

# Latent Variable Methods Workshop

## Longitudinal Data Analysis Models

Tue Sep 2 - Thu Sep 4 2025

# Day 1 : Overview, Path Analysis, Factor Analysis, Structural Equation Modeling, and Effectively Using Mplus Software

[lvmworkshop.org](http://lvmworkshop.org)

Zoom

Brown University

September 2, 2025

# Day 1

- Orientation, overview and objectives
- Longitudinal data analysis (LDA)
- Structural equation modeling (SEM)
- Using Mplus & R/lavaan
- Lab exercise: CFA in Mplus & R/lavaan

# Workshop Objectives

- ① Introduce LDA in LVM (Mplus, R/lavaan) framework
  - ▶ A range of questions addressed by LDA
  - ▶ How approached using Mplus, R/lavaan
- ② Experience applied LDA using Mplus, or R/lavaan
  - ▶ A range of models
    - ★ Latent growth curve
    - ★ latent change score models
    - ★ survival analysis (Mplus only)
    - ★ Mixture models (Mplus only)

# Workshop Structure

- Morning Sessions

- ▶ Lecture
- ▶ Lab sessions
- ▶ Run through examples
- ▶ Ask questions at any time

- Afternoon Sessions

- ▶ More lecture
- ▶ More time to run through examples
- ▶ Or with your own data
- ▶ Individual assistance is the goal, emphasis on provided examples

## Content Covered - Day 1

- Introduction to Longitudinal Data Analysis

- ▶ Overview
- ▶ Data Management for LDA
- ▶ Introduction to Structural Equation Modeling (with Mplus)
- ▶ Nuts and Bolts of Using Mplus
  - ★ How to communicate about models
  - ★ How to get data into Mplus
  - ★ The Statistical model in general (arrays)

## Content Covered - Day 2

- Example data set
- Latent Growth Curve Models (LGCM)
  - ▶ Unconditional LGCM
  - ▶ Conditional LGCM
  - ▶ Non-Linear LGCM
- Random Effects Models for growth with Mplus
- Multilevel model approach for growth modeling with Mplus
- Growth Mixture Models
- Multiple indicator LGCM
- Modeling Retest, Practice and other methods artifacts

# Content Covered - Day 3

- Advanced Topics
  - ▶ LCGM
    - ★ Multiple Group Models
    - ★ Randomized Controlled Trials
    - ★ Bivariate, Parallel, and Sequential Process Models
  - ▶ Latent Change Score (Dual Change Score Models)
    - ★ Dual Change Score Models (DSCM), Bivariate DSCM
  - ▶ Bayesian Data Analysis and LGCM
  - ▶ Survival Analysis with Mplus
    - ★ Discrete time, Continuous time
    - ★ Simultaneous survival and growth modeling
  - ▶ Modeling with complex sample weights

## Where to go for more information

- see links at
- [https://www.statmodel.com/courses\\_openenroll.shtml](https://www.statmodel.com/courses_openenroll.shtml)
- The Center for Statistical Training by Curran-Bauer Analytics  
<https://centerstat.org/>
- <https://www.lvmworkshop.org/home/---deprecated-pages/usefullinks>

## Other Resources

- What is longitudinal data analysis?
  - ▶ Singer JD & Willett JB. Applied longitudinal data analysis: Modeling change and event occurrence. 2003, New York: Oxford University Press.
- How do I do latent growth curve modeling?
  - ▶ Duncan TE, Duncan SC, & Strycker LA. An introduction to latent variable growth curve modeling: concepts, issues and applications. Second ed. 2006, Mahwah, New Jersey: LEA, Inc.
  - ▶ Newsom, J., Jones, R., & Hofer, S. (Eds.). (2011). Longitudinal Data Analysis: A Practical Guide for Researchers in Aging, Health and Social Sciences. New York: Routledge.
- Tell me more about the math behind latent curve methods
  - ▶ Bollen KA & Curran PJ. Latent curve models: a structural equation perspective. Wiley series in probability and statistics. 2006, Hoboken, N.J.: Wiley-Interscience.
- More about structural equation modeling
  - ▶ Bryne, B. (2011). Structural Equation Modeling with Mplus: Basic Concepts, Applications and Programming. New York: Routledge.

## Other Key References

- McArdle J. Five steps in latent curve modeling with longitudinal life-span data. In: Levy R, Ghisletta P, Le Gorr J, Spini D, Widmer E, eds. Towards an Interdisciplinary Perspective on the Life Course (10). San Diego, CA: JAI Press; 2005
- McArdle J, Epstein D. Latent growth curves within developmental structural equation models. *Child Development*. 1987;58:110-133.
- McArdle JJ. Latent variable modeling of differences and changes with longitudinal data. *Annual review of psychology*. 2009;60:577-605.
- Curran, PJ, & Muthén, BO. Testing developmental theories in intervention research: Latent growth analysis and power estimation. *American Journal of Community Psychology*. 1999; 27:567-595.
- Jung T, & Wickrama, KAS. An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*. 2008;2:302-317.

# Introduction

Longitudinal Data Analysis

# This Workshop

- Applied statistical data analysis of longitudinal data

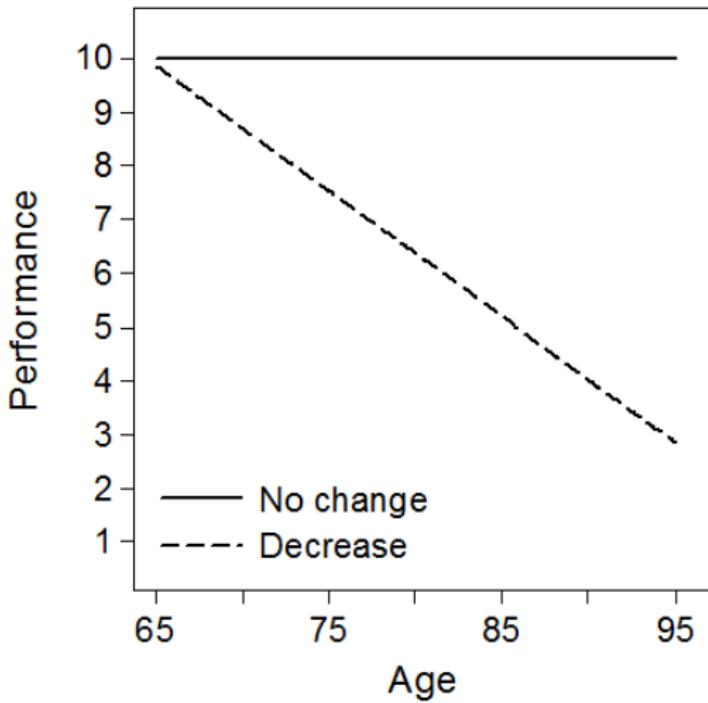
- ▶ Latent growth curve (LGC) or latent growth model (LGM)
- ▶ Survival models
- ▶ Latent change score models
- ▶ Mixture models
  - ★ growth mixture
  - ★ change score mixture

## The Motivating Questions

- How does  $Y$  change over time?  $Y$  could be
  - ▶ cognitive functioning, neuropsych performance
  - ▶ mood, depression
  - ▶ weight
  - ▶ number of falls
  - ▶ etc.
- What factors predict who changes slower or faster on  $Y$

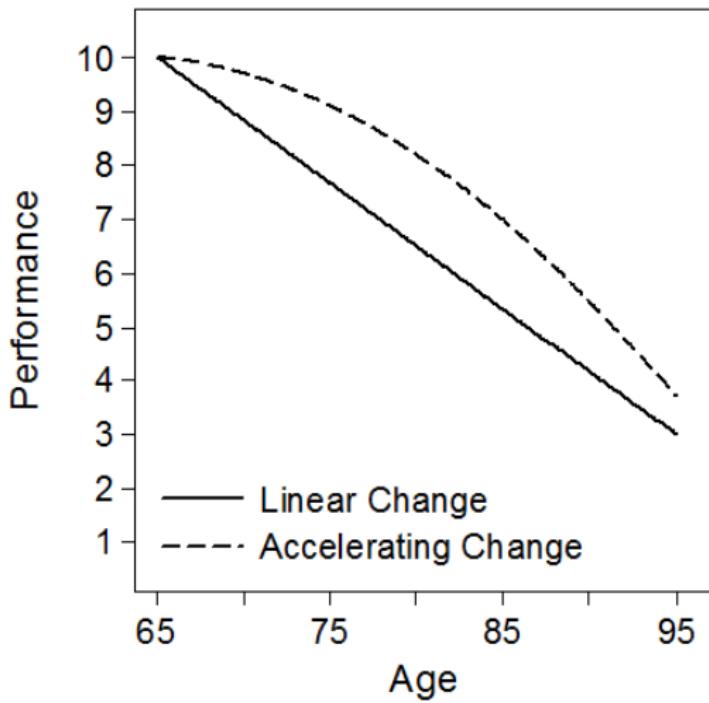
# About Change

Figure 1



# About Change

Figure 2



# The M in LGM

Why modeling is important

## Change Scores

*Since life-span research is, in essence, concerned with change over time, it is tempting to assume that the basic psychometric datum with which one works is the change score. Thus, if one is making comparisons of intelligence test scores at two points in time (symbolized by  $X_1$  and  $X_2$ ), it is assumed that the interest is in  $(X_2 - X_1)$  and that such scores should be computed as one of the first steps in performing analyses. Actually, both the history of the problem and the logic of investigation indicate that the last thing one wants to do is to think in terms of or compute such change scores unless the problem makes it absolutely necessary. (p87, emphasis added)*

Nunnally, J. C. (1973). Research Strategies and Measurement Methods for Investigating Human Development. In J. R. Nesselroade & H. W. Reese (Eds.), Life-Span Developmental Psychology: Methodological Issues (pp. 87-88). New York: Academic Press.

# The Problems with $Y_t - Y_{(t-1)}$

- Classical
  - ▶ Reliability of Change Scores
  - ▶ Regression to the Mean
  - ▶ "Law" of Initial Values
- Empirical
  - ▶  $\text{VAR}(\text{observed change}) > \text{VAR}(\text{true change})$
  - ▶ Floor/Ceiling Effects
  - ▶ Limited information / assumption of linear change

# Change Scores

- Difference between current and prior state
- Easy to interpret
- Easy to communicate meaning
- Really only useful in completely randomized designs when comparison and control groups are (or can be assumed to be) equivalent at baseline
- ...and even still may not be most powerful approach

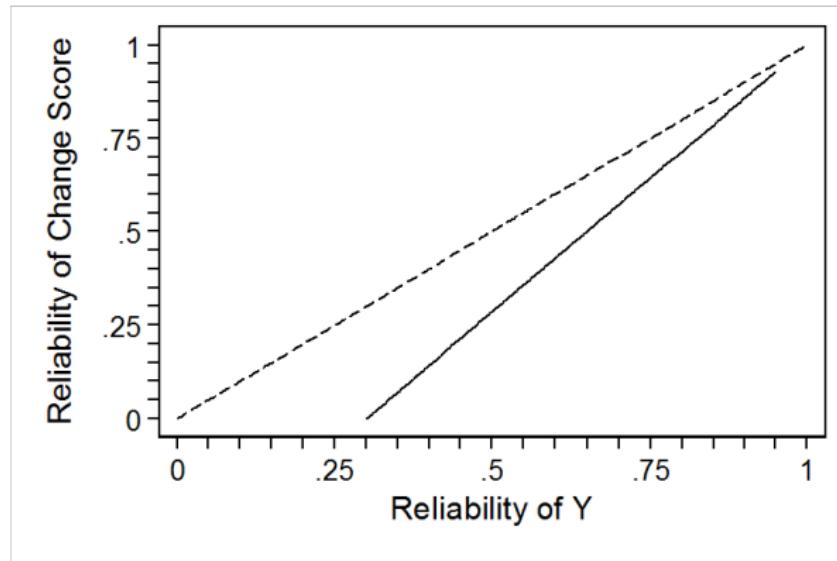
# (Un)Reliability of Change Scores

- The first line of criticism of change scores
  - ▶ Cronbach, L., & Furby, L. (1970). How we should measure "change" - or should we? *Psyc Bull* 74; 68-80.
  - ▶ Lord FM. The measurement of growth. *Educ Psychol Meas* 1956;16:421-437.
  - ▶ Lord FM. Further problems in the measurement of growth. *Educ Psychol Meas* 1958;18:437-454.
- Aging research example
  - ▶ van Belle, G., Uhlmann, R. F., Hughes, J. P., & Larson, E. B. (1990). Reliability of estimates of changes in mental status test performance in senile dementia of the Alzheimer type. *J Clin Epi*, 43(6), 589-95.
- See Also
  - ▶ Galton, F. (1886). Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute*, 15, 246-263.

# (Un)Reliability of Change Scores

- Premise is not universally accepted.
  - ▶ Rogosa D: Myths about longitudinal research. In Schaie K, Cambell R, Meredith W, Rawlings S (eds), Methodological issues in aging research. New York, Springer Publishing Company, 1988, 171-209
  - ▶ Nesselroade JR: Application of multivariate strategies to problems of measuring and structuring long-term change. In Goulet LR, Baltes PB (eds), Life-span developmental psychology: Research and theory. New York, Academic Press, 1970, 193-207
- Detailed review
  - ▶ Wittmann, W. (1997). The reliability of change scores: many misinterpretations of Lord and Cronbach by many others; revisiting some basics for longitudinal research. Methodology conference: Evaluation of Change in Longitudinal Data, Nürnberg.  
<http://www.psychologie.uni-mannheim.de/psycho2/publi/papers/reliab.pdf>

# (Un)Reliability of Change Scores



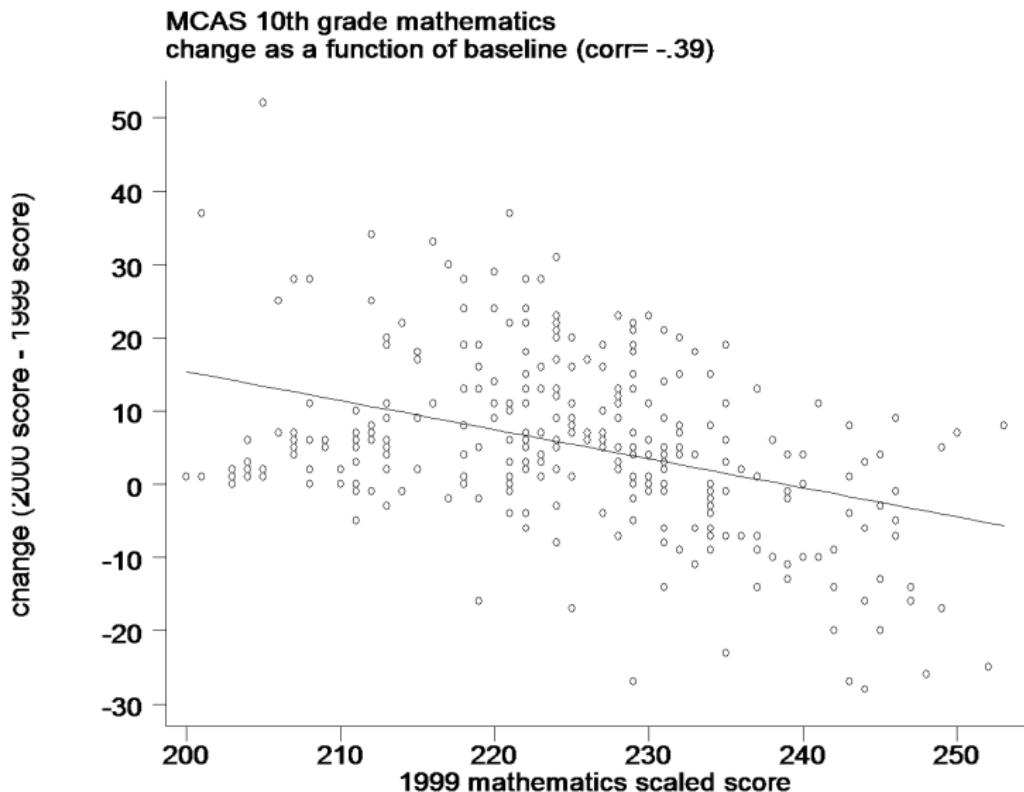
$$r_{\Delta} = \frac{r_{Y(t)} + r_{Y(t-1)} - 2 \times r_{Y(t), Y(t-1)}}{2 - 2 \times r_{Y(t), Y(t-1)}}$$

$r_{YY}$  = reliability = 0.80;  $r_{Y(t), Y(t-1)}$  = 0.3

Streiner DL & Norman GR (1995). *Health Measurement scales*. New York. Oxford.

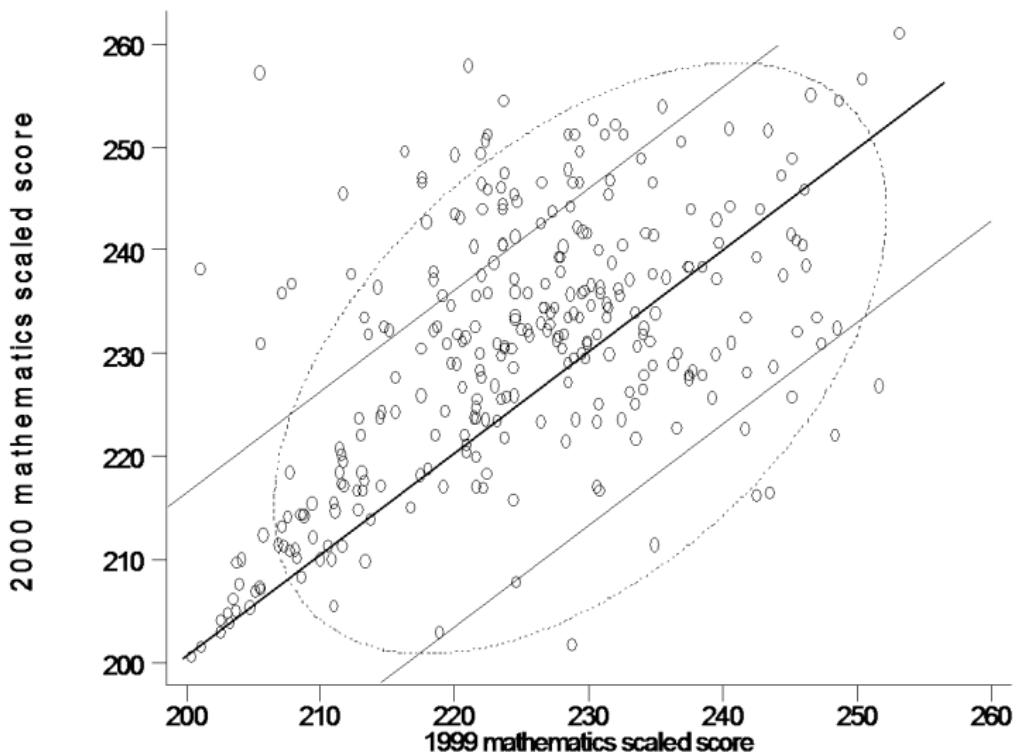
# What Is Regression to the Mean?

- The only thing consistent about regression to the mean is *the absence of explicit, defensible definitions of the phenomenon* (p186).
- Rogosa, D. (1988). Myths about longitudinal research. In K. Schaie, R. Campbell, W. Meredith, & S. Rawlings (Eds.), Methodological issues in aging research (pp. 171-209). New York: Springer Publishing Company.



$$COR(y_2, y_1) = 0.54; \quad COR(y_2 - y_1, y_1) = -0.39$$

MCAS results, regular students  
correlation of mathematics scores 1999-2000 = 0.54



$$COR(y_2, y_1) = 0.54; \quad COR(y_2 - y_1, y_1) = -0.39$$

## Some top-scoring schools faulted

Many questioning MCAS assessment

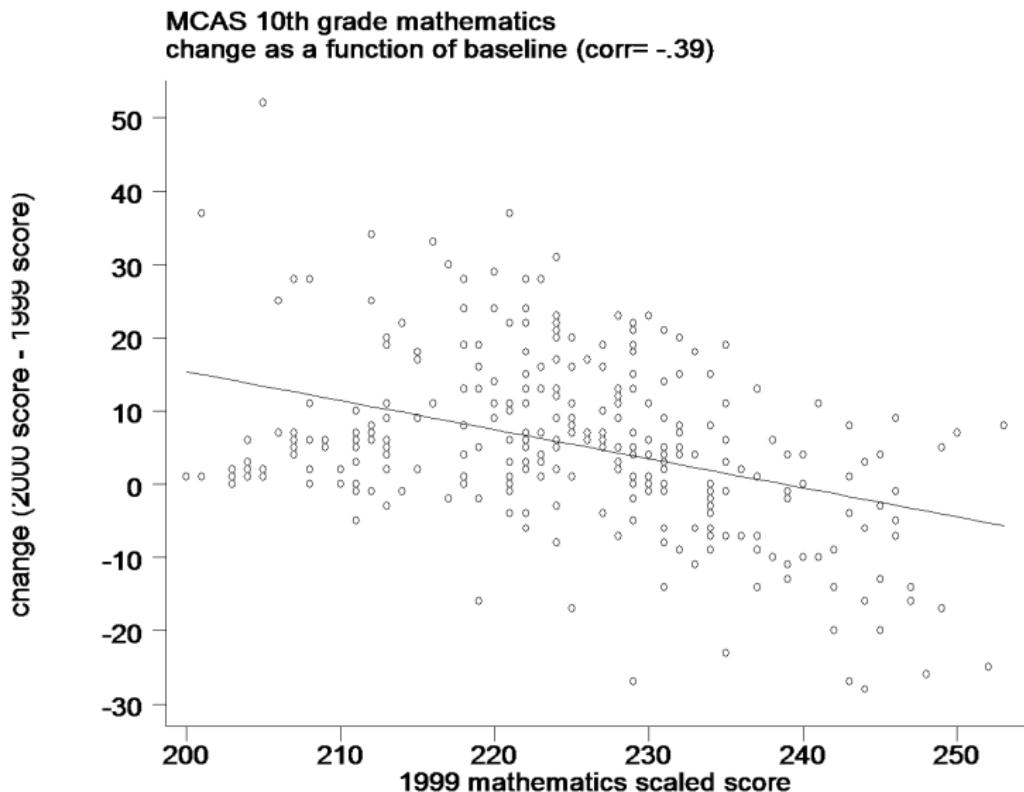
By Anand Vaishnav, Globe Staff, 1/10/2001

A state report rating public schools on three years of MCAS results shows some of Massachusetts's highest-performing schools "failed to meet" expectations, a description that angered many educators across the Commonwealth.

In the state's first attempt to give schools targets to reach on the Massachusetts Comprehensive Assessment System exams, 56 percent of 1,539 schools did not improve enough to meet those goals. The remaining 44 percent "exceeded, met, or approached" their growth targets, according to the report released yesterday.

The results largely bode well for urban districts that had the most ground to make up. But the state's top public schools - in wealthier districts such as Harvard, Newton, Wellesley - performed so well in 1998 that boosting scores any higher was nearly impossible. As a result, many of those schools received an "F," for "failed to meet" expectations.

Boston Globe.



$$COR(y_2, y_1) = 0.54; \quad COR(y_2 - y_1, y_1) = -0.39$$

$\text{VAR}(\text{observed change}) > \text{VAR}(\text{true change})$

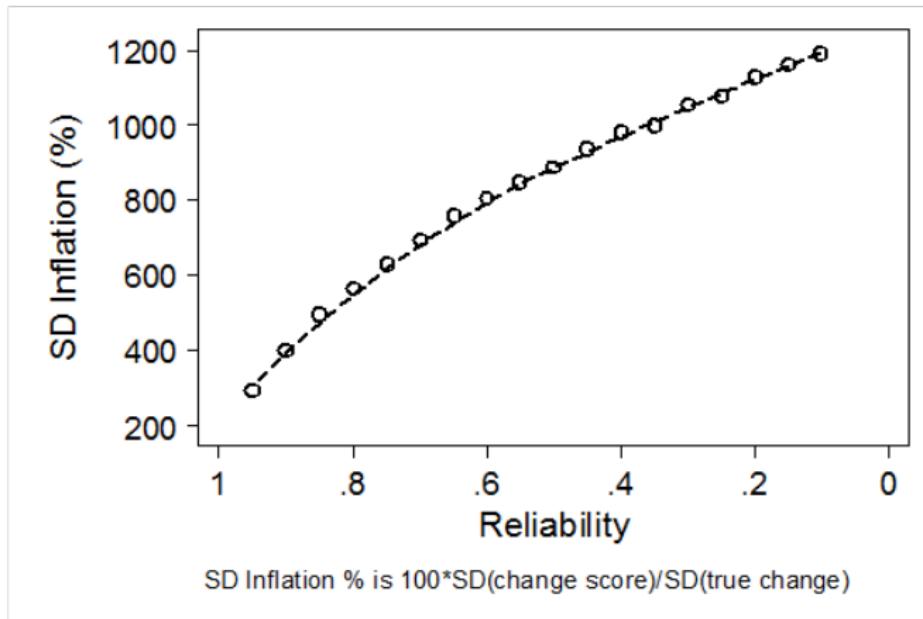
$$Y = T + e$$

$$Y_2 - Y_1 = (T_2 - T_1) + (e_2 - e_1)$$

$$\begin{aligned}\text{VAR}(Y_2 - Y_1) &= \text{VAR}(Y_2) + \text{VAR}(Y_1) - 2\text{COV}(Y_2, Y_1) \\ &= \text{VAR}(T_2) + \text{VAR}(T_1) - 2\text{COV}(T_2, T_1) + \\ &\quad \text{VAR}(e_2) + \text{VAR}(e_1) - 2\text{COV}(e_2, e_1)\end{aligned}$$

- Even if errors are uncorrelated, the variance of observed change score compounds measurement errors
- The lower the  $\text{COR}(Y_t, Y_{t-1})$ , the more variable is the change score
- Observed change score variance can be greater than baseline variance

$\text{VAR}(\text{observed change}) > \text{VAR}(\text{true change})$

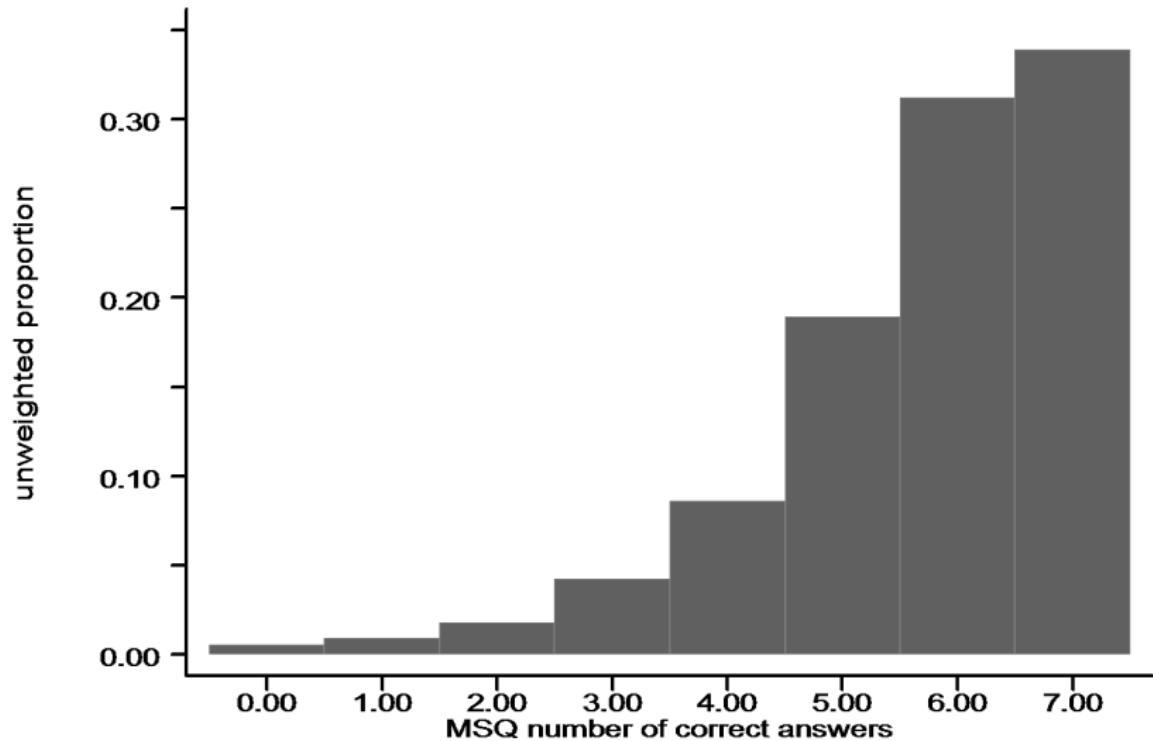


Change is random, additive, and correlated with baseline ( $r = -0.40$ )  
 $\text{SD}(\text{true change}) = 0.1\text{SD}(\text{baseline})$

