

MASTERCLASS SERIES 2026/27

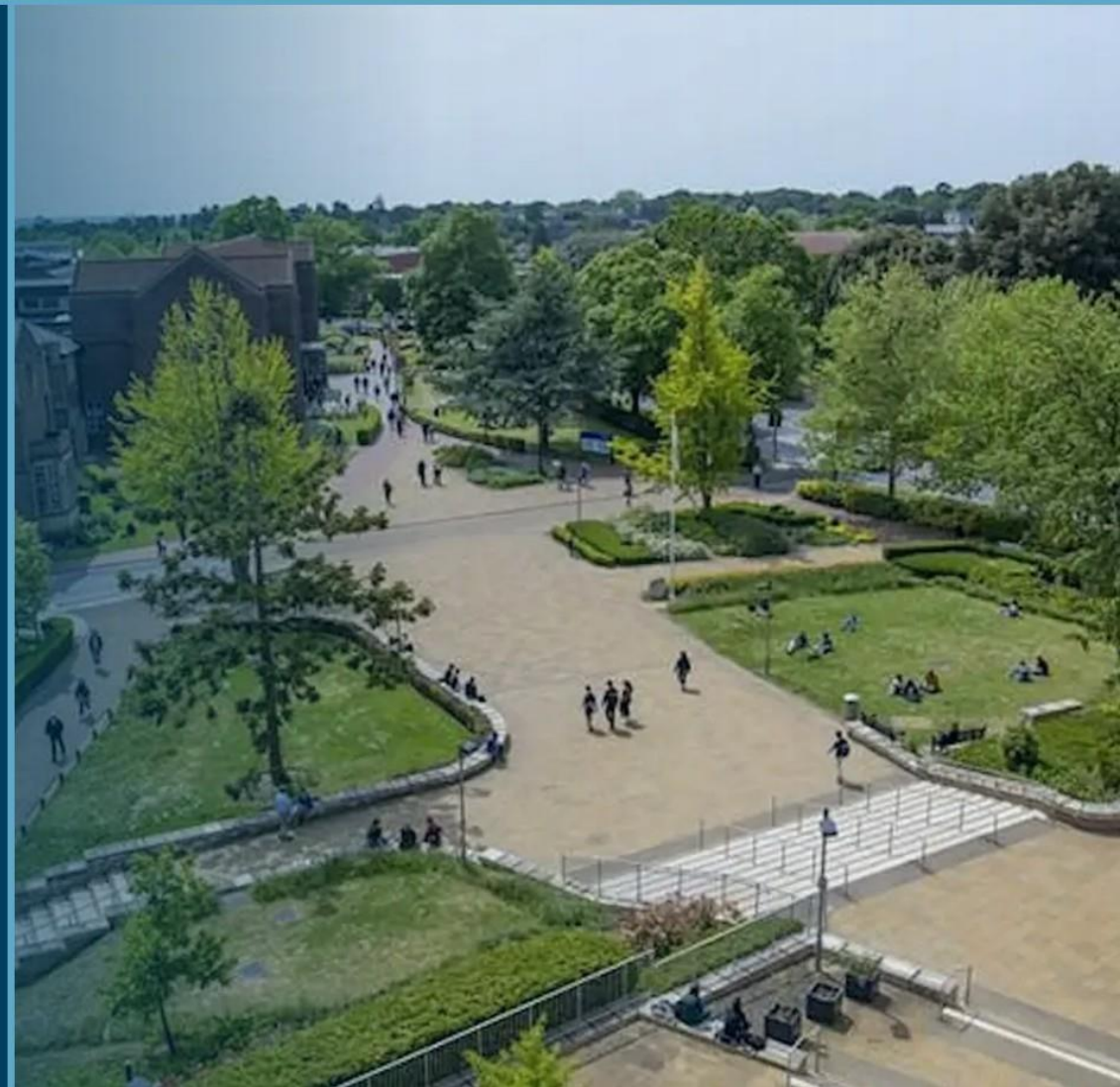
Longitudinal Growth Modelling with Structural Equation Models

James Hall

Associate Professor of Psychology of Education

Southampton Education School, University of Southampton

21 October 2026 ▪ 1–3pm UK time ▪ Online ▪ J.E.Hall@Soton.ac.uk



Today: Longitudinal Growth Modelling with Structural Equation Modelling (SEM)



Aims:

To develop attendees' understanding of:

1. What Longitudinal Growth Modelling is conceptually
2. The two major statistical approaches to growth modelling
 1. Via Multilevel Modelling (MLM)
 2. Via Structural Equation Modelling (SEM)
3. How each approach to growth modelling treats data and data structure
4. How modern SEM software can model growth using both multilevel (nested) and latent growth approaches
5. The unique features of the SEM approach to growth modelling and why these features can make this approach useful
6. How to carry out a basic SEM growth model in the R software package
7. Options for where to go next

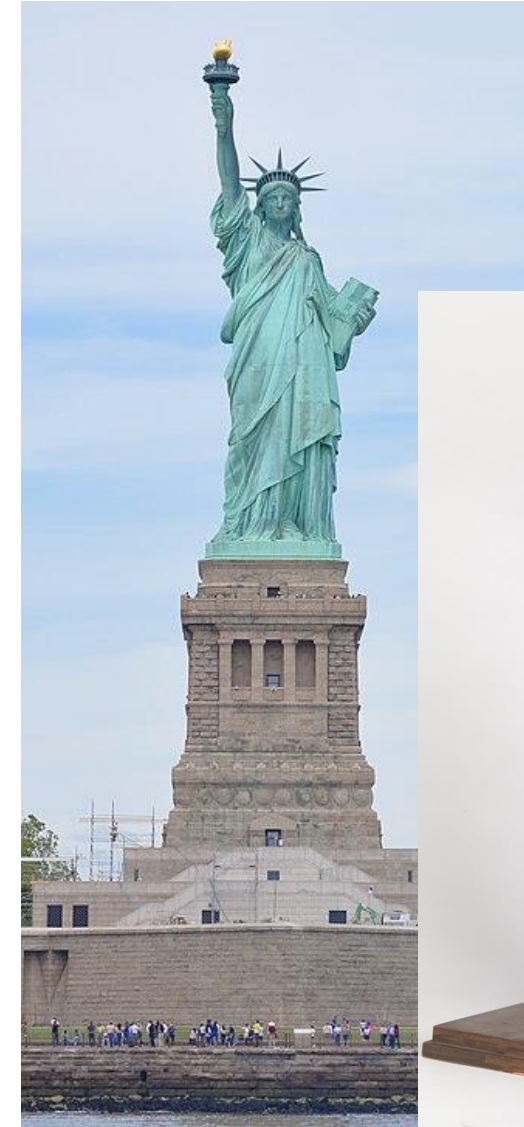
Content

- Part 1: Concepts
 - Defining Growth Modelling
 - Two Approaches to Growth Modelling
 - Growth Modelling and Data Structure
 - Growth Modelling via Multilevel Modelling (MLM) and Structural Equation Modelling (SEM)
 - Why Focus on SEM for Growth Modelling?
 - Extensions to the SEM approach to Growth Modelling
 - Further Learning
- Part 2: Guided walkthrough using R
 - Setup & Overview
 - Fit a Basic SEM Growth Model
 - Inspect Model Output
 - Modify & Extend
 - Wrap-Up & Q&A

Part 1: SEM for Growth Modelling

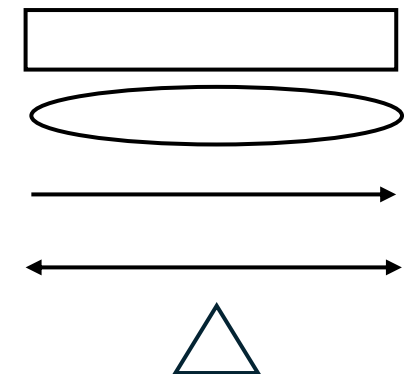
Structural Equation Modelling (SEM) - refresher

- A family of statistical modelling techniques.
- ‘Modelling’ – Recreating a structure and/or process in the world. In this case with numeric data about that structure/process.
 1. Judge if the model is a good replica
 2. Then use a good replica to better understand the structure and/or process
- SEM is a multivariate technique that can be viewed as extending and integrating both general linear modelling (e.g. multiple regression; ANOVA) and latent variable statistics (e.g. factor analysis; cluster analysis)
- Theory-driven and brings close together conceptual models with corresponding empirical testing



