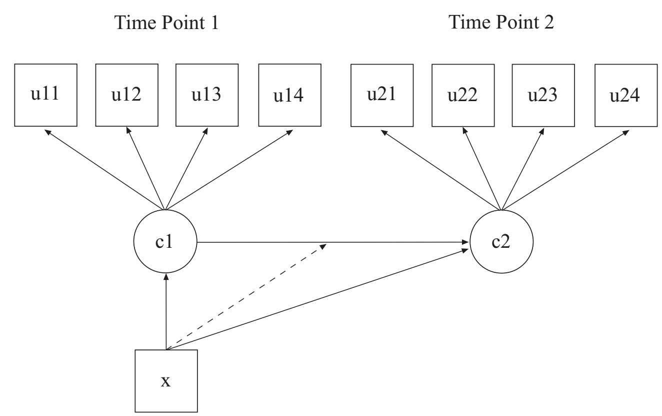

A Second Look at Distal 3-Step

- In some examples the Asparouhov-Muthén distal 3-step method in Mplus Web Note 15 leads to changes in latent class formation between Step 1 and Step 3 - warning given in Mplus Version 7.1
- Lanza et al. (2013) in the SEM journal propose a different distal 3-step method that avoids changes in class formation. Included in Mplus Version 7.1 (DCON/DCAT).
- Future research needed to evaluate which method, including Most Likely Class and Pseudo-class, is least sensitive to violations of assumptions such as no direct effects

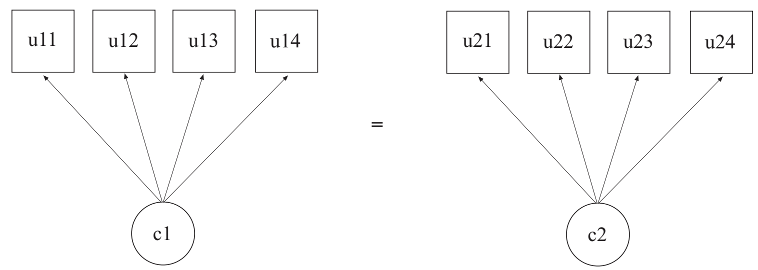
Manual 3-Step Mixture Modeling For Special Models: A Regression Mixture Example



3-Step Latent Transition Analysis

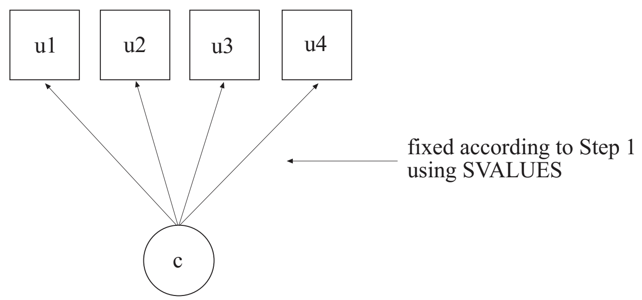


LTA: Step 1



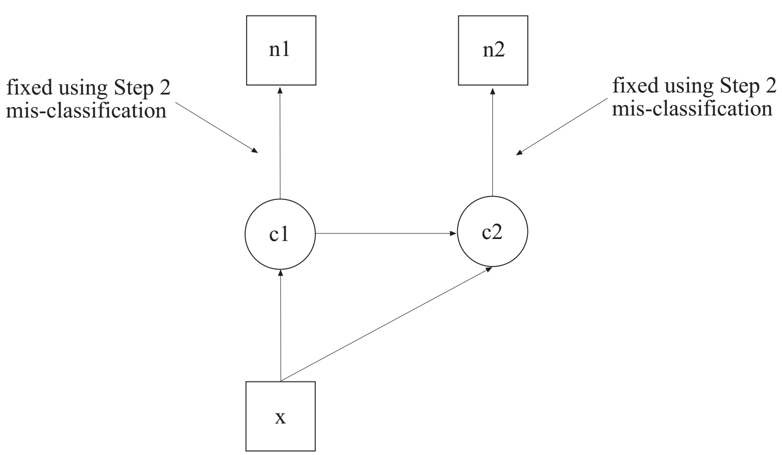
LTA: Step 2

For each time point:

