EARA 2009
Methodology Workshop

New methods for the analysis of change and development

Growth curves and beyond

Prof. Wim Beyers
Ghent University, Belgium
wim.beyers@ugent.be

With special thanks to...

Luc Goossens

Tod Little

Patrick Curran

Karl Jöreskog   Bengt & Linda Muthén
(no pics available)

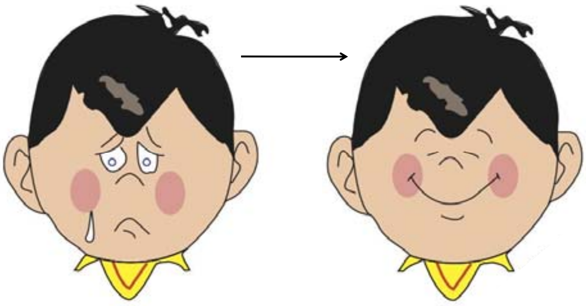
CHANGE

CHANGE

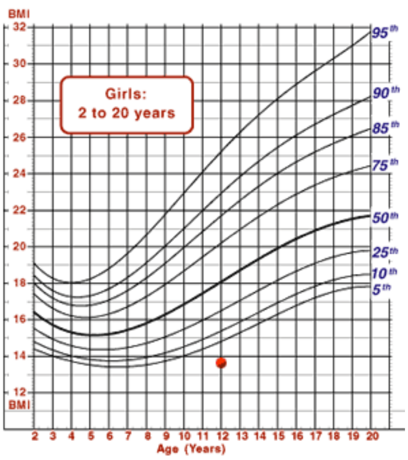
CHANGE

Adolescence is a time of change!

But, how do we measure this?



DEVELOPMENT



- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

- General measurement and design issues
  - Time and intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Absolute change: the difference score
  - Absolute change+: repeated measures ANOVA
  - Relative change: autoregression
  - Relative change+: cross-lagged models
- New and better methods
  - With 2+ waves: LCM
  - With 3+ waves: LGCM
  - With 4+ waves: LCGA
  - With 4+ waves: LGMM

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

The example DATA:

- *N* = 405 adolescents (M_{age} = 13 at start) + mothers
- Measures:

  Age 13:
  - **Gender**: 203 boys and 202 girls (A-report)
  - **Support** from mother (A-report)

  Age 13, 14, 15 and 16:
  - **Antisocial behavior** (M-report)
  - School **GPA** (A-report)
- Missing data:
  - dropout and nonresponse from Age 14 onwards!
  - 7 (ANTI-3) to 34% (GPA-4)
  - 13.5% overall

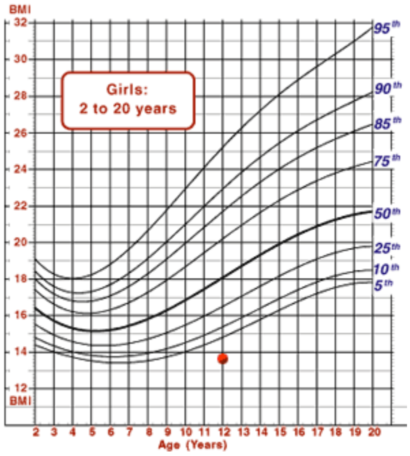
EARADATA.SAV
EARADATA.DAT
EARADATA.XLS
EARADATA.TXT

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

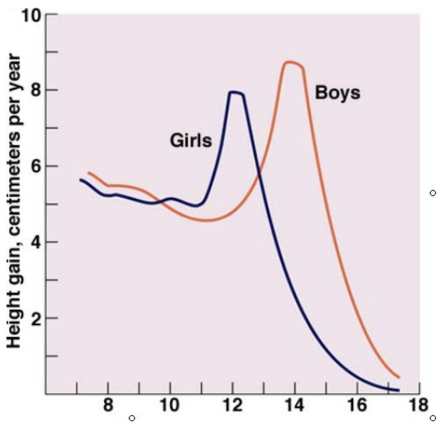
- When measuring CHANGE, how can we define TIME?
  - Age in years, months, days.
  - Experiential time: Amount of time something is experienced
    - Years of schooling (grade), length of relationship, amount of practice
    - Calibrate on beginning of event, measure time experienced
  - Episodic time: Time of onset of a life event
    - Age, toilet trained, driver license, puberty, birth of child, retirement
    - Early onset, on-time, late onset: used to classify or calibrate
    - Time since onset or time from normative or expected occurance.
- What measurement Intervals should we take?
  - How fast is the developmental process?
  - Intervals must be equal to or less than expected processes of change (e.g., schooling studies at half-year intervals)
  - If too short: too sensitive to measurement error
  - If too long: insensitive to change and variability in change

  - An example article: Rueter & Conger (1998)

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !