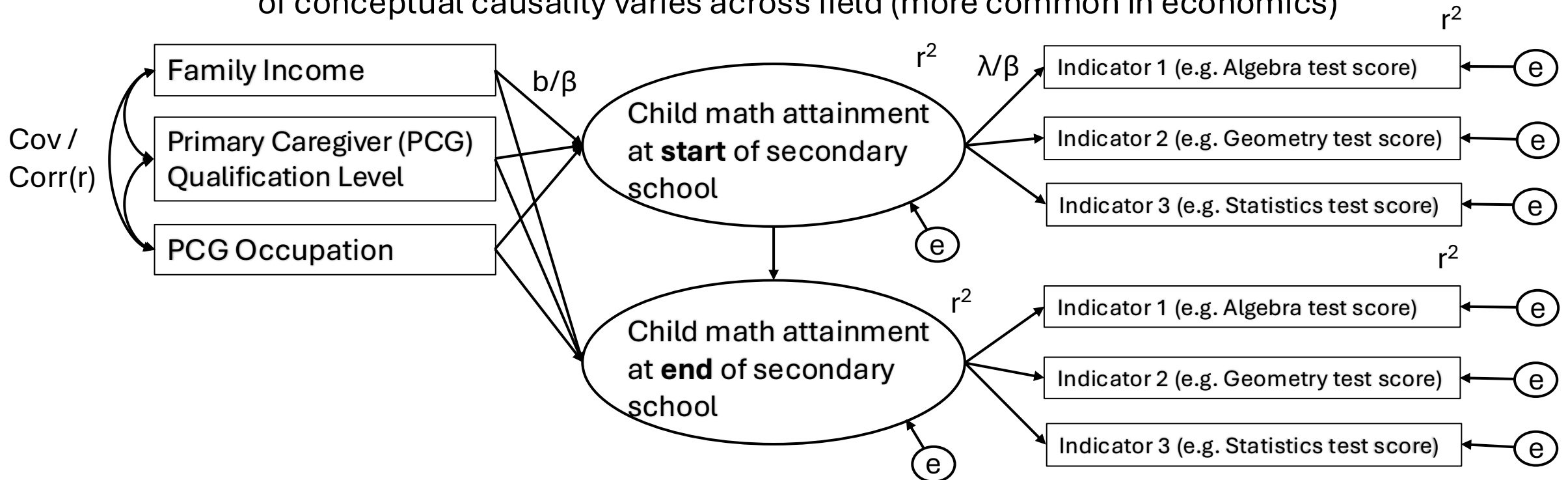
Structural Equation Modelling (SEM) - refresher

- A family of statistical modelling techniques
- SEM is comprised of two parts:
 - A structural model (building on general linear modelling)
 - A measurement model (building on latent variable statistics)
- Simple SEMs can be represented with diagrams.
 - These diagrams blur the lines between a conceptual model explaining how concepts relate and a statistical (SEM) model that is developed to test that conceptual model.
 - Rectangles represent directly 'observed' concepts
 - Ovals represent indirectly 'observed'/'latent' concepts
 - Single-headed arrows represent statistical dependence
 - Double-headed arrows represent correlation/covariance
 - Triangles are the means of variables



Example Structural Equation Model

- Featuring both a Structural and Measurement Model
 - The “e” is a regression residual for each dependent variable (DV)
 - Note the directional arrows implementing statistical dependency
 - How much one also views these statistical dependencies as engaging with notions of conceptual causality varies across field (more common in economics)



Model Fit Estimators: RMSEA, CFI, and SRMR

Fit Index	(Aside: What It Tells You)	Cutoff Guidelines	Type
RMSEA	<i>(How well the model approximates the population data per degree of freedom)</i>	< 0.06 = good 0.06 – 0.08 = okay	Absolute fit (parsimony adjusted)
CFI	<i>(How much better your model fits than a null model)</i>	> 0.95 = good > 0.90 = acceptable	Comparative fit
SRMR	<i>(Average difference between observed and model-implied correlations)</i>	< 0.08 = good	Absolute fit (residual-based)

- RMSEA = Root Mean Square Error of Approximation
- CFI = Comparative Fit Index
- SRMR = Standardized Root Mean Square Residual

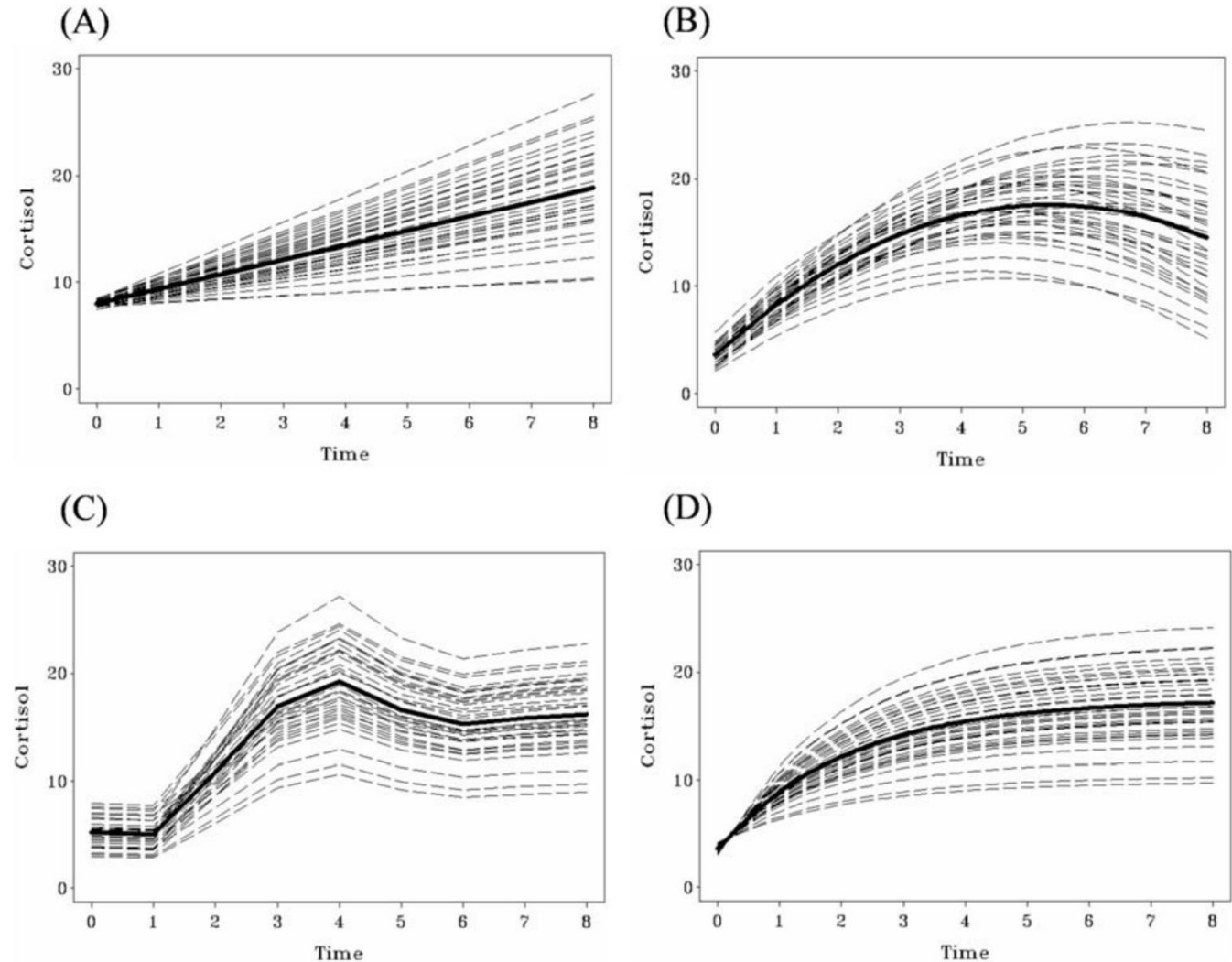
Questions?

Growth Modelling

- The statistical modelling of within-unit change over time.
 - “Within-unit”? Individuals, groups, institutions — any repeated measurement.
- Typically used to estimate:
 - A point in time from which to see the growth (intercept; often some kind of starting point)
 - The rate of change (slope)
 - The variation in intercepts/slopes across individuals
- Used in education research, developmental psychology, health research, etc.
- The underlying maths is that of any straight line: $Y = mX + C + e$
 - Where: m is the slope (rate of change),
 - C is the intercept (starting point),
 - and e is residual (unexplained variation)
 - Alternatively: $Y = \text{intercept} + \text{slope} \times \text{time} + e$

Example growths (from Ram & Grimm, 2007)

- Predicted average (bold lines) and predicted intraindividual changes in cortisol.
 - (A) Linear,
 - (B) quadratic,
 - (C) latent basis,
 - (D) exponential
- Note, for modelling cubic and higher power, you always need 1 more measurement point than the intended power
 - 3+ data points for quadratic
 - 4+ data points for cubic



Why estimate a Growth Model?