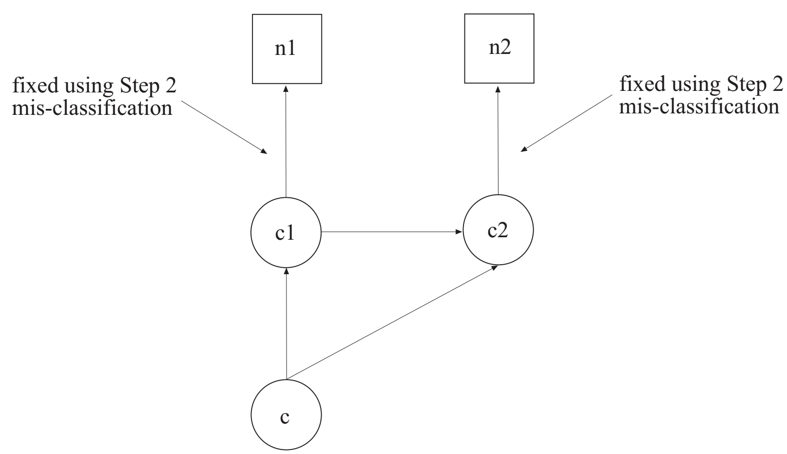


Save most likely class

LTA: Step 3



3-Step Mover-Stayer LTA



7.3 Latent Transition Analysis Developments

New developments in Version 7:

- TECH15 output with conditional class probabilities useful for studying transition probabilities varying as a function of an observed binary or nominal covariate such as treatment/control, ethnicity, or a latent class covariate
- LTA transition probability calculator for continuous covariates
- Probability parameterization to simplify input for Mover-Stayer LTA and other models with restrictions on the transition probabilities
- New User's Guide examples
  - 8.13: LTA for two time points with a binary covariate influencing the latent transition probabilities
  - 8.14: LTA for two time points with a continuous covariate influencing the latent transition probabilities
  - 8.15: Mover-stayer LTA for three time points using a probability parameterization

8. Multilevel Modeling

- Within-cluster multiple-group modeling (Web Note 16)
- Advantages of Bayesian analysis
- Meta analysis (2-level random). See Topic 9
- Two-level latent class analysis (within and between classes). See Topic 7
- 3-level analysis
- Cross-classified analysis
- 3-level and cross-classified multiple imputation
- Applications to Item Response Theory modeling

8.1 Advantages of Bayesian Multilevel Analysis

With carefully chosen priors, Bayes allows a smaller number of level 2 or level 3 units. See, e.g.,

Muthén (2010). Bayesian analysis in Mplus: A brief introduction. Technical Report. Version 3.

8.2 Three-Level Analysis

Continuous outcomes: ML and Bayesian estimation

Categorical outcomes: Bayesian estimation (Bayes uses probit)

Count and nominal outcomes: Not yet available

Types Of Observed Variables In 3-Level Analysis

Each Y variable is decomposed as

$$Y_{ijk} = Y_{1ijk} + Y_{2jk} + Y_{3k},$$

where Y_{1ijk}, Y_{2jk}, and Y_{3k} are components of Y_{ijk} on levels 1, 2, and 3. Here, Y_{2jk}, and Y_{3k} may be seen as random intercepts on respective levels, and Y_{1ijk} as a residual

- Some variables may not have variation over all levels. To avoid variances that are near zero which cause convergence problems specify/restrict the variation level
- WITHIN=Y, has variation on level 1, so Y_{2jk} and Y_{3k} are not in the model
- WITHIN=(level2) Y, has variation on level 1 and level 2
- WITHIN=(level3) Y, has variation on level 1 and level 3
- BETWEEN= Y, has variation on level 2 and level 3
- BETWEEN=(level2) Y, has variation on level 2
- BETWEEN=(level3) Y, has variation on level 3

