

An introduction to contemporary methods for the analysis of longitudinal data

Tim Windsor

Tim.windsor@flinders.edu.au



Flinders
UNIVERSITY

inspiring achievement

Overview

- Why are longitudinal studies important?
- Longitudinal analysis using multilevel models
 - Description of MLMs
 - Example MLM (with SPSS syntax)
- Longitudinal analysis using SEM (latent growth curve models)
- MLM vs. LGM: Compare and contrast
- Some extensions of LGM
- Software for longitudinal analysis
- References, textbooks and resources for getting started

Background to contemporary methods for longitudinal analysis

- Longitudinal research central to the study of human development
- Cross-sectional age comparisons confound developmental and cohort differences
 - E.g., young-old adults express less negative emotion relative to younger adults
 - **Developmental changes** (better skills at regulating emotions with age)
 - or
 - **Cohort differences** (changes in 20th century child rearing practices)?

Background to contemporary methods for longitudinal analysis

- Longitudinal methods permit integration of multiple levels of analysis: between-person differences, and **within-person changes**
 - Average patterns of growth/change over time
 - Heterogeneity in growth trajectories
 - Shapes of growth trajectories (linear vs. non-linear)
 - Predictors of individual differences in rates of change
 - And more...
 - **Be guided by key research questions in deciding on the best approach to analysis**

Background to contemporary methods for longitudinal analysis

- **Multilevel models**
- **Latent growth models**
- Developed over previous 20 to 40 years
- Computer intensive ... we have the power!

Multilevel models (MLMs)

- Conceptually similar to standard (ordinary least squares or OLS) regression
- To what extent does a linear combination of predictor variables ($X_1 \dots X_K$) account for variance in a dependent variable (Y)?
- OLS regression - One variance term for Y , partitioned into variance accounted for by model (R^2) and variance unaccounted for (residual variance)

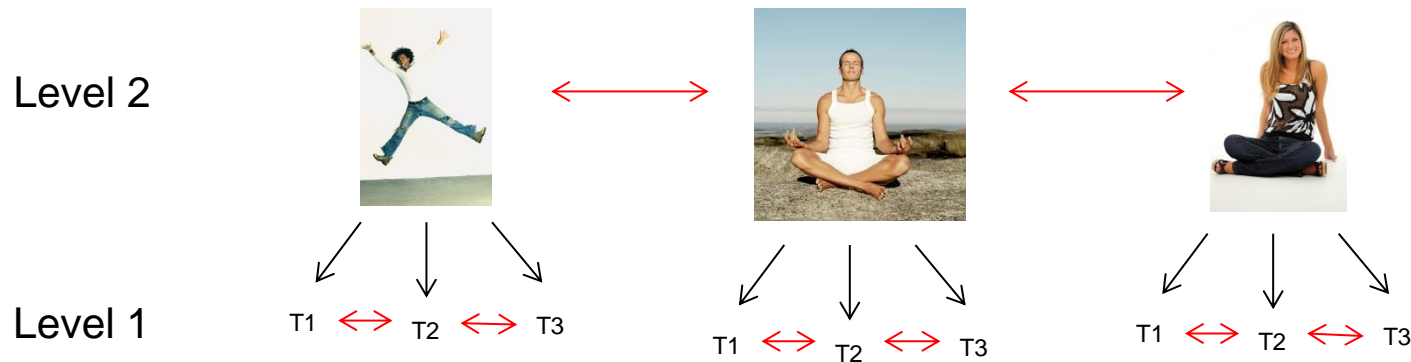
MLMs

- Multilevel models simultaneously analyse variance in the dependent variable at more than one level
- In the typical longitudinal case, this translates to two levels of analysis:
- Level 2
 between-person variance, or variance of the intercept
- Level 1
 within-person, or residual variance

Background to contemporary methods for longitudinal analysis

Multilevel models

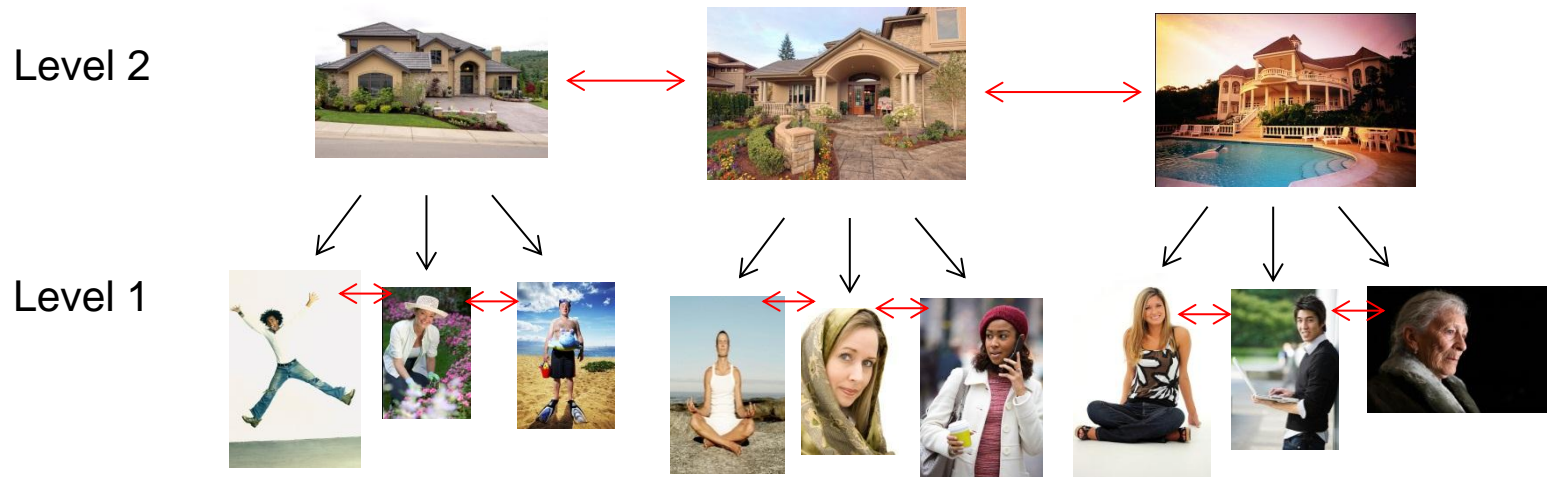
- Variance in the dependent variable analysed at multiple levels
- Longitudinal = measurement occasions (Level 1) nested within individuals (Level 2)



Background to contemporary methods for longitudinal analysis

Multilevel models

Equivalent to individuals (Level 1) nested within clusters (Level 2: e.g., schools, geographic areas, households etc.)




Traditional versus contemporary methods for longitudinal analysis

	Repeated measures ANOVA	Multilevel models
Treatment of missing data	Listwise deletion	Use of all available data in estimation
Participants measured at different time points?	No	Yes
Estimation of individual trajectories	No	Yes
Time-varying continuous predictors	No	Yes
Interactions involving continuous predictors	No	Yes

Adapted from Gueorguieva (2004)

A note on terminology

- Random coefficients models
- Hierarchical linear models
- Multilevel models
- Mixed models



Refer to the
same types
of model

- Covariates = predictor variables

- MLM parameters are referred to in terms of **fixed** and **random** components
 - **Fixed component** = population average
 - **Random effect** = variance component
- A **variance components model** (empty model with no predictors – also called ‘null’ model) can be used to determine the proportions of variance in the dependent variable that occur between- and within-individuals

$$Y_{ti} = \beta_{0i} + r_{ti}$$

1. Variance components model

$$Y_{ti} = \beta_{0i} + r_{ti}$$

Y_{ti} = DV score for person i with t measurement occasions

$$\beta_{0i} = Y_{00} + U_{0i}$$

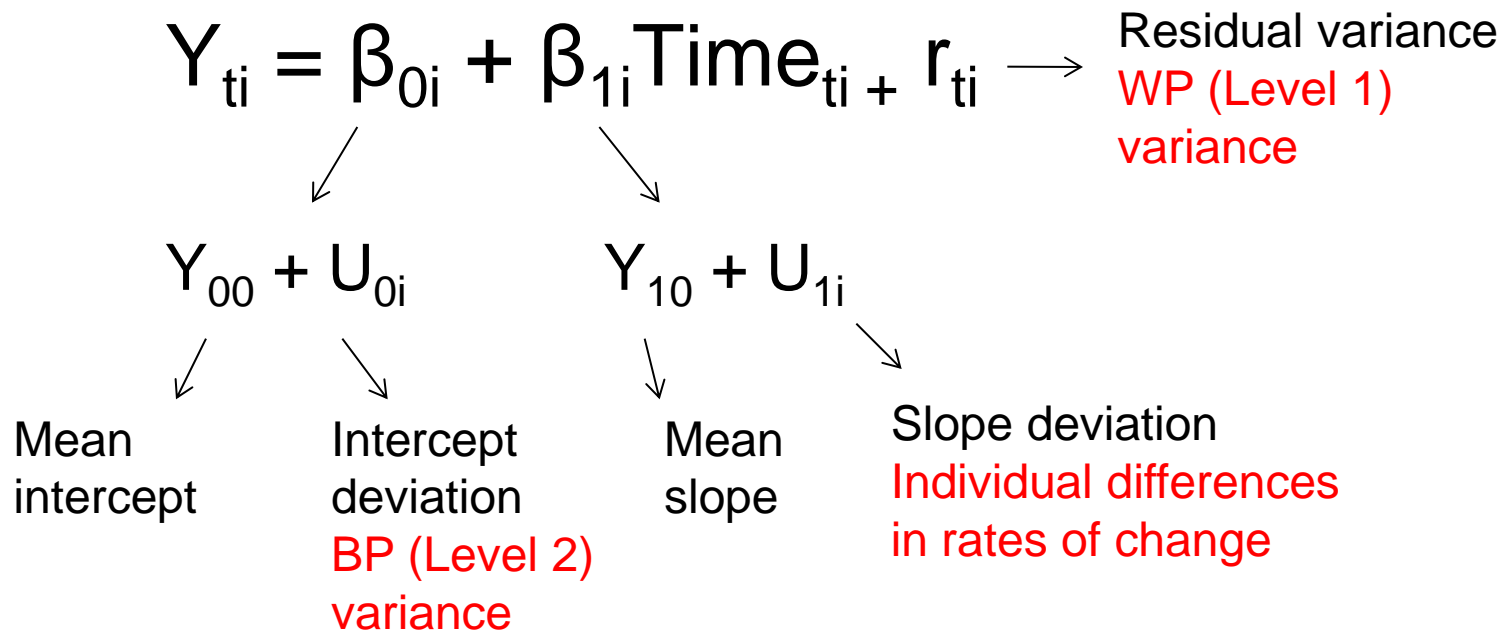
Sample grand mean – intercept **FIXED EFFECT**

Individual deviations from intercept :
between-person variance **RANDOM EFFECT**

r_{ti} = residual variance (**within-person** variance)

- Two variance components: intercept (between-person) and residual (within-person)

2. Unconditional growth model



6 model parameters

- Fixed effects: Intercept (Y_{00}), Slope (Y_{10})
- Random effects: Intercept variance (U_{0i}), slope variance (U_{1i}), Intercept-slope covariance
- Residual variance (r_{ti})

Example

- Longitudinal analysis of delayed recall performance in young-old adults
- Research questions
 - Does recall performance decline over time?
 - Do individuals show significant differences in their rates of change in recall?
 - Is older age associated with poorer recall performance?
 - Do rates of change in recall vary as a function of education (i.e., is more education related to slower rates of decline)?