- Repeated measures ANOVA can only model *average* change
- Growth modelling allows us to:
 - Capture individual trajectories
 - Estimate variance in change (some improve, some worsen, some don't change)
 - Model predictors of growth (e.g. what explains who improves more?)
- Repeated measurement is a strict requirement:
 - Same measurement process
 - Same metric
 - Same substantive meaning
 - Not meaningful: growth modeling across z-scores from different measures

Two approaches to growth modelling

- Most statistical growth models fall into one of two camps:
 - Multilevel Modelling (MLM; AKA ‘hierarchical linear models’)
 - Structural Equation Modelling (SEM; latent growth curve models)
- Both estimate:
 - Intercept (e.g. starting point)
 - Slope (rate of change)
 - Variance in both
- They differ in assumptions, data structure, and flexibility
- Making matters a little more complicated: modern SEM software packages can carry out growth modelling in both the MLM and SEM ways
 - Mplus
 - R (via the OpenMX package)
 - This is a product of SEM being a ‘family’ of approaches under continual development

Growth Modelling and Data Structure

- Growth modelling requires us to store repeated measures data in a correct format
 - “Long format” (MLM) vs. “Wide format” (SEM).
 - Most modern software packages can “restructure” your data files to move between these formats

Long format (MLM):

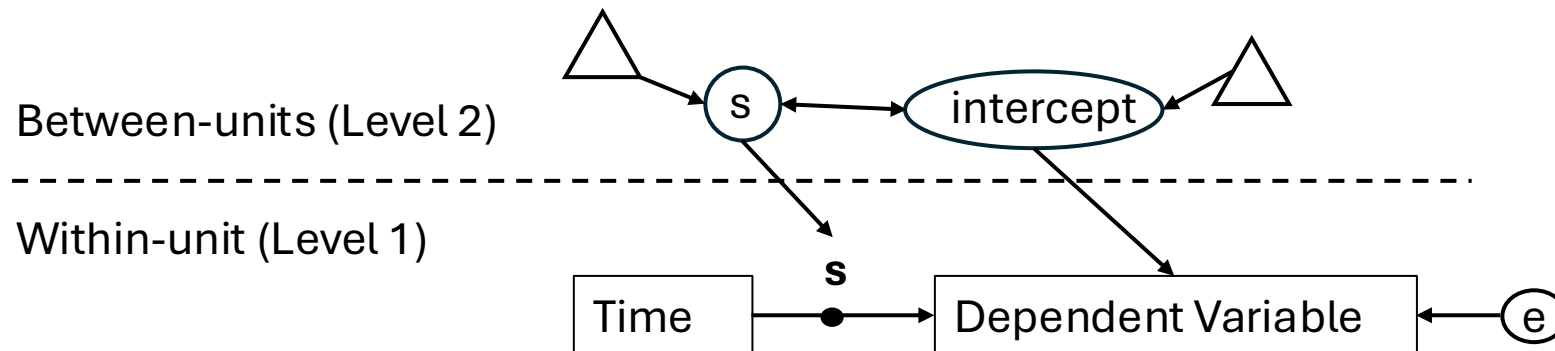
Person	Timepoint	Reading
A	1	88
A	2	92
A	3	96
B	1	91
B	2	97
B	3	106
C	1	99
C	2	100
C	3	101
D	1	94
D	2	95
D	3	95

Wide format (SEM):

Person	T1_Reading	T2_Reading	T3_Reading
A	88	92	96
B	91	97	106
C	99	100	101
D	94	95	95

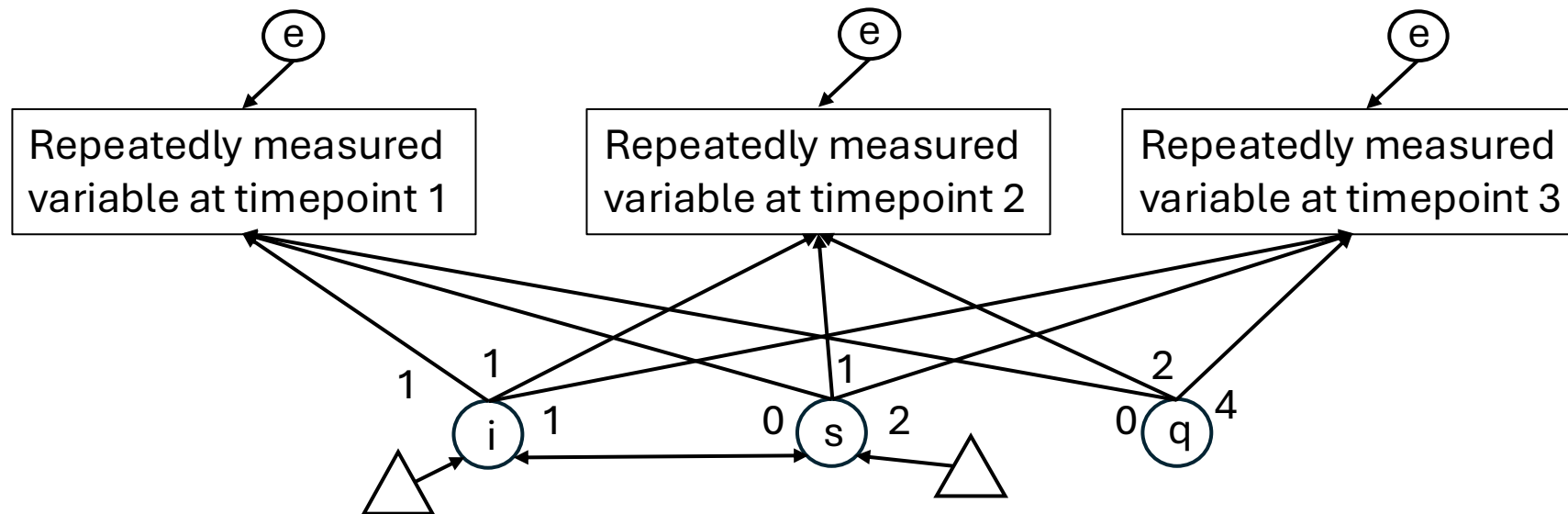
Growth Modelling via Multilevel Modelling

- Observations are nested within units (e.g., repeated measures within people)
- Time is a within-unit predictor of a dependent variable
- Versus the SEM approach, this approach can handle:
 - Unbalanced data (participants don't need the same number of time points)
 - Flexible time structures (inconsistent time intervals between participants are okay)
- Because this approach is also available via modern SEM software packages, we can illustrate it using the SEM diagram approach (inc. covariance & means):



Growth Modelling via Structural Equation Modelling

- Time-specific outcomes are modelled as indicators of:
 - A latent intercept factor
 - A latent slope factor. You estimate the mean and variance of these latent variables
- SEM includes measurement error and supports multiple growth processes





Quantitative
Methods in
Education Group

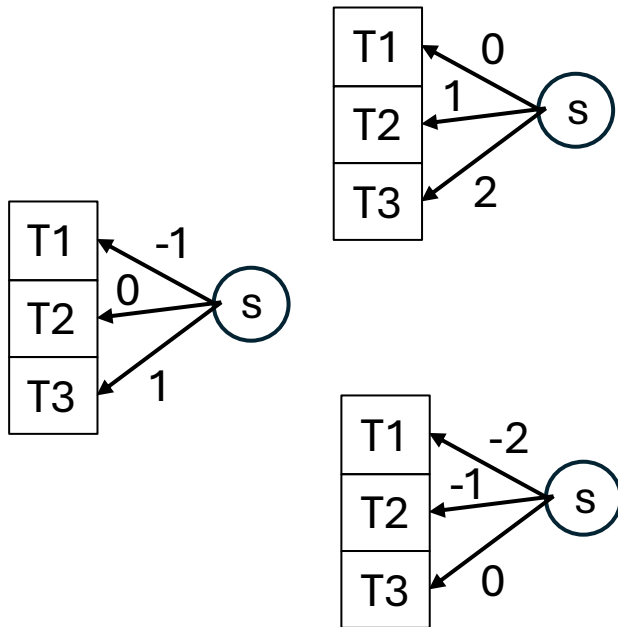
Questions?