Types Of Random Slopes In 3-Level Analysis

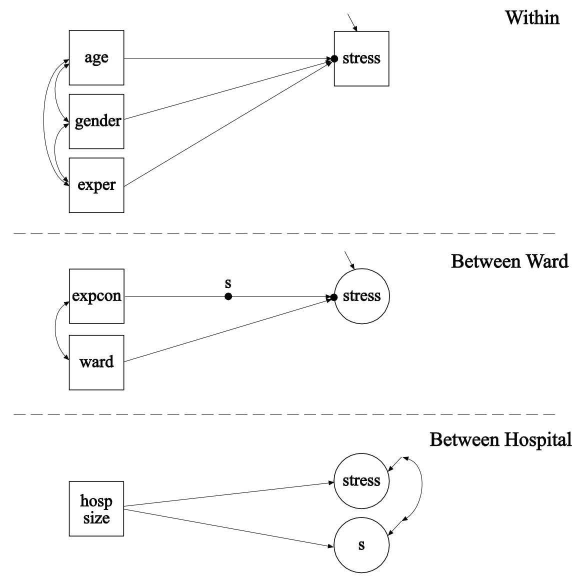
- Type 1: Defined on the level 1
  %WITHIN%
  s | y ON x;
  The random slope s has variance on level 2 and level 3
- Type 2: Defined on the level 2
  %BETWEEN level2%
  s | y ON x;
  The random slope s has variance on level 3 only
- The dependent variable can be an observed Y or a factor. The covariate X should be specified as WITHIN= for type 1 or BETWEEN=(level2) for type 2, i.e., no variation beyond the level it is used at

Three-Level Regression: Nurses Data

Source: Hox (2010). Multilevel Analysis. Hypothetical data discussed in Section 2.4.3

- Study of stress in hospitals
- Reports from nurses working in wards nested within hospitals
- In each of 25 hospitals, 4 wards are selected and randomly assigned to experimental or control conditions (cluster-randomized trial)
- 10 nurses from each ward are given a test that measures job-related stress
- Covariates are age, experience, gender, type of ward (0=general care, 1=special care), hospital size (0=small, 1=medium, 2=large)
- Research question: Is the experimental effect different in different hospitals? - Random slope varying on level 3

3-Level Regression Example: Nurses Data



Input For Nurses Data

```
TITLE:      Nurses data from Hox (2010)
DATA:       FILE = nurses.dat;
VARIABLE:   NAMES = hospital ward wardid nurse age gender
            experience stress wardtype hospsize expcon zage
            zgender zexperience zstress zwardtyi zhospsize
            zexpcon cexpcon chospsize;
            CLUSTER = hospital wardid;
            WITHIN = age gender experience;
            BETWEEN = (hospital) hospsize (wardid) expcon wardtype;
            USEVARIABLES = stress expcon age gender experience
            wardtype hospsize;
            CENTERING = GRANDMEAN(expcon hospsize);
ANALYSIS:   TYPE = THREELEVEL RANDOM;
            ESTIMATOR = MLR;
```


Input For Nurses Data, Continued

```
MODEL:      %WITHIN%
            stress ON age gender experience;
            %BETWEEN wardid%
            s | stress ON expcon;
            stress ON wardtype;
            %BETWEEN hospital%
            s stress ON hospsize;
            s; s WITH stress;
OUTPUT:     TECH1 TECH8;
SAVEDATA:   SAVE = FSCORES;
            FILE = fs.dat;
PLOT:       TYPE = PLOT2 PLOT3;
```


Model Results For Nurses Data

	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
WITHIN Level				
stress ON				
age	0.022	0.002	11.911	0.000
gender	-0.455	0.032	-14.413	0.000
experience	-0.062	0.004	-15.279	0.000
Residual Variances				
stress	0.217	0.011	20.096	0.000
BETWEEN wardid Level				
stress ON				
wardtype	0.053	0.076	0.695	0.487

Model Results For Nurses Data, Continued

	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
Residual Variances				
stress	0.109	0.033	3.298	0.001
BETWEEN hospital Level				
s ON				
hospsize	0.998	0.191	5.217	0.000
stress ON				
hospsize	-0.041	0.152	-0.270	0.787
s WITH				
stress	-0.036	0.058	-0.615	0.538

Model Results For Nurses Data, Continued

	Estimates	S.E.	Est./S.E.	Two-Tailed P-Value
Intercepts				
stress	5.753	0.102	56.171	0.000
s	-0.699	0.111	-6.295	0.000
Residual Variances				
stress	0.143	0.051	2.813	0.005
s	0.178	0.087	2.060	0.039

8.3 Cross-Classified Analysis

Students are cross-classified by school and neighbourhood at level 2. An example with 33 students:

	School 1	School 2	School 3	School 4
Neighbourhood 1	XXXX	XX	X	X
Neighbourhood 2	X	XXXXX	XXX	XX
Neighbourhood 3	XX	XX	XXXX	XXXXXX

Source: Fielding & Goldstein (2006). Cross-classified and multiple membership structures in multilevel models: An introduction and review. Research Report RR 791, University of Birmingham.