Study – The PATH Through life project

- ANU Cohort study of young (aged 20-24 at baseline), midlife (aged 40-44 at baseline) and older (aged 60-64 at baseline) adults interviewed every four years
- Data from oldest PATH cohort (N=2511)
- To date, 3 waves of data available (measurement interval = 8 years)

Data – wide form (multivariate)

*PATH_IAGG_wide.sav [DataSet5] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

	id	age_c63	female	yred_c14	srh1	srh2	srh3	recall1	recall2	recall3	var
1	60002	.0	1.00	-3.00	44	.	50	4	4	3	
2	60003	2.00	1.00	-3.00	53	.	.	4	4	.	
3	60008	2.00	1.00	.	44	44.00	53	7	5	4	
4	60010	.0	.0	-3.50	47	53.00	50	3	3	4	
5	60011	1.00	1.00	.0	56	53.00	53	11	10	11	
6	60015	-3.00	1.00	1.00	35	46.00	44	7	5	3	
7	60017	2.00	.0	-3.00	44	48.00	41	4	4	5	
8	60021	2.00	.0	4.00	51	50.00	44	8	7	8	
9	60023	1.00	.0	-5.00	31	.	.	3	.	.	
10	60024	1.00	1.00	.0	56	55.00	52	1	3	1	
11	60025	-1.00	1.00	1.00	27	48.00	44	7	8	5	
12	60027	1.00	1.00	-2.00	32	32.00	25	6	7	7	
13	60028	-2.00	1.00	-5.00	51	20.00	.	6	9	.	
14	60030	1.00	.0	-10.00	23	20.00	20	5	6	4	
15	60039	1.00	.0	-3.00	38	37.00	54	4	5	3	
16	60043	2.00	.0	-1.00	56	.	.	5	.	.	
17	60047	1.00	1.00	-4.00	52	51.00	48	10	10	9	
18	60049	2.00	.0	4.00	44	50.00	36	5	8	5	
19	60051	.0	.0	3.00	59	59.00	37	3	4	7	
20	60054	-1.00	1.00	-3.00	23	.	.	2	.	.	
21	60055	-1.00	.0	2.00	41	51.00	54	7	10	8	


Data – convert to long form

- SPSS
 - Data > Restructure > Restructure selected variables into cases
- Stata
 - ‘Reshape’

Data – long form (stacked)

PATH_IAGG_long_1.sav [DataSet3] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help



	id	age_c63	female	yred_c14	srh	recall	Index1	Time	v
1	60002	.0	1.00	-3.00	44	4	1	.0	
2	60002	.0	1.00	-3.00	.	4	2	4.00	
3	60002	.0	1.00	-3.00	50	3	3	8.00	
4	60003	2.00	1.00	-3.00	53	4	1	.0	
5	60003	2.00	1.00	-3.00	.	4	2	4.00	
6	60003	2.00	1.00	-3.00	.	.	3	8.00	
7	60008	2.00	1.00	.	44	7	1	.0	
8	60008	2.00	1.00	.	44	5	2	4.00	
9	60008	2.00	1.00	.	53	4	3	8.00	
10	60010	.0	.0	-3.50	47	3	1	.0	
11	60010	.0	.0	-3.50	53	3	2	4.00	
12	60010	.0	.0	-3.50	50	4	3	8.00	
13	60011	1.00	1.00	.0	56	11	1	.0	
14	60011	1.00	1.00	.0	53	10	2	4.00	
15	60011	1.00	1.00	.0	53	11	3	8.00	
16	60015	-3.00	1.00	1.00	35	7	1	.0	

Data – long form (stacked)

PATH_IAGG_long_1.sav [DataSet3] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

Each individual (defined by a unique identifier) has multiple rows, with each row representing a different measurement occasion

	id	age_c63	female	yred_c14	srh	recall	Index1	Time	v
1	60002	n	1.00	-3.00	44	4	1	.0	
2	60002						2	4.00	
3	60002						3	8.00	
4	60003						1	.0	
5	60003						2	4.00	
6	60003						3	8.00	
7	60008	2.00	1.00	.	44	7	1	.0	
8	60008	2.00	1.00	.	44	5	2	4.00	
9	60008	2.00	1.00	.	53	4	3	8.00	
10	60010	.0	.0	-3.50	47	3	1	.0	
11	60010	.0	.0	-3.50	53	3	2	4.00	
12	60010	.0	.0	-3.50	50	4	3	8.00	
13	60011	1.00	1.00	.0	56	11	1	.0	
14	60011	1.00	1.00	.0	53	10	2	4.00	
15	60011	1.00	1.00	.0	53	11	3	8.00	
16	60015	-3.00	1.00	1.00	35	7	1	.0	

Data – long form (stacked)

PATH_IAGG_long_1.sav [DataSet3] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

Dependent variables vary between individuals and over time (within individual)

	id	age_c63	female	yred_c14	srh	recall	Index1	Time	v
1	60002	.0	1.00	-3.00	44	4	1	.0	
2	60002	.0	1.00	-3.00	.	4	2	4.00	
3	-----	-----	-----	-----	-----	3	3	8.00	
4						4	1	.0	
5						4	2	4.00	
6						.	3	8.00	
7	60008	2.00	1.00	.	44	7	1	.0	
8	60008	2.00	1.00	.	44	5	2	4.00	
9	60008	2.00	1.00	.	53	4	3	8.00	
10	60010	.0	.0	-3.50	47	3	1	.0	
11	60010	.0	.0	-3.50	53	3	2	4.00	
12	60010	.0	.0	-3.50	50	4	3	8.00	
13	60011	1.00	1.00	.0	56	11	1	.0	
14	60011	1.00	1.00	.0	53	10	2	4.00	
15	60011	1.00	1.00	.0	53	11	3	8.00	
16	60015	-3.00	1.00	1.00	35	7	1	.0	

Data – long form (stacked)

PATH_IAGG_long_1.sav [DataSet3] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

	id	age_c63	female	yred_c14	srh	recall	Index1	Time	v
1	60002	.0	1.00	-3.00	44	4	1	.0	
2	60002	.0	1.00	-3.00	.	4	2	4.00	
3	60002	.0	1.00	-3.00	50	3	3	8.00	
4	60003	2.00	1.00	-3.00	53	4	1	.0	
5	60003	2.00	1.00	-3.00	.	4	2	4.00	
					.	.	3	8.00	
					44	7	1	.0	
					44	5	2	4.00	
					53	4	3	8.00	
					47	3	1	.0	
					53	3	2	4.00	
					50	4	3	8.00	
					56	11	1	.0	
					53	10	2	4.00	
15	60011	1.00	1.00	.0	53	11	3	8.00	
16	60015	-3.00	1.00	1.00	35	7	1	.0	

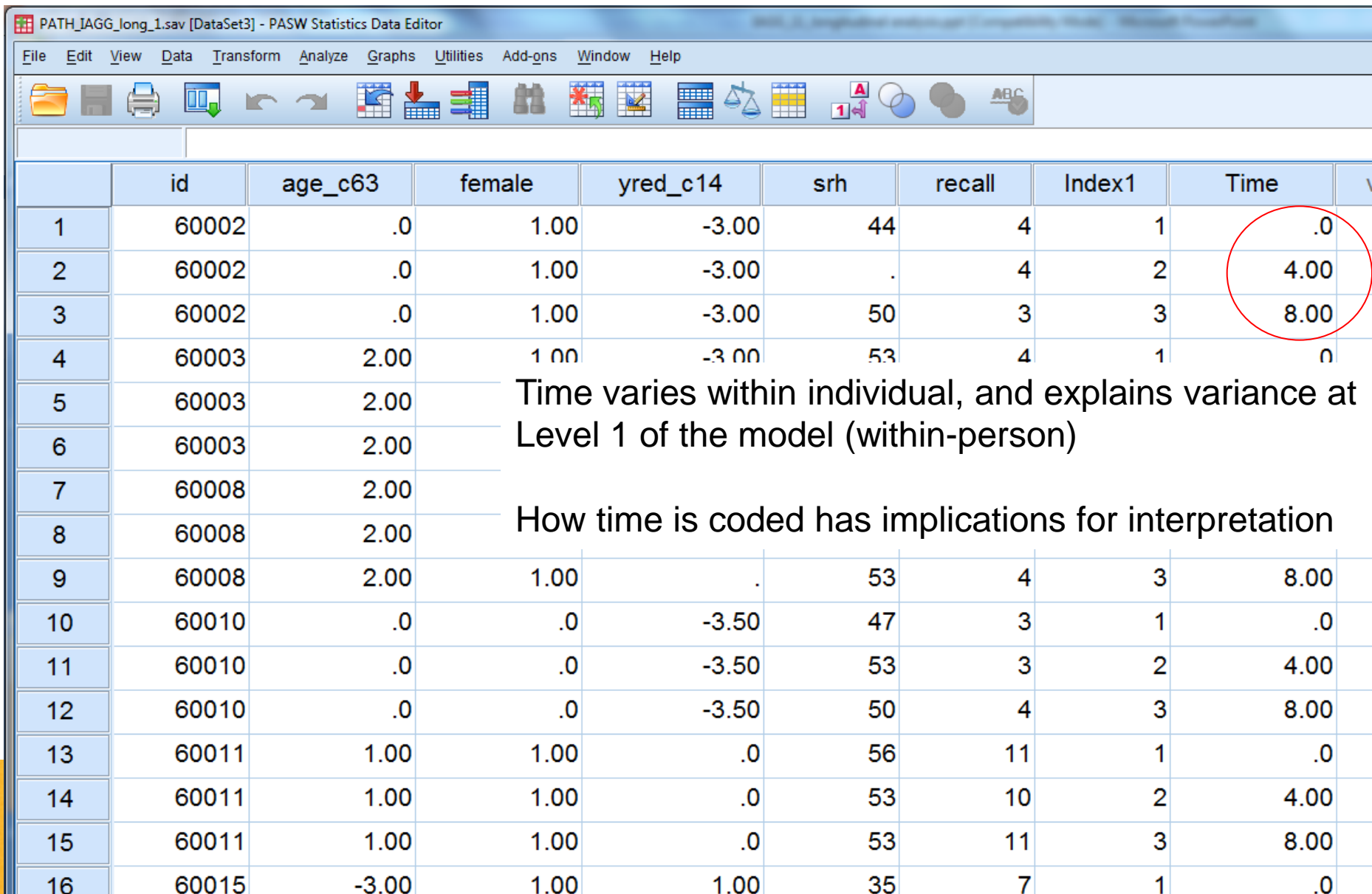
'Fixed' or 'time-invariant' predictors remain constant over time, and potentially account for variance in the DV at Level 2 (between person)

Time-varying predictors (e.g., self-rated health, depressive symptoms) can also be modelled (though not included in this example)

Data – long form (stacked)

PATH_IAGG_long_1.sav [DataSet3] - PASW Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help