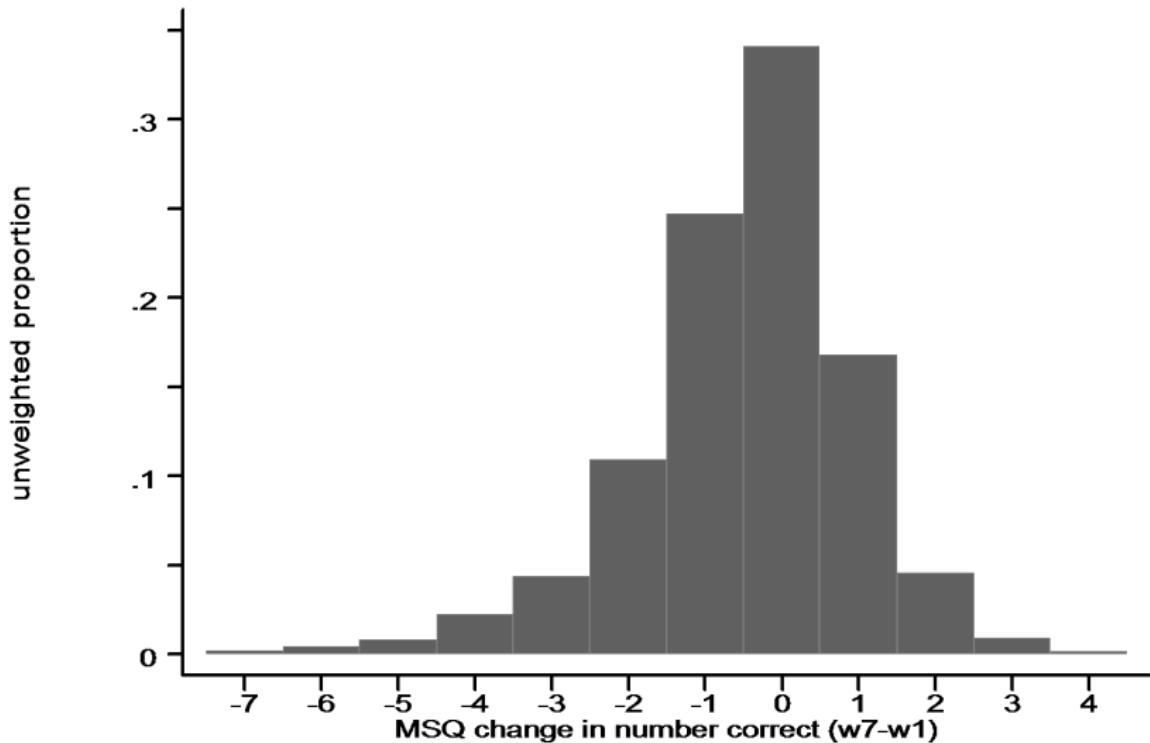
## Floor/Ceiling Effect

- Many measures in health research suffer floor/ceiling effects: a large proportion of the sample score at lowest or highest possible level
- For Example
  - ▶ Cognitive function
  - ▶ Depressive symptoms
  - ▶ Physical functioning

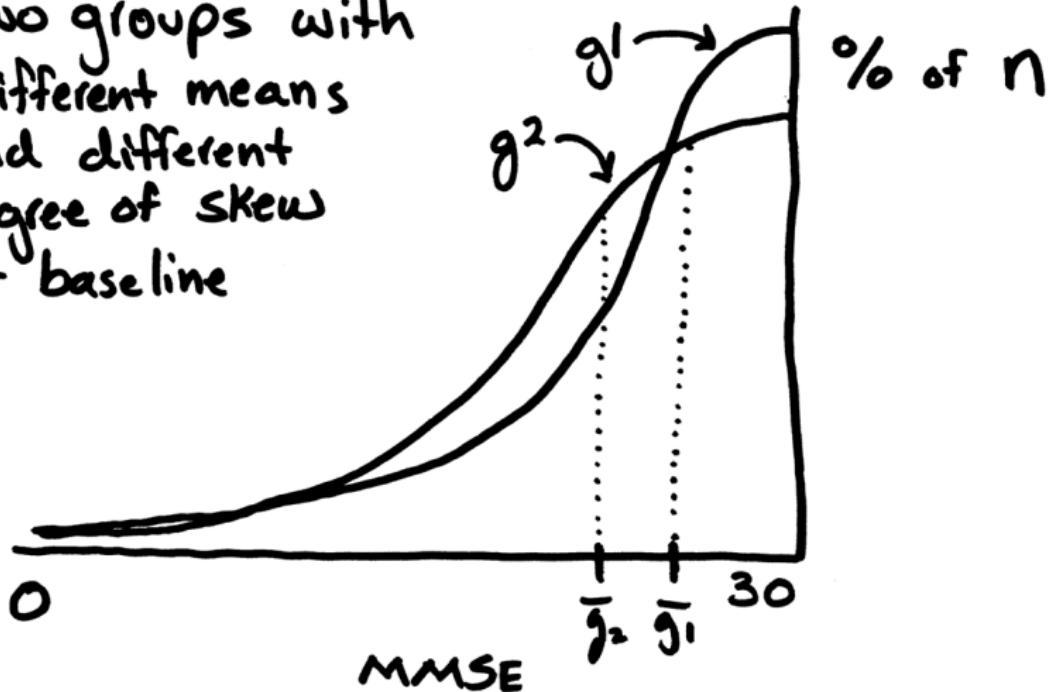
# Distribution of Cognitive Functioning in EPESE



# Distribution of Cognitive Functioning ( $w7-w1$ ) in EPESE

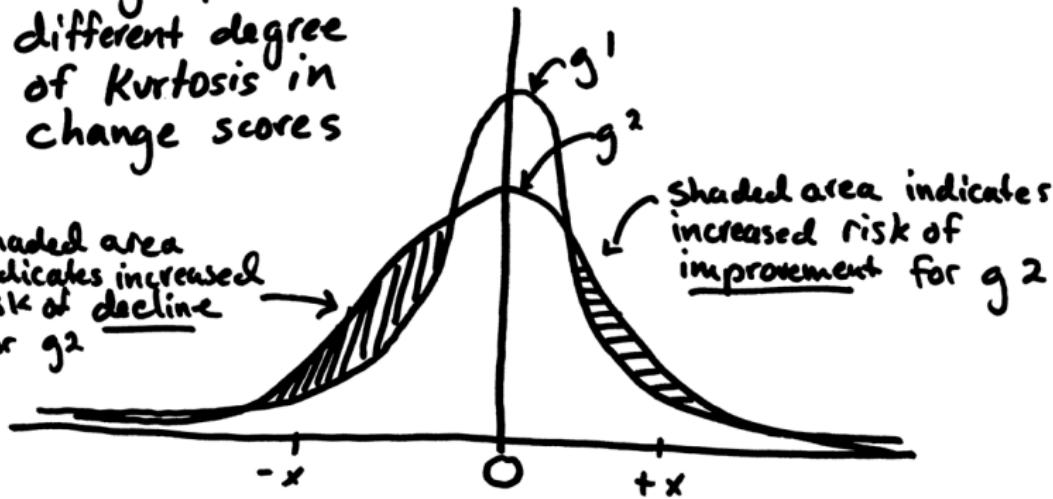


Two groups with different means and different degree of skew at baseline

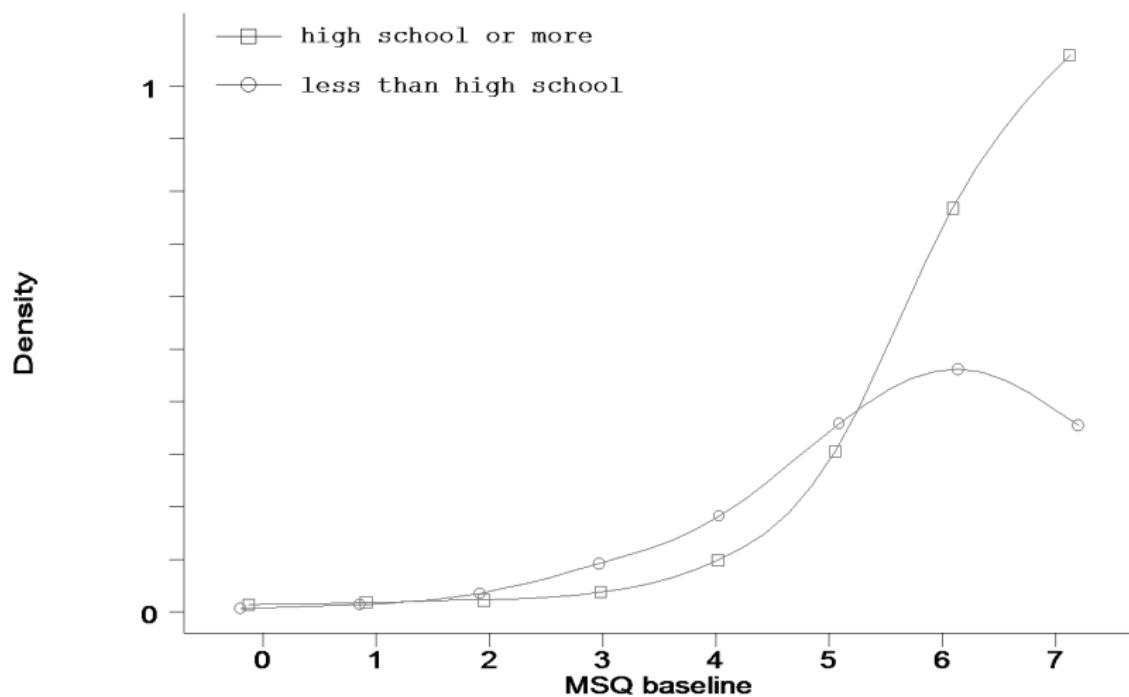


Two groups with different degree of kurtosis in change scores

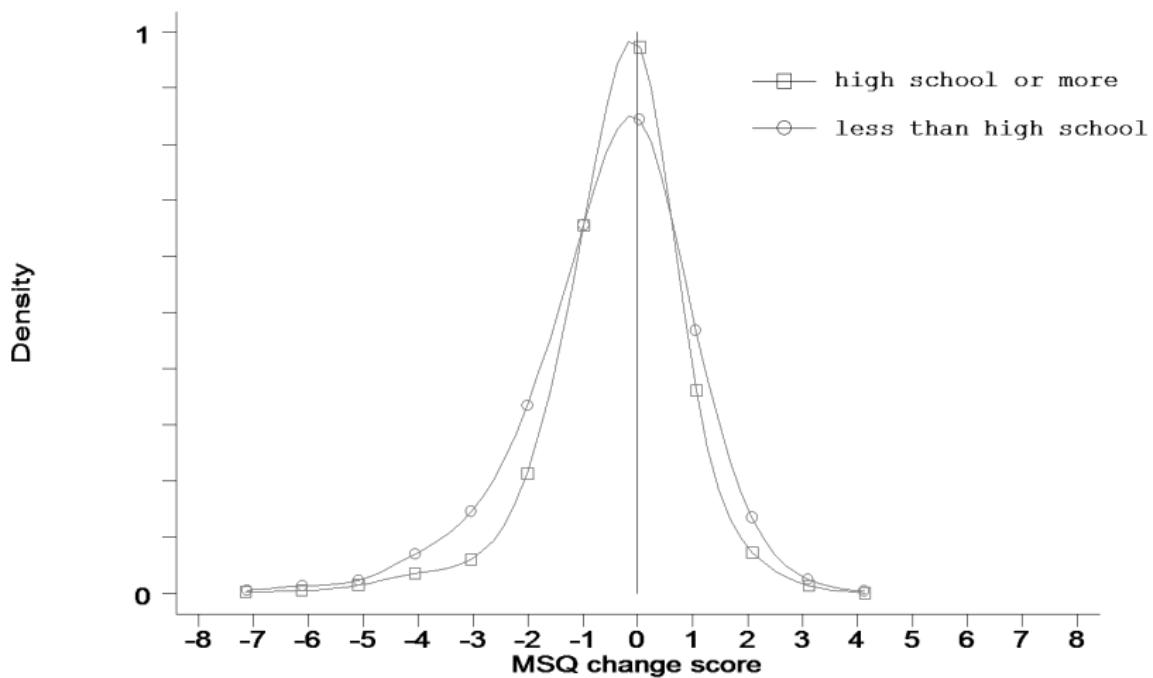
Shaded area indicates increased risk of decline for  $g_2$



Change  
( $t_2 - t_1$ )



Density plot. MSQ scores at year 1. EPESE (N=13,697) by education group

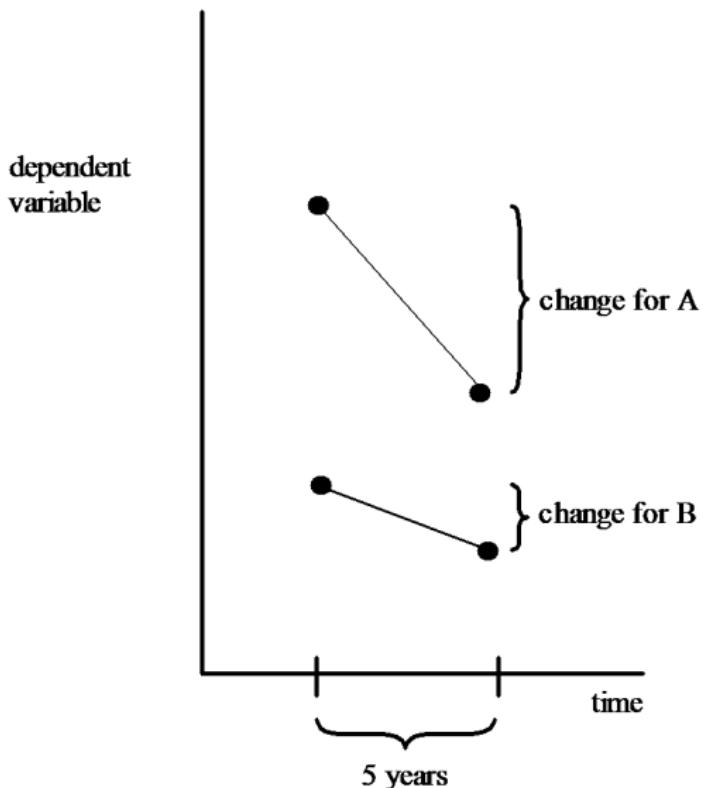


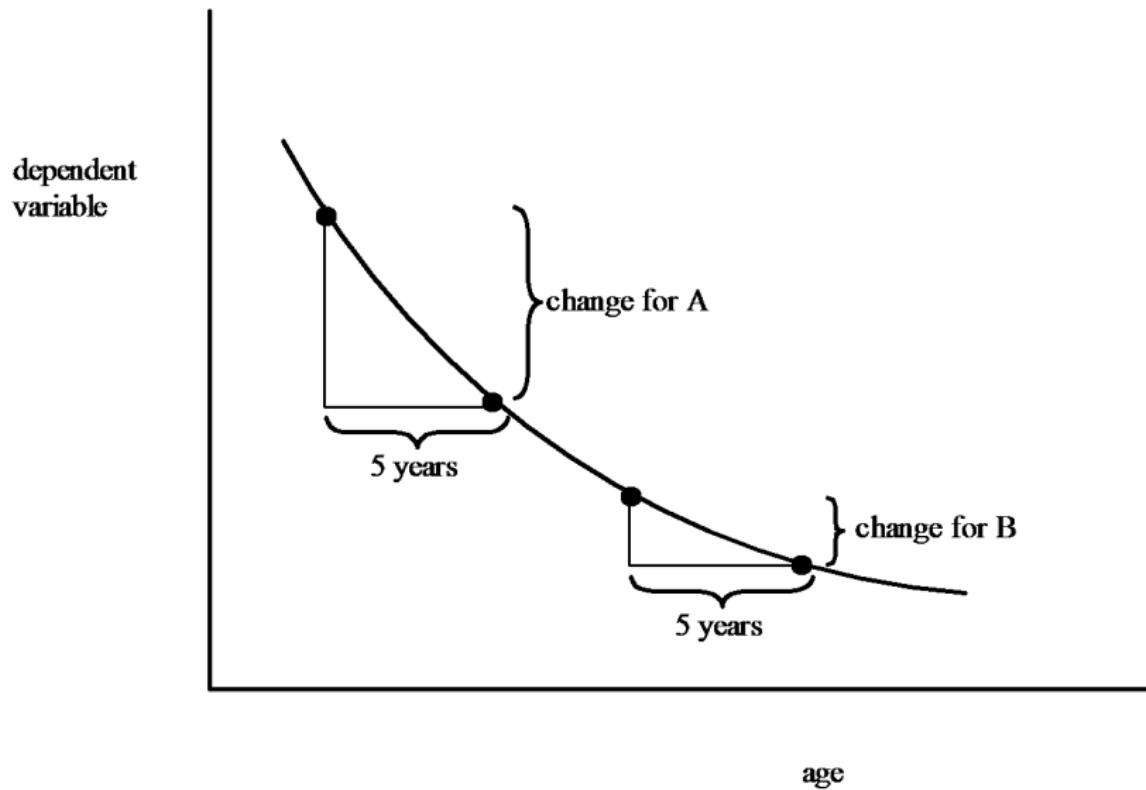
Density plot. MSQ change scores year 7 - year 1. EPESE  
(N=7,939) by education group

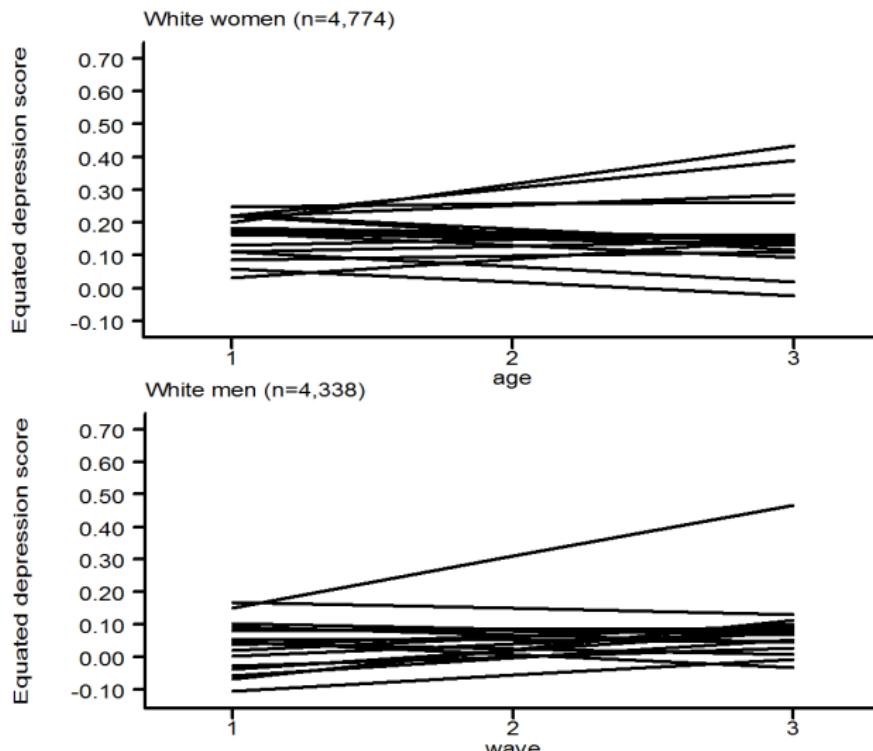
Table. Cognitive decline (2 or more points lower at follow-up) and cognitive improvement (2 or more points higher at 6th annual follow-up) among EPESE participants, by level of educational attainment.

Level of Education (years)	% OR
<b>Cognitive Decline</b>	
less than 12	22 1.7 (1.5, 2.0)
12 or more	14 1.0
<b>Cognitive Improvement</b>	
less than 12	6 1.7 (1.4, 2.1)
12 or more	4 1.0

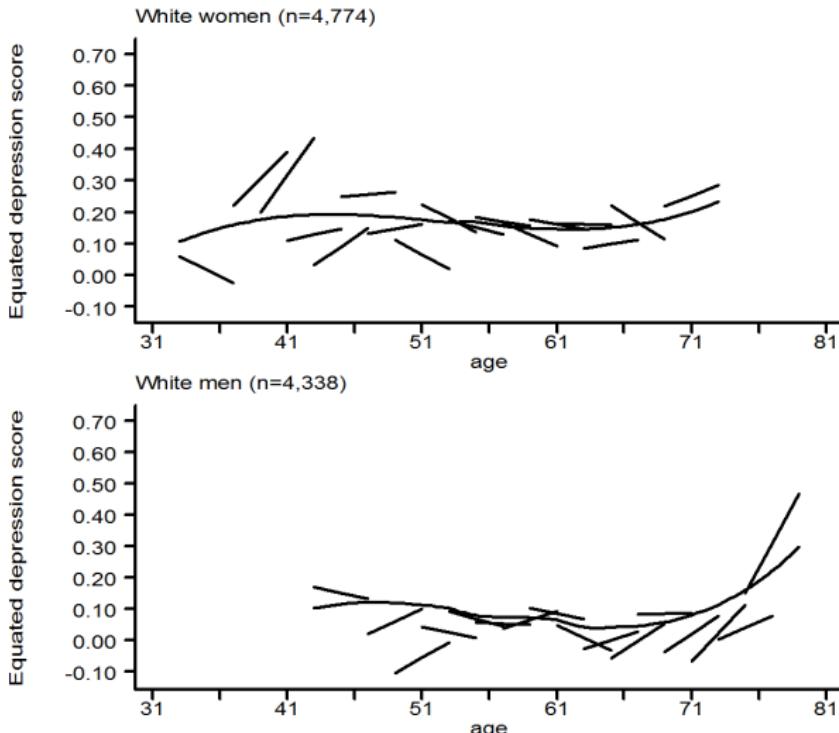
# Assumption of Linear Change







Birth cohort-specific depression symptom trajectories, Health and Retirement Study waves 1,2,3



Birth cohort-specific depression symptom trajectories, Health and Retirement Study waves 1,2,3

# An Approach

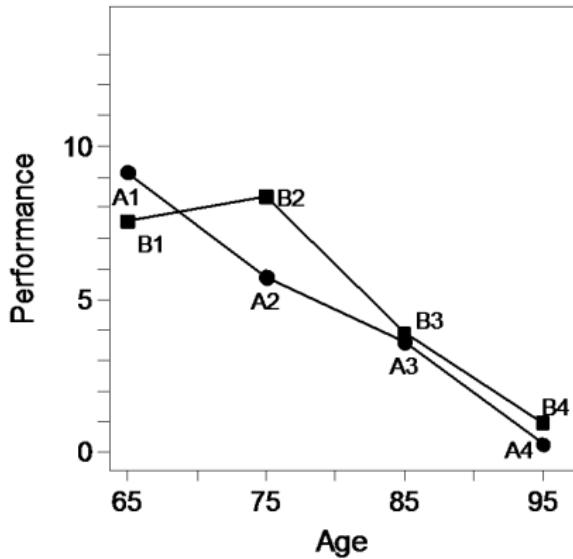
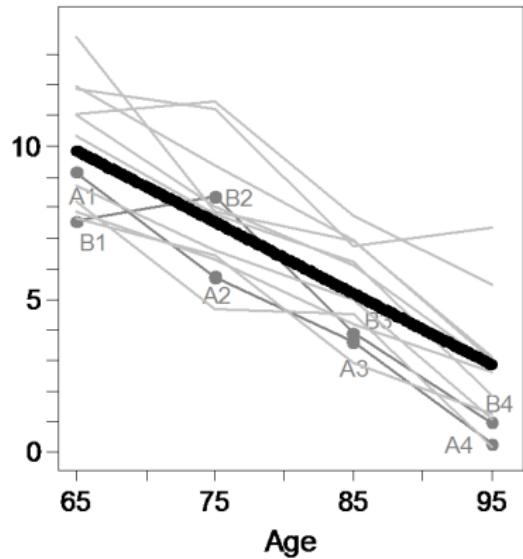
Model, Rather than Calculate, Change

# Latent Growth Curve Models

- Appropriate when outcome
  - ▶ changes systematically over time, on a time interval relevant to the study
  - ▶ when there at least three, preferably four repeated observations
  - ▶ is discrete or continuous
- Alternatives
  - ▶ Survival models (for "absorbing events": events that can only be observed once)
  - ▶ Latent change score models
  - ▶ Autoregressive models

# Latent Growth Curve Models

- Sometimes equivalent to
  - ▶ mixed effect models
  - ▶ random coefficient regression models
  - ▶ hierarchical or multilevel models
  - ▶ HLM, Mx, PROC Mixed, xtmixed
- Can be different
  - ▶ how time is handled (data vs. model parameter)
  - ▶ how variances are handled
  - ▶ how flexible the modeling framework is

**Figure 3a****Figure 3b**

Observed scores and an overall sample mean trajectory

Panel 3a shows two persons (A and B) and their performance at four ages (A1=person A at age 65, A4=person A at age 95). Panel B shows how person A and B and their performance trajectory are distributed in a larger sample of persons. Panel 3b also shows a group average trajectory (heavy line).