Why carry out Growth Modelling via the SEM approach rather than the MLM approach?

- There are unique advantages to both approaches.
 - We've already covered the advantages of the MLM approach
- Latent Growth Modelling via SEM uniquely enables:
 - Growth in latent constructs (not just observed scores)
 - E.g. estimate latent reading scores, then estimate growth across these
 - Testing of longitudinal measurement consistency
 - Nonlinear or custom growth via flexible factor loadings on slope variables
 - Full integration of structural pathways (allowing tests of mediation, moderation etc)

Growth Modelling via the SEM approach

- We can set the intercept parameter at whatever point in time we want.
- The intercept of the slope is wherever we set our zero factor loading on a slope latent variable.
 - This creates opportunities for new research questions
 - E.g. We can have a “launch” intercept statistically predict subsequent growth, or growth statistically predict the “landing” intercept



Model Name	Intercept Placement	Use Case
Launch Model	At the start of growth	Baseline or pre-treatment
Landmark Model	At a theoretically meaningful point within the growth period (e.g., midpoint)	During an intervention
Landing Model	At the end of growth	Post-treatment, follow-up, outcome-focused

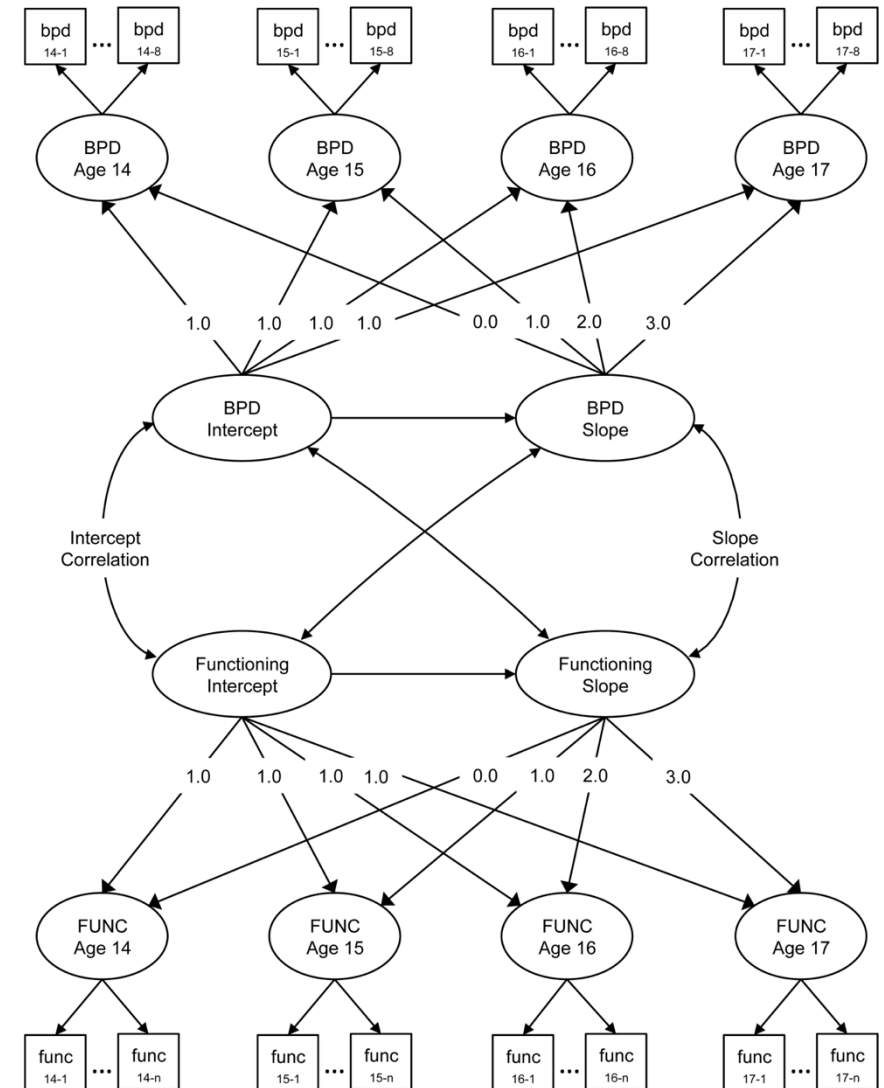
Extensions to the SEM approach to Growth Modelling – parallel growth models

- It would take another 1-2 hours to cover the many extensions that exist to both types of growth model so far considered...
- E.g. Estimating parallel growth models is one popular extension
- Here, each construct has its own intercept and slope (latent growth factors)
- The model estimates:
 - Correlation between slopes (e.g., does increasing motivation relate to decreasing school detentions?)
 - Correlation between intercepts
- Supports testing of theory-driven developmental dynamics (e.g., co-morbidity, compensation, mutual reinforcement)

Parallel growth modelling

- Example

- Wright, A. G., Zalewski, M., Hallquist, M. N., Hipwell, A. E., & Stepp, S. D. (2016). Developmental trajectories of borderline personality disorder symptoms and psychosocial functioning in adolescence. *Journal of Personality Disorders*, 30(3), 351-372.
- At-risk community sample of girls (n=2,450)... across ages 14–17
- Bivariate latent growth curve models were used to analyze the co-development of **BPD symptoms** with eight domains of **psychosocial functioning** (e.g., social skills, sexual behavior)



Extensions to the SEM approach to Growth Modelling – Growth Mixture Models

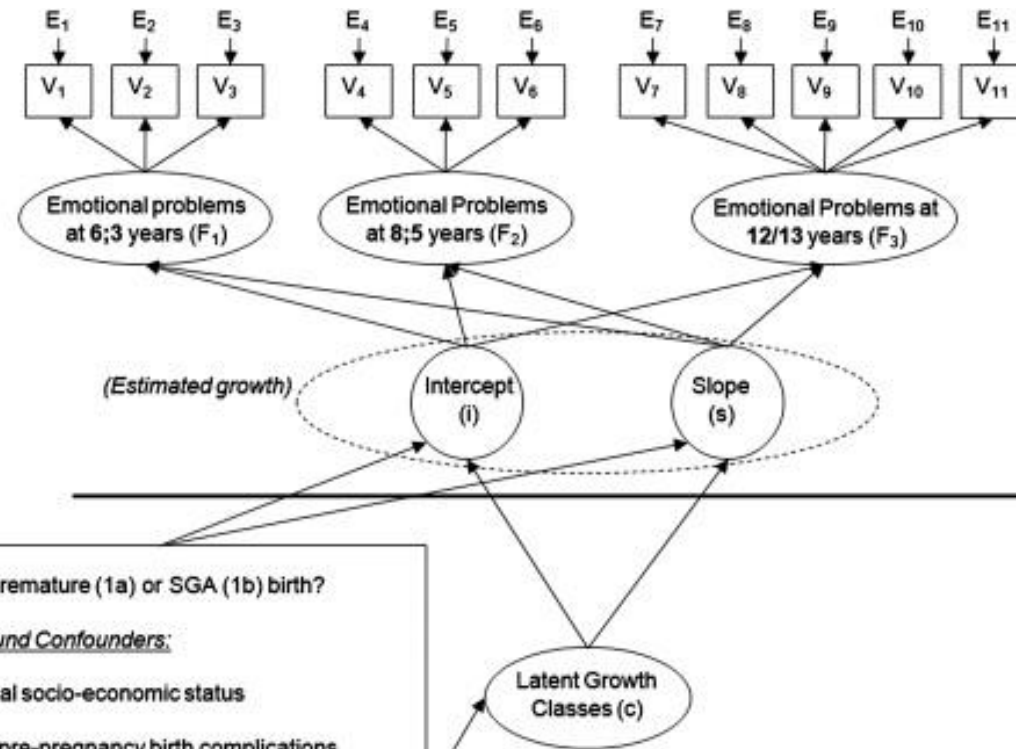
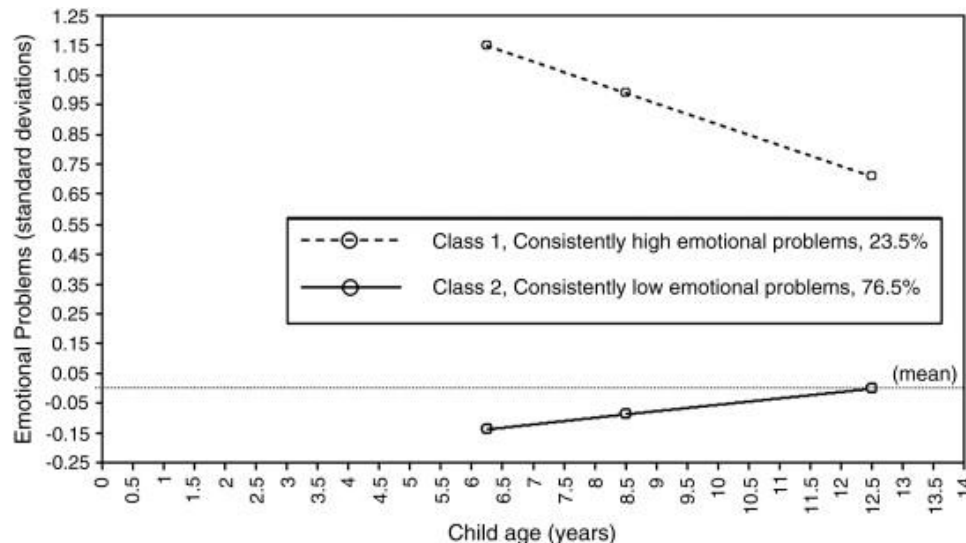