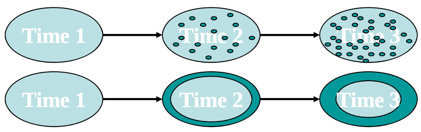
- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !



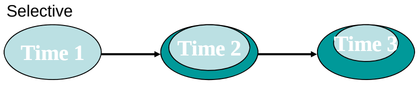
- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

DROPOUT

- Random



- Selective



- How to test for?
  - MVA (SPSS) →
  - Mean ànd Covariances (Structure Analysis; MACS; Little, Lindenberger, & Maier, 2000) Example (MACS1.SPL)

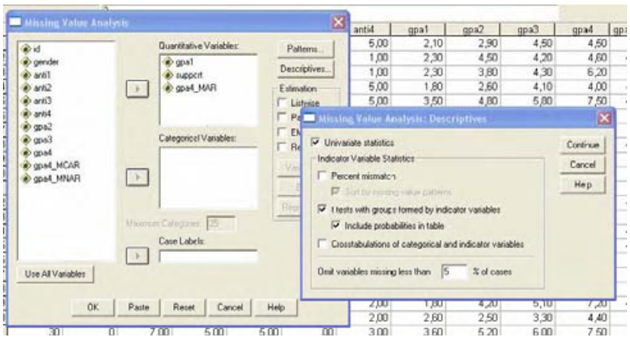


Table with "Separate Variance t Test' in output, with in rows all variables with (+5%) missings, and in colums all variables in dataset. Each cell contains a t-value (with a p-value) indicating whether or not missingness in the row variable is correlated significantly with the values of the column variable, and therefore is selective.

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

TYPES of MISSING DATA

- **Missing Completely at Random (MCAR)**
  - No association with unobserved variables (selective process) and no association with observed variables
- **Missing at Random (MAR)**
  - No association with unobserved variables, but maybe related to observed variables
- **Non-random Missing (NMAR)**
  - Some association with unobserved variables and maybe with observed variables

Little & Rubin (1987); Shafer & Graham (2002)

	Y (wave 2)			
X (wave1)	Compleet	**MCAR**	**MAR**	**MNAR**
130	101	101	**101**	101
145	155			
136	140	140	**140**	
146	134	134		134
111	129		**111**	129
134	124		**124**	124
153	112			112
137	122	122	**122**	122
118	118	118	**118**	118

gpa4	gpa4_MCAR	test_MCAR	gpa4_MAR	test_MAR	gpa4_MNAR	test_MNAR
4,50	4,50	0	-999999	1	-999999,0	1
4,60	-999999,0	1	4,60	0	-999999,0	1
6,20	6,20	0	-999999	1	6,20	0
4,00	-999999,0	1	4,00	0	-999999,0	1
7,50	-999999,0	1	7,50	0	7,50	0
6,90	6,90	0	6,90	0	6,90	0
6,10	6,10	0	6,10	0	-999999,0	1
4,20	-999999,0	1	-999999	1	-999999,0	1
6,30	6,30	0	6,30	0	6,30	0
7,20	7,20	0	7,20	0	7,20	0
5,80	5,80	0	5,80	0	-999999,0	1
5,80	-999999,0	1	5,80	0	5,80	0

		gpa4	gpa1	support
test_ MCAR	Pearson Correlation	-,038	,017	-,077
	Sig. (2-tailed)	,576	,798	,252
	N	221	221	221

		support	gpa1	gpa4
test_ MAR	Pearson Correlation	**,266**	-,079	,009
	Sig. (2-tailed)	,000	,241	,889
	N	221	221	221

		support	gpa1	gpa4
test_ MNAR	Pearson Correlation	-,067	**-,188**	**-,399**
	Sig. (2-tailed)	,320	,005	,000
	N	221	221	221

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

PREVENTING!