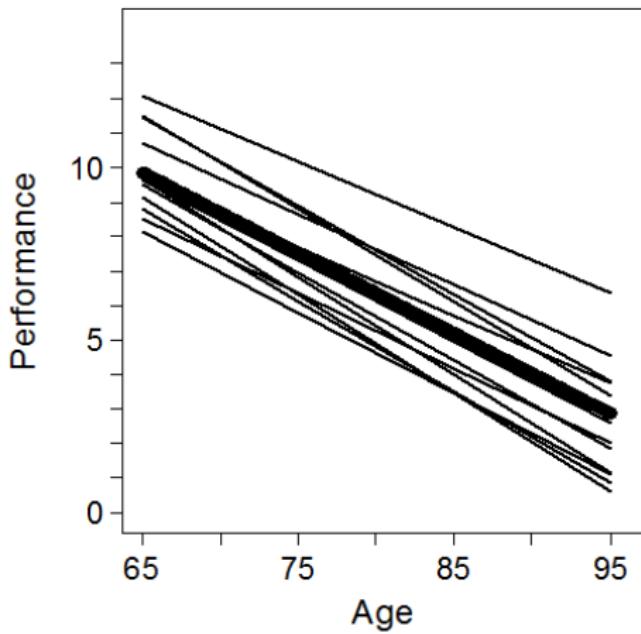
**Figure 4**

Figure 4

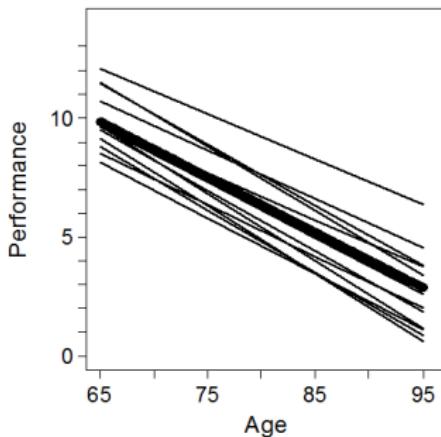
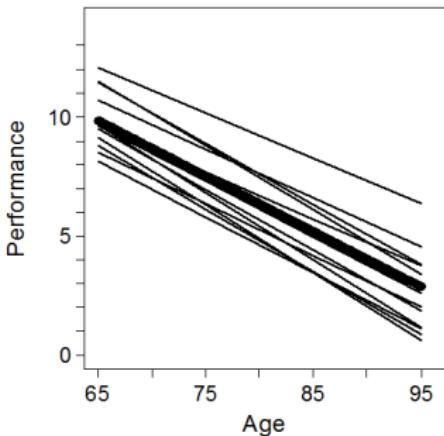

$$\text{Performance} = \text{Intercept} + \text{Slope} \times \text{Age} + \text{Residual}$$
$$\text{Intercept} = \text{Mean}_{65} + \text{Disturbance}$$
$$\text{Slope} = \text{Mean Change} + \text{Disturbance}$$

Figure 4

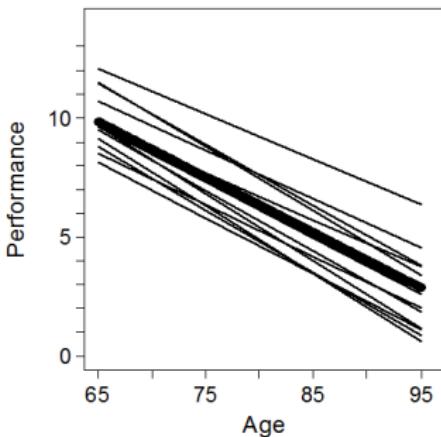


$$y_{it} = \eta_{i1} + \eta_{i2} \times \text{Age}_{it} + \epsilon_{it}$$

$$\eta_{i1} = \alpha_1 + \zeta_{i1}$$

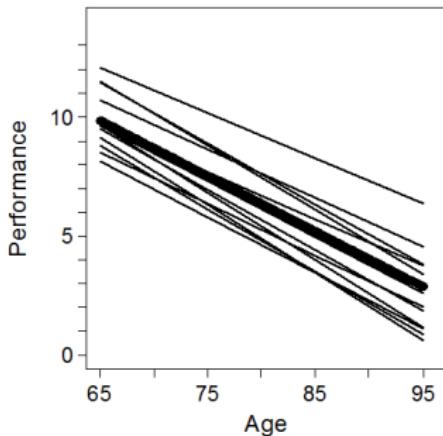
$$\eta_{i2} = \alpha_2 + \zeta_{i2}$$

Figure 4



$$y_{it} = (\alpha_1 + \zeta_{i1}) + (\alpha_2 + \zeta_{i2}) \times \text{Age}_{it} + \epsilon_{it}$$

Figure 4



$$\begin{aligned}y_{it} &= (\alpha_1 + \alpha_2 \times \text{Age}_{it}) + (\zeta_{i1} + \zeta_{i2} \times \text{Age}_{it} + \epsilon_{it}) \\y_{it} &= (\text{fixed part}) + (\text{random part})\end{aligned}$$

# Time

- Time as data ( $Age_{it}$ )
  - ▶ Random effects models, Mixed effect models, Multilevel/hierarchical models
  - ▶ Accommodates irregular assessment schedule
  - ▶ Time varying covariate effects vary over people
- Time as model parameter ( $Age_t$ )
  - ▶ LGCM
  - ▶ Assumes regular reassessments
    - ★ not necessarily evenly spaced
  - ▶ Allows flexible curve shape
  - ▶ Time varying covariate effects vary over time
- Mplus will do both

# Summary

- Change scores have issues
- Modeling doesn't remove those issues but can accommodate (estimate) them
- There are many paths to enlightenment (hierarchical, multilevel, mixed, random, latent growth)
- Different approaches can be equivalent, mostly depending upon how time is handled

Don't get too hung up on the validity/invalidity of one model over another

- LGCM  $\approx$  Mixed/Random Hierarchical
  - ▶ Curran PJ: Have Multilevel Models Been Structural Equation Models All Along? *Multivariate Behavioral Research* 38:529-569, 2003
- LGCM  $\approx$  paired t-tests, RM-ANOVA
  - ▶ Voelkle MC: Latent growth curve modeling as an integrative approach to the analysis of change. *Psychology Science* 49:375, 2007
- But, mixed, random, hierarchical, change score, repeated measures ANOVA, all can be viewed as special cases of SEM models of change
- So SEM model is a flexible, extensible and general approach to change and worth mastering.

# Questions

# Structural Equation Modeling

Communicating and expressing modeling ideas  
using latent variables

# Learning Note

- Gist
  - ▶ What am I doing?
  - ▶ How do I explain it to my colleagues?
- Statistical concepts
  - ▶ What am I really doing?
  - ▶ How do I try to explain it to my statistician colleagues?
- Person-machine interface
  - ▶ How do I do this?

# Covariance Structure Model

		(Continuous) Latent Variables	
		Yes	No
Regressions among dependent (or involving latent variables)	Yes	Structural Equation Modeling	Path Analysis
	No	Factor Analy- sis	Multivariate Regression

# How to Learn about SEM

## How

- Brown T. Confirmatory factor analysis for applied research. (2006). Guilford Press
- Jöreskog and Sörbom, LISREL user's reference guide. (from version 7)
- Raykov and Marcoulides, A first course in structural equation modeling. 2006
- Raykov T, Tomer A, Nesselroade JR. Reporting structural equation modeling results in Psychology and Aging: Some proposed guidelines. *Psychol Aging.* 1991;6:499-503.
- Resources to help you use Mplus. UCLA: Institute for Digital Research & Education (was: Academic Technology Services, Statistical Consulting Group).  
<https://stats.idre.ucla.edu/mplus/>

## Why

- De Stavola B et al. Statistical issues in life course epidemiology. *Am J Epidemiol.* 2006;163:84-96.
- Hoyle RH, Smith GT. Formulating clinical research hypotheses as structural equation models. *J Consult Clin Psychol.* 1994;62:429-440.
- Muthén L B. Psychometric modeling in epidemiology with an emphasis on longitudinal analysis of alcohol use. Alcoholic Beverage Medical Research Foundation. 1994.

## Latent Variable Models

Latent variables are variables that are not directly observed

- Presumed to exist
- Presence inferred from observed variables
- Account for covariation among observed variables

Latent and observed variables can be continuous or categorical

# Ridiculously Brief and Highly Selected History of LV Models

- Spearman (1904) - Factor Analysis
- Wright (1934) - Path Analysis
- Lazarsfeld (1950) - Latent Class
- Lord (1952) - Item Response Theory
- Jöreskog (1973) - LISREL model
- Muthén (1987) - LISCOMP model
- Muthén (1998) - General LVM (Mplus)
- Asparouhov + Muthén (2008) - Exploratory SEM
- Rosseel (2010) - R/lavaan

# Other Latent Variables

- Measurement errors
- Random effects
- Frailties, liabilities
- Variance components
- Missing data

see [www.statmodel.com/download/Topic 3.pdf](http://www.statmodel.com/download/Topic 3.pdf)

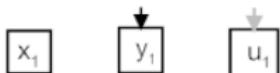
## Path diagrams

Path analysis was invented by the geneticist Sewall Wright in 1918. Wright also made use of path diagrams. However, the diagrams used here follow a set of rules [loosely] for drawing path diagrams that has been described as the LISREL model (specifically, LISREL submodel 2). LISREL was developed in the 1960s and released widely in the 1970s by Karl G. Joreskog.

Challenges with path diagrams in LDA is that conventions for describing means and thresholds of observed (and latent) variables are not well developed. This ows to the historical fact that most early applications of covariance structure analysis (path analysis, factor analysis) concern the covariances among observed data, and variables were often standardized or centered (to mean 0).

The most widely used convention for describing threshold and mean structures is McArdle's RAM notation (Reticular Action Model).

Finally, it is worth mentioning that no specific knowledge of or ability to use and interpret path diagrams is required for using Mplus.

Observed variables in rectangles

Exogenous variables without residuals: covariate, cause

Endogenous variables ( $y, u$ ) usually with residuals (errors): item, indicator, outcome

Sometimes  $y$  implies continuous,  $u$  implies discrete

Latent variables in ovals

Sometimes  $f$  and  $\eta$  imply continuous latent variables,  $c$  categorical

Continuous latent variables with disturbance terms

factors, random effects, classes, clusters, mixtures



## Constant (=1)

Used to indicate means, thresholds, intercepts in regression relationships. Several other conventions are used for this purpose.



Straight arrows indicate regression relationships including regression and factor loadings. Sometimes labeled with a regression coefficient ( $b, \beta, \gamma, \kappa, \lambda$ ; \*; \* implies freely estimated). If not labeled and others in model are, might be constrained to be equal to 1. If labeled, subscripts in the order to-from ( $b_{yx}$ ).



Curved arrows indicate covariance (correlation) relationships.

Often omitted from among exogenous variables, even when included in the model.

## Notes:

There is no formal and universally accepted language or syntax rules for path diagrams.

Different authors have their own conventions.

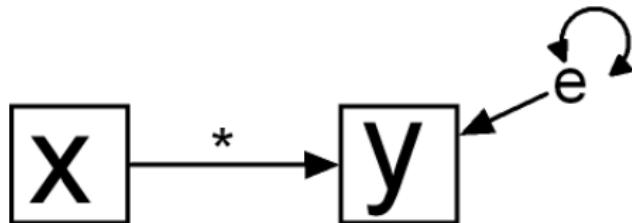
Path diagrams in publications are rarely complete images of the statistical model.

Path diagrams do not always translate directly into the syntax of an estimation program.

Path diagrams can be helpful tools. They can also be confusing and distracting.

Path diagrams as used in SEM are in general not the same as directed acyclic graphs (DAGs).

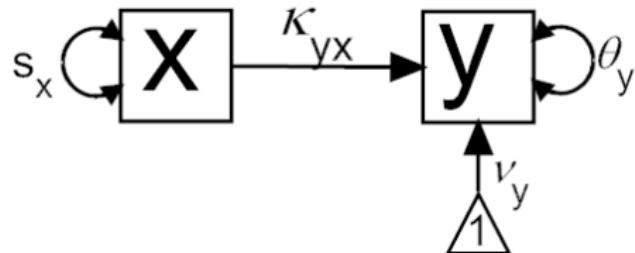
# Simple Regression



$$y = \nu + kx + \epsilon$$

$$E(y) = \nu + kx$$

# Simple Regression

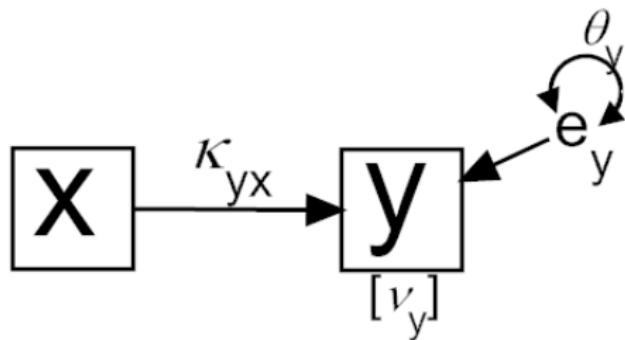


$$y = \nu + kx + \epsilon$$

$$E(y) = \nu + kx$$

RAM path diagram conventions described (James Steiger, Vanderbilt University)

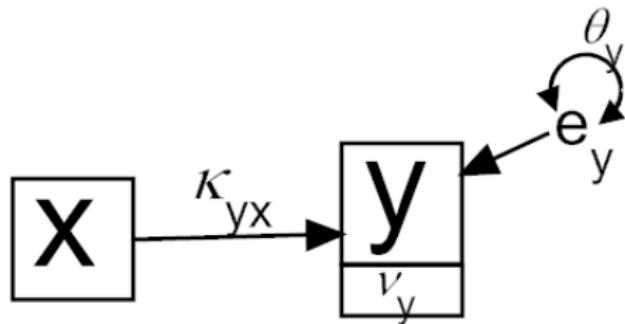
# Simple Regression



$$y = \nu + kx + \epsilon$$

$$\mathbb{E}(y) = \nu + kx$$

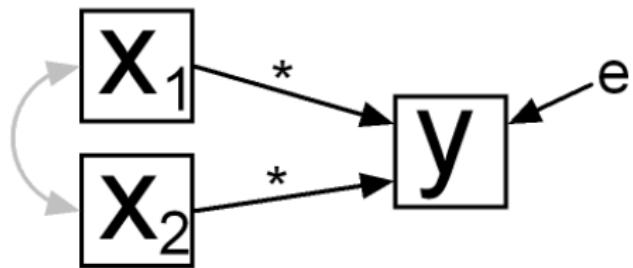
# Simple Regression



$$y = \nu + kx + \epsilon$$

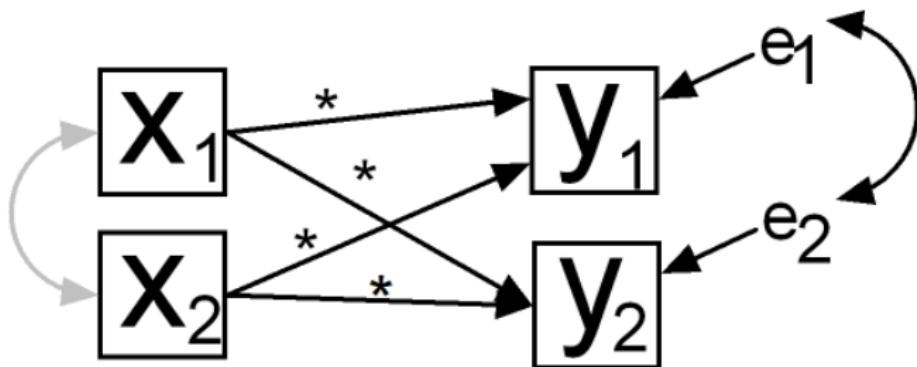
$$E(y) = \nu + kx$$

# Multivariable Regression



$$y = \nu + k_1x_1 + k_2x_2 + \epsilon$$

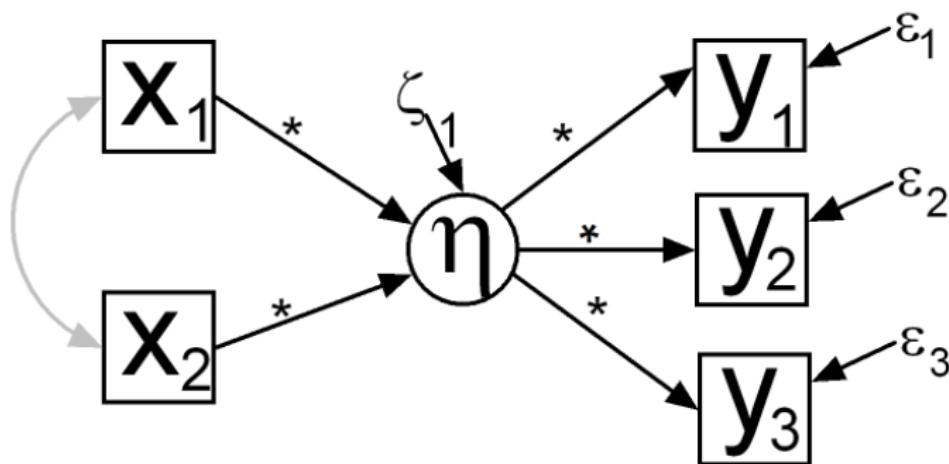
# Multivariate Regression



$$y = Kx + \epsilon$$

## (Confirmatory) Factor Analysis with Covariates

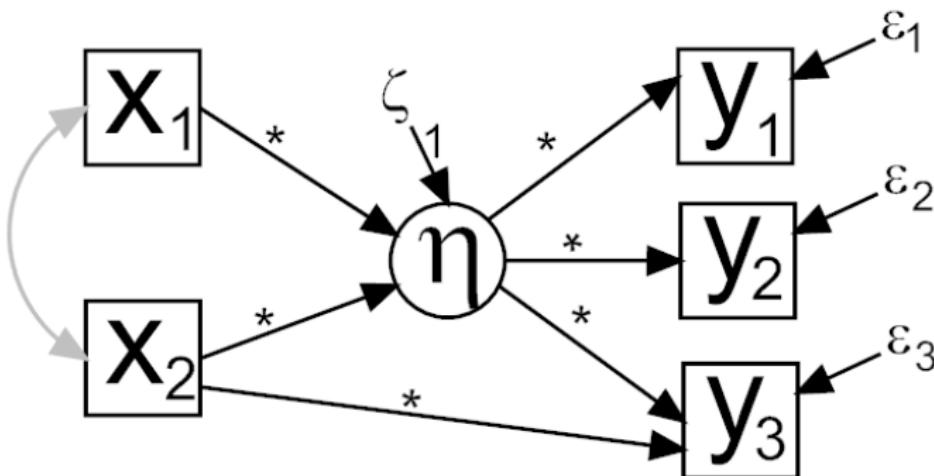
(MIMIC = Multiple Indicators Multiple Causes)



$$y = \nu + \Lambda\eta + Kx + \epsilon$$

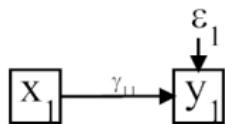
$$\eta = \alpha + \Gamma x + \zeta$$

## MIMIC Model with Direct Effect

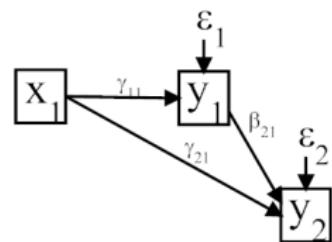


$$y = \nu + \Lambda\eta + Kx + \epsilon$$

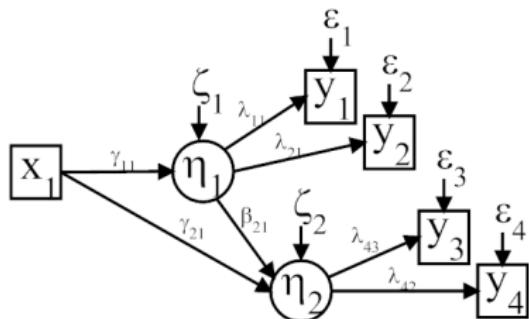
$$\eta = \alpha + \Gamma x + \zeta$$



regression

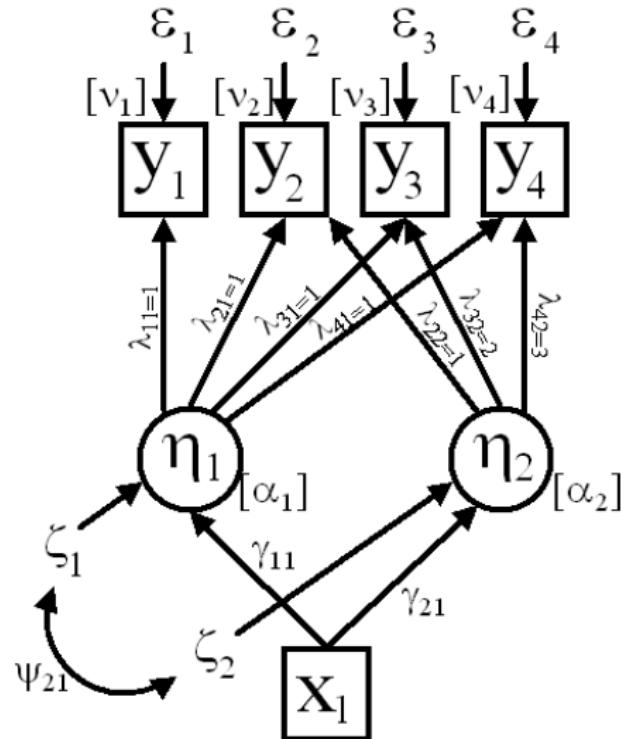


path analysis



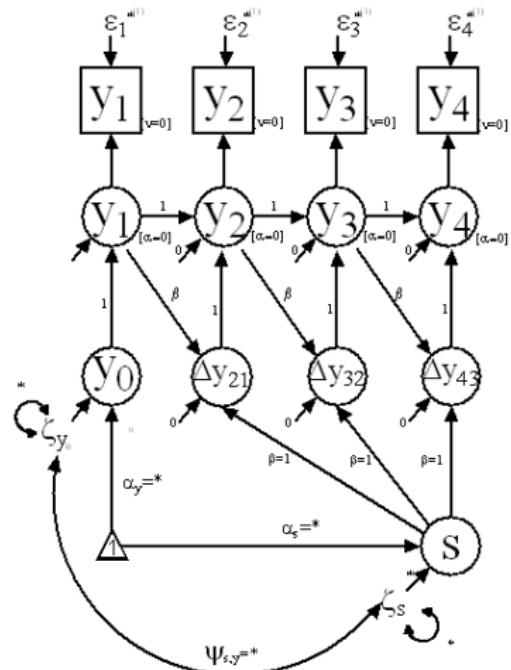
Structural Equation Model

# Latent Growth Model (Linear Change)



$$\begin{aligned}
 y_{it} &= \nu + \eta_1 \lambda_{t1} + \eta_2 \lambda_{t2} + \epsilon_{it} \\
 \eta_1 &= \alpha_1 + \gamma_{11} x_{1i} + \zeta_1 \\
 \eta_2 &= \alpha_2 + \gamma_{21} x_{1i} + \zeta_2
 \end{aligned}$$

# Dual Change Score Model



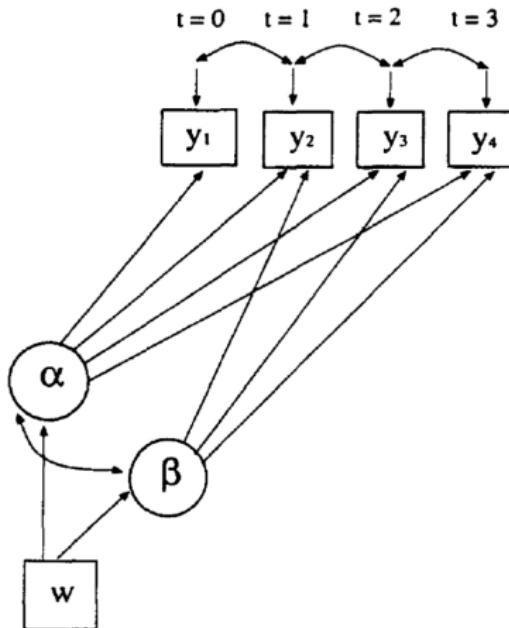
$$y_{it} = y_0 + \Delta y_{t(t-1)} + \epsilon$$

$$\Delta y_{t(t-1)} = \beta y_{(t-1)} + s$$

$$y_0 = \alpha_y + \zeta_{y_0}$$

$$s = \alpha_s + \zeta_s$$

**FIGURE 3**  
Graphical representation of a growth model for four time points.



Muthén BO, Khoo ST. Longitudinal studies of achievement growth using latent variable modeling. Learning and individual differences. 1998;10(2):73-101.

## The Basic Structure of Growth Models 455

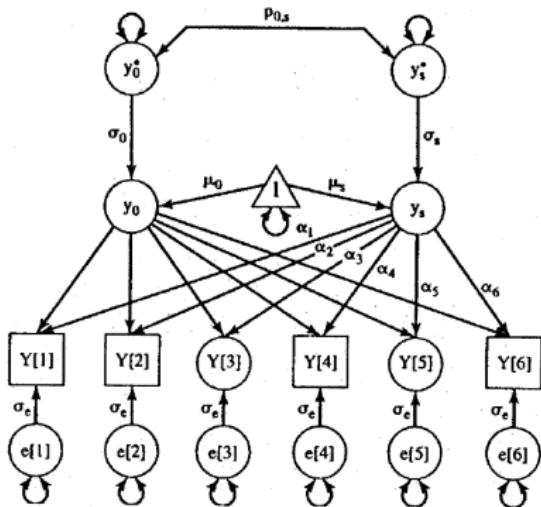


Figure 18.7 The basic latent growth structural model as a path diagram from McArdle & Epstein (1987) and McArdle & Hamagami (1992).

McArdle J, Nesselroade J. Growth curve analysis in contemporary psychological research. In: Schinka JA, Velicer W, eds. Comprehensive Handbook of Psychology, Volume Two: Research Methods in Psychology. New York: Wiley; 2002.

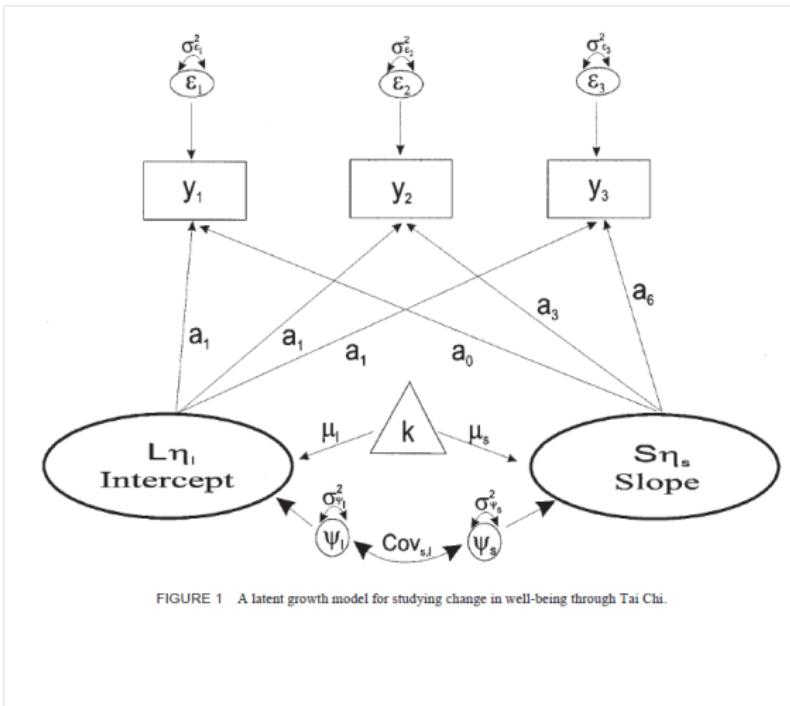
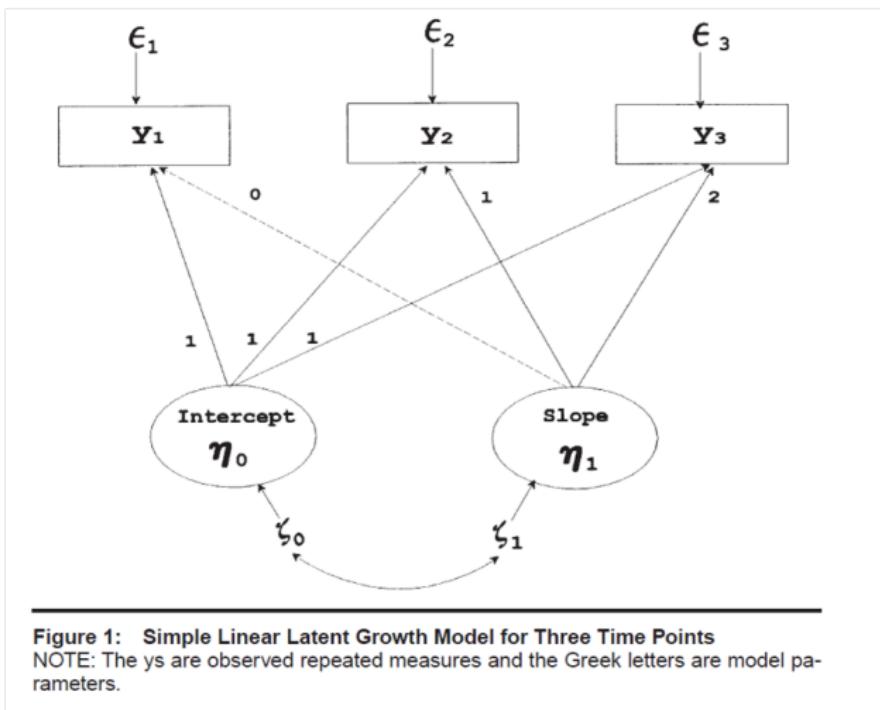


FIGURE 1 A latent growth model for studying change in well-being through Tai Chi.