[Quantitative
Methods in
Education Group](#)

- ...Estimating whether there are homogenous patterns of growth within a growth model is another popular example extension of growth modelling
- In Growth Mixture Modelling (GMM), the sample is assumed to contain latent classes, each with its own:
 - Intercept and slope parameters
 - Within-class variation in growth (random effects)
 - (This variation is optional to estimate. Without it we have “Latent Class Growth Models”).
- Allows for:
 - Different patterns of change across subgroups
 - Identification of distinct developmental pathways
- Class membership can be: Unexplained (purely latent), or statistically predicted by covariates

Growth mixture modelling - Example

- Hall, J., & Wolke, D. (2012). A comparison of prematurity and small for gestational age as risk factors for age 6–13 year emotional problems. *Early Human Development*, 88(10), 797-804.
- 654 children who were followed from birth to age 12/13 years.
- Emotional problems at ages 6.3 and 8.5 years were measured via the CBCL, at age 12/13 years via the SDQ.
- Latent class growth analysis revealed:
 - A low and stable level of emotional problems in 76% of children;
 - A high and stable level of emotional problems in 24% of children.
- The high and stable pattern of emotional problems was significantly associated with a VP but not an SGA birth.



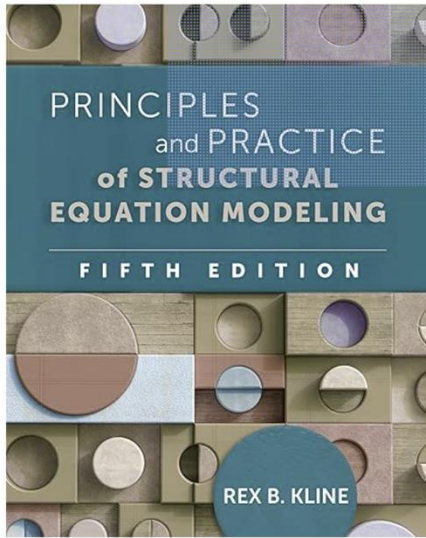
Software for Growth Modelling SEM

- Dedicated software:
 - **AMOS (part of SPSS): Great for learners** as very intuitive. Less suited specifying SEM Growth Models compared to R or Mplus. Expensive to purchase -- though many organisations have a site license.
 - **R** (Most often via the **laavan** package): **Great for doing SEM**. A slower learning curve and slower to implement latest SEM developments (vs Mplus) but free.
 - **Mplus: Great for power users**. Has most/all of the latest SEM developments. Very easy to use. Expensive to purchase – organisations unlikely to have a site license.

Suggested follow-on reading

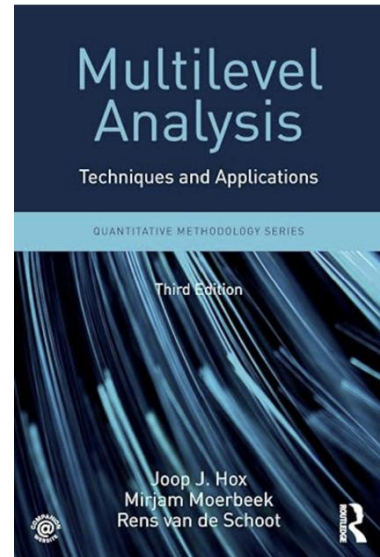


[Quantitative
Methods in
Education Group](#)



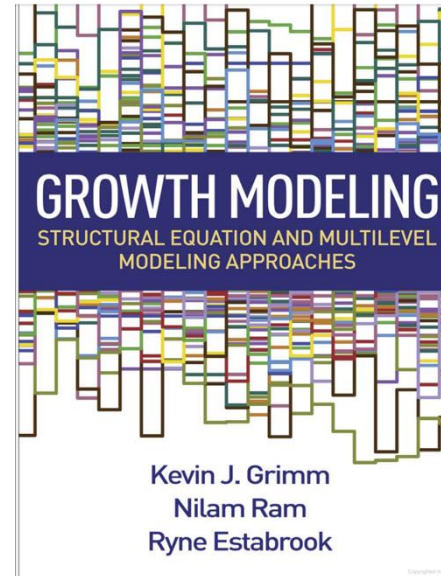
Kline, R. B. (2023). *Principles and practice of structural equation modeling*. (5th Edition). Guilford publications.

Great starter SEM textbook – lots of direct instruction on concepts and on different simpler models



Hox, J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.

Great starter MLM textbook – includes growth modelling via MLM



Grimm, K. J., Ram, N., & Estabrook, R. (2016). *Growth modeling: Structural equation and multilevel modeling approaches*. Guilford Publications.

Expands on all the content of this talk. Includes introductions to MLM and SEM

Continuing problems with SEM textbooks:

1. Most focus too much on how to run models, not why the model makes sense/is appropriate for use in different contexts
2. They often skip the complex stuff, like moderation, nonlinear effects, & theory-building
3. They rarely connect measurement, theory, and inference in a meaningful way

Continuing problem this fosters:

SEM often gets taught as a technical tool, not as a way to think clearly about research questions.

Lingering question: Given SEM is so widely popular – why is there still no good textbook? Is it just too difficult/impractical?



Quantitative
Methods in
Education Group

Questions?

5-minute break

Part 2: SEM Growth Modelling in the R Software Package - A guided walkthrough

Aims for this guided walkthrough

To show how to:

1. Load the required SEM packages

Example:

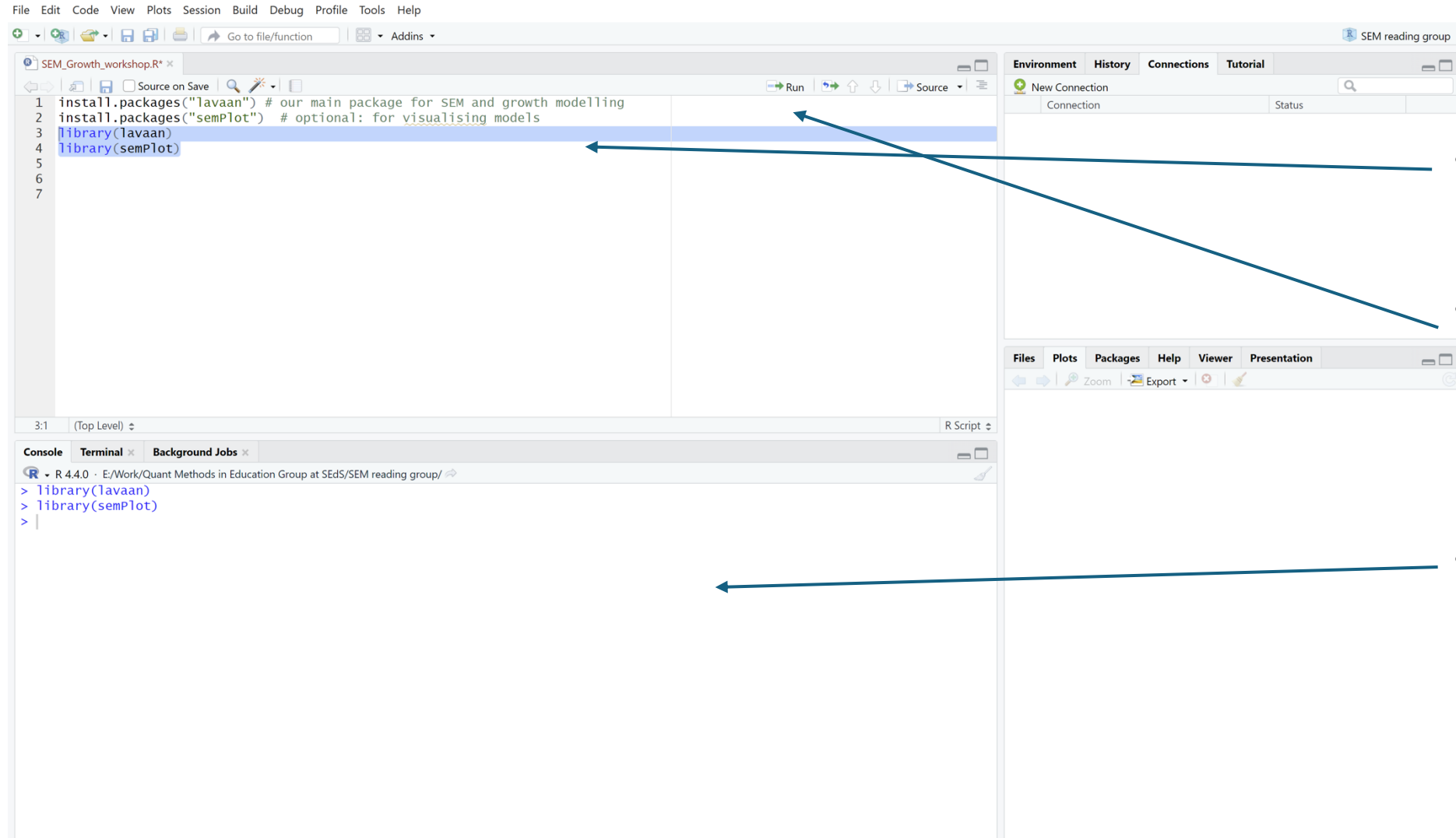
2. Setup & Overview
3. Look at the repeated measures data
4. Fit a Basic SEM Growth Model
5. Inspect Model Output
6. Modify & Extend
7. Discussion of possible next steps