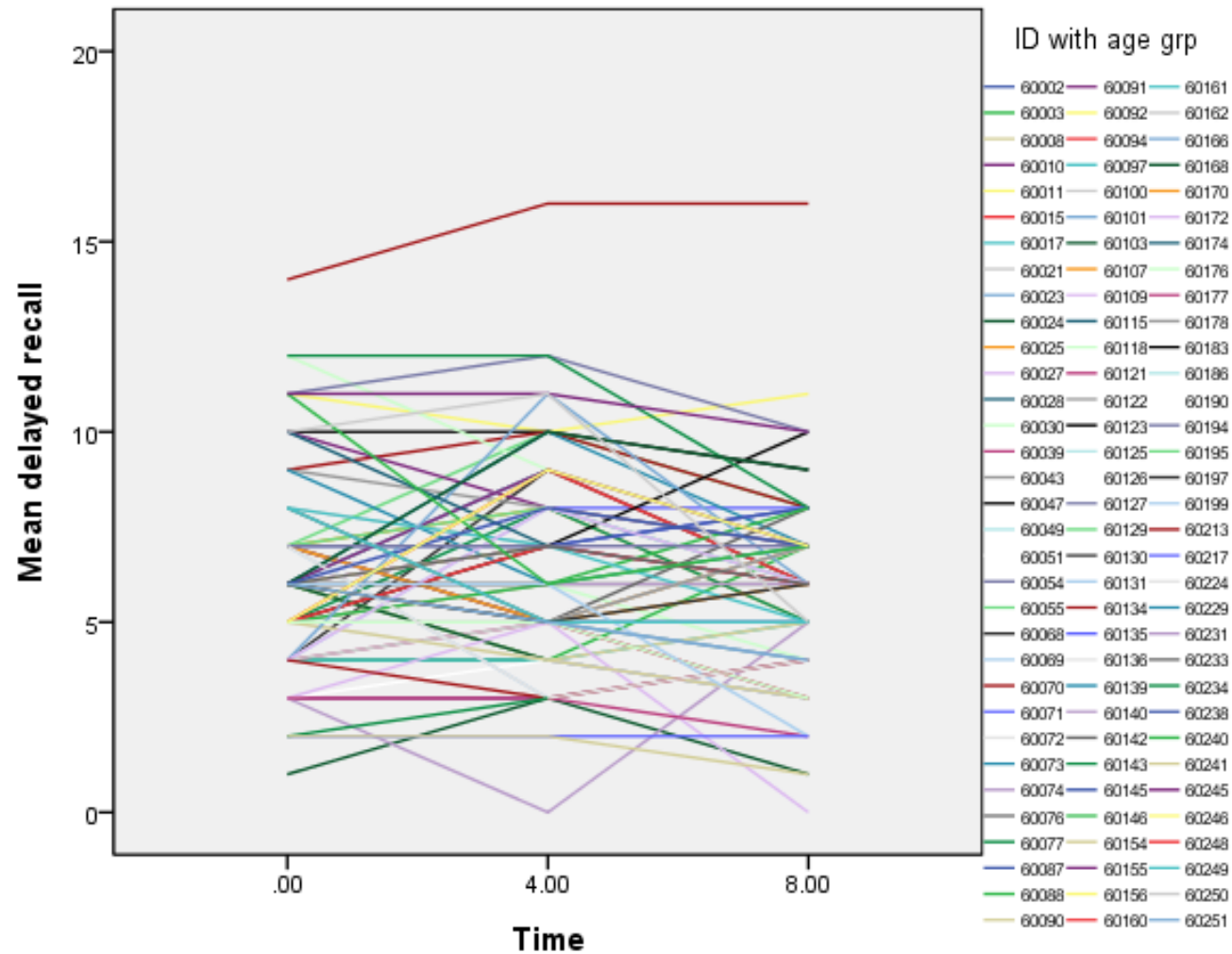


	id	age_c63	female	yred_c14	srh	recall	Index1	Time	v
1	60002	.0	1.00	-3.00	44	4	1	.0	
2	60002	.0	1.00	-3.00	.	4	2	4.00	
3	60002	.0	1.00	-3.00	50	3	3	8.00	
4	60003	2.00	1.00	-3.00	53	4	1	.0	
5	60003	2.00							
6	60003	2.00							
7	60008	2.00							
8	60008	2.00							
9	60008	2.00	1.00	.	53	4	3	8.00	
10	60010	.0	.0	-3.50	47	3	1	.0	
11	60010	.0	.0	-3.50	53	3	2	4.00	
12	60010	.0	.0	-3.50	50	4	3	8.00	
13	60011	1.00	1.00	.0	56	11	1	.0	
14	60011	1.00	1.00	.0	53	10	2	4.00	
15	60011	1.00	1.00	.0	53	11	3	8.00	
16	60015	-3.00	1.00	1.00	35	7	1	.0	

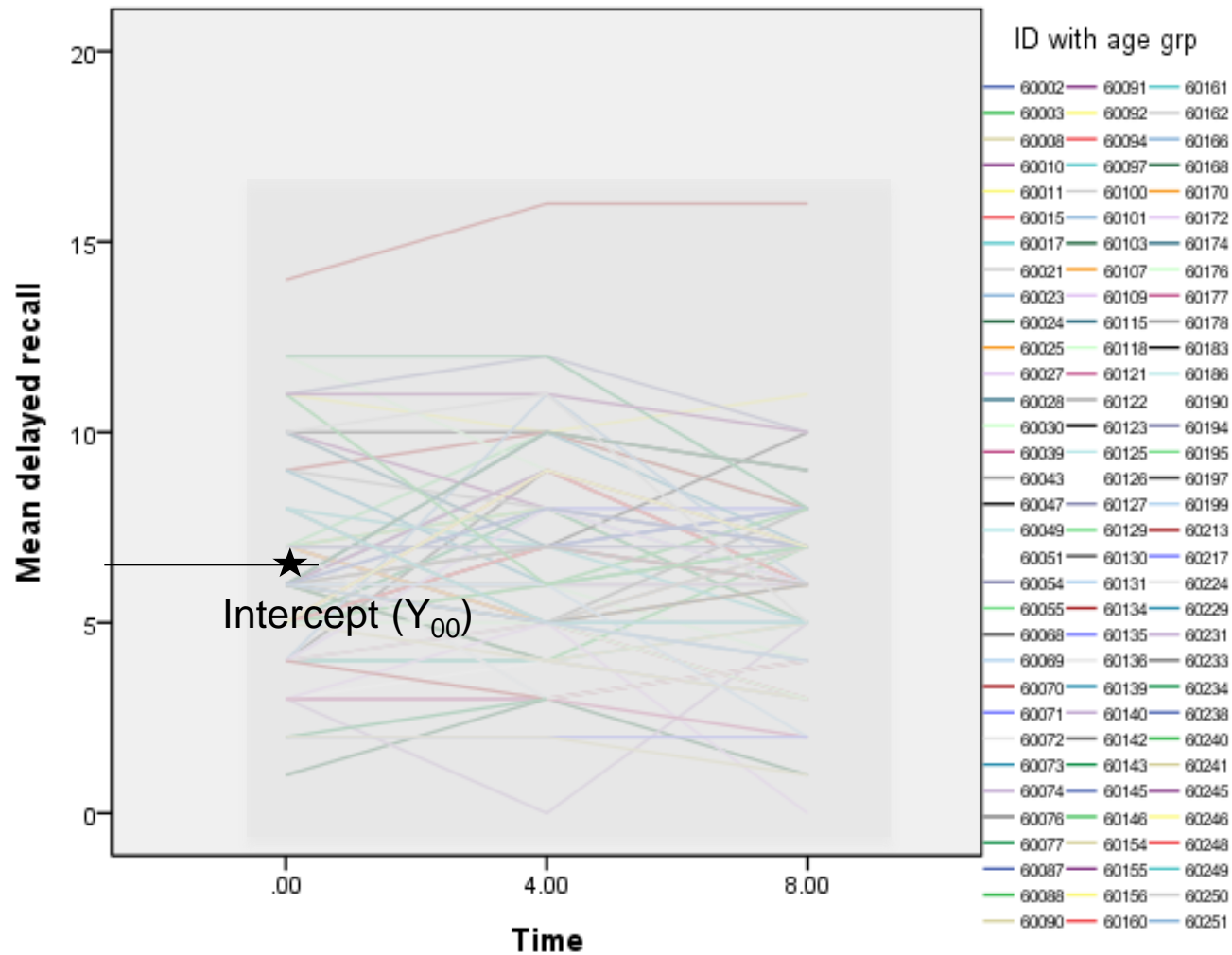
Time varies within individual, and explains variance at Level 1 of the model (within-person)

How time is coded has implications for interpretation

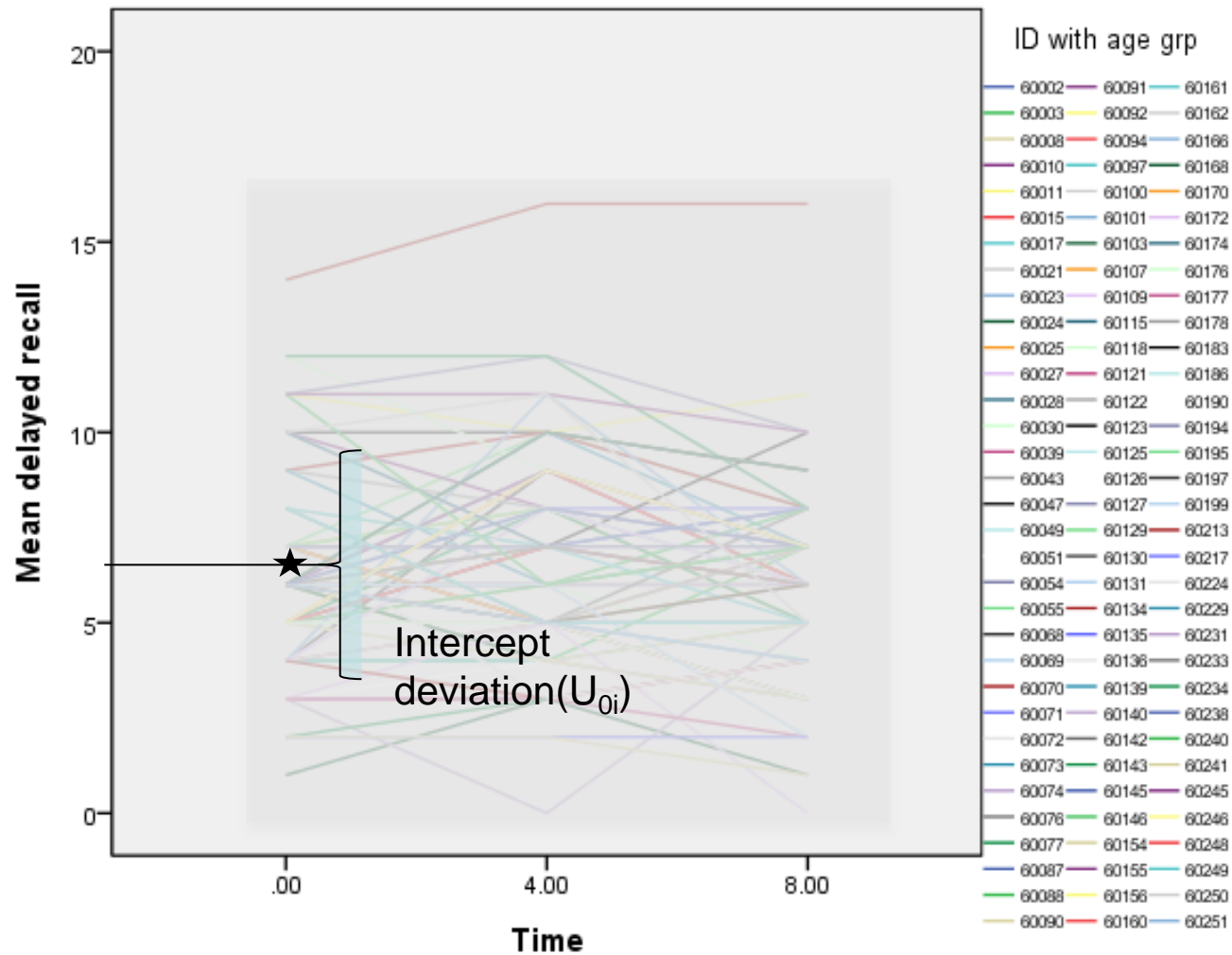
Individual Trajectories for Recall - first 100 pps