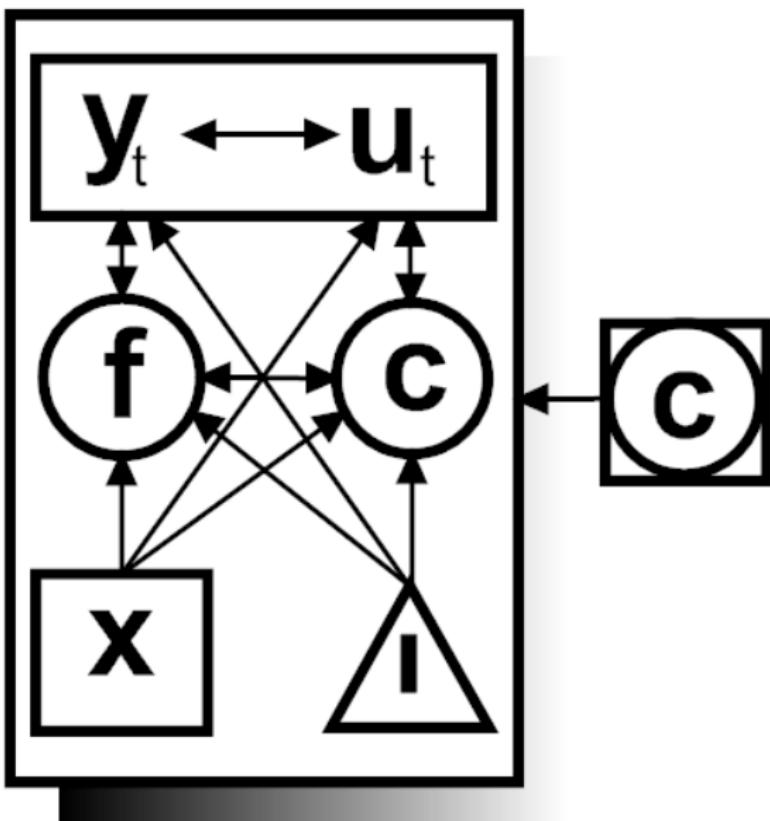
Li et al. Enhancing the psychological well-being of elderly individuals through Tai Chi exercise: A latent growth curve analysis. Structural Equation Modeling. January 2001;8(1):53-83.



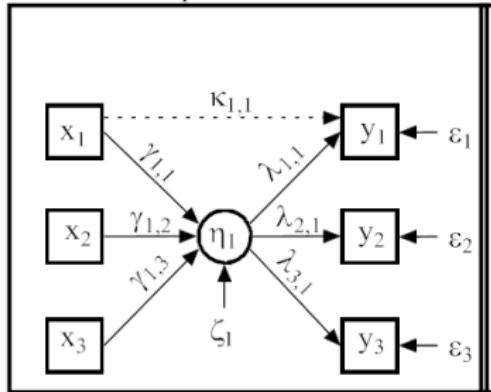
**Figure 1: Simple Linear Latent Growth Model for Three Time Points**  
 NOTE: The ys are observed repeated measures and the Greek letters are model parameters.

Wang et al. Evaluation of Hiv Risk Reduction Intervention Programs Via Latent Growth Model.  
 Eval Rev 1999 23: 648-662

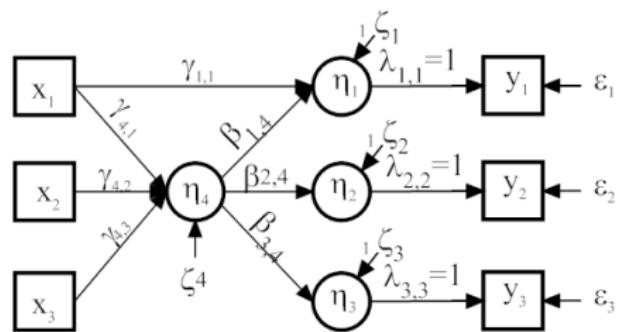
# Mplus Framework



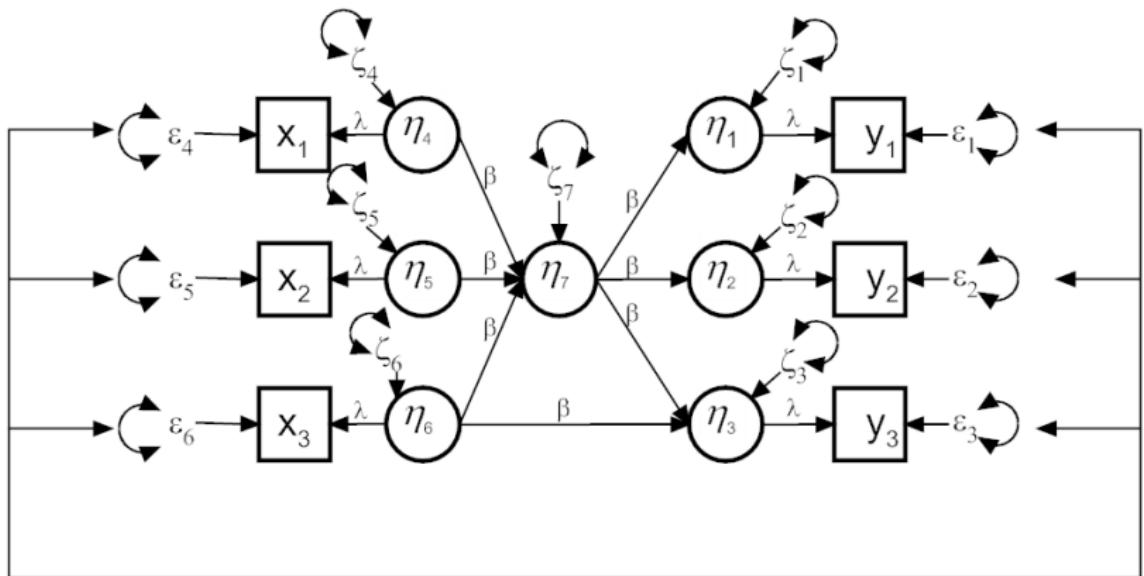
Mplus Model



LISCOMP Model



## Mplus Model, All Y continuous



# Mplus and LISCOMP Parameter Arrays

## Mplus Variable and Parameter Arrays

$K$  (kappa) a  $p \times q$  matrix of slopes for the regressions of the  $p$  observed response variables ( $y$ 's) on the  $q$   $x$ -variables

## LISCOMP Variable and Parameter Arrays

- $x$  (exogenous variables) a  $q$ -vector, i.e.,  $x_1, x_2, x_3 \dots x_q$ ,
- $y$  (observed response variables) a  $p$ -vector i.e.,  $y_1, y_2 \dots y_p$
- $\tau$  (tau) a vector of length ( $c-1$ ), equal to the number of thresholds; where  $c$  is number of categories
- $v$  (nu) a  $p \times 1$  vector of measurement intercepts, where  $p$  is the number of response variables ( $y_1, y_2 \dots y_p$ )
- $\Lambda$  (lambda) a  $p \times m$  matrix of measurement slopes/loadings, where  $p$  is equal to the number of  $y$ 's, and  $m$  is equal to the number of latent constructs ( $\eta$ 's).
- $I$  (identity matrix), a  $m \times m$  identity matrix
- $\Theta$  (theta) a  $p \times p$  covariance matrix for the residuals in the measurement relations
- $\alpha$  (alpha) a  $m \times 1$  vector of the structural regression latent variable ( $\eta$ 's) means and intercepts. *May be fixed to zero in the for categorical variables.*
- $B$  (beta) a  $m \times m$  matrix of slopes for the regressions among the latent variables ( $\eta$ 's) with zeros as it's diagonal elements.
- $\Gamma$  (gamma) a  $m \times q$  matrix of slopes for the regressions of the  $m$  latent variables ( $\eta$ 's) on the  $q$   $x$ -variables
- $\Psi$  (psi) a  $m \times m$  covariance matrix for the latent variables ( $\eta$ 's) and the residuals in the latent variable relations
- $\Delta$  (delta) diagonal of a  $p \times p$  matrix of scaling factors. Only the  $p$  diagonal elements are parameters.

## Measurement Model

$$y^* = v + \Lambda \eta + Kx + \varepsilon$$

## Structural Regression model:

$$\eta = \alpha + B\eta + \Gamma x + \zeta$$

## Measurement Model

$$y^* = v + \Lambda \eta + \varepsilon$$

## Structural Regression model:

$$\eta = \alpha + B\eta + \Gamma x + \zeta$$

[Link to PDF handout version of this slide](#)

$$y = \nu + \Lambda \eta + Kx + \epsilon$$

$$\eta = \alpha + B\eta + \Gamma x + \zeta$$

$$V(y) = \Lambda \Psi \Lambda^T + \Theta$$

$$V(\epsilon) = \Theta$$

$$V(\zeta) = \Psi$$

$$V(\eta) = \ddot{B}^{-1} (\Gamma \Phi \Gamma^T + \Psi) \ddot{B}^{T-1}$$

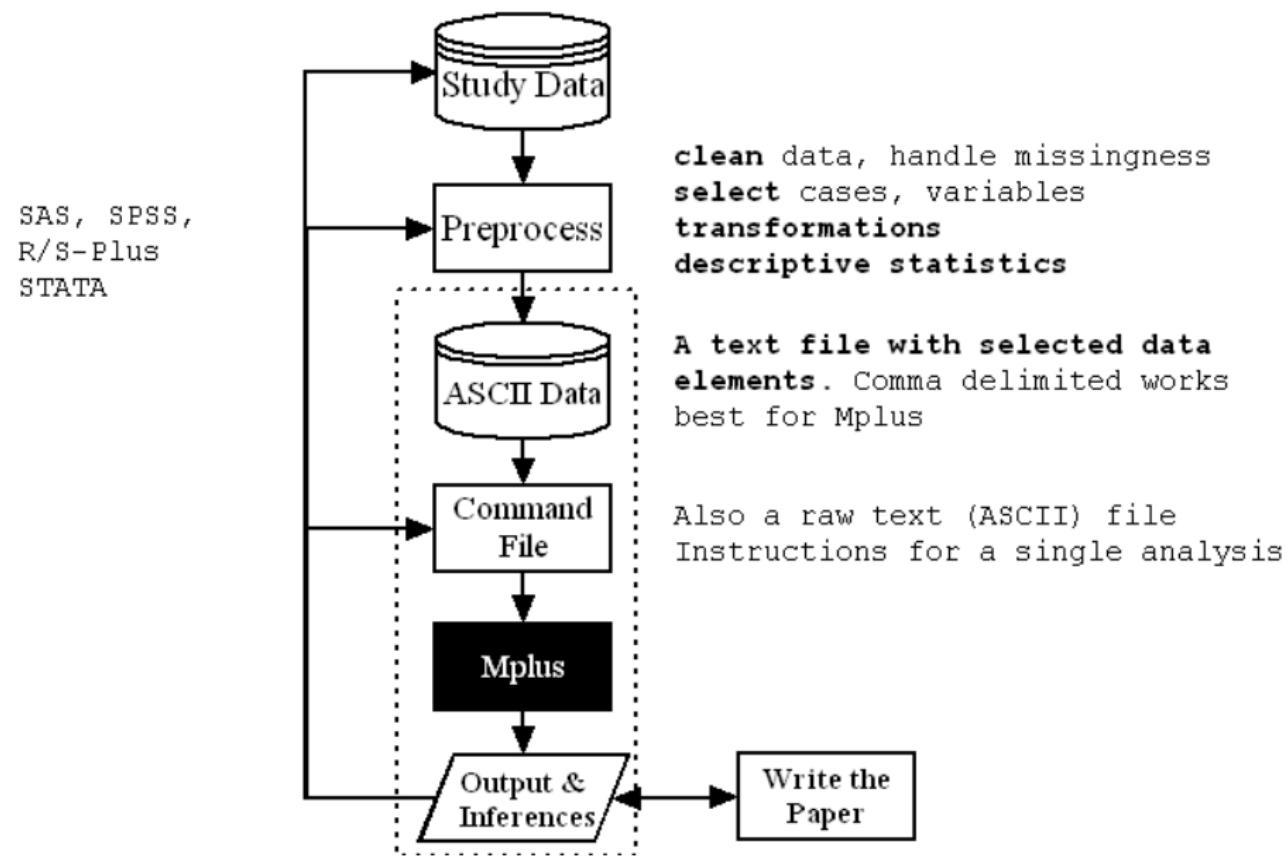
where

$$\ddot{B} = (I - B)$$

$$\Phi = COV(x)$$

# Person-Machine Interface

How do I do this?



# How to Write a Mplus Command File

- Read the Users Manual
  - ▶ (at least once all the way through)
- Keep it close to your computer
- Find a similar example
  - ▶ (Mplus provides input programs and data for all examples)
- Hack the example to suit your problem

# Mplus Commands

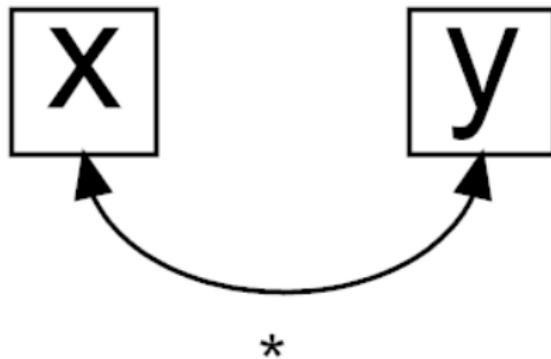
- Title
- Data
- Variable
- Analysis
- Model
- Output
- Define
- Plot
- Savedata
- Montecarlo

# How to Write a Mplus Command File

- Read the Users Manual
  - ▶ (at least once all the way through)
- Keep it close to your computer
- Find a similar example
  - ▶ (Mplus provides input programs and data for all examples)
- Hack the example to suit your problem

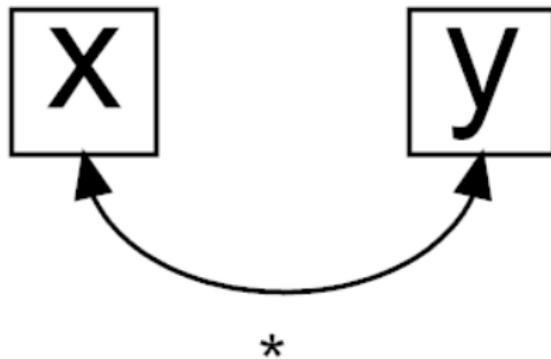
# Understanding Mplus Syntax

# Correlation



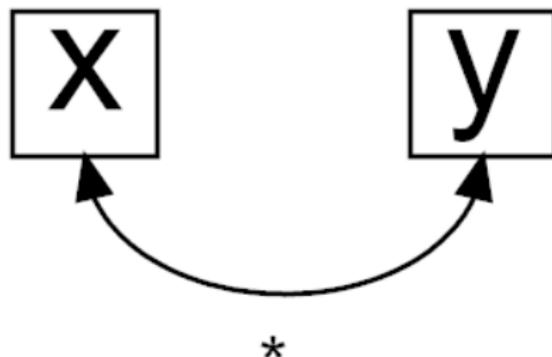
MODEL: x with y ;

# Covariance



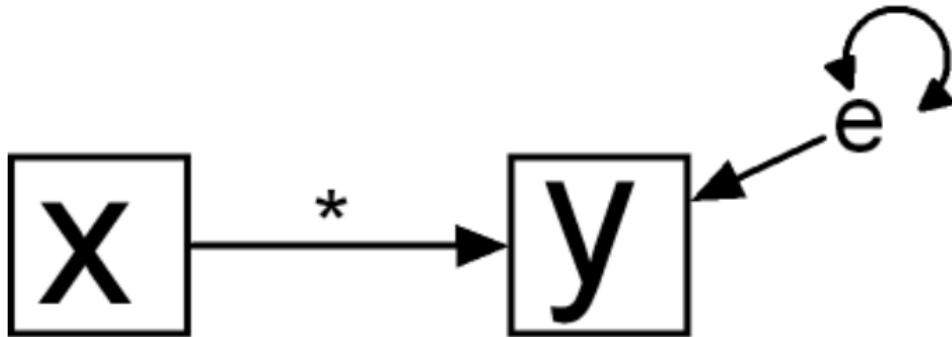
```
TITLE: A Covariance
DATA: FILE = temp.dat ;
VARIABLE: NAMES = x y ;
MODEL: x with y ;
```

# Correlation



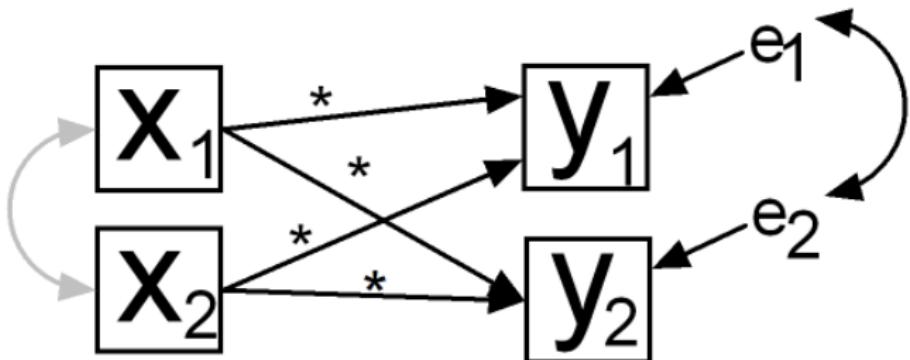
```
TITLE:      A Covariance
DATA:       FILE = temp.dat ;
VARIABLE:   NAMES = x y ;
OUTPUT:     STANDARDIZED ;
MODEL:      x with y ;
```

# Simple Regression



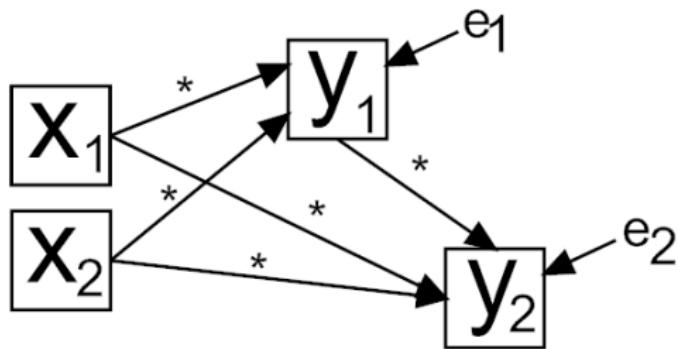
```
MODEL: y on x * ;  
        y * ; ! error (residual) term  
[y*] ; ! and intercept are  
        ! estimated by default  
! so "MODEL: y on x* ;" would be  
! sufficient
```

# Multivariate Regression



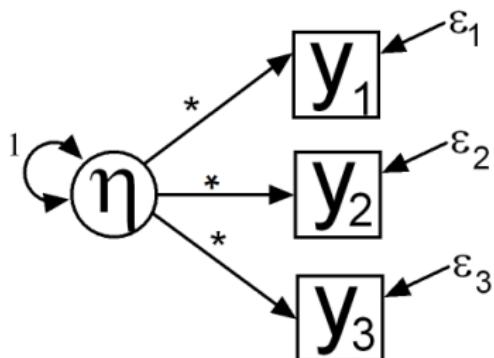
```
MODEL: y1 y2 on x1 * x2 * ;  
! default intercepts estimated -> [y1 * y2*] ;  
! default resid. var. estimated -> y1 * y2 * ;  
! default resid. cov. estimated -> y2 with y1 * ;  
! default covariance of x's -> x1 with x2 * ;
```

# Path Analysis



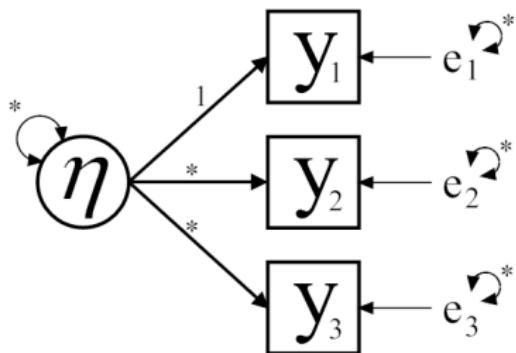
```
MODEL:  y1 y2 on x1 x2 * ;  
        y2 on y1 * ;
```

# Confirmatory Factor Analysis



```
TITLE: Confirmatory factor analysis  
Alternative specification  
DATA: FILE = trash.dat ;  
VARIABLE: NAMES = y1-y3 ;  
OUTPUT: STANDARDIZED ;  
MODEL: eta by y1-y3* ;  
eta@1 ;
```

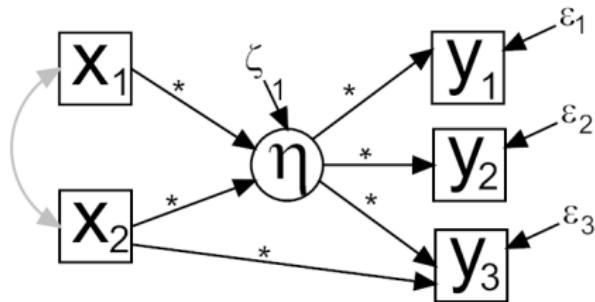
# Confirmatory Factor Analysis



```
TITLE: Confirmatory factor analysis  
Default specification  
DATA: FILE = trash.dat ;  
VARIABLE: NAMES = y1-y3 ;  
OUTPUT: STANDARDIZED ;  
MODEL: eta by y1-y3 ; ! NB: no *
```

# MIMIC

## Multiple Indicators Multiple Causes



```

TITLE: MIMIC
DATA: FILE = trash.dat ;
VARIABLE: NAMES = y1-y3 x1 x2 ;
MODEL: eta by y1-y3* ;
eta@1 ;
eta on x1-x2 ; ! indirect effects
y3 on x2 ; ! direct effect
  
```

# Getting Data into Mplus

# What Mplus Eats

- Individual-level data
- Summary Data
  - ▶ correlations, covariances,
  - ▶ means, standard deviations
- ASCII Data
  - ▶ Raw text
  - ▶ Fixed, Free format

https://stats.idre.ucla.edu/mplus/faq/



Institute for Digital Research & Education

HOME

SOFTWARE ▾

RESOURCES ▾

SERVICES ▾

ABOUT US

## MPLUS FREQUENTLY ASKED QUESTIONS

---

- [How can I convert a Stata file to an Mplus file?](#)
- [How can I convert an SPSS file to an Mplus file?](#)
- [How can I convert a SAS file to an Mplus file?](#)

<https://stats.idre.ucla.edu/mplus/faq/>

## General Idea

- ① Use a general purpose stat package for EDA and data preparation
- ② Always use scripts/syntax/programs
- ③ Understand your data before you put move to Mplus
- ④ Never overwrite original data
- ⑤ Handle missing values
  - ▶ Replace with very large negative number (e.g., -9999)
  - ▶ Address missing values among INDEPENDENT variables
    - ★ Multiple Imputation (consider Mplus Bayes [Day 3])
    - ★ Declare independent variables dependent variables, assume a distribution, and use ML techniques
- ⑥ Output selected variables to ASCII text file (or CSV file)

# Converting Microsoft Excel File into raw data file for Mplus

- ① Click on "file", select "save as"
- ② Under "save as type" select CSV MS-DOS, click save
- ③ A warning will pop up: click "OK"
  - ▶ "The selected file type does not support workbooks..."
- ④ Then another warning will pop up: click "yes"
  - ▶ "\*.csv may contain features..."
- ⑤ Close out of Excel
- ⑥ Find the \*.csv file in your directory/file and right click
- ⑦ Open with text editor (e.g., Notepad, TextWrangler) and examine
  - ▶ First row of variable names and
  - ▶ Commas delimiting the variables
- ⑧ Copy variable names (for Mplus command file) and delete this line

## Stata program

```
prepare_csv.do  
use msq1.dta, clear  
summarize  
* Missing values in dependent variables (-9999)  
foreach var of varlist msq3-msq9 {  
    replace 'var' = -9999 if inlist('var',0,1)^=1  
}  
foreach var of varlist male age educ {  
    drop if 'var'==. // one way to handle missingness in x's  
}  
* assume anyone not identified as Black is White  
generate white=black^=1  
order msq3-msq9 site id male age white educ  
keep msq3-msq9 site id male age white educ  
outsheet using msq1.csv , nonames noquote nolabel comma
```

## Selected lines from data file

msq1.csv

```
2,1,1,1,1,2,3,300065,0,67,1,1  
1,1,1,1,1,1,3,300066,0,67,0,4  
2,1,1,2,1,2,3,300067,1,87,0,1  
1,2,1,1,1,2,3,300068,1,82,0,1  
2,1,1,2,1,2,3,300069,1,82,0,2  
2,1,1,1,1,2,3,300070,0,77,0,2  
2,1,1,2,2,2,3,300071,0,72,0,2  
2,1,1,1,1,1,3,300072,1,72,0,4  
1,1,1,1,1,1,3,300073,1,67,0,4  
2,1,-9999,-9999,2,-9999,3,300074,1,67,0,1
```

# SPSS program

```
prepare_csv.sps
```

```
get file= 'msq1.sav' .
execute .

compute n = 1 .
desc var all .

do repeat X = male age educ / .
  if missing(X) n=0 .
end repeat .

select if n=1 .

do repeat X = msq3 msq4 msq5 msq6 msq7 msq9 / .
  if missing(X) X=-9999 .
end repeat .

compute white=1 .
if black=1 white=0 .
save outfile ='trash.sav'

/keep=msq3 msq4 msq5 msq6 msq7 msq9 site id male age white educ .
get file ='trash.sav' .
save translate outfile='msq_from_spss.dat'
  /type=tab /map /replace.
execute .
```

## Selected lines from data file

msq\_from\_spss.dat

2	1	1	1	1	2	3	300065	0	67	1	1
1	1	1	1	1	1	3	300066	0	67	0	4
2	1	1	2	1	2	3	300067	1	87	0	1
1	2	1	1	1	2	3	300068	1	82	0	1
2	1	1	2	1	2	3	300069	1	82	0	2
2	1	1	1	1	2	3	300070	0	77	0	2
2	1	1	2	2	2	3	300071	0	72	0	2
2	1	1	1	1	1	3	300072	1	72	0	4
1	1	1	1	1	1	3	300073	1	67	0	4
2	1	-9999	-9999	2	-9999	3	300074	1	67	0	1

## SAS procedure

from <http://www.ats.ucla.edu/stat/mplus/faq/sas2mplus.htm>

```
data sample2;
set "sample";
array allvars _numeric_ ;
do over allvars;
  if missing(allvars) then allvars = -9999 ;
end;
run;
proc export data=sample2 outfile='sample.dat'
  dbms=dlm replace ;
run;
```

# Make your life easy

- Stata

stata2mplus.ado - format and export data, start an inp. From Stata:

```
ssc install stata2mplus
```

runmplus.ado - format and export data, write inp, run Mplus, display results to Stata results window, save estimation results.