Load required SEM packages

- Open Rstudio
- Open a new R Script (File > New File > R Script)
- Save it as sem_growth_workshop.R
 - This is your working file
- Loading R SEM packages (text that gets entered into the “script window” within RStudio):

```
install.packages("lavaan") # our main package for SEM and growth modelling
install.packages("semPlot") # optional: for visualising models
library(lavaan)
library(semPlot)
```
- One highlights each of those lines at a time (from the top) and then, for each line, one then hits the “run” button on the top right-hand side of the script window
- What does this look like in practice? See next slide.

Screenshot



• Script Window (top-left)

• Button to press to either run:

- All script in script window
- Just highlighted script

• Console Window (bottom-left)

Setup and overview

- What you'll see in this guided walkthrough:
 - Exploration of a repeated measure (**GPA over 5 timepoints from 200 US Highschoolers**)
 - (From a free-to-access online datafile shared by UCLA from a 2017 Joop Hox textbook)
 - Fit of an SEM growth model (A “latent growth curve model [LGCM]”) using lavaan
 - Interpretation:
 - Intercept (starting point)
 - Slope (rate of change)
 - Model fit
 - Visualisation of the model using semPlot
- Things to know:
 - We're working in wide format: one row per person
 - No latent measurement model here — just a basic SEM growth model

Look at repeated measures data

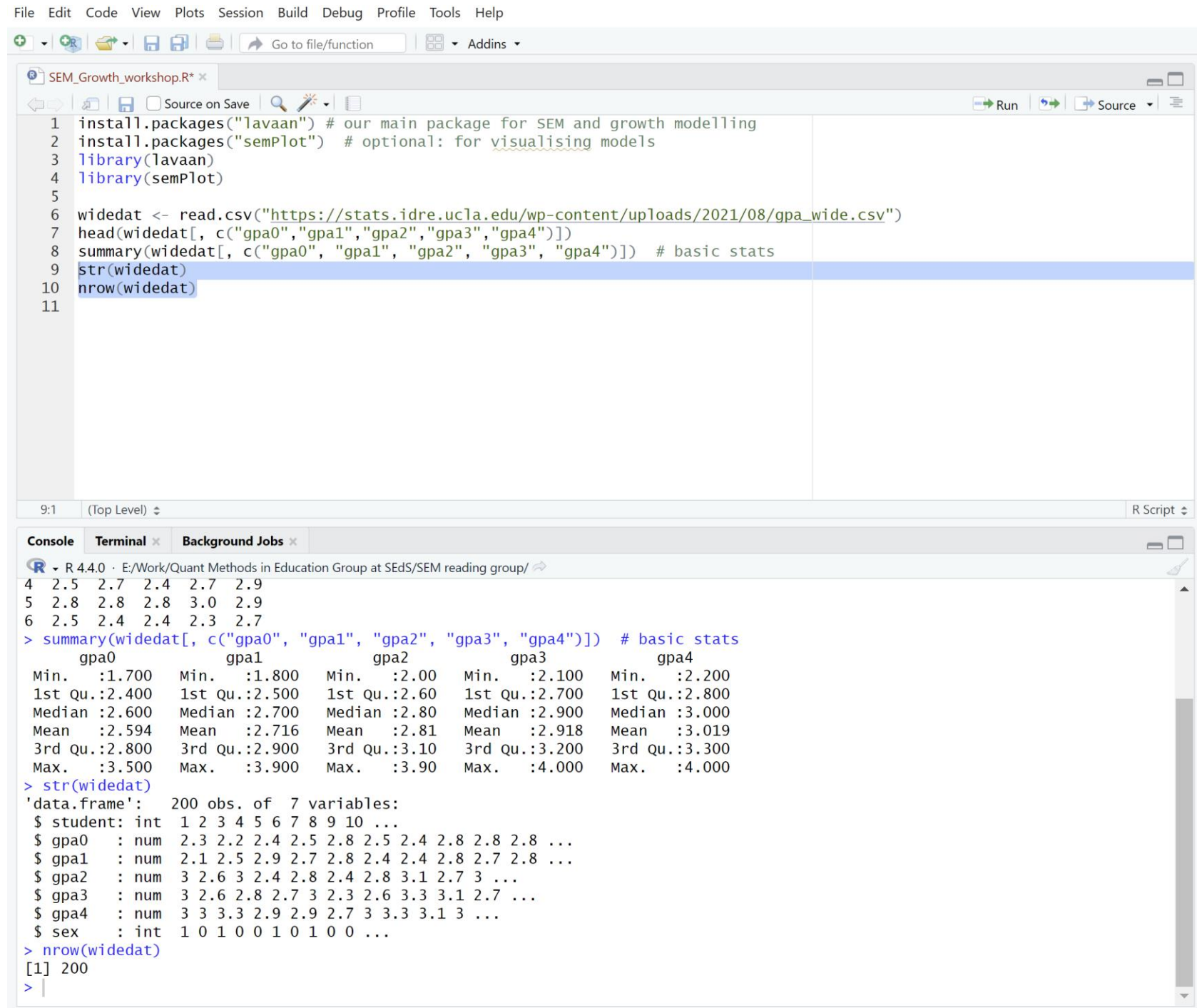
- We're using the gpa_wide.csv dataset freely shared by UCLA from Hox (2017).
- It contains 200 students, including 5 repeated measures of GPA
- First, we load and inspect the data. Text to enter into Rstudio (into the “script window”):

```
widedat <- read.csv("https://stats.idre.ucla.edu/wp-content/uploads/2021/08/gpa_wide.csv")
head(widedat[, c("gpa0","gpa1","gpa2","gpa3","gpa4")]) #top of variables
summary(widedat[, c("gpa0", "gpa1", "gpa2", "gpa3", "gpa4")]) # summary stats
str(widedat) # snapshot of datafile to allow structure to be seen
nrow(widedat) # number of rows (thus also number of participants)
```

- **Important note:**

- R needs *straight* quote marks ' and " and not *curly* quote marks ‘ and “.