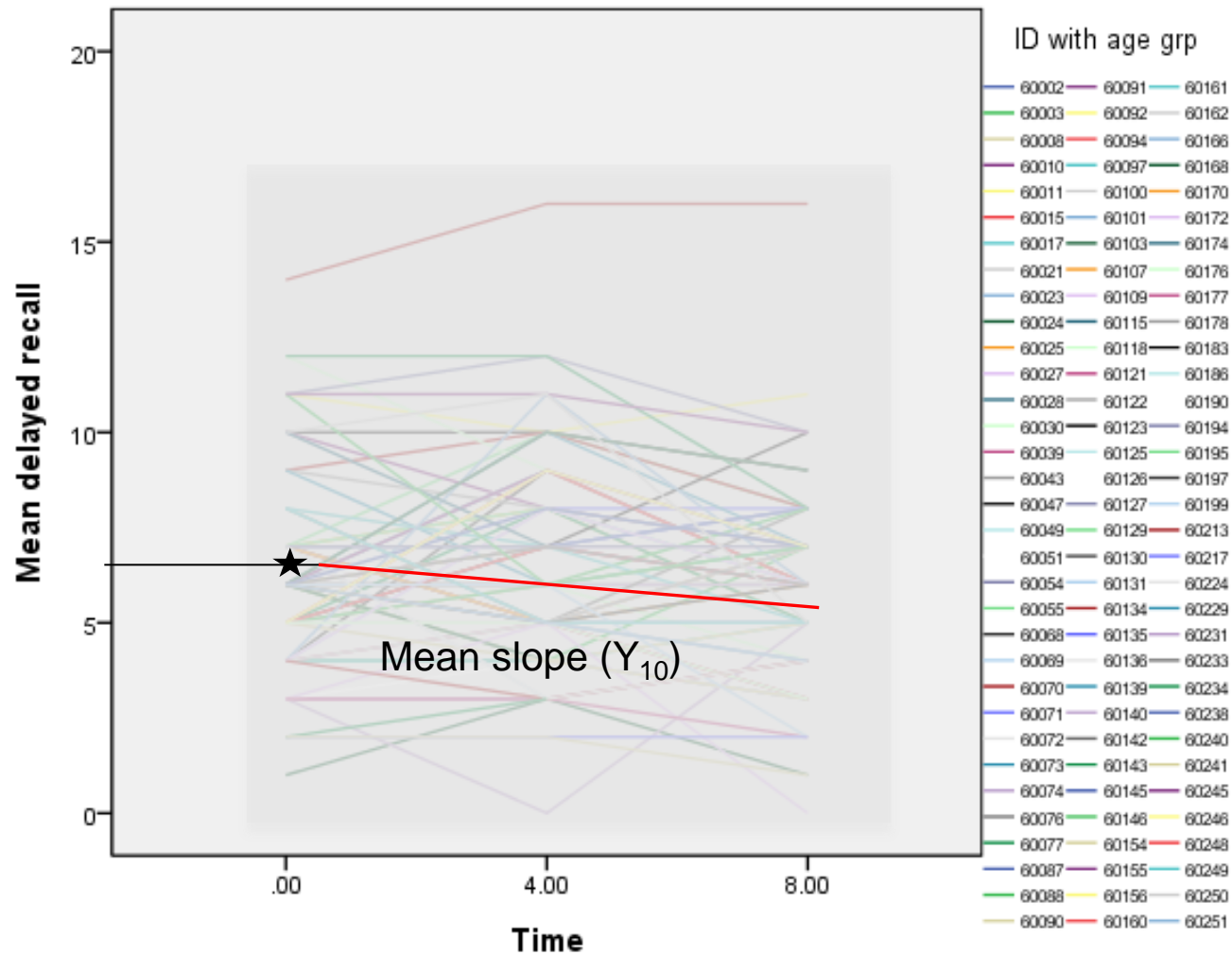
Individual Trajectories for Recall - first 100 pps



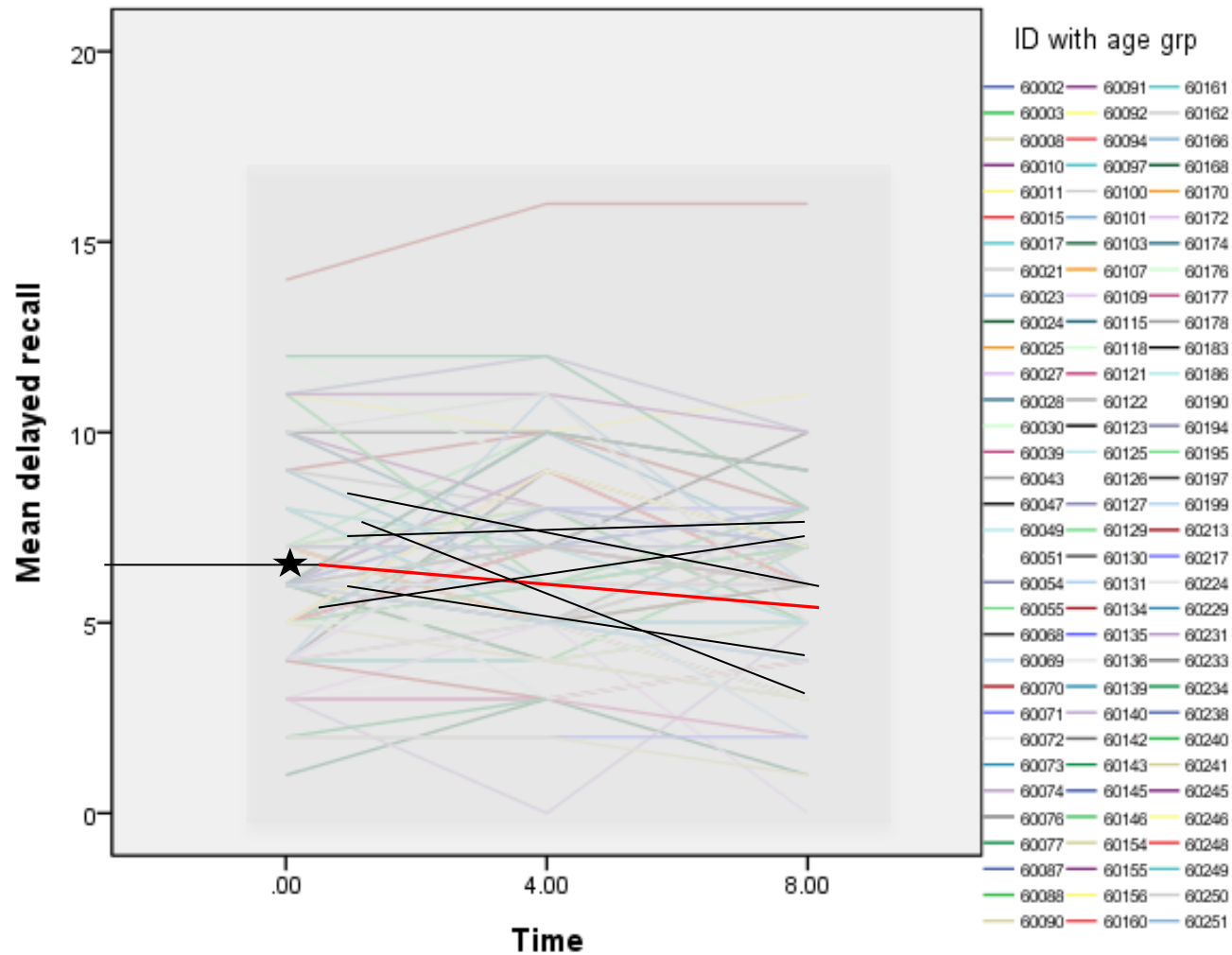
Individual Trajectories for Recall - first 100 pps



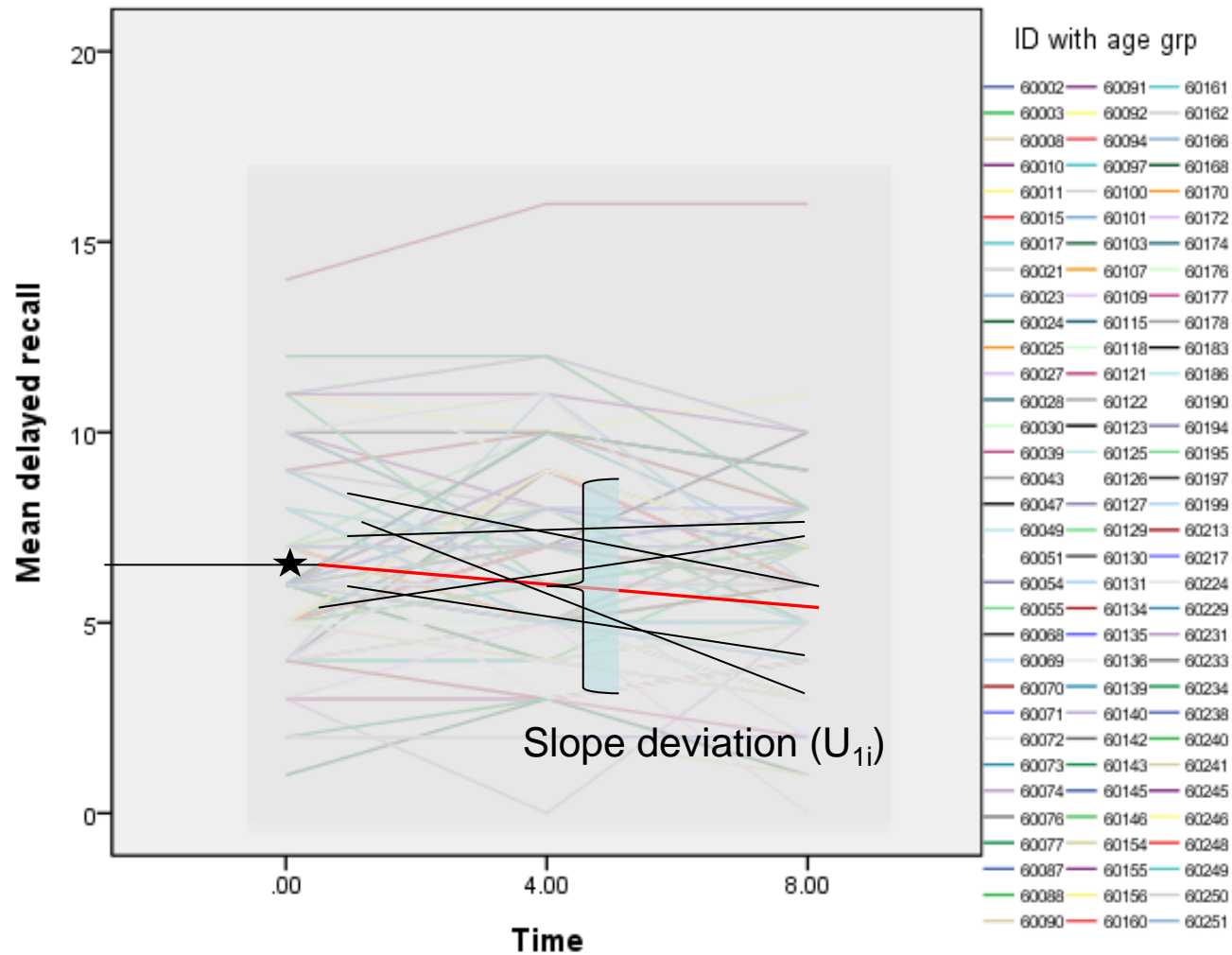
Individual Trajectories for Recall - first 100 pps



Individual Trajectories for Recall - first 100 pps



Individual Trajectories for Recall - first 100 pps



Analysis

- General recommendation- start simple and build the model up according to hypotheses/theoretical considerations

Singer & Willett, 2003

Variance components model

Selected SPSS output

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	6.000504	.041652	2502.570	144.062	.000	5.918828	6.082180

a. Dependent Variable: recall recall.

Intercept

Estimates of Covariance Parameters^a

Parameter	Estimate	Std. Error
Residual	2.492070	.054834
Intercept [subject = id] Variance	3.384340	.125955

a. Dependent Variable: recall recall.

Residual variance
(WP variance)

Intercept variance
(BP variance)

Variance components model

Selected SPSS output

Estimates of Covariance Parameters^a

Parameter	Estimate	Std. Error
Residual	2.492070	.054834
Intercept [subject = id] Variance	3.384340	.125955

a. Dependent Variable: recall recall.

BP variance in recall = $(3.38 / (3.38 + 2.49)) \times 100 = 58 \%$

WP variance in recall = $(2.49 / (3.38 + 2.49)) \times 100 = 42 \%$

Adding predictor variables

Does recall performance decline over time?

- Unconditional growth model- add Time as a Level 1 predictor (fixed effect of time)
- Selected SPSS output

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	6.160467	.046594	3786.397	132.215	.000	6.069115	6.251820
Time	-.047211	.006054	4358.183	-7.798	.000	-.059081	-.035341

a. Dependent Variable: recall recall.

With each 1 year increase in time, recall scores on average decline by .05 units

Significant linear fixed effect for time

- Selected SPSS output (continued)

Estimates of Covariance Parameters^a

Parameter	Estimate	Std. Error
Residual	2.446600	.053911
Intercept [subject = id] Variance	3.432909	.127072

a. Dependent Variable: recall recall.