- ***Dillman (1978)***
  - ✓ *Intensive follow-up and tracking of subjects*
  - ✓ *Repeated invitations to participation, reminders*
  - ✓ *Repeated sending of the measurements*
  - ✓ *Do everything to prevent large dropout!*
- ***Planned missingness***
  - ✓ *Do not measure all variables in all participants at all times.*
- ***Cohort-sequential design***
  - ✓ *Let new persons come in at each wave of the study,*
  - ✓ *This way you create different patterns of missingness, not only dropout!*

= Different ways to increase the chances of MAR or MCAR!

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

CURING!

- ***Purpose is NOT to fill in empty cells in the data!***
- ***Purpose IS to estimate the population parameters as good as possible, using a sample with missing data!***

Which methods can help us in this challenging task?

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

BAD ways to deal with missing data

- **List-wise Deletion**
  - If a single data point is missing, delete subject
  - N is uniform but small
  - Variances biased, means biased
  - *Acceptable only if power is not an issue and the incomplete data is MCAR*
- **Pair-wise Deletion**
  - If a data point is missing, delete paired data points when calculating the correlation
  - N varies per correlation
  - Variances biased, means biased
  - Matrix often non-positive definite
  - *Acceptable only if power is not an issue and the incomplete data is MCAR*

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

BAD ways to deal with missing data

- **Sample-wise Mean Substitution**
  - Use the mean of the sample for any missing value of a given individual
  - Variances reduced
  - Correlations biased
  - *Never acceptable!*
- **Subject-wise Mean Substitution**
  - Use the mean score of other items for a given missing value
    - Depends on the homogeneity of the items used
    - Is like regression imputation with regression weights fixed at 1.0
    - *Acceptable only if set of items is homogeneous and only few missings!*

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

QUESTIONABLE ways to deal with missing data

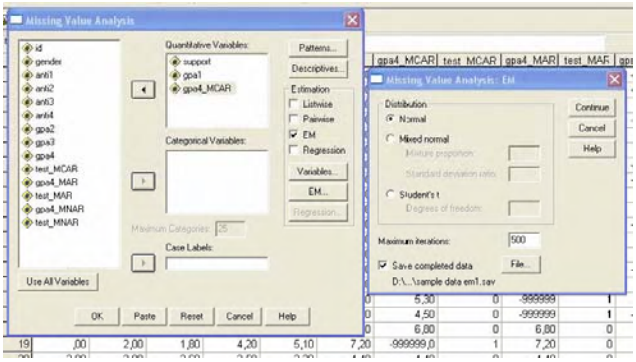
- **Regression Imputation**
  - Regress the variable with missing data on to other variables in the dataset
  - All subjects with same values on IV get the same estimated value on the DV.
  - Variances reduced
  - Better if much predictors are used
  - *Assumes MCAR*
- **Stochastic Regression Imputation**
  - Same as above but a random error component is added to the imputed value to reduce the loss in variance
  - *Still assumes MCAR*

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

GOOD ways to deal with missing data

- **But only if enough variables related to missingness are included in analysis (MAR), or missingness is MCAR**
- **EM Imputation**
  - Imputes the missing data values in an iterative way, starting with the E step
  - The E(stimation)-step is a stochastic regression-based imputation for each variable.
  - The M(aximization)-step is to calculate a complete covariance matrix based on the estimated values.
  - The E-step is repeated for each variable but the regression is now on the covariance matrix estimated in the previous M-step.
  - The EM-steps are repeated until the imputed estimates don't differ from one iteration to the other
  - Available in SPPS 12+ (output MVA.spo; resulting data after 500 iterations: EARADATA MVA EM.SAV)

EM in SPPS (assuming MCAR)



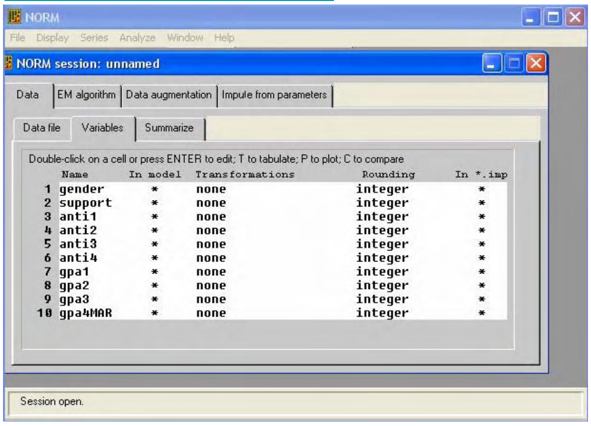
Also available in Lisrel (Prelis), SAS (PROC MI, with OUTEM= option), etc.