Screenshot



The screenshot displays the RStudio environment. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for saving, running, and other functions. The main editor window shows a script named 'SEM_Growth_workshop.R' with the following code:

```
1 install.packages("lavaan") # our main package for SEM and growth modelling
2 install.packages("semPlot") # optional: for visualising models
3 library(lavaan)
4 library(semPlot)
5
6 widedat <- read.csv("https://stats.idre.ucla.edu/wp-content/uploads/2021/08/gpa_wide.csv")
7 head(widedat[, c("gpa0", "gpa1", "gpa2", "gpa3", "gpa4")])
8 summary(widedat[, c("gpa0", "gpa1", "gpa2", "gpa3", "gpa4")]) # basic stats
9 str(widedat)
10 nrow(widedat)
11
```

The console window at the bottom shows the output of the executed code:

```
R 4.4.0 · E:/Work/Quant Methods in Education Group at SEDS/SEM reading group/
4 2.5 2.7 2.4 2.7 2.9
5 2.8 2.8 2.8 3.0 2.9
6 2.5 2.4 2.4 2.3 2.7
> summary(widedat[, c("gpa0", "gpa1", "gpa2", "gpa3", "gpa4")]) # basic stats
      gpa0      gpa1      gpa2      gpa3      gpa4
Min.   :1.700 Min.   :1.800 Min.   :2.00  Min.   :2.100 Min.   :2.200
1st Qu.:2.400 1st Qu.:2.500 1st Qu.:2.60  1st Qu.:2.700 1st Qu.:2.800
Median :2.600 Median :2.700 Median :2.80  Median :2.900 Median :3.000
Mean   :2.594 Mean   :2.716 Mean   :2.81  Mean   :2.918 Mean   :3.019
3rd Qu.:2.800 3rd Qu.:2.900 3rd Qu.:3.10  3rd Qu.:3.200 3rd Qu.:3.300
Max.   :3.500 Max.   :3.900 Max.   :3.90  Max.   :4.000 Max.   :4.000
> str(widedat)
'data.frame': 200 obs. of 7 variables:
 $ student: int  1 2 3 4 5 6 7 8 9 10 ...
 $ gpa0    : num  2.3 2.2 2.4 2.5 2.8 2.5 2.4 2.8 2.8 2.8 ...
 $ gpa1    : num  2.1 2.5 2.9 2.7 2.8 2.4 2.4 2.8 2.7 2.8 ...
 $ gpa2    : num  3 2.6 3 2.4 2.8 2.4 2.8 3.1 2.7 3 ...
 $ gpa3    : num  3 2.6 2.8 2.7 3 2.3 2.6 3.3 3.1 2.7 ...
 $ gpa4    : num  3 3 3.3 2.9 2.9 2.7 3 3.3 3.1 3 ...
 $ sex     : int  1 0 1 0 0 1 0 1 0 0 ...
> nrow(widedat)
[1] 200
>
```

Specify our Basic SEM Growth Model

- Text to enter into Rstudio's Script Window:

```
model <- '
# Intercept and slope factors
i =~ 1*gpa0 + 1*gpa1 + 1*gpa2 + 1*gpa3 + 1*gpa4
s =~ 0*gpa0 + 1*gpa1 + 2*gpa2 + 3*gpa3 + 4*gpa4

# Variances and covariance
i ~~ i
s ~~ s
i ~~ s

# Residual variances
gpa0 ~~ gpa0
gpa1 ~~ gpa1
gpa2 ~~ gpa2
gpa3 ~~ gpa3
gpa4 ~~ gpa4
'
```

Notes:

- This text has created our growth model (and called it "model") but has NOT YET applied it to our data
- The i and s are latent variables (intercept and slope)
- The loadings on s define time spacing (0 to 4)
- This is a "launch model" as intercept is at first timepoint
- This model assumes linear change over time

Fit our Basic SEM Growth Model

- Text to enter into Rstudio's Script Window:
fit <- growth(model, data = widedat)
summary(fit, fit.measures = TRUE, standardized = TRUE)

Screenshot

The screenshot displays the RStudio environment with a script editor and a console window. The script defines a growth model and fits it to data. The console output shows the model's fit statistics and parameter estimates.

```
12 model <- '  
13   # Intercept and slope factors  
14   i =~ 1*gpa0 + 1*gpa1 + 1*gpa2 + 1*gpa3 + 1*gpa4  
15   s =~ 0*gpa0 + 1*gpa1 + 2*gpa2 + 3*gpa3 + 4*gpa4  
16  
17   # Variances and covariance  
18   i =~ i  
19   s =~ s  
20   i =~ s  
21  
22   # Residual variances  
23   gpa0 =~ gpa0  
24   gpa1 =~ gpa1  
25   gpa2 =~ gpa2  
26   gpa3 =~ gpa3  
27   gpa4 =~ gpa4  
28 ,  
29  
30 fit <- growth(model, data = widedat)  
31 summary(fit, fit.measures = TRUE, standardized = TRUE)  
32
```

Console Output:

```
R 4.4.0 · E:/Work/Quant Methods in Education Group at SEdS/SEM reading group/  
gpa4 4.000 0.309 0.865  
  
Covariances:  
      Estimate Std.Err z-value P(>|z|) Std.lv Std.all  
i =~  
s      -0.000 0.002 -0.073 0.942 -0.010 -0.010  
  
Intercepts:  
      Estimate Std.Err z-value P(>|z|) Std.lv Std.all  
i      2.602 0.018 140.832 0.000 14.129 14.129  
s      0.104 0.006 16.444 0.000 1.351 1.351  
  
Variances:  
      Estimate Std.Err z-value P(>|z|) Std.lv Std.all  
i      0.034 0.007 4.577 0.000 1.000 1.000  
s      0.006 0.001 5.880 0.000 1.000 1.000  
.gpa0 0.069 0.009 7.263 0.000 0.069 0.670  
.gpa1 0.068 0.008 8.628 0.000 0.068 0.631  
.gpa2 0.056 0.006 9.195 0.000 0.056 0.496  
.gpa3 0.036 0.005 7.997 0.000 0.036 0.294  
.gpa4 -0.001 0.005 -0.121 0.903 -0.001 -0.004  
  
> |
```

Inspect Model output

- n=200 participants
- Model fit: excellent from CFI and RMSEA, and SRMR

ConsoleTerminal xBackground Jobs x

R 4.4.0 · E:/Work/Quant Methods in Education Group at SEdS/SEM reading group/ ↗

> summary(fit, fit.measures = TRUE, standardized = TRUE)

lavaan 0.6-19 ended normally after 62 iterations

EstimatorML

Optimization methodNLMINB

Number of model parameters10

Number of observations200

Model Test User Model:

Test statistic12.202

Degrees of freedom10

P-value (Chi-square)0.272

Model Test Baseline Model:

Test statistic482.370

Degrees of freedom10

P-value0.000

User Model versus Baseline Model:

Comparative Fit Index (CFI)0.995

Tucker-Lewis Index (TLI)0.995

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)-110.013

Loglikelihood unrestricted model (H1)-103.912

Akaike (AIC)240.026

Bayesian (BIC)273.009

Sample-size adjusted Bayesian (SABIC)241.328

Root Mean Square Error of Approximation:

RMSEA0.033

90 Percent confidence interval - lower0.000

90 Percent confidence interval - upper0.087

P-value H_0: RMSEA <= 0.0500.631

P-value H_0: RMSEA >= 0.0800.086

Standardized Root Mean Square Residual:

SRMR0.057

Inspect Model output

- Estimates are b unstandardised regression coefficients
- Std.all are Beta standardized regression coefficients
- $P(>|z|)$ is the typical p-value
- Variances refers to the residual variance. If significant more variance in the DV could still be accounted for (in theory)
- r^2 calculated from $(1 - \text{Std.all})$ for the variance of the DV.
- Intercept mean = 2.602 (average GPA at time 0; $p < 0.001$)
- Slope mean = 0.104 (average increase in GPA per time unit; $p < 0.001$ [sig.different from zero/nothing])
- Intercept variance = 0.034. people vary in where they start
- Slope variance = 0.006. people vary in how fast they improve
 - Both $p < 0.001$ showing these variations to be significantly different from zero/nothing
- Residuals decrease over time — suggests increased stability or less error in later GPA scores
- Intercept-slope covariance is approximately estimated to be Zero
 - Suggests no relationship between starting GPA and change rate
 - E.g. no “Matthew Effect” in the GPA scores

Console

Terminal

Background Jobs

R 4.4.0 · E:/Work/Quant Methods in Education Group at SEDS/SEM reading group/

Parameter Estimates:

Standard errors

Information

Information saturated (h1) model

Standard

Expected

Structured

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
i =~						
gpa0	1.000				0.184	0.575
gpa1	1.000				0.184	0.562
gpa2	1.000				0.184	0.547
gpa3	1.000				0.184	0.526
gpa4	1.000				0.184	0.516
s =~						
gpa0	0.000				0.000	0.000
gpa1	1.000				0.077	0.236
gpa2	2.000				0.154	0.458
gpa3	3.000				0.232	0.661
gpa4	4.000				0.309	0.865