Types of Cross-Classified Analyses in Mplus

- Regression analysis
- Path analysis (both subject and context are random modes)
  - Gonzalez, de Boeck, & Tuerlinckx (2008). A double-structure structural equation model for three-mode data. Psychological Methods, 13, 337-353
- SEM
- Longitudinal analysis (both subject and time are random modes)
- Random items (both subject and item are random modes)
- General idea: Two random modes
- Limited forms of multiple membership modeling (see Day 3 of Utrecht and the article Jeon & Rabe-Hesketh (2012). Profile-Likelihood Approach for Estimating Generalized Linear Mixed Models With Factor Structures. JEBS)

8.4 Three-Level and Cross-Classified Multiple Imputation

New Multiple Imputation Methods

- Multiple imputations for three-level and cross-classified data
- Continuous and categorical variables
- H0 imputations. Estimate a three-level or cross-classified model with the Bayes estimator. Not available as H1 imputation where the imputation model is setup as unrestricted model.
- The imputation model can be an unrestricted model or a restricted model. Restricted models will be easier to estimate especially when the number of clustering units is not large
- In the input file simply add the DATA IMPUTATION command

9. Mplus Strengths For IRT And Categorical Factor Analysis

- High-dimensional analysis using WLSMV, Bayes, and ML two-tier
- Bi-factor EFA
- Modification indices, correlated residuals
- Multiple-group analysis
- Mixtures*
- Complex survey data handling: Stratification, weights
- Multilevel: two-level, three-level, and cross-classified
- Random loadings (discrimination) using Bayesian analysis
- Random item IRT
- Random subjects and contexts

* Muthén, B. (2008). Latent variable hybrids: Overview of old and new models. In Hancock, G. R., & Samuelsen, K. M. (Eds.), Advances in latent variable mixture models, pp. 1-24. Charlotte, NC: Information Age Publishing, Inc.

Random items, Generalizability Theory

- Items are random samples from a population of items
- The same or different items may be administered to individuals
- Suited for computer generated items and adaptive testing
- 2-parameter IRT model

$$P(Y_{ij} = 1) = \Phi(a_j \theta_i + b_j)$$

- $a_j \sim N(a, \sigma_a)$, $b_j \sim N(b, \sigma_b)$: random discrimination and difficulty parameters
- The ability parameter is $\theta_i \sim N(0,1)$
- Cross-classified model. Nested within items and individuals. 1 or 0 observation in each cross-classified cell
- Interaction of two latent variables: a_j and θ_i
- The model has only 4 parameters - much more parsimonious than regular IRT models

Version 7 User's Guide ex9.26

10. Recent Mplus Papers

- Asparouhov & Muthén (2010). Computing the strictly positive Satorra-Bentler chi-square test in Mplus. Mplus Web Notes: No. 12
- Muthén (2010). Bayesian analysis in Mplus: A brief introduction. Technical Report. Version 3
- Muthén (2011). Applications of Causally Defined Direct and Indirect Effects in Mediation Analysis using SEM in Mplus.
- Muthén & Asparouhov (2011). LTA in Mplus: Transition probabilities influenced by covariates. Mplus Web Notes: No. 13
- Asparouhov & Muthén (2012). Using Mplus TECH11 and TECH14 to test the number of latent classes. Mplus Web Notes: No. 14
- Asparouhov & Muthén (2012). Comparison of computational methods for high dimensional item factor analysis
- Asparouhov & Muthén (2012). General random effect latent variable modeling: Random subjects, items, contexts, and parameters
- Asparouhov & Muthén (2012). Auxiliary variables in mixture modeling: A 3-step approach using Mplus. Mplus Web Note 15
- Asparouhov & Muthén (2012). Multiple group multilevel analysis. Web Note 16
- Muthén & Asparouhov (2012). Bayesian SEM: A more flexible representation of substantive theory. Psychological Methods, 17, 313-335
- Sobel & Muthén (2012). Compliance mixture modelling with a zero effect complier class and missing data. Biometrics, 68, 1037-1045
- Muthén & Asparouhov (2013). Item response modeling in Mplus: A multi-dimensional, multi-level, and multi-timepoint example
- Muthén & Asparouhov (2013). BSEM measurement invariance analysis. Web Note 17
- Asparouhov & Muthén (2013). Multiple-group rotational alignment. In preparation