Visit <http://www.lvmworkshop.org/home/runmplus-stuff>

mplus.ado - similar to runmplus

- ▶ net from <http://www.soc-research.net/ado/mplus>

- R

- ▶ MplusAutomation package

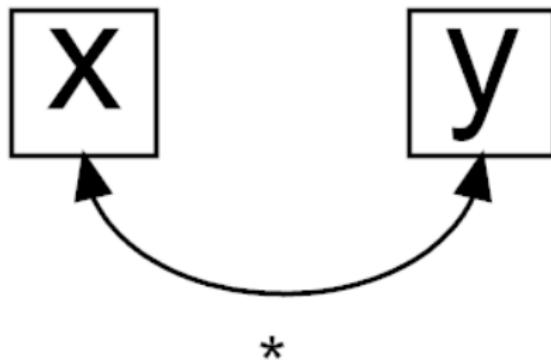
- SPSS, SAS

- ▶ Write your own macros

## Stata/runmplus and R/MplusAutomation

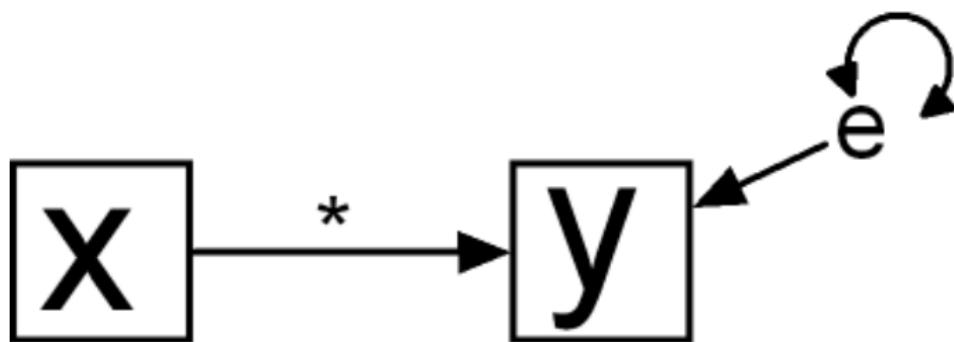
- Integration of Mplus with Stata and R
- Results stored as objects (estimates, fits)
- Facilitates
  - ▶ Reproducible data analysis
  - ▶ Simulation, bootstrap, multiple imputation
  - ▶ Automated model building
  - ▶ Learning about models
- Eliminates
  - ▶ Copying and pasting of output (and associated errors)
  - ▶ Other drudgery (making DAT or CSV files)
  - ▶ Clutter (inp, out, and dat, files of unknown origin)
- But,
  - ▶ Risk not learning about Mplus software
  - ▶ No Mplus plots

# Correlation



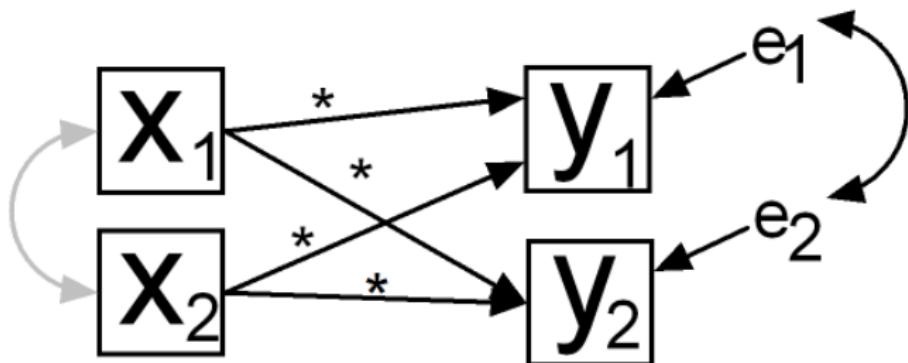
```
runmplus x y , model(x with y;) standardized
```

# Simple Regression



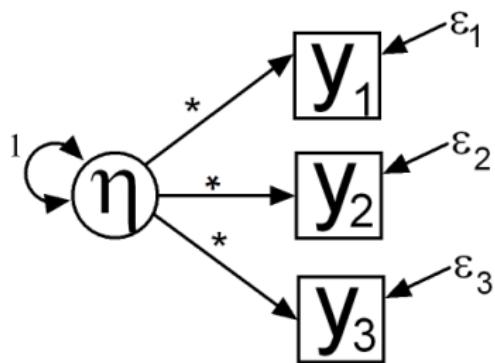
```
runmplus x y , model(y on x *;)
```

# Multivariate Regression



```
runmplus y1-y2 x1-x2 , model(y1 y2 on x1 x2; )
```

# Confirmatory Factor Analysis



```
runmplus y1-y3 , model(eta by y1-y3* ; eta@1;)
```

## R/MplusAutomation

- Run Mplus from R
- Requires separately constructed INP file
- Good parameter and fit statistic extraction
- Nice utilities for SWEAVEing (output to LaTeX)
- Ideally suited for simulations (running lots of INPs sequentially)
- Can be tricked into being a general purpose Mplus launcher

# Using Mplus via R

- MplusAutomation R package
- Designed to automate three major aspects of latent variable modeling in Mplus:
  - ▶ Creating related groups of models
  - ▶ Running Batches
  - ▶ Extracting and tabulating model parameters and test statistics
- <http://statmodel.com/usingmplusviar.shtml>
- <http://cran.r-project.org/web/packages/MplusAutomation/index.html>
- Written by [Michael.Hallquist@gmail.com](mailto:Michael.Hallquist@gmail.com)

Hallquist, M. N., & Wiley, J. F. (2018). Mplus Automation: An R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling*, 25(4), 621-638.

## Step 1: Install R Packages

```
# Step 1
# Select your CRAN Mirror
# done already-> chooseCRANmirror()
# Install the MplusviaR package
# done already-> install.packages("MplusAutomation", .Library)
library(MplusAutomation)
```

## Step 2: Set your work directory

```
# Step 2  
# Set working directory  
# This is where your data and input files should be  
setwd("C:/work/shortcourse/posted/analysis/learnMPA")
```

## Step 3: Getting Data into R

```
# Step 3
# Input data from a .dat file into R
EPESE_data <- read.table("ex01-02.dat",
  header=FALSE, sep=",")

# Declare column names
colnames(EPESE_data) <- c("sad", "blues", "depress",
  "happy", "enjoy", "hopeful", "age", "male")

# Attach dataframe (allows you to call variables by name)
attach(EPESE_data)
```

## Optional Step 4: Create your variables

```
# Step 4
# Create a binary variable for age, which equals
# 0 when a participant's age is less than the mean
# age in the sample, and 1 if it is more
age_mean <- mean(age)
age_high <- ifelse(age < age_mean, 0, 1)

# Exchange binary age variable for continuous
# in EPESE dataset
EPESE_data_cat <- data.frame(sad, blues, depress,
    happy, enjoy, hopeful, age_high, male)

# Export data from R to Mplus-friendly format
prepareMplusData(EPESE_data_cat,
    "epesecfa.dat")
```

## Step 5: Prepare your input file for Mplus

```
# Step 5
# Prepare Mplus INP file
cfa <- mplusObject(
  TITLE = "EX01-02 - CFA with covariates EPESE CESD;",
  VARIABLE = "CATEGORICAL  = sad blues depress
               happy enjoy hopeful;",
  MODEL = "f by sad-hopeful;
            f on age_high male ;",
  OUTPUT = "STDYX;
            MODINDICES (ALL -0);",
  usevariables = c("sad","blues","depress",
                  "happy","enjoy", "hopeful",
                  "age_high","male"),
  rdata = EPESE_data_cat)
```

## Step 6: Running the Model via Mplus

```
result <- mplusModeler(cfa,  
"EPESE_data_cat.dat",  
modelout = "ex0102.inp",  
run = 1L)
```

## Step 7a: Show output from Mplus in R

```
# Step 7a
# Show all output
model_output <-readModels(
  "C:/work/shortcourse/posted/analysis/learnMPA")
model_output

# Extract summary statistics
summaryStats <- extractModelSummaries(
  "C:/work/shortcourse/posted/analysis/learnMPA")
summaryStats
```

## Step 7b: Show table of output from Mplus in R

```
# Step 7b
# Show relevant summary statistics in a table
summary_table <- cbind(summaryStats$Observations,
                        summaryStats$ChiSqM_Value,
                        summaryStats$ChiSqM_DF,
                        summaryStats$ChiSqM_PValue,
                        summaryStats$CFI,
                        summaryStats$TLI,
                        summaryStats$RMSEA_Estimate)

colnames(summary_table) <- cbind(
  "N", "Chi-square", "(DF)", "(p)", "CFI", "TLI", "RMSEA")
summary_table
```

## Step 7c: Show table of other output from Mplus in R

```
# Step 7c
# Show modification indices
mod_indices <- extractModIndices(
  "C:/work/shortcourse/posted/analysis/learnMPA")
mod_indices

# Show parameter estimates
model_results <- extractModelParameters(
  "C:/work/shortcourse/posted/analysis/learnMPA")
model_results
```

# R lavaan Package

## lavaan

- R package for structural equation modeling and more
- Yves Rosseel, PhD, Ghent University, Belgium
- <http://lavaan.ugent.be/>
- Replicates Mplus estimators and output
- Supports some models and features (categorical data, sampling weights) but not all (multilevel, mixture, Bayesian) that Mplus does.
- GNU General Public License

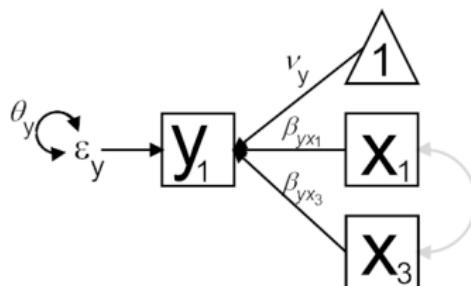
## lavaan vs Stata/runmplus and R/MplusAutomation

- Stata/runmplus and R/MplusAutomation require Mplus
- lavaan is its own software
  - ▶ Have to learn new model syntax in lavaan
  - ▶ Have to learn something new regardless

# lavaan vs. Mplus

	<i>Mplus</i>	<i>R/lavaan</i>
TITLE:	Multivariable regression	# Multivariable Regression
DATA:	FILE = ex3.1.dat ;	library(lavaan) setwd("c:/whereneverdatais ") data <- read.csv("ex3.1.dat", head=FALSE)
VARIABLE:	NAMES = y1 x1 x3 ;	names(data) <- c("y1","x1","x3")
MODEL:	y1 on x1 x3 ;	fit <- sem('y1 ~ 1 + x1 + x3', std.lv=TRUE, data=data)
OUTPUT:	STANDARDIZED ;	standardizedsolution(fit)

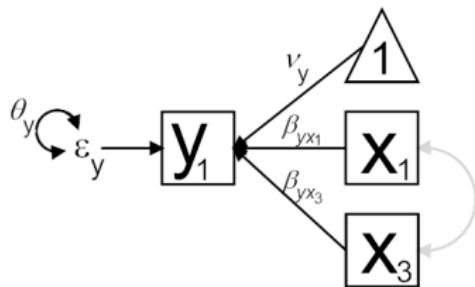
# Mplus Users Guide Ex 3.1: Multivariable Regression



Mplus Users Guide Example 3.1

```
TITLE: Multivariable regression
DATA: FILE = ex3.1.dat ;
VARIABLE: NAMES = y1 x1 x3 ;
MODEL: y1 on x1 x3 ;
OUTPUT: STANDARDIZED ;
```

# Mplus Users Guide Ex 3.1: Multivariable Regression (lavaan)



[Mplus Users Guide Example 3.1](#)

```

# Mplus Users Guide Ex 3.1: Multivariable Regression
library(lavaan)
setwd("c:/work/shortcourse/private/rnj")
data <- read.csv("ex3.1.dat",head=FALSE)
names(data) <- c("y1","x1","x3")
fit <- sem('y1 ~ 1 + x1 + x3', std.lv=TRUE, data=data )
summary(fit, standardized=TRUE, fit.measures = TRUE)
standardizedsolution(fit)
  
```

## Extra credit

Astute observers will note that Mplus and lavaan provide different loglikelihood and AIC values. This is because Mplus and lavaan allow different parts of the model into the likelihood function. To get Mplus to produce the same loglikelihood, use this model statement:

MODEL:

```
y1s by y1@1 ;  
x1s by x1@1 ;  
x3s by x3@1 ;  
y1@0 ;  
x1@0 ;  
x3@0 ;  
y1s on x1s x3s ;  
x1s with x3s ;
```

Refer back to the all-y model case (slide 85) to understand why.

# Questions

# Footnote on Data Management for LDA

Some approaches to LDA require data in "wide" format, others "long" format

LONG				WIDE						
Vertical Multiple record				Horizontal Single record						
<b>id</b>	<b>time</b>	<b>y1</b>	<b>x1</b>	<b>id</b>	<b>y1</b>	<b>y2</b>	<b>y3</b>	<b>x1</b>	<b>x2</b>	<b>x3</b>
1	1	y11	x11	1	y11	y12	y13	x11	x12	x13
1	2	y12	x12	2	y21	y22	y23	x21	x22	x23
1	3	y13	x13	...						
2	1	y21	x21	n	yn1	yn2	yn3	xn1	xn2	xn3
2	2	y22	x22							
2	3	y23	x23							
3	1	y31	x31							
...										
n	p	ynp	xnp							

Most Mplus LDA models require WIDE format

# Skills Challenge

- Invest time learning how to move from wide to long easily with your favorite general purpose statistics package
- Why
  - ▶ Wide
    - ★ Mplus LDA models want wide
  - ▶ Long
    - ★ Storage and data management more efficient
    - ★ Better for exploratory graphics
    - ★ Enables checking your growth models with mixed effects or HLM models

# Resources

- UCLA IDRE Learning Modules on reshaping data
  - ▶ Stata [https://stats.idre.ucla.edu/stata/faq/  
how-can-i-reshape-doubly-or-triply-wide-data-to-long/](https://stats.idre.ucla.edu/stata/faq/how-can-i-reshape-doubly-or-triply-wide-data-to-long/)
  - ▶ SAS [https://stats.idre.ucla.edu/sas/modules/  
how-to-reshape-data-long-to-wide-using-proc-transpose/](https://stats.idre.ucla.edu/sas/modules/how-to-reshape-data-long-to-wide-using-proc-transpose/)
  - ▶ SPSS [https://stats.idre.ucla.edu/spss/modules/  
reshaping-data-wide-to-long/](https://stats.idre.ucla.edu/spss/modules/reshaping-data-wide-to-long/)
  - ▶ R [https://stats.idre.ucla.edu/r/faq/  
how-can-i-reshape-my-data-in-r/](https://stats.idre.ucla.edu/r/faq/how-can-i-reshape-my-data-in-r/)
- Other Resources
  - ▶ R (S-plus) [https://stat.ethz.ch/pipermail/r-help/  
2009-February/189684.html](https://stat.ethz.ch/pipermail/r-help/2009-February/189684.html)
- Mplus will do it with built-in functionality
  - ▶ But it is not the right tool for data management

# Questions

# Examples

Latent growth curve modeling is a highly constrained form of confirmatory factor analysis (CFA). The goal of this session is to provide examples of specifying and estimating a CFA model in Mplus, as a basic Mplus skills-building exercise.

Time permitting, we will look at growth modeling examples from Singer and Willet (2003).

# Example Data Set

Established Populations for the Epidemiologic  
Study of the Elderly

## EPESE

- N=14,456
- General population survey of older adults
- Carried out at four US sites:
  - ▶ East Boston, MA (N=3,809)
  - ▶ Iowa and Washington Counties, IA (N=3,673)
  - ▶ New Haven, CT(N=2,812)
  - ▶ Durham-Piedmont area of NC (N=4,162)
- Field work began 1981, 1985 in NC
- Data are in the public domain (ICPSR)
- Results we derive have not been published

age		Freq.	Percent	Cum.
67: 65-69		4,457	30.83	30.83
72: 70-74		3,970	27.46	58.29
77: 75-79		2,872	19.87	78.16
82: 80-84		1,878	12.99	91.15
87: 85+		1,279	8.85	100.00
<hr/>				
Total		14,456	100.00	

male		Freq.	Percent	Cum.
0		8,960	61.98	61.98
1		5,496	38.02	100.00
Total		14,456	100.00	

## Black or African-American

site	0	1	.	Total
1 EB-MA	0	0	3,809	3,809
2 IW-IO	0	0	3,673	3,673
3 NH-CT	2,283	529	0	2,812
4 DP-NC	1,901	2,261	0	4,162
Total	4,184	2,790	7,482	14,456

educ		Freq.	Percent	Cum.
1 <8		4,056	28.06	28.06
2 8-11		5,686	39.33	67.39
3 12		2,552	17.65	85.04
4 >12		1,863	12.89	97.93
.		299	2.07	100.00
<hr/>				
Total		14,456	100.00	

# Cognition

## Short Portable Mental Status Questionnaire

- Conventionally, 10 items
- We will use six items from EPESE
  - ▶ msq3 what is date
  - ▶ msq4 what is the day of week
  - ▶ msq5 who is the president
  - ▶ msq6 who was the prior president
  - ▶ msq7 what is your Mothers maiden name
  - ▶ msq9 Count back from 20 by 3s

EPESE

Mental Status Questionnaire

Wave 1 Item Response Frequency

item		Response			Total
		Error	NoErr	miss	
		0	1	.	
msq3 what is the date...		3907	9788	761	14456
msq4 day of week.....		772	12916	768	14456
msq5 who is president...		1239	12436	781	14456
msq6 who was prior pres.		4302	9355	799	14456
msq7 what is M Maiden...		539	13075	842	14456
msq9 back from 20 by 3..		5194	7768	1494	14456

# Depression

- CES-D Centers for Epidemiologic Studies Depression Scale
- We will use six items from New Haven site, which ask if *During the past week* ...

SAD: ...I felt sad

BLUES: ...I felt I could not  
shake off the blues

DEPRESS: ... I felt depressed

- 0: Rarely or None of the time
- 1: Some of the time
- 2: Much of the time
- 3: Most or all of the time

HAPPY: ...I was happy

ENJOY: ...I enjoyed life

HOPEFUL: ...I felt hopeful  
about the future

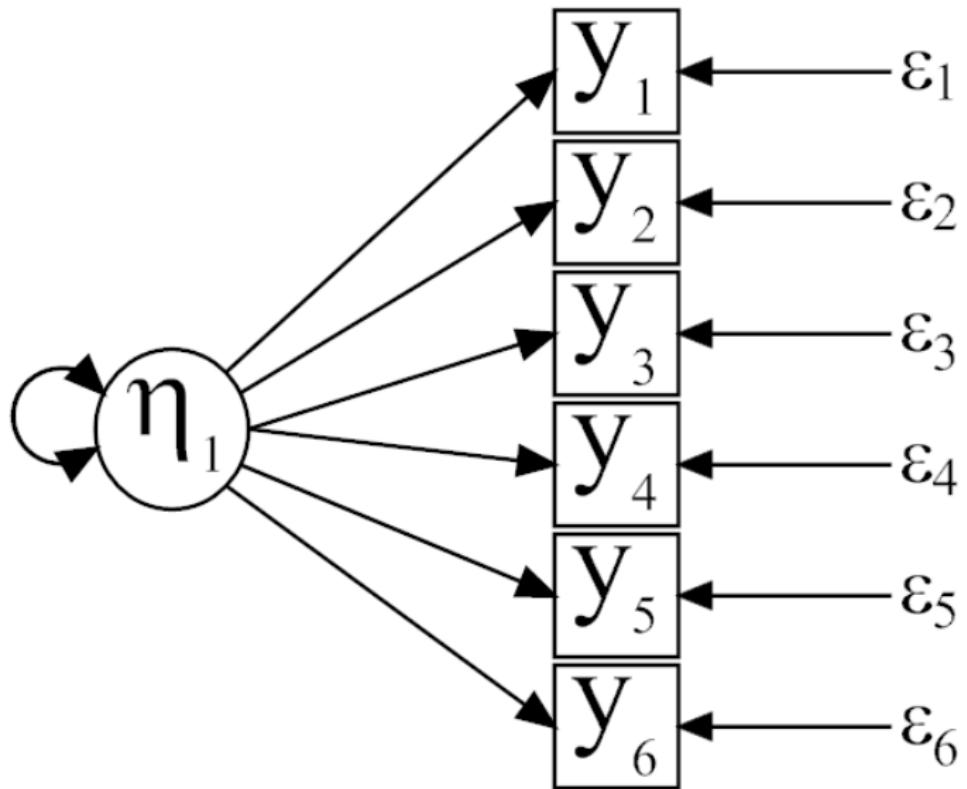
- 0: Most or all of the time
- 1: Much of the time
- 2: Some of the time
- 3: Rarely or None of the time

Item Response Frequency  
(entries are percentages)

item	Response				N
	0	1	2	3	
01 sad.....	66	27	4	4	2442
02 blues....	78	15	3	4	2442
03 depress..	64	26	4	6	2442
04 happy....	67	10	16	7	2442
05 enjoy....	76	8	11	6	2442
06 hopeful..	58	7	16	20	2442

## Confirmatory Factor Analysis Example [ex0101.inp]

Confirmatory Factor Analysis  
New Haven EPESE  
Six-item CES-D  
Continuous Factor Indicators



```
1 Mplus VERSION 8.8 DEMO (Mac)
2 MUTHEN & MUTHEN
3 07/11/2022    2:07 PM
4
5 INPUT INSTRUCTIONS
6
7 TITLE:    EX0101 CFA EPESE CESD ;
8
9 DATA:    FILE = ex0101.dat ;
10
11 VARIABLE: NAMES = sad blues depress happy enjoy hopeful age male ;
12 USEVARIABLES = sad blues depress happy enjoy hopeful ;
13 MISSING ARE ALL (-9999) ;
14
15 OUTPUT: STANDARDIZED TECH1 ;
16
17 MODEL:   f1 by sad-hopeful;
18
19
20
21
22 INPUT READING TERMINATED NORMALLY
23
24
25
26 EX0101 CFA EPESE CESD ;
27
28 SUMMARY OF ANALYSIS
29
30 Number of groups                                1
31 Number of observations                          2442
32
```

```
33 Number of dependent variables 6
34 Number of independent variables 0
35 Number of continuous latent variables 1
36
37 Observed dependent variables
38
39 Continuous
40   SAD      BLUES      DEPRESS     HAPPY      ENJOY      HOPEFUL
41
42 Continuous latent variables
43   F1
44
45
46 Estimator                      ML
47 Information matrix              OBSERVED
48 Maximum number of iterations    1000
49 Convergence criterion          0.500D-04
50 Maximum number of steepest descent iterations 20
51 Maximum number of iterations for H1 2000
52 Convergence criterion for H1 0.100D-03
53
54 Input data file(s)
55   ex0101.dat
56
57 Input data format  FREE
58
59
60 SUMMARY OF DATA
61
62   Number of missing data patterns 1
63
64
```

```
65 COVARIANCE COVERAGE OF DATA
66
67 Minimum covariance coverage value 0.100
68
69
70      PROPORTION OF DATA PRESENT
71
72
73      Covariance Coverage
74          SAD           BLUES        DEPRESS       HAPPY        ENJOY
75          -----        -----        -----        -----
76 SAD      1.000
77 BLUES    1.000      1.000
78 DEPRESS  1.000      1.000      1.000
79 HAPPY    1.000      1.000      1.000      1.000
80 ENJOY    1.000      1.000      1.000      1.000      1.000
81 HOPEFUL 1.000      1.000      1.000      1.000      1.000
82
83
84      Covariance Coverage
85          HOPEFUL
86          -----
87 HOPEFUL 1.000
88
89
90
91 UNIVARIATE SAMPLE STATISTICS
92
93
94      UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS
95
```

		Variable/ Percentiles	Mean/ Variance	Skewness/ Kurtosis	Minimum/ Maximum	% with Min/Max	20%/60%	40%
		Sample Size 80% Median						
96								
97	SAD	0.000 0.000 2442.000 1.000	0.454 0.548	1.831 3.185	0.000 3.000	65.77% 3.81%	0.000 0.000	
98	BLUES	0.000 0.000 2442.000 1.000	0.337 0.545	2.417 5.259	0.000 3.000	78.05% 4.34%	0.000 0.000	
99	DEPRESS	0.000 0.000 2442.000 1.000	0.510 0.674	1.749 2.457	0.000 3.000	64.41% 5.81%	0.000 0.000	
100	HAPPY	0.000 0.000 2442.000 2.000	0.630 0.979	1.241 0.069	0.000 3.000	67.08% 7.25%	0.000 0.000	
101	ENJOY	0.000 0.000 2442.000 1.000	0.465 0.816	1.742 1.624	0.000 3.000	75.96% 5.94%	0.000 0.000	
102	HOPEFUL	0.000 0.000 2442.000 2.000	0.973 1.517	0.681 -1.263	0.000 3.000	57.70% 19.57%	0.000 1.000	
103								
104								
105								
106								
107								
108								
109								
110								
111								
112								
113	THE MODEL ESTIMATION TERMINATED NORMALLY							

```
114  
115  
116  
117 MODEL FIT INFORMATION  
118  
119 Number of Free Parameters 18  
120  
121 Loglikelihood  
122  
123      H0 Value -17278.438  
124      H1 Value -17036.545  
125  
126 Information Criteria  
127  
128      Akaike (AIC) 34592.876  
129      Bayesian (BIC) 34697.286  
130      Sample-Size Adjusted BIC 34640.096  
131      (n* = (n + 2) / 24)  
132  
133 Chi-Square Test of Model Fit  
134  
135      Value 483.785  
136      Degrees of Freedom 9  
137      P-Value 0.0000  
138  
139 RMSEA (Root Mean Square Error Of Approximation)  
140  
141      Estimate 0.147  
142      90 Percent C.I. 0.136 0.158  
143      Probability RMSEA <= .05 0.000  
144  
145 CFI/TLI
```

146  
147       CFI                                   0.883  
148       TLI                                   0.805  
149  
150 Chi-Square Test of Model Fit for the Baseline Model  
151  
152       Value                               4064.686  
153       Degrees of Freedom               15  
154       P-Value                             0.0000  
155  
156 SRMR (Standardized Root Mean Square Residual)  
157  
158       Value                               0.057  
159  
160  
161  
162 MODEL RESULTS  
163  
164   Two-Tailed  
165   Estimate   S.E.   Est./S.E.   P-Value  
166  
167 F1       BY  
168     SAD                                1.000    0.000   999.000   999.000  
169     BLUES                              0.963    0.035   27.751   0.000  
170     DEPRESS                            1.193    0.040   30.075   0.000  
171     HAPPY                             1.230    0.049   24.975   0.000  
172     ENJOY                             1.075    0.045   23.950   0.000  
173     HOPEFUL                          0.883    0.057   15.489   0.000  
174  
175 Intercepts  
176     SAD                                0.454    0.015   30.325   0.000  
177     BLUES                              0.337    0.015   22.553   0.000

178 DEPRESS 0.510 0.017 30.685 0.000  
179 HAPPY 0.630 0.020 31.451 0.000  
180 ENJOY 0.465 0.018 25.421 0.000  
181 HOPEFUL 0.973 0.025 39.018 0.000  
182  
183 Variances  
184 F1 0.260 0.015 17.461 0.000  
185  
186 Residual Variances  
187 SAD 0.288 0.011 27.263 0.000  
188 BLUES 0.304 0.011 27.677 0.000  
189 DEPRESS 0.304 0.013 23.883 0.000  
190 HAPPY 0.586 0.021 27.902 0.000  
191 ENJOY 0.516 0.018 28.700 0.000  
192 HOPEFUL 1.315 0.039 33.536 0.000  
193  
194  
195 QUALITY OF NUMERICAL RESULTS  
196  
197 Condition Number for the Information Matrix 0.141E-01  
198 (ratio of smallest to largest eigenvalue)  
199  
200  
201 STANDARDIZED MODEL RESULTS  
202  
203  
204 STDYX Standardization  
205  
206  
207 Estimate S.E. Est./S.E. Two-Tailed P-Value  
208  
209 F1 BY

210	SAD	0.689	0.014	50.000	0.000
211	BLUES	0.665	0.015	45.558	0.000
212	DEPRESS	0.741	0.013	56.771	0.000
213	HAPPY	0.634	0.016	40.045	0.000
214	ENJOY	0.606	0.016	36.810	0.000
215	HOPEFUL	0.366	0.020	18.142	0.000
216					
217	Intercepts				
218	SAD	0.614	0.022	27.819	0.000
219	BLUES	0.456	0.021	21.463	0.000
220	DEPRESS	0.621	0.022	28.096	0.000
221	HAPPY	0.636	0.022	28.680	0.000
222	ENJOY	0.514	0.022	23.890	0.000
223	HOPEFUL	0.790	0.023	34.068	0.000
224					
225	Variances				
226	F1	1.000	0.000	999.000	999.000
227					
228	Residual Variances				
229	SAD	0.526	0.019	27.698	0.000
230	BLUES	0.558	0.019	28.739	0.000
231	DEPRESS	0.451	0.019	23.359	0.000
232	HAPPY	0.598	0.020	29.838	0.000
233	ENJOY	0.632	0.020	31.669	0.000
234	HOPEFUL	0.866	0.015	58.825	0.000
235					
236					
237	STDY Standardization				
238					
239				Two-Tailed	
240		Estimate	S.E.	Est./S.E.	P-Value
241					

```

242 F1      BY
243   SAD          0.689    0.014    50.000    0.000
244   BLUES        0.665    0.015    45.558    0.000
245   DEPRESS      0.741    0.013    56.771    0.000
246   HAPPY        0.634    0.016    40.045    0.000
247   ENJOY         0.606    0.016    36.810    0.000
248   HOPEFUL      0.366    0.020    18.142    0.000
249
250 Intercepts
251   SAD          0.614    0.022    27.819    0.000
252   BLUES        0.456    0.021    21.463    0.000
253   DEPRESS      0.621    0.022    28.096    0.000
254   HAPPY        0.636    0.022    28.680    0.000
255   ENJOY         0.514    0.022    23.890    0.000
256   HOPEFUL      0.790    0.023    34.068    0.000
257
258 Variances
259   F1           1.000    0.000    999.000    999.000
260
261 Residual Variances
262   SAD          0.526    0.019    27.698    0.000
263   BLUES        0.558    0.019    28.739    0.000
264   DEPRESS      0.451    0.019    23.359    0.000
265   HAPPY        0.598    0.020    29.838    0.000
266   ENJOY         0.632    0.020    31.669    0.000
267   HOPEFUL      0.866    0.015    58.825    0.000
268
269
270 STD Standardization
271
272
273   Estimate     S.E.    Est./S.E.  Two-Tailed P-Value