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

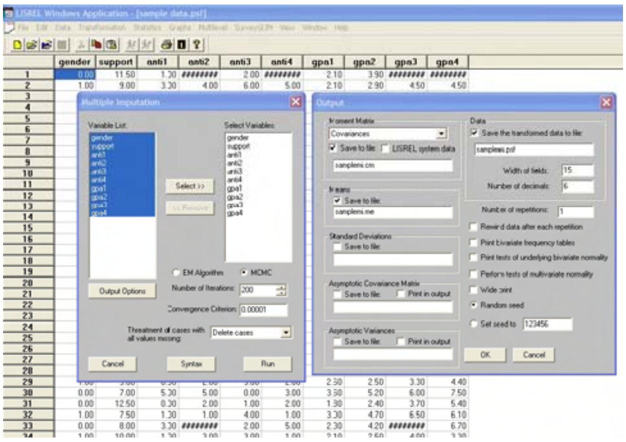
GOOD ways to deal with missing data

- **But only if enough variables related to missingness are included in analysis (MAR), or missingness is MCAR**
- **Multiple Imputation**
  - Estimate *N* (e.g., 5) datasets using the EM algorithm
  - Each dataset is based on a kind of resampling of the original sample (equivalent to a random selection of a different sample from the population)
  - Possible way 1:
    - Run the analyses *N* times
    - Summarize the results of these *N* analyses using the formulas of Rubin (1987)
  - Possible way 2:
    - Collapse the *N* samples to one dataset and do the analyses.

MI in NORM (Shafer,1997) freely available at
http://www.stat.psu.edu/~jls/misoftwa.html



MI in Lisrel



- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

GOOD ways to deal with missing data

- **But only if enough variables related to missingness are included in analysis (MAR), or missingness is MCAR**

- **Full Information Maximum Likelihood**
  - Sufficient statistics (means, covariances) are estimated with the Expectation Maximization (EM) algorithm
  - Those estimates then serve as the start values for the Maximum Likelihood model estimation
  - Does not impute the missing values.
  - Can only be used when testing a SEM-model.
  - Available in Lisrel, AMOS, Mplus, EQS, etc.

EXAMPLE comes with LGC models.

- General measurement and design issues
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

GOOD ways to deal with missing data

- **But only if enough variables related to missingness are included in analysis (MAR), or missingness is MCAR**
- **Treat missingness as a case of individual-varying intervals!**
  - Specific for longitudinal studies!
  - Available in Multilevel approaches of longitudinal data
  - ~ FIML

'Missing' Websites

- *http://www2.chass.ncsu.edu/garson/pa765/missing.htm*
- *http://www.stat.psu.edu/~jls/missing_data/index.html*
- *http://www.utexas.edu/its/rc/answers/general/gen25.html*
- *http://www.stat.psu.edu/~jls/mifaq.html*
- *http://www.stat.psu.edu/~jls/misoftwa.html (NORM)*

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

$CHANGE_{14} = ANTI_4 - ANTI_1$

- Though many problems with it, still popular (e.g., intervention, pretest-posttest, or clinical studies)
- Most cited problem: Unreliability of the difference score
  - when the component measures comprising the difference are only modestly reliable and positively correlated
  - which is typically the case in longitudinal research!
- And therefore also: lack of valididity
- Change is measured without taking level into account!

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

r_2		0.60	0.70	0.80	0.90	1.00
				r_1 = 0.60		
r_{12}	0.60	0.00	0.13	0.25	0.38	0.50
	0.70	-0.33	-0.17	0.00	0.17	0.33
	0.80	-1.00	-0.75	-0.50	-0.25	0.00
	0.90	-	-	-	-	-1.00
	0.95	-	-	-	-	-
				r_1 = 0.70		
	0.60	0.13	0.25	0.38	0.50	0.63
	0.70	-0.17	0.00	0.17	0.33	0.50
	0.80	-0.75	-0.50	-0.25	0.00	0.25
	0.90	-	-	-	-1.00	-0.50
	0.95	-	-	-	-	-
				r_1 = 0.80		
	0.60	0.25	0.38	0.50	0.63	0.75
	0.70	0.00	0.17	0.33	0.50	0.67
	0.80	-0.50	-0.25	0.00	0.25	0.50
	0.90	-	-	-	-1.00	-0.50
	0.95	-	-	-	-	-
				r_1 = 0.90		
	0.60	0.38	0.50	0.63	0.75	0.88
	0.70	0.17	0.33	0.50	0.67	0.83
	0.80	-0.25	0.00	0.25	0.50	0.75
	0.90	-	-	-1.00	0.00	0.50
	0.95	-	-	-	-1.00	0.00
				r_1 = 1.00		
	0.60	0.50	0.63	0.75	0.88	1.00
	0.70	0.33	0.50	0.67	0.83	1.00
	0.80	0.00	0.25	0.50	0.75	1.00
	0.90	-	-1.00	0.00	0.50	1.00
	0.95	-	-	-1.00	0.00	1.00