- Inclusion of Time (Level 1 predictor) accounts for variance at Level 1 of the model (i.e., residual variance = WP variance)
- As a result, residual variance estimate decreases (from 2.49 in variance components model to 2.45)
- Proportion change in variance after inclusion of predictors (Level 1 or Level 2) can be expressed as Pseudo R^2 change (Singer & Willett, 2003). ~ 2%

- Do individuals show significant differences in their rates of change in recall?
- Include random effect of time
- Selected SPSS output

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	6.159430	.048011	2543.232	128.292	.000	6.065285	6.253574
Time	-.046168	.006277	2201.954	-7.355	.000	-.058477	-.033859

a. Dependent Variable: recall recall.

Estimates of Covariance Parameters^a

Parameter		Estimate	Std. Error
Residual		2.234215	.069627
Intercept + Time [subject = id]	UN (1,1)	3.956218	.175299
	UN (2,1)	-.089542	.019275
	UN (2,2)	.013531	.003465

a. Dependent Variable: recall recall.

Intercept-slope covariance

Slope variance

Does addition of a random slope for time contribute significantly to model fit?

- Compare nested models using likelihood ratio test
- Assess difference in log likelihood against chi-square distribution with df = difference in number of parameters (here $df = 2$; slope variance + intercept-slope covariance)
- This example $\Delta\chi^2(2) = 23.3, p < .001$
- Indicates presence of between individual heterogeneity in rates of change- retain random slope in the model

E.g., Singer and Willett (2003), Snijders, & Bosker (2011).

Is older age associated with poorer recall performance?

- Add level 2 (time-invariant) predictors
- Selected SPSS output

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	5.586726	.062159	2761.919	89.878	.000	5.464843	5.708608
Time	-.050479	.006406	2101.947	-7.880	.000	-.063042	-.037916
age_c63	-.046523	.026677	2364.136	-1.744	.081	-.098835	.005790
female	1.239626	.081142	2367.772	15.277	.000	1.080510	1.398742
yred_c14	.203497	.014401	2436.972	14.130	.000	.175257	.231737

a. Dependent Variable: recall n

Age not significantly associated with initial levels of recall

- Add level 2 (time-invariant) predictors
- Selected SPSS output

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	5.586726	.062159	2761.919	89.878	.000	5.464843	5.708608
Time	-.050479	.006406	2101.947	-7.880	.000	-.063042	-.037916
age_c63	-.046523	.026677	2364.136	-1.744	.081	-.098835	.005790
female	1.239626	.081142	2367.772	15.277	.000	1.080510	1.398742
yred_c14	.203497	.014401	2436.972	14.130	.000	.175257	.231737

a. Dependent Variable: recall recall.

Women have higher recall scores relative to men

- Add level 2 (time-invariant) predictors
- Selected SPSS output

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	5.586726	.062159	2761.919	89.878	.000	5.464843	5.708608
Time	-.050479	.006406	2101.947	-7.880	.000	-.063042	-.037916
age_c63	-.046523	.026677	2364.136	-1.744	.081	-.098835	.005790
female	1.239626	.081142	2367.772	15.277	.000	1.080510	1.398742
yred_c14	.203497	.014401	2436.972	14.130	.000	.175257	.231737

a. Dependent Variable: recall recall.

Years of education is related to better initial recall performance

Do rates of change in recall vary as a function of education?

- Test cross-level interaction:
Years education (Level 2) by Time (Level 1)

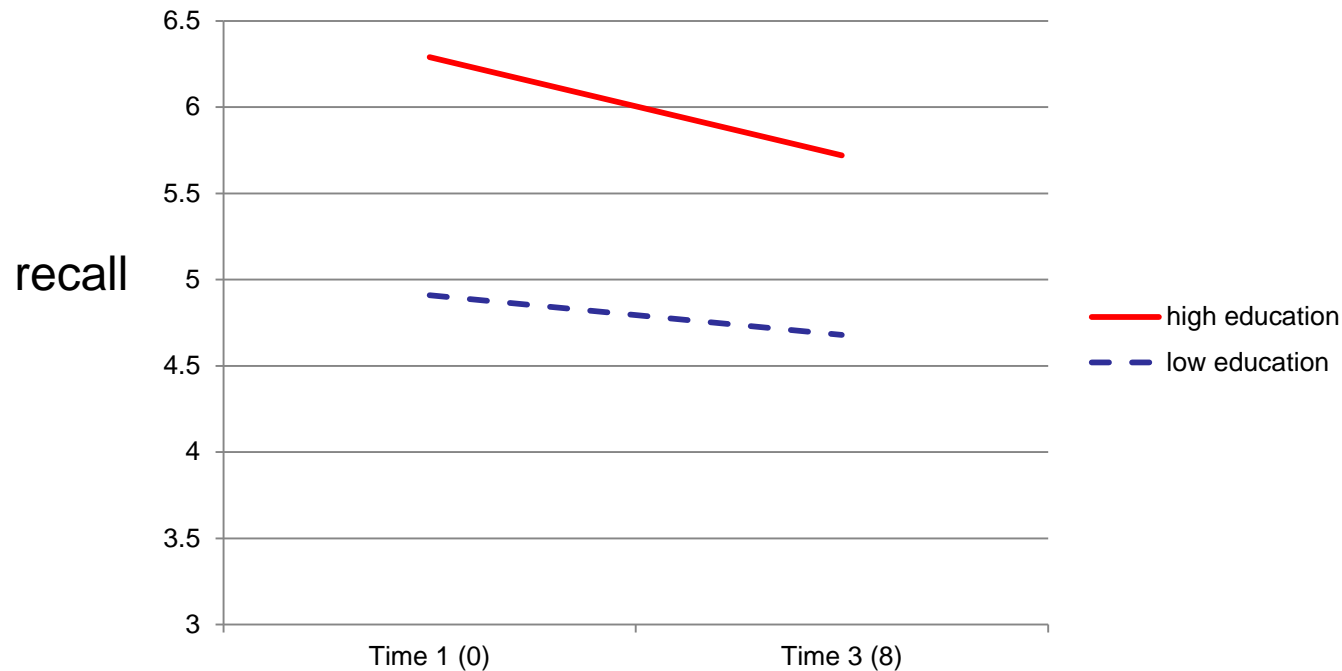
Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	5.591379	.062158	2762.117	89.955	.000	5.469499	5.713260
Time	-.050746	.006399	2100.179	-7.930	.000	-.063295	-.038196
age_c63	-.046304	.026666	2364.775	-1.736	.083	-.098594	.005987
female	1.239644	.081107	2368.344	15.284	.000	1.080596	1.398692
yred_c14	.227220	.016575	2450.775	13.709	.000	.194719	.259722
Time * yred_c14	-.006705	.002319	2164.774	-2.891	.004	-.011253	-.002157

a. Dependent Variable: recall recall.

Significant Education x Time interaction. Average rates of change in recall vary according to level of education

Display Education x time interaction by solving the regression equation (based on values of fixed effects) for hypothetical individuals with low (-1 SD) and high (+1 SD) education at Time 1, and Time 3



More education = better performance, marginally steeper rate of decline

Can MLM incorporate time-varying predictors?

- Yes!
- But may need to partition TV predictors into between- and within-person components to fit interpretable models
- Consult
 - Singer and Willett (2003)
 - Hoffman and Stawski (2009)
 - Bauer & Curran (2011)

Other issues for MLM

- Assumptions
 - Functional form (i.e., linearity)
 - Normality of residuals
 - Homoscedasticity
- Appropriate error covariance matrix
 - ‘unstructured’ assumes no set pattern of correlations of residuals over time
 - Alternative covariance structures could improve model fit
 - Singer & Willett (2003)

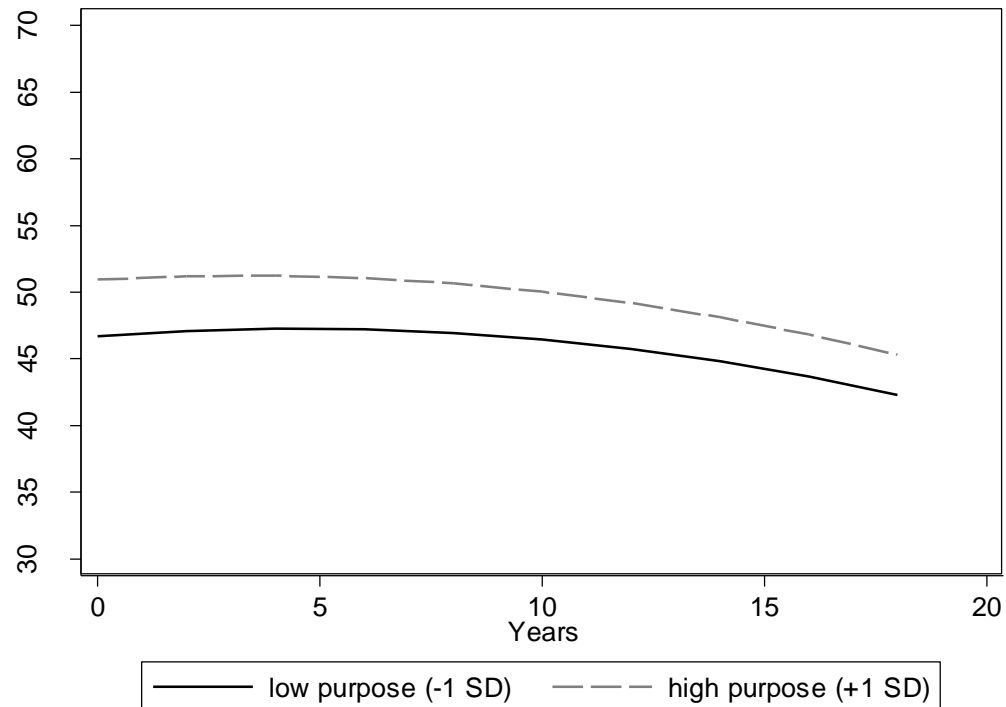
Other issues for MLM

- Modelling non-linear growth
 - Flexible treatment of time (e.g., Time^2 , Time^3)
 - Discontinuity in change (e.g., distinct trajectories for time before and after an event- 'spline' models)

Other issues for MLM

- Modelling non-linear growth
- Australian Longitudinal Study of Ageing (ALSA)
- Quadratic change in social activity for hypothetical individuals high and low in sense of purpose

Intercept	48.81*
Linear slope (Time)	0.22*
Quadratic slope (Time ²)	-0.03*
Purpose	0.61*
Purpose x Time	-0.01



Other issues for MLM (continued)

- Methods of estimation
 - Maximum Likelihood (ML)
 - Restricted ML (REML)
 - Iterative Generalized Least Squares (MLWin)
 - And more...
- Generally produce similar estimates, but be aware of what method applies (i.e., the default) in the software you are using and whether this has implications for your analysis (Singer & Willett, 2003)

Other issues for MLM (continued)

- Variance explained
 - Pseudo R^2 (Singer & Willett, Snijders & Bosker)
- Missing data
 - MLM uses all available data at Level 1 (under Missing at Random assumption), thereby accounting for missingness due to attrition
 - Participants with missing data on Level 2 predictors are excluded

Longitudinal analysis for binary and categorical outcomes

- Principles of MLM can be extended to analysis of binary and categorical outcomes using Generalised Linear Mixed Models (GLMM)
 - Random coefficients logistic regression
 - Random coefficients multinomial logistic regression

Longitudinal analysis for binary and categorical outcomes (continued)

- Specify link function that is appropriate for distribution of outcome variable
 - E.g., Binary data (binomial distribution) – logit link
- Same principles for analysis as MLM, except parameters are on a different scale

Longitudinal analysis for binary and categorical outcomes (continued)

- As for ordinary logistic regression, interpretation of random coefficients logistic regression facilitated by estimating Odds Ratios

Recall as a binary outcome

0 = 5 – 16 correct (good); 1 = 0 – 4 correct (poor)

Results of random coefficients logistic regression
(random intercept only) in Stata

r_bin	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
Time	1.039156	.0114168	3.50	0.000	1.017019	1.061776
age_c63	1.04944	.0378975	1.34	0.181	.9777299	1.126409
female	.3842147	.043063	-8.53	0.000	.3084402	.4786048

- Odds of being in the poor performance group increase by 1.04 per year
- Women 2.63 times ($1/0.38$) more likely to be in the good performance group relative to men