```

274  
275 F1 BY  
276 SAD 0.510 0.015 34.922 0.000  
277 BLUES 0.491 0.015 33.059 0.000  
278 DEPRESS 0.608 0.016 37.799 0.000  
279 HAPPY 0.627 0.021 30.522 0.000  
280 ENJOY 0.548 0.019 28.846 0.000  
281 HOPEFUL 0.450 0.027 16.672 0.000  
282  
283 Intercepts  
284 SAD 0.454 0.015 30.325 0.000  
285 BLUES 0.337 0.015 22.553 0.000  
286 DEPRESS 0.510 0.017 30.685 0.000  
287 HAPPY 0.630 0.020 31.451 0.000  
288 ENJOY 0.465 0.018 25.421 0.000  
289 HOPEFUL 0.973 0.025 39.018 0.000  
290  
291 Variances  
292 F1 1.000 0.000 999.000 999.000  
293  
294 Residual Variances  
295 SAD 0.288 0.011 27.263 0.000  
296 BLUES 0.304 0.011 27.677 0.000  
297 DEPRESS 0.304 0.013 23.883 0.000  
298 HAPPY 0.586 0.021 27.902 0.000  
299 ENJOY 0.516 0.018 28.700 0.000  
300 HOPEFUL 1.315 0.039 33.536 0.000  
301  
302  
303 R-SQUARE  
304  
305 Observed Two-Tailed

	Variable	Estimate	S.E.	Est./S.E.	P-Value
306	SAD	0.474	0.019	25.000	0.000
307	BLUES	0.442	0.019	22.779	0.000
308	DEPRESS	0.549	0.019	28.385	0.000
309	HAPPY	0.402	0.020	20.022	0.000
310	ENJOY	0.368	0.020	18.405	0.000
311	HOPEFUL	0.134	0.015	9.071	0.000
312					
313					
314					
315					
316	TECHNICAL 1 OUTPUT				
317					
318					
319	PARAMETER SPECIFICATION				
320					
321					
322	NU				
323	SAD		BLUES		DEPRESS
324		-----		-----	
325		1		2	
326					
327					
328	NU				
329	HOPEFUL				
330		-----			
331		6			
332					
333					
334	LAMBDA				
335	F1				
336		-----			
337	SAD		0		

338	BLUES	7			
339	DEPRESS	8			
340	HAPPY	9			
341	ENJOY	10			
342	HOPEFUL	11			
343					
344					
345	THETA				
346	SAD	BLUES	DEPRESS	HAPPY	ENJOY
347	-----	-----	-----	-----	-----
348	SAD	12			
349	BLUES	0	13		
350	DEPRESS	0	0	14	
351	HAPPY	0	0	0	15
352	ENJOY	0	0	0	0
353	HOPEFUL	0	0	0	0
354					
355					
356	THETA				
357	HOPEFUL				
358	-----				
359	HOPEFUL	17			
360					
361					
362	ALPHA				
363	F1				
364	-----				
365	0				
366					
367					
368	BETA				
369	F1				

370  
371 F1 ----- 0  
372  
373  
374 PSI  
375 F1  
376 -----  
377 F1 ----- 18  
378  
379  
380 STARTING VALUES  
381  
382  
383 NU  
384 SAD BLUES DEPRESS HAPPY ENJOY  
385 -----  
386 0.454 0.337 0.510 0.630 0.465  
387  
388  
389 NU  
390 HOPEFUL  
391 -----  
392 0.973  
393  
394  
395 LAMBDA  
396 F1  
397 -----  
398 SAD 1.000  
399 BLUES 0.991  
400 DEPRESS 1.218  
401 HAPPY 1.176

```
402 ENJOY      1.031
403 HOPEFUL    0.847
404
405
406     THETA
407         SAD          BLUES        DEPRESS      HAPPY       ENJOY
408         -----      -----      -----
409 SAD      0.274
410 BLUES    0.000      0.273
411 DEPRESS   0.000      0.000      0.337
412 HAPPY    0.000      0.000      0.000      0.490
413 ENJOY    0.000      0.000      0.000      0.000      0.408
414 HOPEFUL  0.000      0.000      0.000      0.000      0.000
415
416
417     THETA
418         HOPEFUL
419         -----
420 HOPEFUL  0.759
421
422
423     ALPHA
424         F1
425         -----
426         0.000
427
428
429     BETA
430         F1
431         -----
432 F1       0.000
433
```

```
434
435      PSI
436          F1
437          -----
438 F1          0.050
439
440
441      Beginning Time: 14:07:22
442      Ending Time: 14:07:22
443      Elapsed Time: 00:00:00
444
445
446 Mplus VERSION 8.8 DEMO (Mac) has the following limitations:
447 Maximum number of dependent variables: 6
448 Maximum number of independent variables: 2
449 Maximum number of between variables: 2
450 Maximum number of continuous latent variables in time series analysis: 2
451
452
453 MUTHEN & MUTHEN
454 3463 Stoner Ave.
455 Los Angeles, CA 90066
456
457 Tel: (310) 391-9971
458 Fax: (310) 391-8971
459 Web: www.StatModel.com
460 Support: Support@StatModel.com
461
462 Copyright (c) 1998-2022 Muthen & Muthen
```

## A Note Model Fit

- Model fit is based on how close the model-implied covariance matrix is to the observed covariance matrix
- Chi-Square should be low, P-value high
  - ▶ an *absolute fit* index
  - ▶ very sensitive to sample size
- CFI > 0.95
  - ▶ a *parsimony-adjusted relative fit* index
  - ▶ CFI ranges from 0 to 1
  - ▶ Bentler. *Psych Bull*, 1990; 107:238-46.
- RMSEA < .05
  - ▶ an *absolute fit* index
  - ▶ Hu & Bentler. *Psych Methods*, 1998; 4:424-53.

## A note on Mplus Standardized Estimates

- Standardized with respect to...

**STDYX** ...the variances of the continuous latent variables, background, and outcome variables. Use this if the background variable is continuous (and it otherwise makes sense to do so).

**STDY** ...the variances of the continuous latent variables and the outcome variables. Use this one if the background variable is discrete.

**STD** ... the variances of the continuous latent variables.

- Standard errors and significance testing

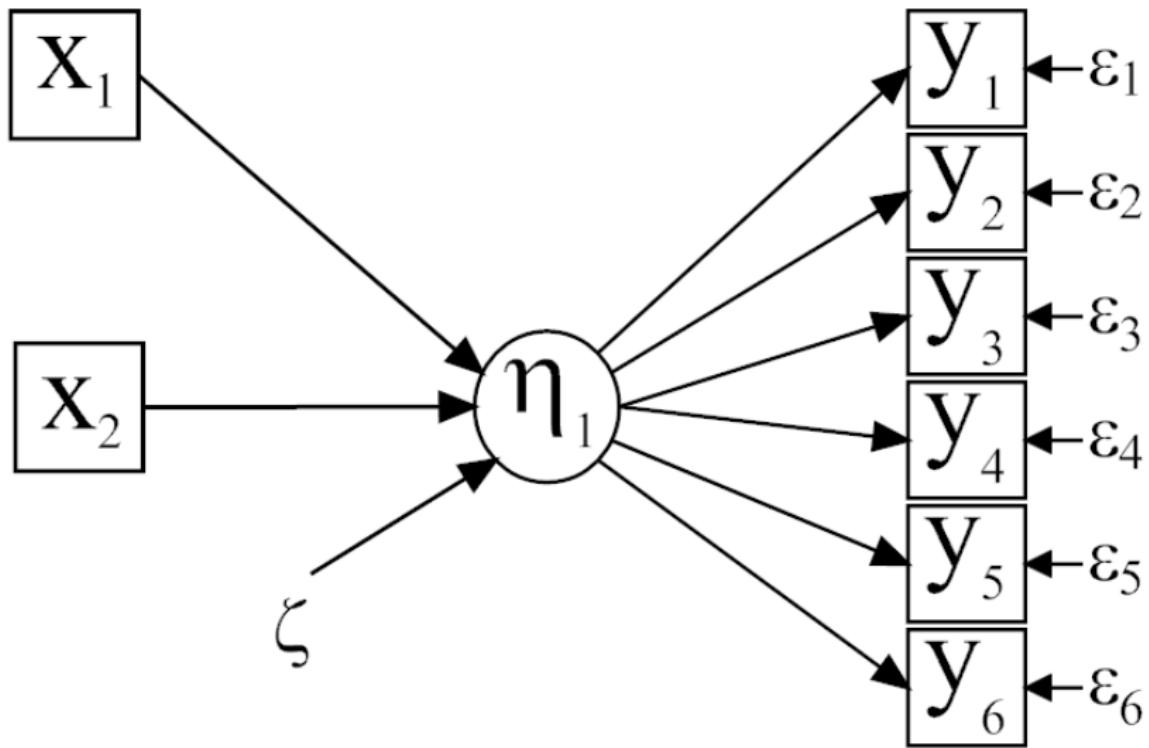
- ▶ Each standardization solution may come with standard errors, z-test, P-value (e.g., STDY not available for categorical dependent variables)
- ▶ Significance tests don't always agree with unstandardized estimates
- ▶ I recommend basing inferences on unstandardized estimates

## Example 1.2

CFA with Covariates  
CES-D Depression, Age, Sex  
New Haven EPESE

## CFA with Covariates Example 1.2 ex01-02.inp

MIMIC - Multiple Indicators Multiple Causes EPESE CES-D (six items)  
Age, Gender



```
1 Mplus VERSION 8.8 DEMO (Mac)
2 MUTHEN & MUTHEN
3 07/11/2022    2:07 PM
4
5 INPUT INSTRUCTIONS
6
7 TITLE:    EX0102 CFA WITH COVARIATES EPESE CESD ;
8
9 DATA:    FILE = ex0101.dat ;
10
11 VARIABLE: NAMES = sad blues depress happy enjoy hopeful age male ;
12           MISSING ARE ALL (-9999) ;
13
14 DEFINE:   CENTER age male (grandmean) ;
15
16 OUTPUT:   TECH1 ; standardized ;
17
18 MODEL:    f by sad-hopeful ;
19           f on age male ;
20
21
22
23
24
25
26 INPUT READING TERMINATED NORMALLY
27
28
29
30 EX0102 CFA WITH COVARIATES EPESE CESD ;
31
32 SUMMARY OF ANALYSIS
```

33  
34 Number of groups 1  
35 Number of observations 2442  
36  
37 Number of dependent variables 6  
38 Number of independent variables 2  
39 Number of continuous latent variables 1  
40  
41 Observed dependent variables  
42  
43 Continuous  
44 SAD BLUES DEPRESS HAPPY ENJOY HOPEFUL  
45  
46 Observed independent variables  
47 AGE MALE  
48  
49 Continuous latent variables  
50 F  
51  
52 Variables with special functions  
53  
54 Centering (GRANDMEAN)  
55 AGE MALE  
56  
57  
58 Estimator ML  
59 Information matrix OBSERVED  
60 Maximum number of iterations 1000  
61 Convergence criterion 0.500D-04  
62 Maximum number of steepest descent iterations 20  
63 Maximum number of iterations for H1 2000  
64 Convergence criterion for H1 0.100D-03

```
65
66 Input data file(s)
67   ex0101.dat
68
69 Input data format  FREE
70
71
72 SUMMARY OF DATA
73
74     Number of missing data patterns          1
75
76
77 COVARIANCE COVERAGE OF DATA
78
79 Minimum covariance coverage value  0.100
80
81
82     PROPORTION OF DATA PRESENT
83
84
85     Covariance Coverage
86           SAD        BLUES      DEPRESS      HAPPY      ENJOY
87           -----      -----      -----      -----
88 SAD       1.000
89 BLUES    1.000      1.000
90 DEPRESS  1.000      1.000      1.000
91 HAPPY    1.000      1.000      1.000      1.000
92 ENJOY    1.000      1.000      1.000      1.000      1.000
93 HOPEFUL 1.000      1.000      1.000      1.000      1.000
94 AGE      1.000      1.000      1.000      1.000      1.000
95 MALE    1.000      1.000      1.000      1.000      1.000
96
```

```

97
98      Covariance Coverage
99          HOPEFUL      AGE      MALE
100
101 HOPEFUL      1.000
102 AGE          1.000      1.000
103 MALE         1.000      1.000      1.000
104
105
106
107 UNIVARIATE SAMPLE STATISTICS
108
109
110 UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS
111
112      Variable/      Mean/      Skewness/      Minimum/ % with
113          Percentiles
114          Sample Size      Variance      Kurtosis      Maximum      Min/Max    20%/60%    40%
115          80%       Median
116
117 SAD          0.454      1.831      0.000      65.77%    0.000
118          0.000      0.000
119          2442.000     0.548      3.185      3.000      3.81%    0.000
120          1.000
121
122 BLUES        0.337      2.417      0.000      78.05%    0.000
123          0.000      0.000
124          2442.000     0.545      5.259      3.000      4.34%    0.000
125          1.000
126
127 DEPRESS      0.510      1.749      0.000      64.41%    0.000
128          0.000      0.000
129          2442.000     0.674      2.457      3.000      5.81%    0.000
130          1.000

```

121	HAPPY		0.630	1.241	0.000	67.08%	0.000
122		0.000 0.000	2442.000	0.979	0.069	3.000	7.25%
			2.000				0.000
123	ENJOY		0.465	1.742	0.000	75.96%	0.000
124		0.000 0.000	2442.000	0.816	1.624	3.000	5.94%
			1.000				0.000
125	HOPEFUL		0.973	0.681	0.000	57.70%	0.000
126		0.000 0.000	2442.000	1.517	-1.263	3.000	19.57%
			2.000				1.000
127	AGE		0.000	0.497	-7.316	28.50%	-7.316
128		-2.316 -2.316	2442.000	40.696	-0.862	12.684	8.39%
			7.684				2.684
129	MALE		0.000	0.307	-0.424	57.58%	-0.424
130		-0.424 -0.424	2442.000	0.244	-1.906	0.576	42.42%
			0.576				0.576

131  
132  
133 THE MODEL ESTIMATION TERMINATED NORMALLY  
134  
135  
136

137 MODEL FIT INFORMATION

138

139 Number of Free Parameters

20

140

141 Loglikelihood

142

143            H0 Value                            -17255.177  
144            H1 Value                            -17006.684  
145  
146 Information Criteria  
147  
148            Akaike (AIC)                        34550.354  
149            Bayesian (BIC)                        34666.365  
150            Sample-Size Adjusted BIC            34602.820  
151            (n\* = (n + 2) / 24)  
152  
153 Chi-Square Test of Model Fit  
154  
155            Value                                496.985  
156            Degrees of Freedom                    19  
157            P-Value                                0.0000  
158  
159 RMSEA (Root Mean Square Error Of Approximation)  
160  
161            Estimate                              0.101  
162            90 Percent C.I.                        0.094  0.109  
163            Probability RMSEA <= .05            0.000  
164  
165 CFI/TLI  
166  
167            CFI                                    0.883  
168            TLI                                    0.834  
169  
170 Chi-Square Test of Model Fit for the Baseline Model  
171  
172            Value                                4124.409  
173            Degrees of Freedom                    27  
174            P-Value                                0.0000

175  
176 SRMR (Standardized Root Mean Square Residual)  
177  
178       Value    0.048  
179  
180  
181  
182 MODEL RESULTS  
183  
184    Two-Tailed  
185    P-Value  
186  
187 F           BY  
188     SAD    1.000        0.000       999.000       999.000  
189     BLUES    0.963        0.035       27.784       0.000  
190     DEPRESS   1.190        0.040       30.103       0.000  
191     HAPPY    1.231        0.049       25.035       0.000  
192     ENJOY    1.075        0.045       24.007       0.000  
193     HOPEFUL   0.884        0.057       15.518       0.000  
194  
195 F           ON  
196     AGE    0.006        0.002       3.234       0.001  
197     MALE   -0.134       0.023       -5.778       0.000  
198  
199 Intercepts  
200     SAD    0.454        0.015       30.494       0.000  
201     BLUES   0.337        0.015       22.670       0.000  
202     DEPRESS   0.510        0.017       30.881       0.000  
203     HAPPY   0.630        0.020       31.599       0.000  
204     ENJOY   0.465        0.018       25.530       0.000  
205     HOPEFUL   0.973        0.025       39.078       0.000  
206

207 Residual Variances  
208 SAD 0.288 0.011 27.278 0.000  
209 BLUES 0.304 0.011 27.713 0.000  
210 DEPRESS 0.306 0.013 24.034 0.000  
211 HAPPY 0.585 0.021 27.942 0.000  
212 ENJOY 0.516 0.018 28.740 0.000  
213 HOPEFUL 1.314 0.039 33.541 0.000  
214 F 0.254 0.015 17.456 0.000  
215  
216

## 217 QUALITY OF NUMERICAL RESULTS

218

219 Condition Number for the Information Matrix 0.568E-03  
220 (ratio of smallest to largest eigenvalue)

221

222

## 223 STANDARDIZED MODEL RESULTS

224

225

## 226 STDYX Standardization

227

		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
231 F	BY				
232 SAD		0.689	0.014	50.114	0.000
233 BLUES		0.665	0.015	45.608	0.000
234 DEPRESS		0.739	0.013	56.729	0.000
235 HAPPY		0.634	0.016	40.219	0.000
236 ENJOY		0.607	0.016	36.956	0.000
237 HOPEFUL		0.366	0.020	18.183	0.000

238

239 F ON  
240 AGE 0.072 0.022 3.249 0.001  
241 MALE -0.130 0.022 -5.880 0.000  
242  
243 Intercepts  
244 SAD 0.614 0.022 27.949 0.000  
245 BLUES 0.456 0.021 21.564 0.000  
246 DEPRESS 0.621 0.022 28.247 0.000  
247 HAPPY 0.636 0.022 28.793 0.000  
248 ENJOY 0.514 0.021 23.980 0.000  
249 HOPEFUL 0.790 0.023 34.108 0.000  
250  
251 Residual Variances  
252 SAD 0.525 0.019 27.705 0.000  
253 BLUES 0.558 0.019 28.782 0.000  
254 DEPRESS 0.453 0.019 23.521 0.000  
255 HAPPY 0.598 0.020 29.868 0.000  
256 ENJOY 0.632 0.020 31.705 0.000  
257 HOPEFUL 0.866 0.015 58.794 0.000  
258 F 0.977 0.007 145.590 0.000  
259  
260  
261 STDY Standardization  
262  
263 Two-Tailed  
264 Estimate S.E. Est./S.E. P-Value  
265  
266 F BY  
267 SAD 0.689 0.014 50.114 0.000  
268 BLUES 0.665 0.015 45.608 0.000  
269 DEPRESS 0.739 0.013 56.729 0.000  
270 HAPPY 0.634 0.016 40.219 0.000

271 ENJOY 0.607 0.016 36.956 0.000  
272 HOPEFUL 0.366 0.020 18.183 0.000  
273  
274 F ON  
275 AGE 0.011 0.003 3.252 0.001  
276 MALE -0.263 0.045 -5.900 0.000  
277  
278 Intercepts  
279 SAD 0.614 0.022 27.949 0.000  
280 BLUES 0.456 0.021 21.564 0.000  
281 DEPRESS 0.621 0.022 28.247 0.000  
282 HAPPY 0.636 0.022 28.793 0.000  
283 ENJOY 0.514 0.021 23.980 0.000  
284 HOPEFUL 0.790 0.023 34.108 0.000  
285  
286 Residual Variances  
287 SAD 0.525 0.019 27.705 0.000  
288 BLUES 0.558 0.019 28.782 0.000  
289 DEPRESS 0.453 0.019 23.521 0.000  
290 HAPPY 0.598 0.020 29.868 0.000  
291 ENJOY 0.632 0.020 31.705 0.000  
292 HOPEFUL 0.866 0.015 58.794 0.000  
293 F 0.977 0.007 145.590 0.000  
294  
295  
296 STD Standardization  
297  
298 Two-Tailed  
299 Estimate S.E. Est./S.E. P-Value  
300  
301 F BY  
302 SAD 0.510 0.015 34.965 0.000

303      BLUES                    0.491      0.015      33.077      0.000  
 304      DEPRESS                0.607      0.016      37.766      0.000  
 305      HAPPY                  0.628      0.021      30.604      0.000  
 306      ENJOY                 0.548      0.019      28.920      0.000  
 307      HOPEFUL               0.451      0.027      16.703      0.000  
 308  
 309      F                      ON  
 310      AGE                    0.011      0.003      3.252      0.001  
 311      MALE                  -0.263     0.045      -5.900     0.000  
 312  
 313      Intercepts  
 314      SAD                    0.454      0.015      30.494      0.000  
 315      BLUES                0.337      0.015      22.670      0.000  
 316      DEPRESS               0.510      0.017      30.881      0.000  
 317      HAPPY                0.630      0.020      31.599      0.000  
 318      ENJOY                0.465      0.018      25.530      0.000  
 319      HOPEFUL              0.973      0.025      39.078      0.000  
 320  
 321      Residual Variances  
 322      SAD                    0.288      0.011      27.278      0.000  
 323      BLUES                0.304      0.011      27.713      0.000  
 324      DEPRESS               0.306      0.013      24.034      0.000  
 325      HAPPY                0.585      0.021      27.942      0.000  
 326      ENJOY                0.516      0.018      28.740      0.000  
 327      HOPEFUL              1.314      0.039      33.541      0.000  
 328      F                      0.977      0.007      145.590     0.000  
 329  
 330  
 331 R-SQUARE  
 332  
 333      Observed  
 334      Variable              Estimate     S.E.      Est./S.E.      Two-Tailed P-Value

335  
336    SAD                0.475        0.019        25.057        0.000  
337    BLUES              0.442        0.019        22.804        0.000  
338    DEPRESS            0.547        0.019        28.364        0.000  
339    HAPPY             0.402        0.020        20.110        0.000  
340    ENJOY             0.368        0.020        18.478        0.000  
341    HOPEFUL           0.134        0.015        9.091        0.000  
342  
343    Latent                                  Two-Tailed  
344    Variable           Estimate        S.E.        Est./S.E.      P-Value  
345  
346    F                    0.023        0.007        3.466        0.001  
347  
348  
349 TECHNICAL 1 OUTPUT  
350  
351  
352 PARAMETER SPECIFICATION  
353  
354  
355    NU  
356        SAD                BLUES                DEPRESS            HAPPY                ENJOY  
357        -----            -----                -----                -----  
358        1                    2                    3                    4                    5  
359  
360  
361    NU  
362        HOPEFUL           AGE                MALE  
363        -----            -----                -----  
364        6                    0                    0  
365  
366

	LAMBDA	AGE	MALE		
367	F				
368		-----	-----		
369					
370	SAD	0	0		
371	BLUES	7	0		
372	DEPRESS	8	0		
373	HAPPY	9	0		
374	ENJOY	10	0		
375	HOPEFUL	11	0		
376	AGE	0	0		
377	MALE	0	0		
378					
379					
380	THETA				
381	SAD	BLUES	DEPRESS	HAPPY	ENJOY
382		-----	-----	-----	-----
383	SAD	12			
384	BLUES	0	13		
385	DEPRESS	0	0	14	
386	HAPPY	0	0	0	15
387	ENJOY	0	0	0	0
388	HOPEFUL	0	0	0	0
389	AGE	0	0	0	0
390	MALE	0	0	0	0
391					
392					
393	THETA				
394	HOPEFUL	AGE	MALE		
395		-----	-----		
396	HOPEFUL	17			
397	AGE	0	0		
398	MALE	0	0	0	

399  
400  
401       **ALPHA**  
402           F                   AGE                   MALE  
403           -----  
404            0                  0                  0  
405  
406  
407        **BETA**  
408           F                   AGE                   MALE  
409           -----  
410    F                      0                  18                19  
411    AGE                    0                  0                0  
412    MALE                  0                  0                0  
413  
414  
415        **PSI**  
416           F                   AGE                   MALE  
417           -----  
418    F                      20                    0  
419    AGE                    0                  0  
420    MALE                  0                  0                0  
421  
422  
423        **STARTING VALUES**  
424  
425  
426        **NU**  
427           SAD               BLUES               DEPRESS           HAPPY            ENJOY  
428           -----  
429           0.454             0.337             0.510            0.630            0.465  
430

431  
432       NU  
433       HOPEFUL      AGE      MALE  
434       -----  
435       0.973       0.000     0.000  
436  
437  
438       LAMBDA  
439       F            AGE      MALE  
440       -----  
441 SAD       1.000     0.000     0.000  
442 BLUES     0.991     0.000     0.000  
443 DEPRESS   1.218     0.000     0.000  
444 HAPPY     1.176     0.000     0.000  
445 ENJOY     1.031     0.000     0.000  
446 HOPEFUL  0.847     0.000     0.000  
447 AGE       0.000     1.000     0.000  
448 MALE      0.000     0.000     1.000  
449  
450  
451       THETA  
452       SAD          BLUES     DEPRESS    HAPPY      ENJOY  
453       -----  
454 SAD       0.274     0.000     0.000     0.000     0.000  
455 BLUES     0.000     0.273     0.000     0.000     0.000  
456 DEPRESS   0.000     0.000     0.337     0.000     0.000  
457 HAPPY     0.000     0.000     0.000     0.490     0.000  
458 ENJOY     0.000     0.000     0.000     0.000     0.408  
459 HOPEFUL  0.000     0.000     0.000     0.000     0.000  
460 AGE       0.000     0.000     0.000     0.000     0.000  
461 MALE      0.000     0.000     0.000     0.000     0.000  
462

463  
464       THETA  
465           HOPEFUL      AGE          MALE  
466           -----      -----  
467 HOPEFUL    0.759  
468 AGE        0.000       0.000  
469 MALE       0.000       0.000       0.000  
470  
471  
472       ALPHA  
473           F           AGE          MALE  
474           -----      -----  
475           0.000       0.000       0.000  
476  
477  
478       BETA  
479           F           AGE          MALE  
480           -----      -----  
481 F        0.000       0.000       0.000  
482 AGE     0.000       0.000       0.000  
483 MALE    0.000       0.000       0.000  
484  
485  
486       PSI  
487           F           AGE          MALE  
488           -----      -----  
489 F        0.050  
490 AGE     0.000       40.696  
491 MALE    0.000       -0.190       0.244  
492  
493  
494 Beginning Time: 14:07:22