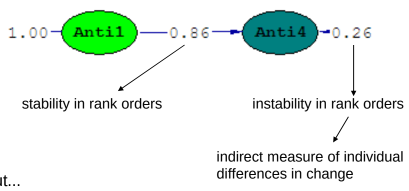
- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

CHANGE = effect of TIME in a repeated ANOVA

- SPSS output: R ANOVA.spo
- So, a good method
  - To describe and test an overall change function, and test for the form of it (linear, quadratic, cubic, etc.)
  - To test for the efffect of covariates on the change function (e.g., support)
  - To test for the effect of between-subject factors on the change fuction (e.g., gender): Time x Gender interactions!
- But
  - Only tests mean change over time in the whole sample and not deviations from that mean change
  - And... Group statistics (e.g., mean) represent everyone, and no one!
  - Equal intervals between measurements are necessary!
  - Change is an outcome of the repeated measures (time) and cannot be used as a predictor of outcomes.

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

- With observed variables: ar anti.spl (Lisrel)
- With latent variables! arsem_anti.spl (Lisrel)

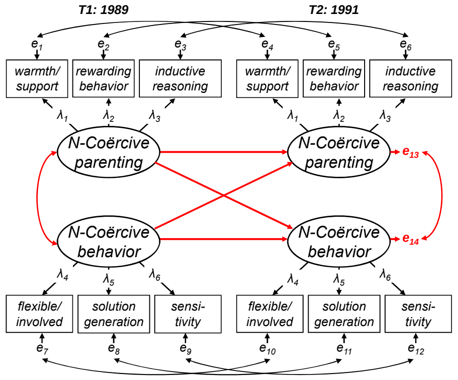


- But...
  - Implicit assumption of decreasing correlations across time!
  - Indifferent to the functional form of change!
  - Change cannot be modeled as a cause of another variable!
  - Only RELATIVE change in terms of rank order!
  - Interpretation?
    - Does ANTI-1 predict (.86) an increase in ANTI-4?
    - Increase only relative to others in the sample!
    - Even if everyone decreases (ar contra example.sav)!
  - No trajectories of individual change over time!

Cross-lagged Model

With Latent Factors (SEM)

SEM = Measurement Model + Structural Model



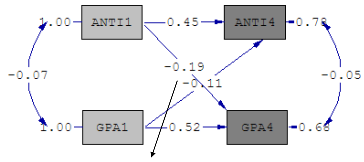
covariances = 12+11+10+9+…+1 = **78**

parameters = $6_{loadings}+14_{errors}+6_{error\text{-}r}+6_{structural}$ = **32**

➔ ***df = 46***

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

- A multivariate extension of the autoregressive model of change: crosslagged.Spl (Lisrel)



Antisocial behavior is 'causing' change in GPA!

- With latent variables: Rueter & Conger (1998) ➔
- Interesting for examining direction of effects, by comparing cross-lagged coefficients (easily done in SEM)
- But
  - Same problem with interpretation of change as in AR model!
  - Also shares all other problems of the AR model!
  - \+ Statistical drawbacks:

    If X is (much) more stable than Y, then Y will have a stronger effect on X than vice versa!

    And, the less reliable X is, the less Y can explain it.

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

So,

What do we want exactly when we try to assess change?

RELIABLE estimates of change!

Estimates of ABSOLUTE change!

INDIVIDUAL estimates of change!