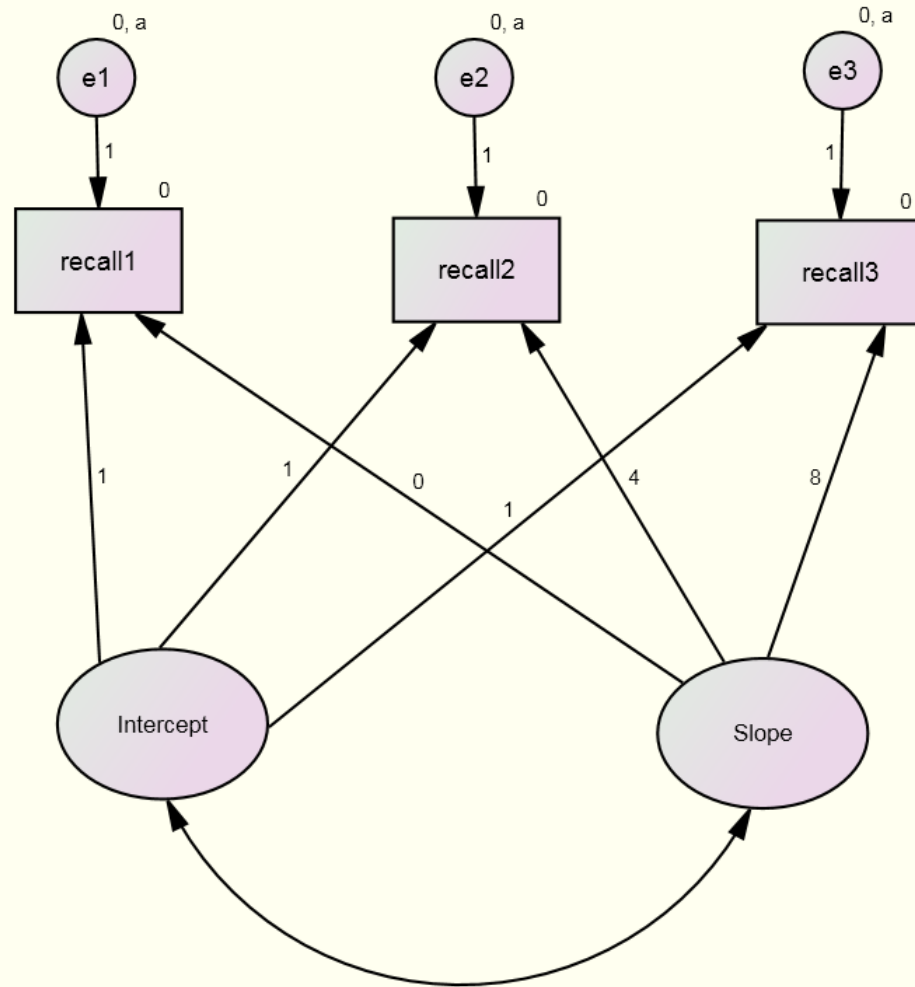
Longitudinal analysis for binary and categorical outcomes

- Generalized Estimating Equations (GEE)
- Alternative to MLM / GLMM
- Parameter estimates often similar....But
- Different implications for interpretation
 - Population-averaged vs. subject specific
- Further information on GEE
 - Consult: Fitzmaurice et al. (2004), Twisk (2006)

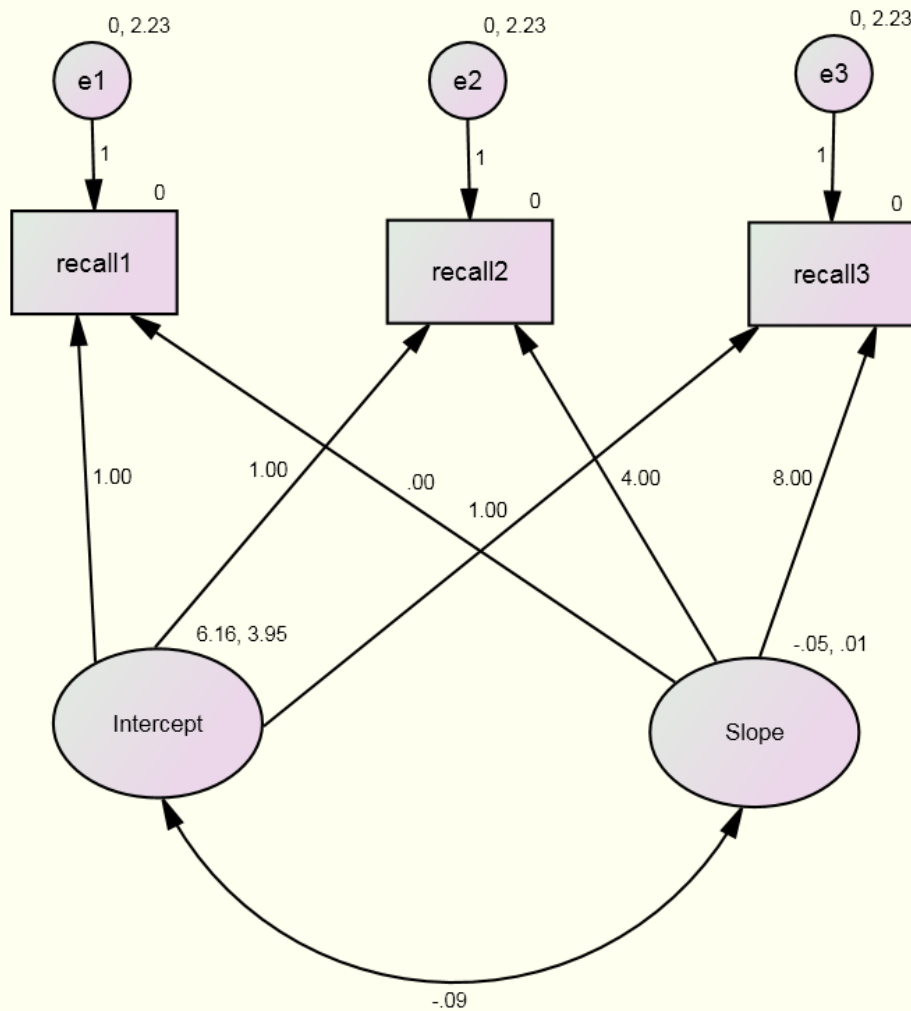
Longitudinal analysis in the Structural Equation Modelling (SEM) context- Latent growth curve models (LGM)

Analysing change in the SEM context

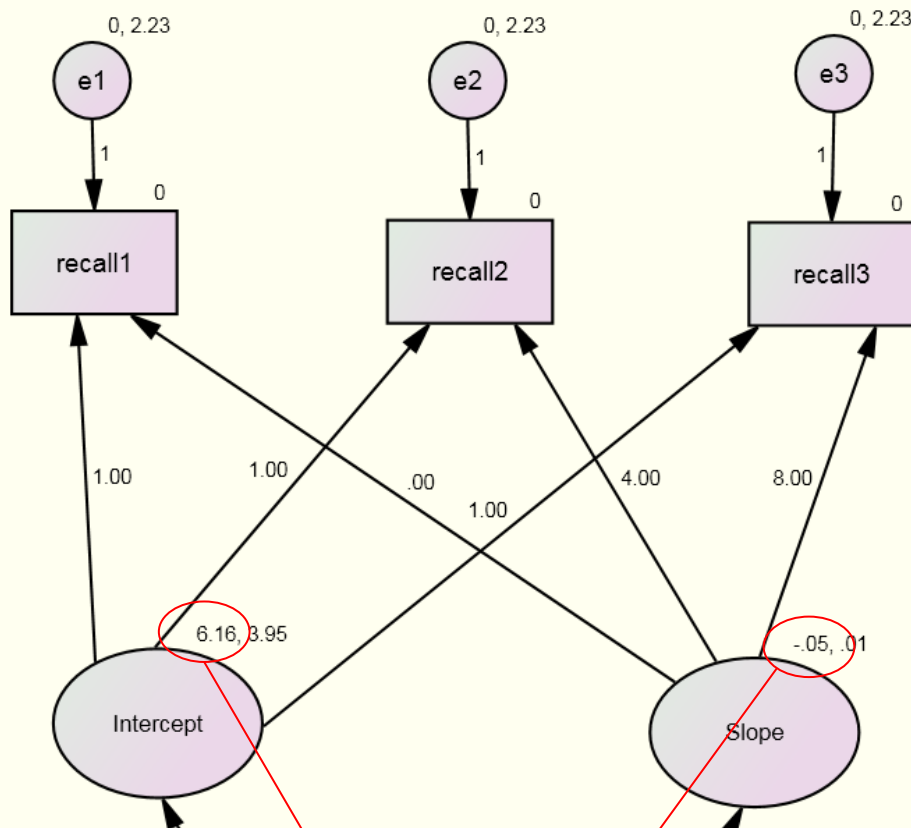
LGM as unconditional growth model



Model results



Fixed effects

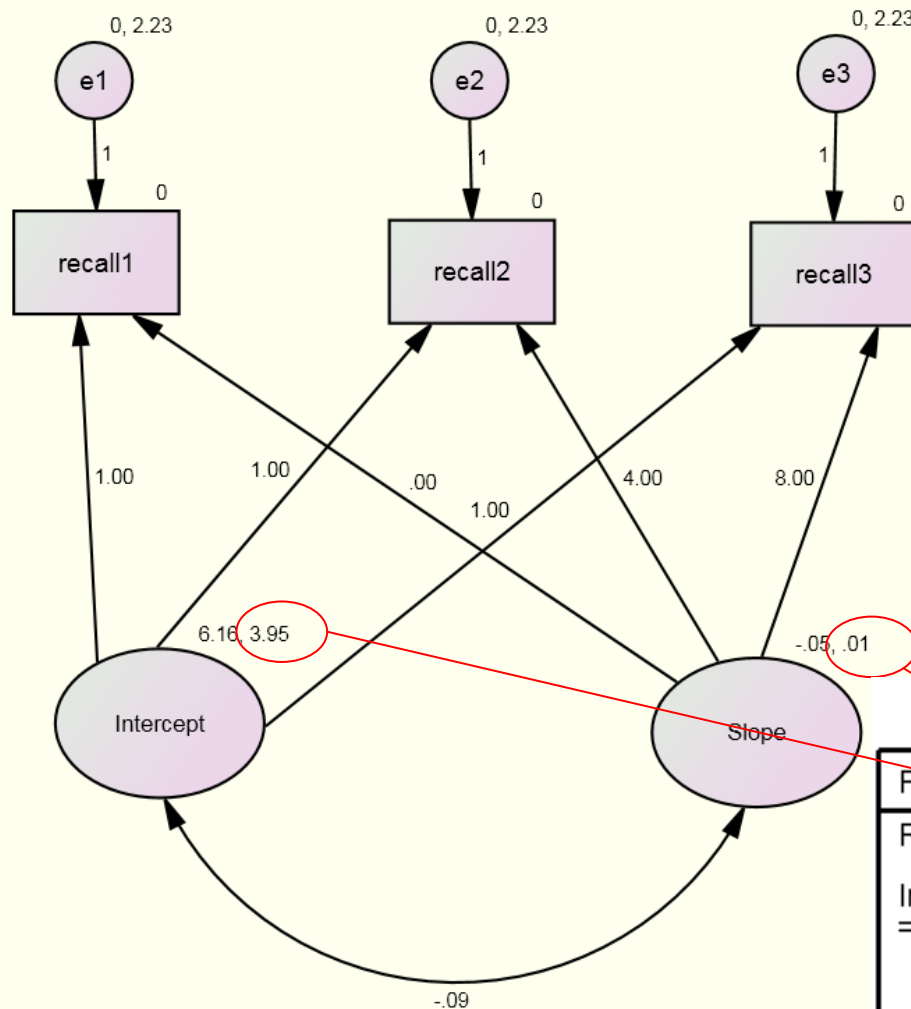


Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	6.159430	.048011	2543.232	128.292	.000	6.065285	6.253574
Time	-.046168	.006277	2201.954	-7.355	.000	-.058477	-.033859

a. Dependent Variable: recall recall.