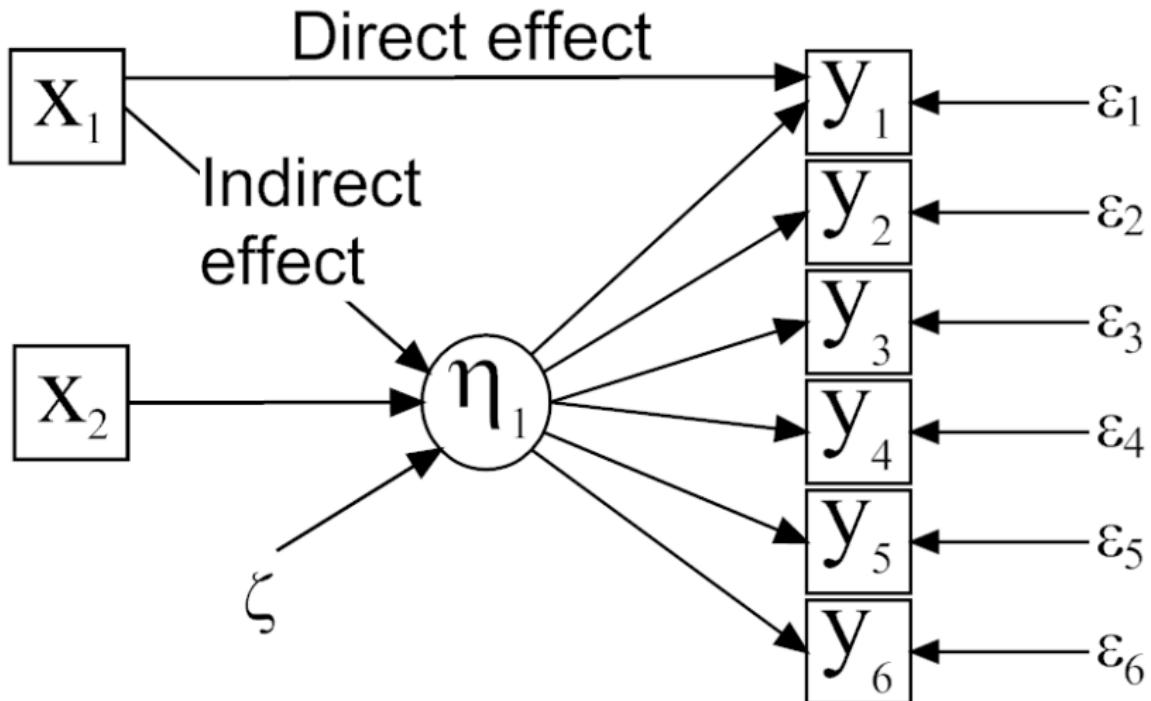
```
495      Ending Time: 14:07:22
496      Elapsed Time: 00:00:00
497
498
499 Mplus VERSION 8.8 DEMO (Mac) has the following limitations:
500 Maximum number of dependent variables: 6
501 Maximum number of independent variables: 2
502 Maximum number of between variables: 2
503 Maximum number of continuous latent variables in time series analysis: 2
504
505
506 MUTHEN & MUTHEN
507 3463 Stoner Ave.
508 Los Angeles, CA 90066
509
510 Tel: (310) 391-9971
511 Fax: (310) 391-8971
512 Web: www.StatModel.com
513 Support: Support@StatModel.com
514
515 Copyright (c) 1998-2022 Muthen & Muthen
```

## ex0102.out

## Example 1.3

# CFA with Covariates and Direct Effects

CES-D Depression, Age, Sex  
New Haven EPESE



```
1 Mplus VERSION 8.8 DEMO (Mac)
2 MUTHEN & MUTHEN
3 07/11/2022    2:07 PM
4
5 INPUT INSTRUCTIONS
6
7 TITLE:    EX0103 CFA WITH COVARIATES AND DIRECT EFFECT EPESE CESD ;
8
9 DATA:    FILE = ex0101.dat ;
10
11 VARIABLE: NAMES = sad blues depress happy enjoy hopeful age male ;
12           MISSING ARE ALL (-9999) ;
13
14 OUTPUT: TECH1 ;
15
16 MODEL:   f by sad-hopeful ;
17           f on age male ;
18           sad-hopeful ON age-male@0 ;
19           hopeful ON age ;
20
21
22
23
24
25 INPUT READING TERMINATED NORMALLY
26
27
28
29 EX0103 CFA WITH COVARIATES AND DIRECT EFFECT EPESE CESD ;
30
31 SUMMARY OF ANALYSIS
32
```

```
33 Number of groups 1
34 Number of observations 2442
35
36 Number of dependent variables 6
37 Number of independent variables 2
38 Number of continuous latent variables 1
39
40 Observed dependent variables
41
42 Continuous
43 SAD      BLUES      DEPRESS     HAPPY      ENJOY      HOPEFUL
44
45 Observed independent variables
46 AGE      MALE
47
48 Continuous latent variables
49 F
50
51
52 Estimator          ML
53 Information matrix OBSERVED
54 Maximum number of iterations 1000
55 Convergence criterion 0.500D-04
56 Maximum number of steepest descent iterations 20
57 Maximum number of iterations for H1 2000
58 Convergence criterion for H1 0.100D-03
59
60 Input data file(s)
61   ex0101.dat
62
63 Input data format FREE
64
```

65  
66 SUMMARY OF DATA  
67  
68 Number of missing data patterns 1  
69  
70  
71 COVARIANCE COVERAGE OF DATA  
72  
73 Minimum covariance coverage value 0.100  
74  
75  
76 PROPORTION OF DATA PRESENT  
77  
78  
79 Covariance Coverage  
80 SAD BLUES DEPRESS HAPPY ENJOY  
81 -----  
82 SAD 1.000  
83 BLUES 1.000 1.000  
84 DEPRESS 1.000 1.000 1.000  
85 HAPPY 1.000 1.000 1.000 1.000  
86 ENJOY 1.000 1.000 1.000 1.000 1.000  
87 HOPEFUL 1.000 1.000 1.000 1.000 1.000  
88 AGE 1.000 1.000 1.000 1.000 1.000  
89 MALE 1.000 1.000 1.000 1.000 1.000  
90  
91  
92 Covariance Coverage  
93 HOPEFUL AGE MALE  
94 -----  
95 HOPEFUL 1.000  
96 AGE 1.000 1.000

97	MALE	1.000	1.000	1.000			
98							
99							
100							
101	UNIVARIATE SAMPLE STATISTICS						
102							
103							
104	UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS						
105							
106	Variable/ Percentiles	Mean/ Variance	Skewness/ Kurtosis	Minimum/ Maximum	% with Min/Max	20%/60%	40%/ 80%
107	Sample Size 80% 0.000	Median 0.000 2442.000 1.000					
108							
109	SAD	0.454 0.000 2442.000 1.000	1.831 3.185	0.000 3.000	65.77% 3.81%	0.000 0.000	
110							
111	BLUES	0.337 0.000 2442.000 1.000	2.417 5.259	0.000 3.000	78.05% 4.34%	0.000 0.000	
112							
113	DEPRESS	0.510 0.000 2442.000 1.000	1.749 2.457	0.000 3.000	64.41% 5.81%	0.000 0.000	
114							
115	HAPPY	0.630 0.000 2442.000 2.000	1.241 0.069	0.000 3.000	67.08% 7.25%	0.000 0.000	
116							
117	ENJOY	0.465 0.000 0.000	1.742	0.000	75.96%	0.000	

118	2442.000 1.000	0.816	1.624	3.000	5.94%	0.000	
119	HOPEFUL 0.000 2442.000 2.000	0.973 0.000	0.681 1.517	0.000 -1.263	57.70% 19.57%	0.000 1.000	
120							
121	AGE 72.000 2442.000 82.000	74.316 72.000 40.696	0.497 0.307	67.000 0.000	28.50% 57.58%	67.000 0.000	
122							
123	MALE 0.000 2442.000 1.000	0.424 0.000	-0.862 0.244	-1.906	87.000 1.000	8.39% 42.42%	77.000 1.000
124							
125							
126							
127	THE MODEL ESTIMATION TERMINATED NORMALLY						
128							
129							
130							
131	MODEL FIT INFORMATION						
132							
133	Number of Free Parameters		21				
134							
135	Loglikelihood						
136							
137	H0 Value		-17251.875				
138	H1 Value		-17006.684				
139							
140	Information Criteria						
141							
142	Akaike (AIC)		34545.751				

```
143      Bayesian (BIC)          34667.563
144      Sample-Size Adjusted BIC 34600.841
145      (n* = (n + 2) / 24)
146
147 Chi-Square Test of Model Fit
148
149      Value                  490.382
150      Degrees of Freedom     18
151      P-Value                0.0000
152
153 RMSEA (Root Mean Square Error Of Approximation)
154
155      Estimate              0.104
156      90 Percent C.I.        0.096  0.112
157      Probability RMSEA <= .05 0.000
158
159 CFI/TLI
160
161      CFI                   0.885
162      TLI                   0.827
163
164 Chi-Square Test of Model Fit for the Baseline Model
165
166      Value                  4124.409
167      Degrees of Freedom     27
168      P-Value                0.0000
169
170 SRMR (Standardized Root Mean Square Residual)
171
172      Value                  0.048
173
174
```

				Two-Tailed	
		Estimate	S.E.	Est./S.E.	P-Value
175					
176	MODEL RESULTS				
177					
178					
179					
180					
181	F BY				
182	SAD	1.000	0.000	999.000	999.000
183	BLUES	0.963	0.035	27.792	0.000
184	DEPRESS	1.191	0.040	30.113	0.000
185	HAPPY	1.230	0.049	25.034	0.000
186	ENJOY	1.074	0.045	24.002	0.000
187	HOPEFUL	0.872	0.057	15.319	0.000
188					
189	F ON				
190	AGE	0.005	0.002	3.007	0.003
191	MALE	-0.134	0.023	-5.779	0.000
192					
193	SAD ON				
194	AGE	0.000	0.000	999.000	999.000
195	MALE	0.000	0.000	999.000	999.000
196					
197	BLUES ON				
198	AGE	0.000	0.000	999.000	999.000
199	MALE	0.000	0.000	999.000	999.000
200					
201	DEPRESS ON				
202	AGE	0.000	0.000	999.000	999.000
203	MALE	0.000	0.000	999.000	999.000
204					
205	HAPPY ON				
206	AGE	0.000	0.000	999.000	999.000

207 MALE 0.000 0.000 999.000 999.000  
 208  
 209 ENJOY ON  
 210 AGE 0.000 0.000 999.000 999.000  
 211 MALE 0.000 0.000 999.000 999.000  
 212  
 213 HOPEFUL ON  
 214 AGE 0.010 0.004 2.572 0.010  
 215 MALE 0.000 0.000 999.000 999.000  
 216  
 217 Intercepts  
 218 SAD 0.110 0.135 0.818 0.413  
 219 BLUES 0.006 0.130 0.047 0.963  
 220 DEPRESS 0.101 0.160 0.627 0.531  
 221 HAPPY 0.207 0.166 1.245 0.213  
 222 ENJOY 0.096 0.145 0.657 0.511  
 223 HOPEFUL -0.037 0.291 -0.126 0.900  
 224  
 225 Residual Variances  
 226 SAD 0.287 0.011 27.267 0.000  
 227 BLUES 0.304 0.011 27.708 0.000  
 228 DEPRESS 0.305 0.013 24.009 0.000  
 229 HAPPY 0.586 0.021 27.941 0.000  
 230 ENJOY 0.516 0.018 28.747 0.000  
 231 HOPEFUL 1.312 0.039 33.574 0.000  
 232 F 0.254 0.015 17.467 0.000  
 233  
 234  
 235 QUALITY OF NUMERICAL RESULTS  
 236  
 237 Condition Number for the Information Matrix 0.178E-06  
 238 (ratio of smallest to largest eigenvalue)

239  
240  
241 TECHNICAL 1 OUTPUT  
242  
243  
244 PARAMETER SPECIFICATION  
245  
246  
247       NU  
248       SAD           BLUES           DEPRESS           HAPPY           ENJOY  
249       -----  
250        0            0              0              0            0  
251  
252  
253       NU  
254       HOPEFUL       AGE           MALE  
255       -----  
256        0            0              0  
257  
258  
259       LAMBDA  
260       F            SAD           BLUES           DEPRESS          HAPPY  
261       -----  
262 SAD        0            0            0              0            0  
263 BLUES      0            0            0              0            0  
264 DEPRESS     0            0            0              0            0  
265 HAPPY      0            0            0              0            0  
266 ENJOY      0            0            0              0            0  
267 HOPEFUL    0            0            0              0            0  
268 AGE        0            0            0              0            0  
269 MALE       0            0            0              0            0  
270

271  
272        LAMBDA  
273        ENJOY           HOPEFUL           AGE           MALE  
274        -----  
275 SAD        0            0            0            0  
276 BLUES       0            0            0            0  
277 DEPRESS      0            0            0            0  
278 HAPPY        0            0            0            0  
279 ENJOY        0            0            0            0  
280 HOPEFUL      0            0            0            0  
281 AGE           0            0            0            0  
282 MALE         0            0            0            0  
283  
284  
285        THETA  
286        SAD            BLUES           DEPRESS          HAPPY           ENJOY  
287        -----  
288 SAD         0            0            0            0            0  
289 BLUES       0            0            0            0            0  
290 DEPRESS      0            0            0            0            0  
291 HAPPY        0            0            0            0            0  
292 ENJOY        0            0            0            0            0  
293 HOPEFUL      0            0            0            0            0  
294 AGE           0            0            0            0            0  
295 MALE         0            0            0            0            0  
296  
297  
298        THETA  
299        HOPEFUL          AGE           MALE  
300  
301 HOPEFUL      0            0            0  
302 AGE           0            0            0

303	MALE	0	0	0	
304					
305					
306	ALPHA				
307	F	SAD	BLUES	DEPRESS	HAPPY
308	-----	-----	-----	-----	-----
309	0	1	2	3	4
310					
311					
312	ALPHA				
313	ENJOY	HOPEFUL	AGE	MALE	
314	-----	-----	-----	-----	
315	5	6	0	0	
316					
317					
318	BETA				
319	F	SAD	BLUES	DEPRESS	HAPPY
320	-----	-----	-----	-----	-----
321	F	0	0	0	0
322	SAD	0	0	0	0
323	BLUES	9	0	0	0
324	DEPRESS	10	0	0	0
325	HAPPY	11	0	0	0
326	ENJOY	12	0	0	0
327	HOPEFUL	13	0	0	0
328	AGE	0	0	0	0
329	MALE	0	0	0	0
330					
331					
332	BETA				
333	ENJOY	HOPEFUL	AGE	MALE	
334	-----	-----	-----	-----	

335	F	0	0	7	8
336	SAD	0	0	0	0
337	BLUES	0	0	0	0
338	DEPRESS	0	0	0	0
339	HAPPY	0	0	0	0
340	ENJOY	0	0	0	0
341	HOPEFUL	0	0	14	0
342	AGE	0	0	0	0
343	MALE	0	0	0	0

344

345

	PSI	F	SAD	BLUES	DEPRESS	HAPPY
349	F	15	-----	-----	-----	-----
350	SAD	0	16			
351	BLUES	0	0	17		
352	DEPRESS	0	0	0	18	
353	HAPPY	0	0	0	0	19
354	ENJOY	0	0	0	0	0
355	HOPEFUL	0	0	0	0	0
356	AGE	0	0	0	0	0
357	MALE	0	0	0	0	0

358

359

	PSI	ENJOY	HOPEFUL	AGE	MALE
361	ENJOY	20	-----	-----	-----
362	HOPEFUL	0	21		
363	AGE	0	0	0	
364	MALE	0	0	0	0

367  
368  
369 STARTING VALUES  
370  
371  
372 NU  
373 SAD BLUES DEPRESS HAPPY ENJOY  
374 -----  
375 0.000 0.000 0.000 0.000 0.000  
376  
377  
378 NU  
379 HOPEFUL AGE MALE  
380 -----  
381 0.000 0.000 0.000  
382  
383  
384 LAMBDA  
385 F SAD BLUES DEPRESS HAPPY  
386 -----  
387 SAD 0.000 1.000 0.000 0.000 0.000  
388 BLUES 0.000 0.000 1.000 0.000 0.000  
389 DEPRESS 0.000 0.000 0.000 1.000 0.000  
390 HAPPY 0.000 0.000 0.000 0.000 1.000  
391 ENJOY 0.000 0.000 0.000 0.000 0.000  
392 HOPEFUL 0.000 0.000 0.000 0.000 0.000  
393 AGE 0.000 0.000 0.000 0.000 0.000  
394 MALE 0.000 0.000 0.000 0.000 0.000  
395  
396  
397 LAMBDA  
398 ENJOY HOPEFUL AGE MALE

399  
400 SAD 0.000 0.000 0.000 0.000  
401 BLUES 0.000 0.000 0.000 0.000  
402 DEPRESS 0.000 0.000 0.000 0.000  
403 HAPPY 0.000 0.000 0.000 0.000  
404 ENJOY 1.000 0.000 0.000 0.000  
405 HOPEFUL 0.000 1.000 0.000 0.000  
406 AGE 0.000 0.000 1.000 0.000  
407 MALE 0.000 0.000 0.000 1.000  
408  
409  
410 THETA  
411 SAD BLUES DEPRESS HAPPY ENJOY  
412 -----  
413 SAD 0.000 -----  
414 BLUES 0.000 0.000 -----  
415 DEPRESS 0.000 0.000 0.000 -----  
416 HAPPY 0.000 0.000 0.000 0.000  
417 ENJOY 0.000 0.000 0.000 0.000 0.000  
418 HOPEFUL 0.000 0.000 0.000 0.000 0.000  
419 AGE 0.000 0.000 0.000 0.000 0.000  
420 MALE 0.000 0.000 0.000 0.000 0.000  
421  
422  
423 THETA  
424 HOPEFUL AGE MALE  
425 -----  
426 HOPEFUL 0.000 -----  
427 AGE 0.000 0.000  
428 MALE 0.000 0.000 0.000  
429  
430

	ALPHA	SAD	BLUES	DEPRESS	HAPPY
431	F				
432		-----	-----	-----	-----
433					
434	0.000	0.454	0.337	0.510	0.630
435					
436					
	ALPHA	HOPEFUL	AGE	MALE	
437	ENJOY				
438		-----	-----	-----	-----
439					
440	0.465	0.973	74.316	0.424	
441					
442					
	BETA	SAD	BLUES	DEPRESS	HAPPY
443	F				
444		-----	-----	-----	-----
445					
446	F	0.000	0.000	0.000	0.000
447	SAD	1.000	0.000	0.000	0.000
448	BLUES	0.991	0.000	0.000	0.000
449	DEPRESS	1.218	0.000	0.000	0.000
450	HAPPY	1.176	0.000	0.000	0.000
451	ENJOY	1.031	0.000	0.000	0.000
452	HOPEFUL	0.847	0.000	0.000	0.000
453	AGE	0.000	0.000	0.000	0.000
454	MALE	0.000	0.000	0.000	0.000
455					
456					
	BETA	HOPEFUL	AGE	MALE	
457	ENJOY				
458		-----	-----	-----	-----
459					
460	F	0.000	0.000	0.000	0.000
461	SAD	0.000	0.000	0.000	0.000
462	BLUES	0.000	0.000	0.000	0.000

463 DEPRESS 0.000 0.000 0.000 0.000  
464 HAPPY 0.000 0.000 0.000 0.000  
465 ENJOY 0.000 0.000 0.000 0.000  
466 HOPEFUL 0.000 0.000 0.000 0.000  
467 AGE 0.000 0.000 0.000 0.000  
468 MALE 0.000 0.000 0.000 0.000

469

470

471           PSI

	F	SAD	BLUES	DEPRESS	HAPPY
474 F	0.050	-----	-----	-----	-----
475 SAD	0.000	0.274			
476 BLUES	0.000	0.000	0.273		
477 DEPRESS	0.000	0.000	0.000	0.337	
478 HAPPY	0.000	0.000	0.000	0.000	0.490
479 ENJOY	0.000	0.000	0.000	0.000	0.000
480 HOPEFUL	0.000	0.000	0.000	0.000	0.000
481 AGE	0.000	0.000	0.000	0.000	0.000
482 MALE	0.000	0.000	0.000	0.000	0.000

483

484

485           PSI

	ENJOY	HOPEFUL	AGE	MALE
486 ENJOY	0.408	-----	-----	-----
487 HOPEFUL	0.000	0.759		
488 AGE	0.000	0.000	40.696	
489 MALE	0.000	0.000	-0.190	0.244

490

491

492

493

494 Beginning Time: 14:07:23

## Example 1.4

# Growth modeling

## Alcohol Use from Singer and Willett (2003)

Model a: Intercept only model

<https://stats.idre.ucla.edu/other/examples/alda/>

The screenshot shows the UCLA Institute for Digital Research & Education website. At the top, there is a header bar with the UCLA logo and the text "Institute for Digital Research & Education". To the right of the header is a search bar labeled "Search this website". Below the header is a navigation bar with links for "HOME", "SOFTWARE", "RESOURCES", "SERVICES", and "ABOUT US". The main content area features a section titled "TEXTBOOK EXAMPLES APPLIED LONGITUDINAL DATA ANALYSIS: MODELING CHANGE AND EVENT OCCURRENCE BY JUDITH D. SINGER AND JOHN B. WILLETT". Below this title, there is a paragraph of text explaining the availability of the book from Academic Technology Services and encouraging users to purchase it from Oxford University Press. It also mentions the permission granted to use the data files. A note below states that the files are for older versions of MLwiN and that newer versions are compatible. Another note for HLM specifies that it consists of two types of files: SPSS format and .mdmt template files.

This is one of the books available for loan from Academic Technology Services (see [Statistics Books for Loan](#) for other such books, and details about borrowing). We encourage you to obtain *Applied Longitudinal Data Analysis*, written by [Judith D. Singer and John B. Willett](#), published by the [Oxford University Press](#), to gain a deeper conceptual understanding of the analysis illustrated (see [Where to buy books](#) for tips on different places you can buy this book). We are very grateful to the authors for granting us permission to create these pages and to distribute the data files via our web pages. (In addition to the data formats below, you can also download the data files as [comma separated text files](#)).

Note: The files for MLwiN here were created for older version of MLwiN (version 2.0 and older). The newer version of MLwiN is compatible with these files. But the newer version of the same files created at MLwiN site will only work for newer version of MLwiN (version 2.10). Please visit [MLwiN site](#) to download the files and for other MLwiN related information.

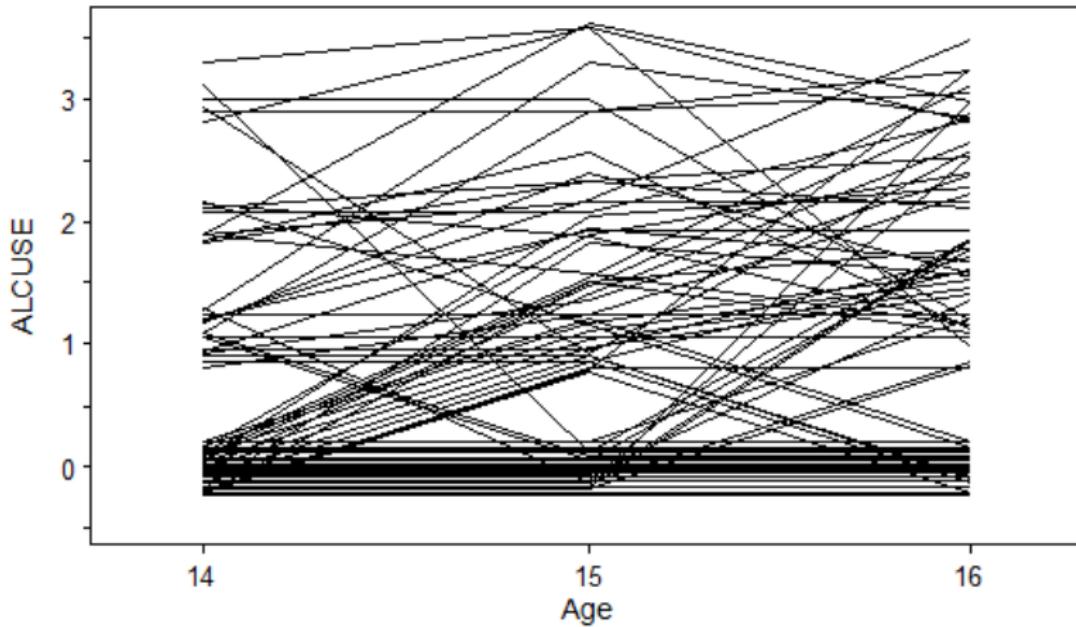
Note: The files for HLM consists of two types of files, the data files in SPSS format and the corresponding .mdmt (template) file for creating the binary HLM files. (Updated in January, 2012).