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 2 WAVES of data

Latent Change Models (LCM)

(McArdle & Nesselroade, 1994; Hertzog & Nesselroade, 2003)

- Change in latent variables, using (CFA)
- First step = Longitudinal invariant factor model (multiple indicators), resulting in reliable scores for Anti-1 and 4.
- Second step = Restructuring Anti-1 and Anti-4 in latent Level and Change factors, using a very simple equation:

  Anti-1 = 1xAnti-Level

  Anti-4 = 1xAnti-Level + 1xAnti-Change

- As a consequence:

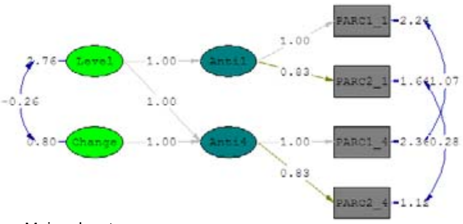
  Anti-Change = Anti-4 – Anti-1

  Anti-Level = Anti-1

  → a reliable difference score !

- LCM_anti.spl (Lisrel)

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !



- Main advantages:
  - Only two waves of data needed!
  - Reliable estimates of change!
  - Change is assessed as a latent factor!
    - With a mean: Mean change in the total sample!
    - With a variance: Individual differences in change!
    - That can be predicted and used as a predictor!
    - Latent factor scores can be estimated!
- But
  - Only linear change fuction!
  - Can be extended to multiple successive change factors!

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)

(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

- Questions:
  - Does an individual characteristic (e.g., antisocial behavior) change over time?
  - Which trajectory is followed?
  - Interindividual differences?
- Step 1: Within-Person
  - Equation for every subject in the sample:

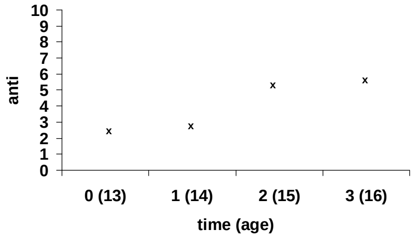
    anti = **intercept** + (**slope** x Time) + error (regression)
  - Growth can be non-linear too!

    anti = **intercept** + (**slope** x Time) + (**curve** x Time²) + error
  - Assumption: Indivuals share the shape of the change function (e.g., linear), but can differ in the amount or rate of change (individual growth parameters: intercept, slope, etc.)
- Step 2: Between-Person
  - Means (fixed) & variances (random) of intercepts, slopes
  - Predictors of change (conditional growth models).

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 3+ WAVES of data

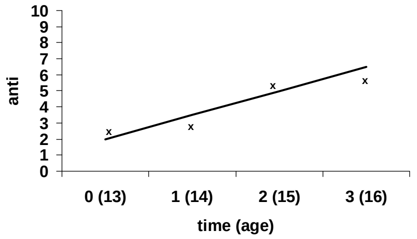
Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)



- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

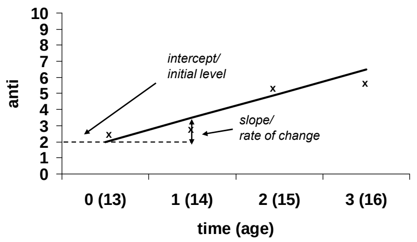


- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

anti = 2 + (1.5 x time) + error

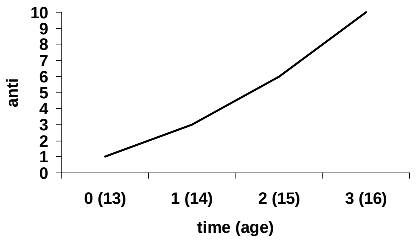


- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

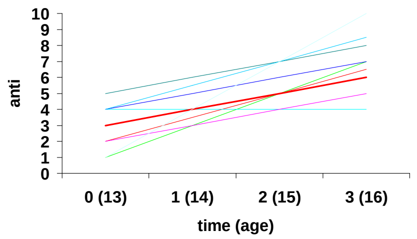
anti = 1 + (1.5 x time) + (.5 x time²) + error



- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

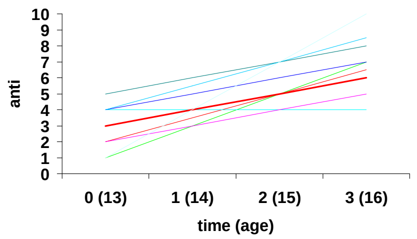


- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 3+ WAVES of data

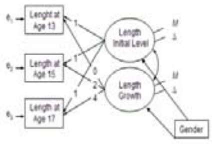
Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

anti = 3 + (1 x time) + error



- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

Two Approaches to GCM

Multilevel	SEM: LGCM
- regression approach	- factor approach
- fixed & random effects	- means & variances
- flexible with missing values & individual-varying intervals	- not flexible with individual-varying intervals
- less flexible in prediction	- very flexible in prediction
length = b_{00} + b_{01}xgender +$error_0$ + (b_{10} + b_{11}xgender +$error_1$)*time + error	
- little attention for model fit	- overload of fit indices
- Software: MLwin, HLM, Mplus SAS PROC MIXED	- Software: LISREL, AMOS, EQS, Mplus