Random effects

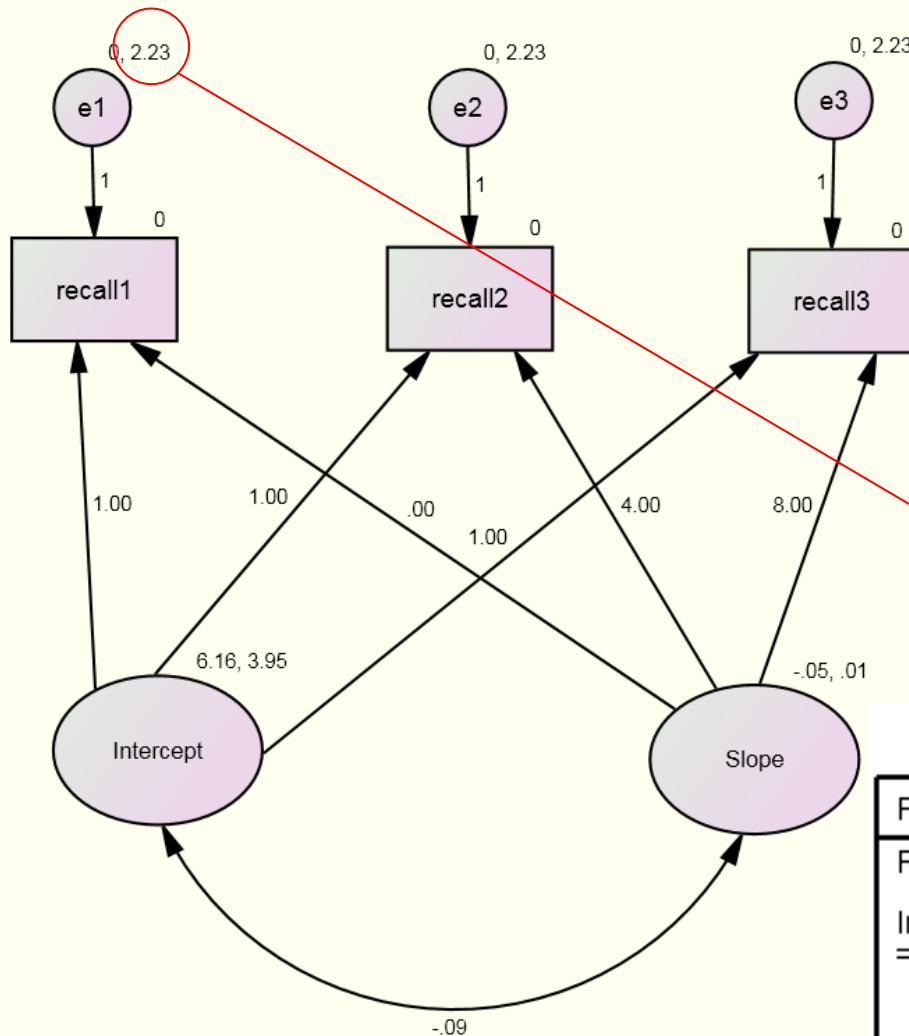


Estimates of Covariance Parameters^a

Parameter	Estimate	Std. Error
Residual	2.234215	.06962
Intercept + Time [subject = id]	3.956218	.17529
	-.089542	.01927
	.013531	.00346

a. Dependent Variable: recall recall.

Residual variance

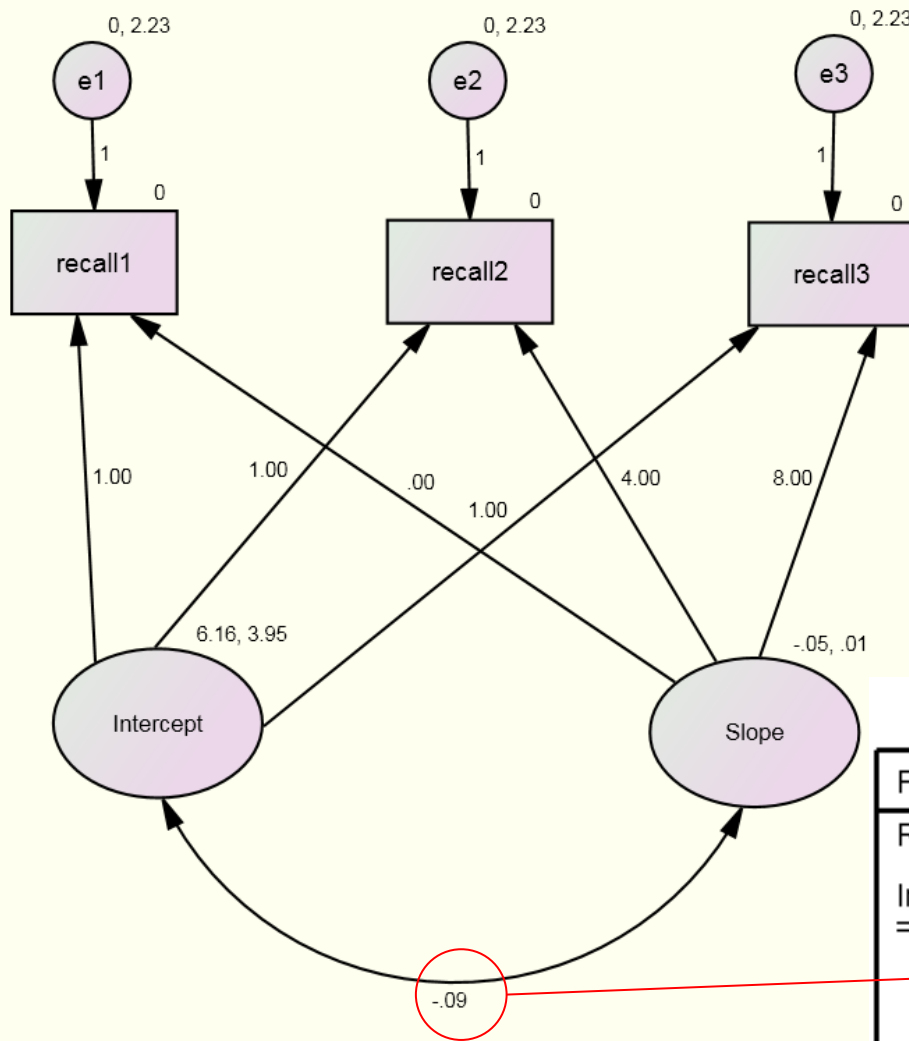


Estimates of Covariance Parameters^a

Parameter		Estimate	Std. Error
Residual		2.234215	.06962
Intercept + Time [subject = id]	UN (1,1)	3.956218	.17529
	UN (2,1)	-.089542	.01927
	UN (2,2)	.013531	.00346

a. Dependent Variable: recall recall.

Intercept-slope covariance



Estimates of Covariance Parameters^a

Parameter	Estimate	Std. Error
Residual	2.234215	.06962
Intercept + Time [subject = id]		
UN (1,1)	3.956218	.17529
UN (2,1)	-0.089542	.01927
UN (2,2)	.013531	.00346

a. Dependent Variable: recall recall.

MLM or LGM?

- Advantages of MLM
 - More readily incorporates additional hierarchies in the data (e.g., 3 level model: occasions (Level 1) nested within individuals (Level 2), nested within schools (Level 3))
 - Accommodates unevenly spaced measurement intervals (i.e. time can be treated more flexibly)
 - Does not require large samples for reliable estimates

MLM or LGM?

- Advantages of LGM

- Allows multiple indicators using common factor measurement model
- Can accommodate missing data on predictor as well as outcome variables via ML estimation
- More flexible treatment of relationships among predictors
- Some better options for limited measurement occasions – e.g., latent difference score models (2-occassions)
- Generalises to multivariate context (i.e., multiple correlated growth processes)

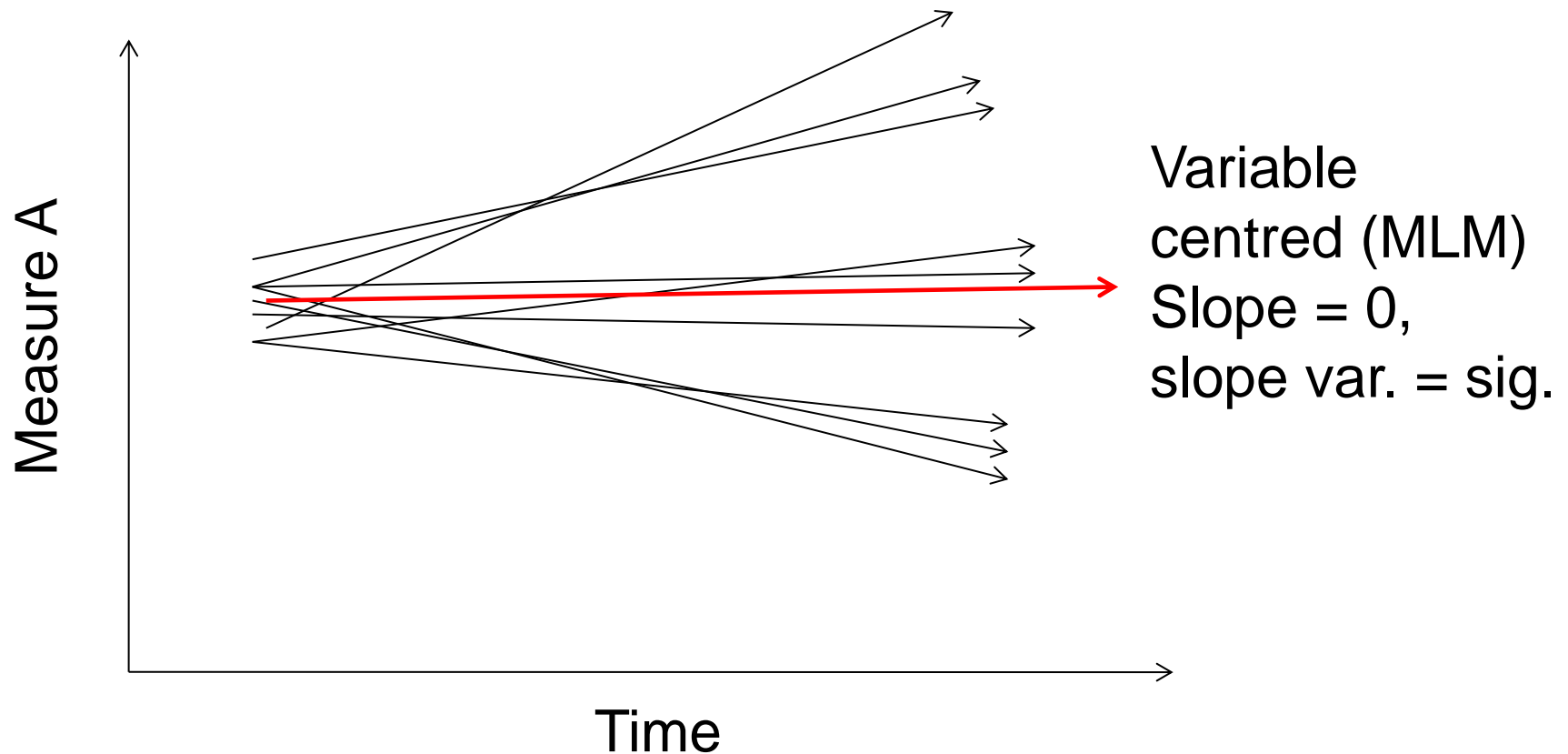
Some extensions of LGM

- **Bivariate dual change score model (BDCSM)**
- Examination of dynamic patterns of development over time
- Do changes in one variable (e.g., well-being) tend to 'lead' changes in another (e.g., cognition)?
- Comparison of overall fit for models representing different 'lead-lag' associations
- Produces stronger evidence for making causal inferences than is often possible in other models
- *Note that lead-lag models can also be fitted in MLM, though less flexibility for comparing fit of different models*