<https://stats.idre.ucla.edu/r/examples/alda/applied-longitudinal-data-analysis-chapter-15-r-textbook-examples/>

	Mplus	SAS	Stata	R	SPSS	Chapter Title
Download Data	Download	<a href="#">Download</a>	<a href="#">Download</a>	<a href="#">Download</a>	<a href="#">Download</a>	<a href="#">Table of Contents</a>
Chapter 1						A framework for investigating change over time
Chapter 2		<a href="#">Chap 2</a>	<a href="#">Chap 2</a>	<a href="#">Chap 2</a>	<a href="#">Chap 2</a>	Exploring longitudinal data on change
Chapter 3		<a href="#">Chap 3</a>	<a href="#">Chap 3</a>	<a href="#">Chap 3</a>	<a href="#">Chap 3</a>	Introducing the multilevel model for change

# Chapter 4. Doing Data Analysis with the Multilevel Model for Change. Singer and Willet Applied Longitudinal Data Analysis. Oxford. (2003)

- Adolescent Alcohol Use Example (Curran et al. 2002).
  - 82 Adolescents, age 14, seen 3 times +1 year
  - 4 item Questionnaire Alcohol use past year
    - ▶ (1) drank beer or wine, (2) drank liquor, (3) had 5+ drinks in a row, (4) got drunk
    - ▶ rated 0-7 (0=not at all, 7=every day)
    - ▶ alcuse = (Sy)0.5
- Covariates/predictors
  - ▶ coa : child of alcoholic (0/1)
  - ▶ peer : alcohol use (alcuse) among peers [0-2.53]



data from Singer and Willet (2003) Chapter 4  
[http://www.ats.ucla.edu/stat/stata/examples/alda/data/alcohol1\\_pp.dta](http://www.ats.ucla.edu/stat/stata/examples/alda/data/alcohol1_pp.dta)  
Random noise added to ALCUSE, between  $\pm 0.5 \times \text{SD}(\text{ALCUSE}|\text{AGE}_14)$

Table 4.1: Alcohol use data (alcohol1\_pp): Fitted multilevel models for change  
**Model A**

## Covariance Parameter Estimates

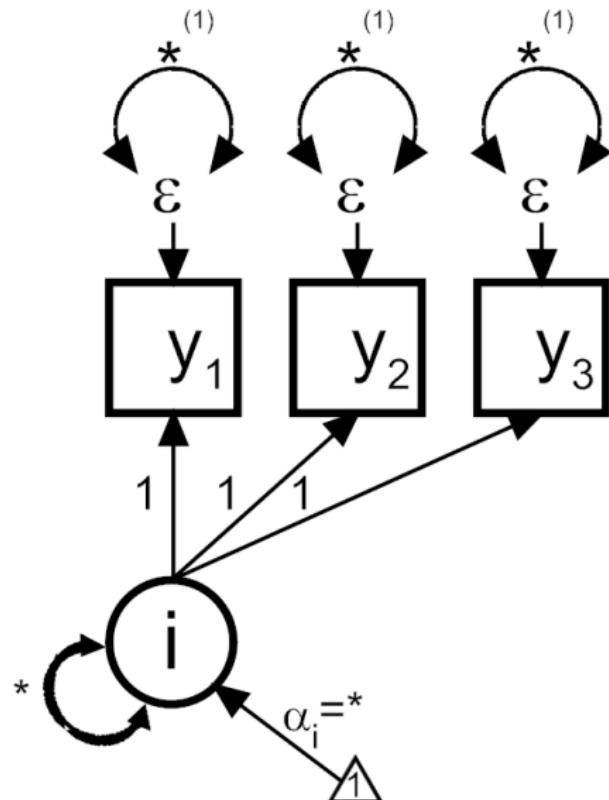
Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
Intercept	ID	0.5639	0.1191	4.73	<.0001
Residual		0.5617	0.06203	9.06	<.0001

## Fit Statistics

-2 Log Likelihood	670.2
AIC (smaller is better)	676.2
AICC (smaller is better)	676.3
BIC (smaller is better)	683.4

## Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value		Pr >  t
				t Value	Pr >  t	
Intercept	0.9220	0.09571	81	9.63	<.0001	



```
1 Mplus VERSION 8.8 DEMO (Mac)
2 MUTHEN & MUTHEN
3 07/11/2022    2:07 PM
4
5 INPUT INSTRUCTIONS
6
7 TITLE:    EX0104 Singer and Willet (2003) Chapter 4 Model A ;
8
9 DATA:    FILE = swch4.dat ;
10
11 VARIABLE: NAMES = alcuse1 alcuse2 alcuse3 ;
12
13 ANALYSIS: ESTIMATOR = mlr ;
14
15 MODEL:    i by alcuse1-alcuse3@1 ;
16          [alcuse1-alcuse3*] ; ! force mean to intercept
17          ![i*] ;           ! intercept
18          alcuse1-alcuse3 (1) ; ! homoscedasticity
19
20
21
22
23
24 INPUT READING TERMINATED NORMALLY
25
26
27
28 EX0104 Singer and Willet (2003) Chapter 4 Model A ;
29
30 SUMMARY OF ANALYSIS
31
32 Number of groups
```

```
33 Number of observations          82
34
35 Number of dependent variables   3
36 Number of independent variables 0
37 Number of continuous latent variables 1
38
39 Observed dependent variables
40
41 Continuous
42   ALCUSE1      ALCUSE2      ALCUSE3
43
44 Continuous latent variables
45   I
46
47
48 Estimator                      MLR
49 Information matrix              OBSERVED
50 Maximum number of iterations    1000
51 Convergence criterion           0.500D-04
52 Maximum number of steepest descent iterations 20
53
54 Input data file(s)
55   swch4.dat
56
57 Input data format   FREE
58
59
60
61 UNIVARIATE SAMPLE STATISTICS
62
63
64     UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS
```

```

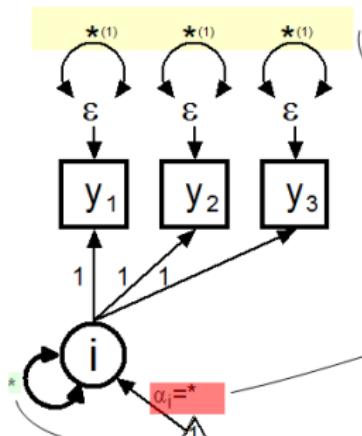
65
66      Variable/          Mean/      Skewness/    Minimum/ % with
67      Percentiles
68      Sample Size       Variance   Kurtosis     Maximum Min/Max 20%/60% 40%
69      80%      Median
70
71      ALCUSE1           0.630      1.320       0.000 62.20% 0.000
72      0.000      0.000
73      82.000      0.871      0.650       3.317 1.22% 0.000
74      1.414
75
76
77 THE MODEL ESTIMATION TERMINATED NORMALLY
78
79
80
81 MODEL FIT INFORMATION
82
83 Number of Free Parameters                   5
84
85 Loglikelihood
86
87      HO Value             -323.402
88      HO Scaling Correction Factor        1.0942

```

```
89      for MLR
90      H1 Value           -316.186
91      H1 Scaling Correction Factor   1.2269
92      for MLR
93
94 Information Criteria
95
96      Akaike (AIC)          656.804
97      Bayesian (BIC)        668.837
98      Sample-Size Adjusted BIC 653.067
99      (n* = (n + 2) / 24)
100
101 Chi-Square Test of Model Fit
102
103      Value             10.361*
104      Degrees of Freedom    4
105      P-Value            0.0348
106      Scaling Correction Factor 1.3928
107      for MLR
108
109 *   The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used
110 for chi-square difference testing in the regular way. MLM, MLR and WLSM
111 chi-square difference testing is described on the Mplus website. MLMV, WLSMV,
112 and ULSMV difference testing is done using the DIFFTEST option.
113
114 RMSEA (Root Mean Square Error Of Approximation)
115
116      Estimate          0.139
117      90 Percent C.I.    0.034  0.247
118      Probability RMSEA <= .05  0.070
119
120 CFI/TLI
```

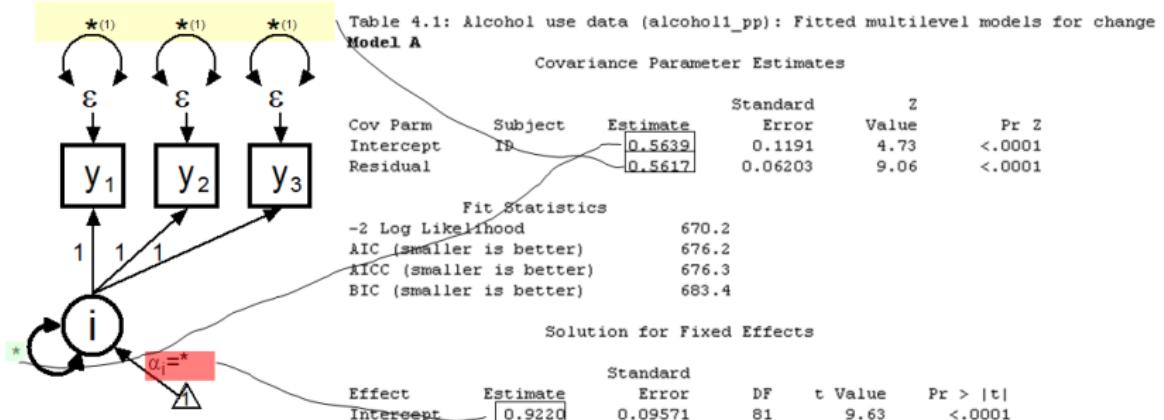
```
121
122      CFI                      0.848
123      TLI                      0.886
124
125 Chi-Square Test of Model Fit for the Baseline Model
126
127      Value                     44.816
128      Degrees of Freedom        3
129      P-Value                   0.0000
130
131 SRMR (Standardized Root Mean Square Residual)
132
133      Value                     0.117
134
135
136
137 MODEL RESULTS
138
139
140      Estimate      S.E.   Est./S.E.   Two-Tailed P-Value
141
142 I      BY
143 ALCUSE1    1.000    0.000    999.000    999.000
144 ALCUSE2    1.000    0.000    999.000    999.000
145 ALCUSE3    1.000    0.000    999.000    999.000
146
147 Intercepts
148 ALCUSE1    0.630    0.103    6.118     0.000
149 ALCUSE2    0.964    0.115    8.383     0.000
150 ALCUSE3    1.172    0.125    9.403     0.000
151
152 Variances
```

```
153      I          0.589      0.103      5.714      0.000
154
155 Residual Variances
156   ALCUSE1          0.487      0.071      6.845      0.000
157   ALCUSE2          0.487      0.071      6.845      0.000
158   ALCUSE3          0.487      0.071      6.845      0.000
159
160
161 QUALITY OF NUMERICAL RESULTS
162
163 Condition Number for the Information Matrix          0.102E+00
164 (ratio of smallest to largest eigenvalue)
165
166
167 Beginning Time: 14:07:23
168 Ending Time: 14:07:23
169 Elapsed Time: 00:00:00
170
171
172 Mplus VERSION 8.8 DEMO (Mac) has the following limitations:
173 Maximum number of dependent variables: 6
174 Maximum number of independent variables: 2
175 Maximum number of between variables: 2
176 Maximum number of continuous latent variables in time series analysis: 2
177
178
179 MUTHEN & MUTHEN
180 3463 Stoner Ave.
181 Los Angeles, CA 90066
182
183 Tel: (310) 391-9971
184 Fax: (310) 391-8971
```



## MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
I BY				
ALCUSE1	1.000	0.000	999.000	999.000
ALCUSE2	1.000	0.000	999.000	999.000
ALCUSE3	1.000	0.000	999.000	999.000
Means I	0.922	0.096	9.633	0.000
Intercepts				
ALCUSE1	0.000	0.000	999.000	999.000
ALCUSE2	0.000	0.000	999.000	999.000
ALCUSE3	0.000	0.000	999.000	999.000
Variances I	0.564	0.106	5.314	0.000
Residual Variances				
ALCUSE1	0.562	0.083	6.788	0.000
ALCUSE2	0.562	0.083	6.788	0.000
ALCUSE3	0.562	0.083	6.788	0.000



Note that the point estimates are the same. Standard errors of covariance parameters are slightly higher in PROC MIXED. It's a difference of N-1 (PROC Mixed) vs N in the denominator.

SAS Proc Mixed  $\approx$  Mplus

intercept  $\approx$  Means(l)

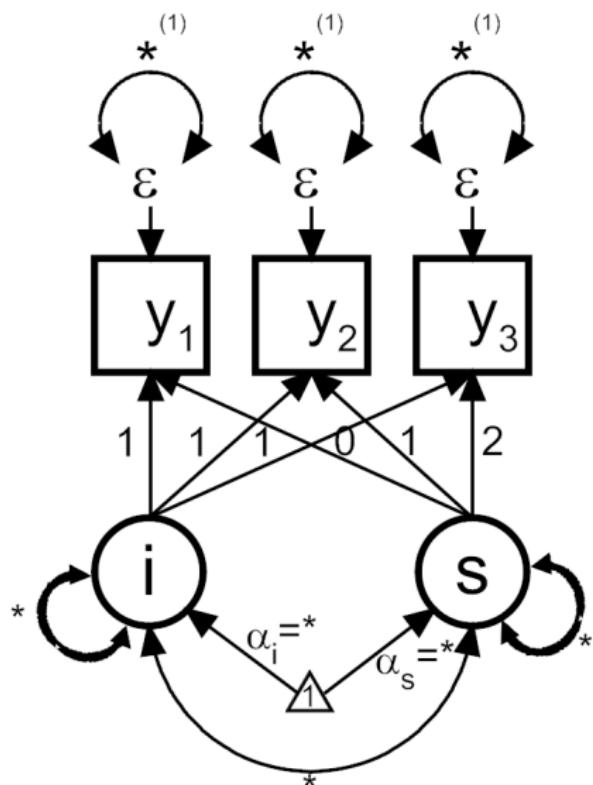
Cov parm intercept  $\approx$  Variance(l)

Cov parm residual  $\approx$  Residual Variances(y)

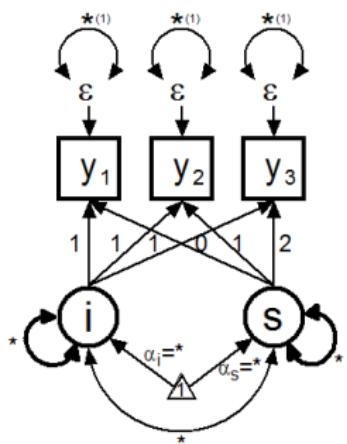
## Example 1.5

# Growth modeling Alcohol Use from Singer and Willet (2003)

Model b: Intercept and Slope model



## Model B (Table 4.1)



```

* SAS Example from http://www.ats.ucla.edu/stat/examples/alda.htm
*-----
* SAS
proc mixed data="c:\alda\alcohol1_pp" method=ml noclprint noinfo covtest;
  title2 "Model B";
  class id;
  model alcuse = age_14/solution nobet;
  random intercept age_14/type=un sub=id;
*-----

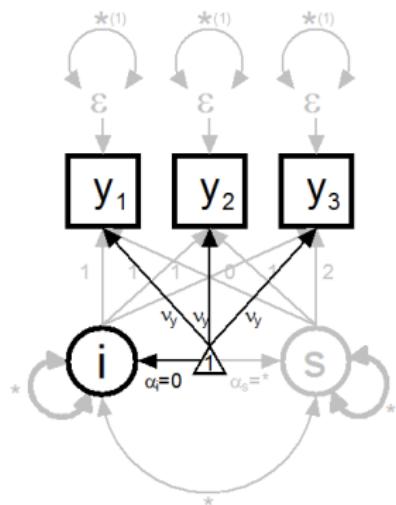
* Mplus (short hand)
DATA:      FILE = C:\work\Shows\SHORTC~1\2009\data\swch4.dat ;
VARIABLE: NAMES = alcuse1 alcuse2 alcuse3 ;
ANALYSIS: ESTIMATOR = mlr ;
MODEL:    i s | alcuse1@0 alcuse2@1 alcuse3@2 ;
           alcuse1-alcuse3 (1) ;

*-----

* Mplus (long hand)
DATA:      FILE = C:\work\Shows\SHORTC~1\2009\data\swch4.dat ;
VARIABLE: NAMES = alcuse1 alcuse2 alcuse3 ;
ANALYSIS: ESTIMATOR = mlr ;
MODEL:    i by alcuse1-alcuse3@1 ;
           s by alcuse1@0 alcuse2@1 alcuse3@2 ;
           [alcuse1-alcuse3@0] ;
           [i* s*] ;
           alcuse1-alcuse3 (1) ;

```

## Model B (Table 4.1) Alternative Mplus Specification

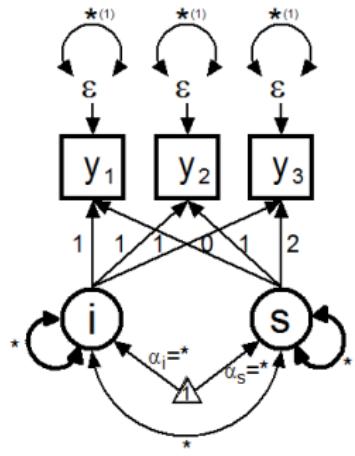


note:  $v_y = v_{y(1)} = v_{y(2)} = v_{y(3)}$

```

* Mplus (alternate specification - long hand)
DATA:      FILE = swch4.dat ;
VARIABLE: NAMES = alcuse1 alcuse2 alcuse3 ;
ANALYSIS: ESTIMATOR = mlr ;
MODEL:      alcuse1-alcuse3 (1) ;
            i by alcuse1-alcuse3@1 ;
            s by alcuse1@0 alcuse2@1 alcuse3@2 ;
            [alcuse1-alcuse3*0] (11) ;
            [i@0 s*] ;

```



$$V(y_{ti}) = V(i)_i + V(s)_i t^2 + 2COV(i,s) + V(\epsilon_{ti})$$

random slope contributes to increasing variance over time even if  $\varepsilon_t = \varepsilon$

```
1 Mplus VERSION 8.8 DEMO (Mac)
2 MUTHEN & MUTHEN
3 07/11/2022    2:07 PM
4
5 INPUT INSTRUCTIONS
6
7 TITLE:    EX0105 Singer and Willet (2003) Chapter 4 Model B ;
8
9 DATA:    FILE = swch4.dat ;
10
11 VARIABLE: NAMES = alcuse1 alcuse2 alcuse3 ;
12           missing are all (-9999) ;
13
14 ANALYSIS: ESTIMATOR = mlr ;
15
16 OUTPUT:   SAMPSTAT ; stdyx ;
17
18 MODEL:
19           ! [alcuse1@0 alcuse2@0 alcuse3@0] ;
20           ! i by alcuse1@1 alcuse2@1 alcuse3@1 ;
21           ! s by alcuse1@0 alcuse2@0 alcuse3@2 ;
22           ! [i*]; [s*];
23           i s | alcuse1@0 alcuse2@1 alcuse3@2 ;
24           alcuse1-alcuse3 (1) ; ! homoscedasticity
25
26
27
28
29
30 INPUT READING TERMINATED NORMALLY
31
32
```

```
33
34 EX0105 Singer and Willet (2003) Chapter 4 Model B ;
35
36 SUMMARY OF ANALYSIS
37
38 Number of groups                                1
39 Number of observations                            82
40
41 Number of dependent variables                  3
42 Number of independent variables                0
43 Number of continuous latent variables          2
44
45 Observed dependent variables
46
47 Continuous
48   ALCUSE1      ALCUSE2      ALCUSE3
49
50 Continuous latent variables
51   I           S
52
53
54 Estimator                                     MLR
55 Information matrix                            OBSERVED
56 Maximum number of iterations                 1000
57 Convergence criterion                         0.500D-04
58 Maximum number of steepest descent iterations 20
59 Maximum number of iterations for H1          2000
60 Convergence criterion for H1                 0.100D-03
61
62 Input data file(s)
63   swch4.dat
64
```

```
65 Input data format   FREE
66
67
68 SUMMARY OF DATA
69
70      Number of missing data patterns           1
71
72
73 COVARIANCE COVERAGE OF DATA
74
75 Minimum covariance coverage value   0.100
76
77
78      PROPORTION OF DATA PRESENT
79
80
81      Covariance Coverage
82          ALCUSE1       ALCUSE2       ALCUSE3
83          -----       -----       -----
84  ALCUSE1       1.000
85  ALCUSE2       1.000       1.000
86  ALCUSE3       1.000       1.000       1.000
87
88
89 SAMPLE STATISTICS
90
91
92      ESTIMATED SAMPLE STATISTICS
93
94
95      Means
96          ALCUSE1       ALCUSE2       ALCUSE3
```

97 -----  
98 0.630 0.964 1.172  
99  
100  
101 Covariances  
102 ALCUSE1 ALCUSE2 ALCUSE3  
103 -----  
104 ALCUSE1 0.871 -----  
105 ALCUSE2 0.620 1.084  
106 ALCUSE3 0.432 0.714 1.273  
107  
108  
109 Correlations  
110 ALCUSE1 ALCUSE2 ALCUSE3  
111 -----  
112 ALCUSE1 1.000 -----  
113 ALCUSE2 0.638 1.000  
114 ALCUSE3 0.411 0.608 1.000  
115  
116  
117 MAXIMUM LOG-LIKELIHOOD VALUE FOR THE UNRESTRICTED (H1) MODEL IS -316.186  
118  
119  
120 UNIVARIATE SAMPLE STATISTICS  
121  
122  
123 UNIVARIATE HIGHER-ORDER MOMENT DESCRIPTIVE STATISTICS  
124  
125 Variable/ Mean/ Skewness/ Minimum/ % with  
Percentiles  
126 Sample Size Variance Kurtosis Maximum Min/Max 20%/60% 40%/  
80% Median

```
127  
128      ALCUSE1           0.630      1.320      0.000    62.20%    0.000  
129          0.000      0.000  
129          82.000      0.871      0.650      3.317    1.22%    0.000  
129          1.414  
130      ALCUSE2           0.964      0.831      0.000    43.90%    0.000  
131          0.000      1.000  
131          82.000      1.084     -0.187      3.606    2.44%    1.000  
131          2.000  
132      ALCUSE3           1.172      0.385      0.000    40.24%    0.000  
133          0.000      1.000  
133          82.000      1.273     -1.196      3.464    1.22%    1.414  
133          2.236  
134  
135  
136 THE MODEL ESTIMATION TERMINATED NORMALLY  
137  
138  
139  
140 MODEL FIT INFORMATION  
141  
142 Number of Free Parameters                      6  
143  
144 Loglikelihood  
145  
146      H0 Value                  -318.306  
147      H0 Scaling Correction Factor      1.1192  
148          for MLR  
149      H1 Value                  -316.186  
150      H1 Scaling Correction Factor      1.2269  
151          for MLR  
152
```

```
153 Information Criteria
154
155     Akaike (AIC)           648.611
156     Bayesian (BIC)         663.051
157     Sample-Size Adjusted BIC
158             (n* = (n + 2) / 24)   644.128
159
160 Chi-Square Test of Model Fit
161
162     Value                  2.939*
163     Degrees of Freedom      3
164     P-Value                 0.4011
165     Scaling Correction Factor
166             for MLR          1.4422
167
168 *  The chi-square value for MLM, MLMV, MLR, ULSMV, WLSM and WLSMV cannot be used
169     for chi-square difference testing in the regular way. MLM, MLR and WLSM
170     chi-square difference testing is described on the Mplus website. MLMV, WLSMV,
171     and ULSMV difference testing is done using the DIFFTEST option.
172
173 RMSEA (Root Mean Square Error Of Approximation)
174
175     Estimate                0.000
176     90 Percent C.I.          0.000  0.185
177     Probability RMSEA <= .05    0.490
178
179 CFI/TLI
180
181     CFI                    1.000
182     TLI                    1.000
183
184 Chi-Square Test of Model Fit for the Baseline Model
```

185  
186       Value    44.816  
187       Degrees of Freedom                            3  
188       P-Value                                        0.0000  
189  
190 SRMR (Standardized Root Mean Square Residual)  
191  
192       Value    0.055  
193  
194  
195  
196 MODEL RESULTS  
197  
198   Two-Tailed  
199   Estimate   S.E.    Est./S.E.   P-Value  
200  
201 I        |  
202    ALCUSE1        1.000        0.000       999.000    999.000  
203    ALCUSE2        1.000        0.000       999.000    999.000  
204    ALCUSE3        1.000        0.000       999.000    999.000  
205  
206 S        |  
207    ALCUSE1        0.000        0.000       999.000    999.000  
208    ALCUSE2        1.000        0.000       999.000    999.000  
209    ALCUSE3        2.000        0.000       999.000    999.000  
210  
211 S        WITH  
212    I              -0.068      0.074       -0.920    0.357  
213  
214 Means  
215    I              0.651       0.105       6.198      0.000  
216    S              0.271       0.062       4.334      0.000

```
217
218  Intercepts
219    ALCUSE1      0.000      0.000    999.000    999.000
220    ALCUSE2      0.000      0.000    999.000    999.000
221    ALCUSE3      0.000      0.000    999.000    999.000
222
223  Variances
224    I            0.624      0.162     3.861     0.000
225    S            0.151      0.056     2.721     0.007
226
227  Residual Variances
228    ALCUSE1      0.337      0.066     5.105     0.000
229    ALCUSE2      0.337      0.066     5.105     0.000
230    ALCUSE3      0.337      0.066     5.105     0.000
231
232
233 QUALITY OF NUMERICAL RESULTS
234
235      Condition Number for the Information Matrix          0.422E-01
236      (ratio of smallest to largest eigenvalue)
237
238
239 STANDARDIZED MODEL RESULTS
240
241
242 STDYX Standardization
243
244
245              Estimate      S.E.   Est./S.E.   Two-Tailed P-Value
246
247  I      |
248  ALCUSE1      0.806      0.046     17.355     0.000
```

249    ALCUSE2            0.800        0.083        9.686        0.000  
250    ALCUSE3            0.695        0.092        7.552        0.000  
251  
252    S                |  
253    ALCUSE1            0.000        0.000        999.000      999.000  
254    ALCUSE2            0.394        0.080        4.938        0.000  
255    ALCUSE3            0.684        0.120        5.679        0.000  
256  
257    S                WITH  
258    I                -0.223        0.193        -1.154      0.249  
259  
260    Means  
261    I                0.824        0.097        8.465        0.000  
262    S                0.696        0.179        3.891        0.000  
263  
264    Intercepts  
265    ALCUSE1            0.000        0.000        999.000      999.000  
266    ALCUSE2            0.000        0.000        999.000      999.000  
267    ALCUSE3            0.000        0.000        999.000      999.000  
268  
269    Variances  
270    I                1.000        0.000        999.000      999.000  
271    S                1.000        0.000        999.000      999.000  
272  
273    Residual Variances  
274    ALCUSE1            0.351        0.075        4.688        0.000  
275    ALCUSE2            0.346        0.062        5.588        0.000  
276    ALCUSE3            0.261        0.058        4.486        0.000  
277  
278  
279 R-SQUARE  
280

	Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
281	ALCUSE1	0.649	0.075	8.678	0.000
282	ALCUSE2	0.654	0.062	10.581	0.000
283	ALCUSE3	0.739	0.058	12.706	0.000
287					
288					
289	Beginning Time:	14:07:23			
290	Ending Time:	14:07:23			
291	Elapsed Time:	00:00:00			
292					
293					
294	Mplus VERSION 8.8 DEMO (Mac) has the following limitations:				
295	Maximum number of dependent variables:	6			
296	Maximum number of independent variables:	2			
297	Maximum number of between variables:	2			
298	Maximum number of continuous latent variables in time series analysis:	2			
299					
300					
301	MUTHEN & MUTHEN				
302	3463 Stoner Ave.				
303	Los Angeles, CA 90066				
304					
305	Tel: (310) 391-9971				
306	Fax: (310) 391-8971				
307	Web: <a href="http://www.StatModel.com">www.StatModel.com</a>				
308	Support: <a href="mailto:Support@StatModel.com">Support@StatModel.com</a>				
309					
310	Copyright (c) 1998-2022 Muthen & Muthen				

Table 4.1: Alcohol use data (alcohol1\_pp): Fitted multilevel models for change  
Model B

## Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	ID	0.6244	0.1481	4.22	<.0001
UN(2,1)	ID	-0.06844	0.07008	-0.98	0.3288
UN(2,2)	ID	0.1512	0.05647	2.68	0.0037
Residual		0.3373	0.05268	6.40	<.0001

## Fit Statistics

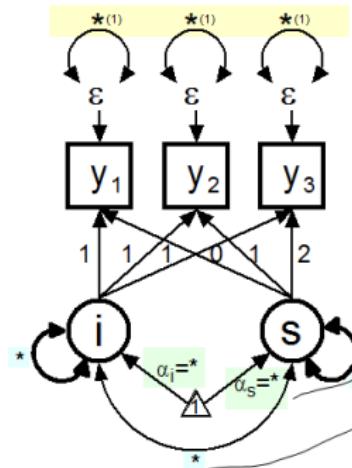
-2 Log Likelihood	636.6
AIC (smaller is better)	648.6
AICC (smaller is better)	649.0
BIC (smaller is better)	663.1

## Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
3	79.70	<.0001

## Solution for Fixed Effects

Effect	Estimate	Error	Standard		
			DF	t Value	Pr >  t
Intercept	0.6513	0.1051	81	6.20	<.0001
AGE_14	0.2707	0.06245	81	4.33	<.0001



## MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
S I WITH	-0.068	0.074	-0.920	0.357
Means I S	0.651 0.271	0.105 0.062	6.198 4.334	0.000 0.000
Intercepts ALCUSE1 ALCUSE2 ALCUSE3	0.000 0.000 0.000	0.000 0.000 0.000	999.000 999.000 999.000	999.000 999.000 999.000
Variances I S	0.624 0.151	0.162 0.056	3.861 2.721	0.000 0.007
Residual Variances ALCUSE1 ALCUSE2 ALCUSE3	0.337 0.337 0.337	0.066 0.066 0.066	5.105 5.105 5.105	0.000 0.000 0.000

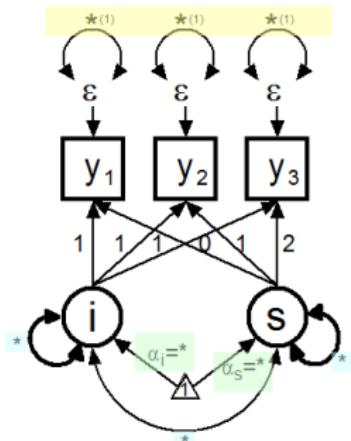


Table 4.1: Alcohol use data (alcohol1\_pp): Fitted multilevel models for change  
Model B

#### Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z
UN(1,1)	ID	0.6244	0.1481	4.22	<.0001
UN(2,1)	ID	-0.06944	0.07008	-0.98	0.3288
UN(2,2)	ID	0.1512	0.05647	2.68	0.0037
Residual		0.3373	0.05268	6.40	<.0001

#### Fit Statistics

-2 Log Likelihood	636.6
AIC (smaller is better)	648.6
AICC (smaller is better)	649.0
BIC (smaller is better)	663.1

#### Null Model Likelihood Ratio Test

DF	Chi-Square	Pr > ChiSq
3	79.70	<.0001

#### Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr >  t
Intercept	0.6513	0.1051	81	6.20	<.0001
AGE_14	0.2707	0.06245	81	4.33	<.0001

SAS Proc Mixed  $\approx$  Mplus

Intercept  $\approx$  Means(I)

AGE\_14  $\approx$  Means(S)

Cov parm UN(1,1)  $\approx$  Variance(I)

Cov parm UN(2,1)  $\approx$  S WITH I

Cov parm UN(2,2)  $\approx$  Variance(S)

Cov parm residual  $\approx$  Residual Variances(y)

# LAB SESSION Alcohol Use LGC Models

# Questions?

# END OF DAY ONE