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

PARAMETERS in the model

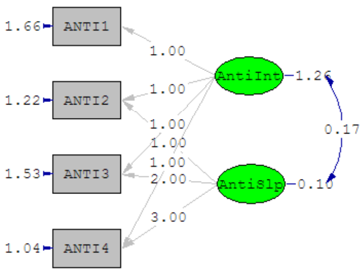
- Mean intercept / fixed effect intercept
  = Mean initial level of all individuals
- Variance intercept / random effect intercept
  = Interindividual differences in initial level
- Mean slope / fixed effect slope
  = Mean rate of growth across individuals
- Variance slope / random effect slope
  = Interindividual differences in rate of change

LGC_anti.spl (with complete 'estimated' data)
LGC_anti-FIML.spl (with missing data, and FIML)
LGC_anti+cov.spl (with gender and support as predictors)
lgc_anti.inp (the same, but now in Mplus)

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)



- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 3+ WAVES of data

Latent Growth Curve Models (LGCM)
(Duncan et al., 1994; McArdle & Nesselroade, 2002; Willet & Sayer, 1994)

- Can be extended to a multvariate LGCM!
  LGC_antiread+FIML.spl
  - Correlated intercepts = cross-sectional association
  - Correlated intercept & slope = level of IV is predicting rates of change in DV!
  - Correlated slopes = common underlying growth in two constructs = change associated with change (causality?)
- But
  - Assumption: same shape of the growth function for all subjects; interindividual differences in change are modeled as deviations from that overall mean.

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)
(Muthén, 2001, 2004)

and a special case of it:
Latent Class Growth Analysis (LCGA)
(Nagin, 1999, 2001)

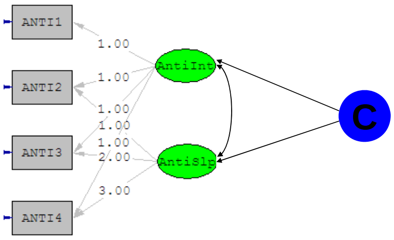
- Cfr. LGCM: heterogenity in the sample is captured by variation around a mean growth function, i.e., it is assumed that all individuals are drawn from the same population.
- LGMM
  - relaxes this single population assumption to allow for parameter differences across unobserved subpopulations
  - by using a combination of continuous and categorical latent variables
  - by introducing a latent class variabele C, a trajectory class variable, representing *k* unobserved subpopulations in the sample (note that in LGCM C=1)
  - by estimating a separate growth model for each of the latent classes

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)
(Muthén, 2001, 2004)

and a special case of it:
Latent Class Growth Analysis (LCGA)
(Nagin, 1999, 2001)



- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)

- LGMM: Class differences
  - Typically between mean intercepts between classes
  - Mean slopes in the classes
  - Variance in intercept and slope
  - Shape of the growth function!
  - Influence of covariates!
- LGMM: Extensions
  - Including covariates of change
  - Including outcomes that are predicted from growth
- LCGA (Nagin)
  - Very similar, only no variances within classes are estimated (therefore a semi-parametric approach)
  - Individuals within classes are treated as homogeneous

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)
Latent Class Growth Analysis (LCGA)

- Estimation: Mplus (Muthén's, www.statmodel.com)
  - Using the EM algorithm
  - Estimation of each individual's probability of membership in each class (conditional probabilities)
  - Measures of fit and classification quality:
    - BIC
      - Small values correspond to a good model with a large likelihood and not too many parameters
      - Look at the big drops in BIC from one solution to another!
      - Sensitive to the number of classes!
      - Less sensitive to differences in growth shape between classes
    - LMR-LR test
      - Test of a solution with k-1 classes against a solution with k classes (e.g. 2 vs 1): low p-value indicates that solution k-1 should be rejected in favor of the solution with k classes.
    - Entropy
      - Measure of classification quality based on the individual class probabilities.
      - High values (closer to 1) indicate good classification.
- Examples (Mplus; 2 classes):
  - lcga_anti2.inp (Nagin approach)
  - lgmm_anti2.inp (with equal variances across classes)
  - lgmm_anti2free.inp (with free variances across classes)