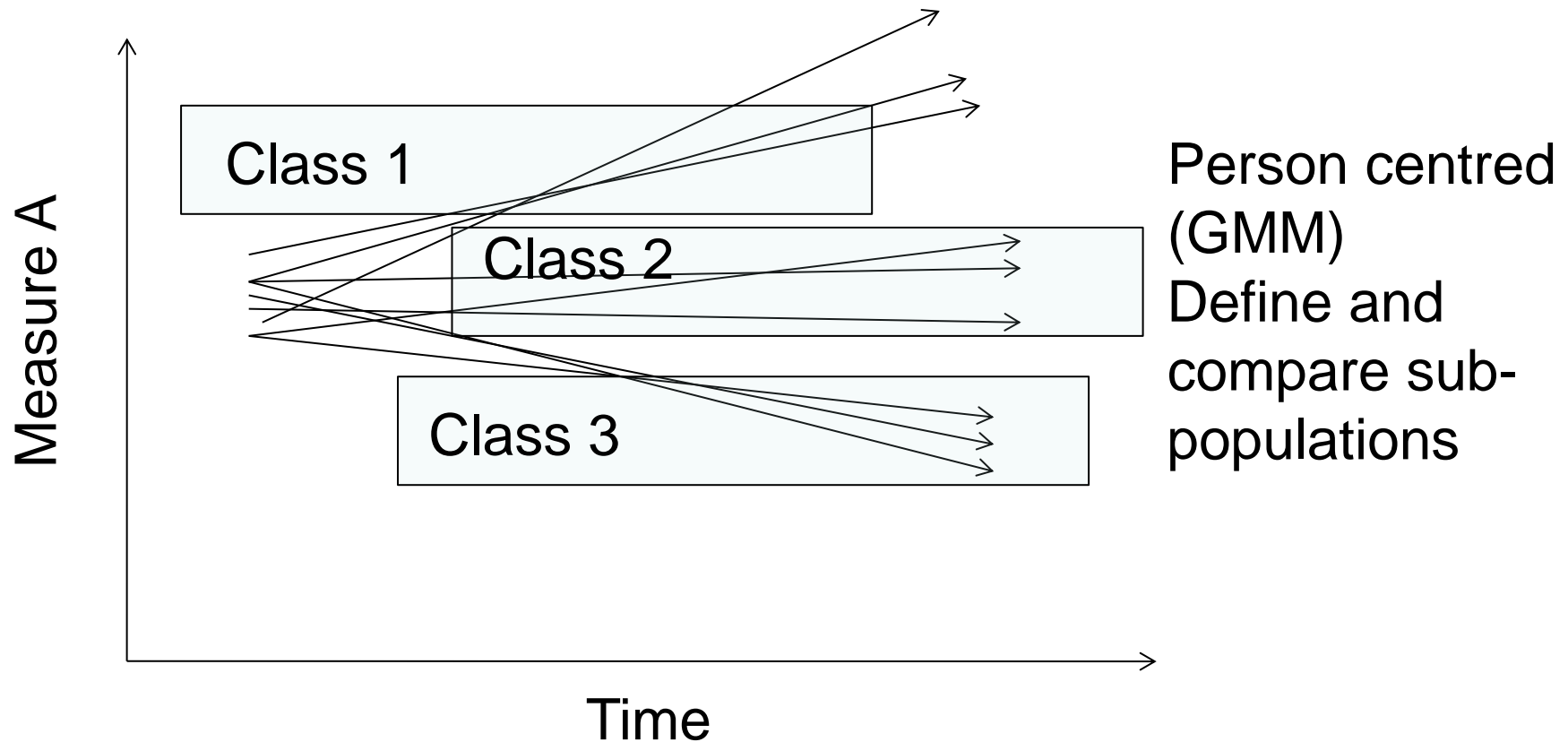
Person-centred approaches

- Conventional growth modelling (e.g., MLM, LGM) assumes that individuals come from a single population, and that a single growth trajectory can adequately approximate development in that population
- Person-centred approaches (e.g., Growth Mixture Models - GMM) identify and compare sub-populations characterised by different patterns of change

Example. Theory suggests that scores on measure A will increase for some, decrease for some, and remain unchanged for others

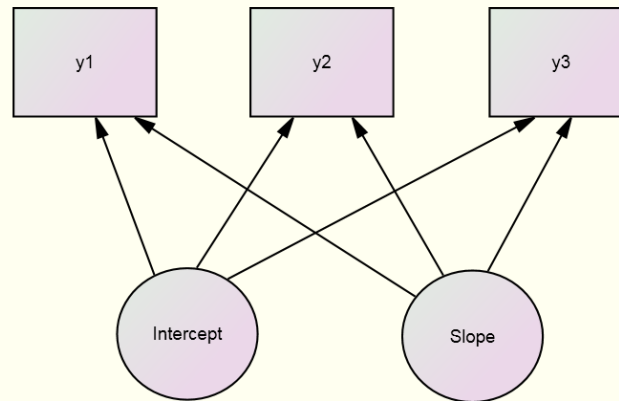


Example. Theory suggests that scores on measure A will increase for some, decrease for some, and remain unchanged for others



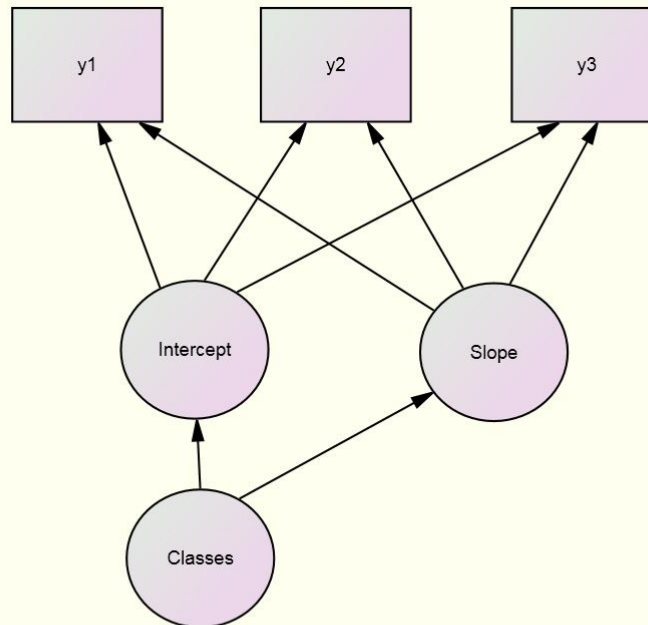
Growth mixture modelling (GMM)

1. Start with
LGM



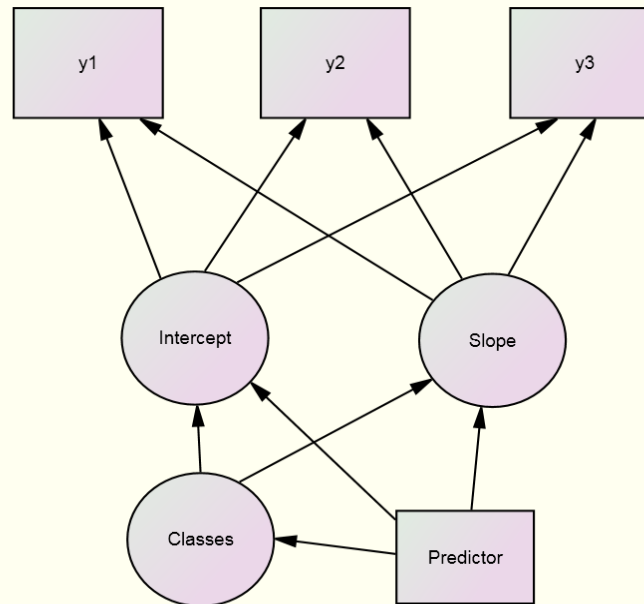
Growth mixture modelling (GMM)

2. Are sub-populations evident based on patterns of trajectories?



Growth mixture modelling (GMM)

3. Do predictor variables explain differences in class membership?



Incomplete overview of software

	SPSS	Stata	Mplus	SAS	R
MLM	Yes	Yes	Yes	Yes	Yes
GLMM (categorical DV)	Yes*	Yes	Yes	Yes	Yes
GEE (categorical DV)	Yes	Yes	No	Yes	Yes
LGM (SEM)	No	Yes**	Yes	No	Yes
GMM	No	No	Yes	Yes	Yes

Other MLM specific software: MLwiN, HLM

Other SEM specific software: Lisrel, AMOS, EQS

*version 19, **version 11

References and resources

Text books

- Multilevel modelling
 - Singer, J.D., & Willett, J.B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
 - Snijders, T.A.B., & Bosker, R.J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). London: SAGE Publications
 - Kreft, I., & De Leeuw, J. (1998). *Introducing multilevel modeling*. London: SAGE Publications
 - Twisk, J.W.R. (2006). *Applied multilevel analysis*. United Kingdom: Cambridge University Press.
 - Fitzmaurice, G.M., Laird, N.M., & Ware, J.H. (2004). *Applied longitudinal analysis*. Hoboken, New Jersey: John Wiley & Sons.
 - Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using Stata* (2nd ed.). Texas: Stata Press.
- Latent growth modelling
 - Duncan, T.E., Duncan, S.C., & Strycker, L.A. (2009). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications* (2nd ed.). Taylor & Francis e-library: www.eBookstore.tandf.co.uk
- General
 - Newsom, J.T., Jones, R.N., & Hofer, S.M. (2012). *Longitudinal data analysis*. New York, Routledge.

- Journal articles
 - Stoel, R.D., van Den Wittenboer, G., & Hox, J. (2003). Analyzing longitudinal data using multilevel regression and latent growth curve analysis. *Metodologia de las Ciencias Del Comportamiento*, 5, 21-42.
 - Collins, L.M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, 57, 505-528.
 - Raudenbush, S.W. (2001). Comparing personal trajectories and drawing causal inferences from longitudinal data. *Annual Review of Psychology*, 52, 501-525.
 - Hertzog, C., & Nesselroade, J.R. (2003). Assessing psychological change in adulthood: An overview of methodological issues. *Psychology and Aging*, 18, 639-657.
 - Jung, T., & Wickrama, K.A.S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, 2, 302-317.
 - Wang, M., & Bodner, T.E. (2007). Growth mixture modeling. *Organizational Research Methods*, 10, 635-656.
 - Curran, P.J., & Bauer, D.J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, 62, 583-619.
- Websites
 - Mplus
<http://www.statmodel.com/>
 - UCLA Statistical computing
<http://www.ats.ucla.edu/stat/>
 - University of Bristol Centre for multilevel modelling
<http://www.bristol.ac.uk/cmm/learning/videos/random-slopes.html>

SPSS syntax for MLM example

* Examine individual growth trajectories for recall*.

GRAPH

```
/LINE(MULTIPLE)MEAN(recall) BY Time BY id  
/TITLE= 'Individual Trajectories for Recall - first 100 pps'.
```

* Variance components model.

MIXED recall

```
/METHOD = REML  
/PRINT = SOLUTION  
/RANDOM = INTERCEPT | SUBJECT(id) COVTYPE(UN).
```


SPSS syntax for MLM example (continued)

* Unconditional growth model fixed linear effect of time.

```
MIXED recall WITH Time  
  /METHOD = REML  
  /PRINT = SOLUTION  
  /FIXED = Time  
  /RANDOM = INTERCEPT | SUBJECT(id) COVTYPE(UN).
```

* Unconditional growth model random linear effect of time.

```
MIXED recall WITH Time  
  /METHOD = REML  
  /PRINT = SOLUTION  
  /FIXED = Time  
  /RANDOM = INTERCEPT Time | SUBJECT (id) COVTYPE (UN).
```

* Include time-invariant (level 2) predictors .

```
MIXED recall WITH Time age_c63 female yred_c14  
  /METHOD = REML  
  /PRINT = SOLUTION  
  /FIXED = Time age_c63 female yred_c14  
  /RANDOM = INTERCEPT Time | SUBJECT (id) COVTYPE (UN).
```

* Test cross level education by time interaction .

```
MIXED recall WITH Time age_c63 female yred_c14  
  /METHOD = REML  
  /PRINT = SOLUTION  
  /FIXED = Time age_c63 female yred_c14 yred_c14*Time  
  /RANDOM = INTERCEPT Time | SUBJECT (id) COVTYPE (UN).
```