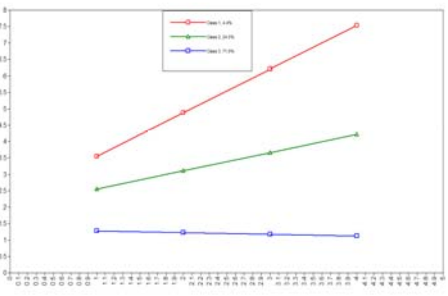
WB1

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)

Antisocial data: Best solution



- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

With 4+ WAVES of data

Latent Growth Mixture Modeling (LGMM)

Antisocial data: Best solution

lgmm_anti3.dat - Notepad

File Edit Format View Help

1.000	1.460	2.000	1.700	0.000	0.024	0.976	3.000
3.000	4.000	6.000	5.000	0.054	0.945	0.000	2.000
0.000	0.000	2.000	1.000	0.000	0.004	0.996	3.000
0.000	1.000	3.000	1.000	0.000	0.017	0.983	3.000
1.000	2.000	1.000	1.000	0.000	0.002	0.998	3.000
3.000	3.000	4.000	5.000	0.004	0.993	0.002	2.000
5.000	5.000	4.000	5.000	0.002	0.995	0.003	2.000
0.000	0.090	0.370	0.000	0.000	0.000	1.000	3.000
0.000	2.000	0.000	1.000	0.000	0.001	0.999	3.000
1.000	2.000	2.000	0.000	0.000	0.000	1.000	3.000
2.000	4.000	3.000	4.000	0.000	0.968	0.031	2.000
0.000	0.000	6.000	3.000	0.001	0.977	0.022	2.000
4.000	2.630	2.000	3.000	0.000	0.178	0.822	3.000
1.000	2.000	0.000	0.000	0.000	0.000	1.000	3.000
0.000	0.000	0.000	0.000	0.000	0.000	1.000	3.000
2.000	3.000	6.000	5.000	0.080	0.920	0.000	2.000
4.000	3.670	4.000	4.000	0.000	0.968	0.031	2.000
3.000	3.000	2.000	4.000	0.000	0.825	0.175	2.000
0.000	1.000	1.000	0.000	0.000	0.000	1.000	3.000
0.000	2.000	1.000	1.000	0.000	0.003	0.997	3.000
0.000	1.000	0.000	0.000	0.000	0.000	1.000	3.000
1.000	0.000	1.000	1.000	0.000	0.001	0.999	3.000
2.000	5.000	6.000	4.000	0.009	0.990	0.001	2.000
3.000	2.000	3.000	9.000	0.986	0.014	0.000	1.000
1.000	2.000	1.390	1.000	0.000	0.003	0.997	3.000

Observed scores *Individual estimated class probabilities* *Classification*

- General measurement and design issues
  - Measurement
  - Time & Intervals
  - Differential growth
  - Missing values
- Classic methods and disadvantages
  - Difference score
  - Repeated ANOVA
  - Autoregression
  - Cross-lagged models
- New(er) methods
  - LCM
  - LGCM
  - LCGA
  - LGMM
- Excercise !

1. Estimate a LGCM of GPA, using the FIML approach for missing data. Interprete the parameters that are found.
   DATA are EARADATA.DAT
   Software is Lisrel
2. Find the optimal LCGA solution of the GPA data. Explain why this solution was chosen and interprete the different classes.
   DATA are GPA.DAT
   Software is Mplus
3. Evaluate the effect of mother support on antisocial behavior (level & change), using a conditional growth model and FIML
   DATA are EARADATA.DAT
   Software is Lisrel

- *Setup and do the analyses using the software!*
- *Present the results to the audience, using a single slide and explain the effects in words!*
- *Ask for help while doing the analyses!*

Change is inevitable. Change is constant.

(Benjamin Disraeli)

Change is the nursery of music, joy, life, and eternity.

(John Donne)

When you're finished changing, you're finished.

(Benjamin Franklin)

We are restless because of incessant change, but we would be frightened if change were stopped.

(Lyman Bryson)

Change is a measure of time.

(Edwin Way Teale)