Covariances:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
i ~~						
s	-0.000	0.002	-0.073	0.942	-0.010	-0.010

Intercepts:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
i	2.602	0.018	140.832	0.000	14.129	14.129
s	0.104	0.006	16.444	0.000	1.351	1.351

Variances:

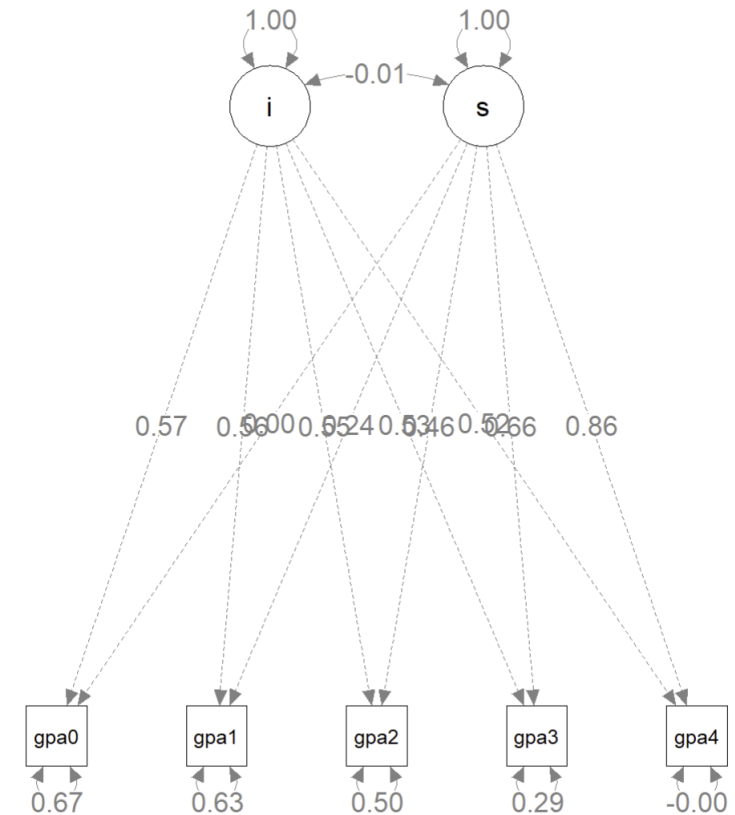
	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
i	0.034	0.007	4.577	0.000	1.000	1.000
s	0.006	0.001	5.880	0.000	1.000	1.000
.gpa0	0.069	0.009	7.263	0.000	0.069	0.670
.gpa1	0.068	0.008	8.628	0.000	0.068	0.631
.gpa2	0.056	0.006	9.195	0.000	0.056	0.496
.gpa3	0.036	0.005	7.997	0.000	0.036	0.294
.gpa4	-0.001	0.005	-0.121	0.903	-0.001	-0.004

>

Visualise the SEM Growth Model via semPlot

- Text to enter into Rstudio's Script Window:


```
semPaths(fit,
        what = "path",      # draw path diagram
        whatLabels = "std",  # show standardized estimates
        layout = "tree",    # tree layout works well for growth models
        edge.label.cex = 1.2, # font size of path coefficients
        sizeMan = 6,        # size of observed variable boxes
        sizeLat = 8,        # size of latent variable circles
        residuals = TRUE,    # include residual variances
        intercepts = FALSE,  # hide intercept arrows for clarity
        nCharNodes = 0,     # do not abbreviate variable names
        title = FALSE       # remove automatic title
      )
```
- How practically useful is this diagram though?



Graph of the growth

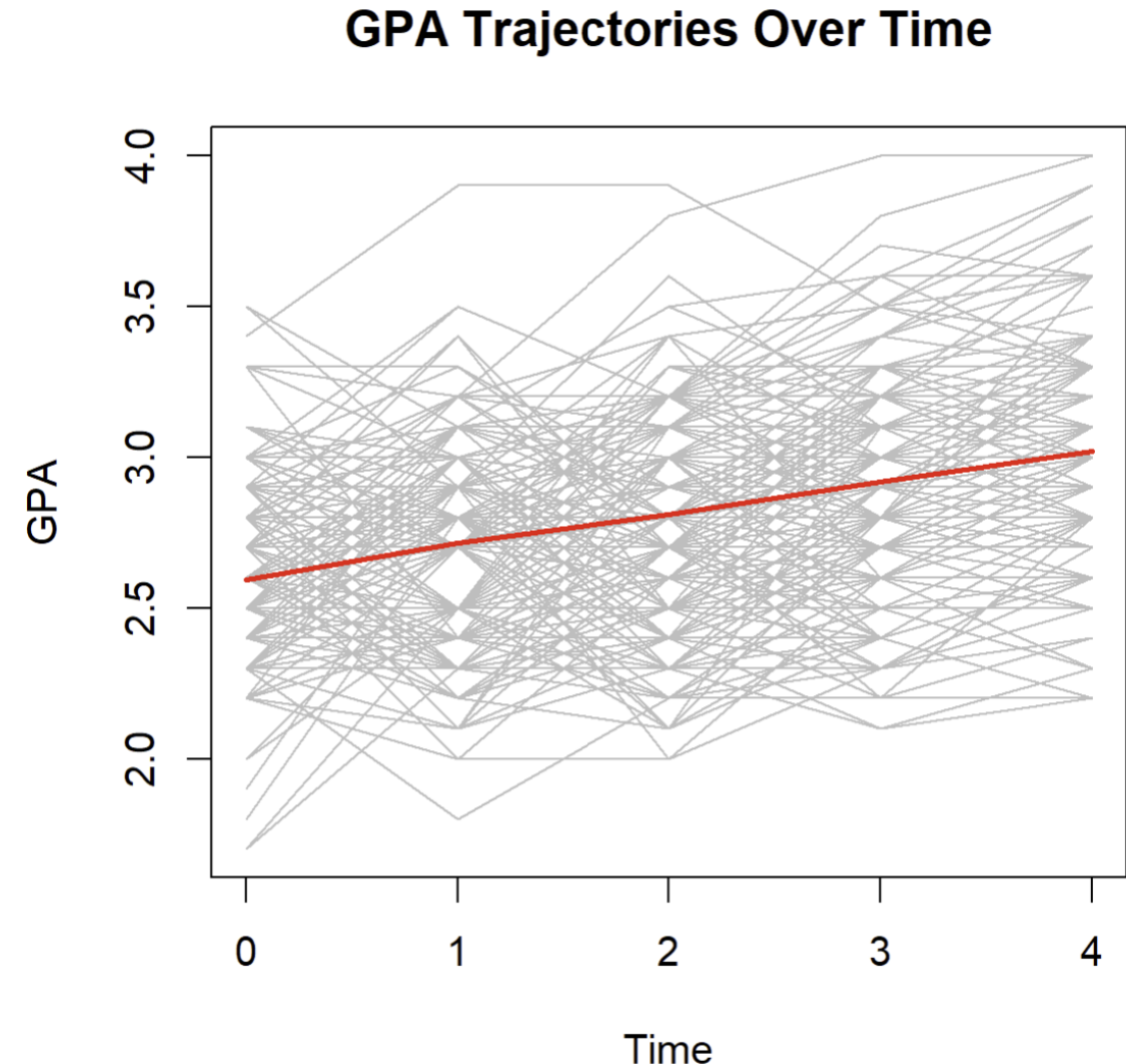
- Text to enter into Rstudio's Script Window:


```
# Create a matrix of GPA scores (columns: gpa0–gpa4)
gpa_matrix <- as.matrix(widedat[, c("gpa0", "gpa1",
  "gpa2", "gpa3", "gpa4")])

# Plot individual trajectories
matplot(t(gpa_matrix), type = "l", lty = 1, col = "grey",
  xlab = "Time", ylab = "GPA", main = "GPA
  Trajectories Over Time",
  xaxt = "n")

# Add time axis labels
axis(1, at = 1:5, labels = 0:4)

# Add average line (mean trajectory)
mean_gpa <- colMeans(gpa_matrix, na.rm = TRUE)
lines(1:5, mean_gpa, col = "red", lwd = 2)
```



Modify and Extend

- Lets consider some modifications of that basic SEM growth model
- First, Lets see how to change the intercept location (e.g. midpoint instead of time 0):

```
model_midpoint <- '
i =~ 1*gpa0 + 1*gpa1 + 1*gpa2 + 1*gpa3 +
1*gpa4
s =~ -2*gpa0 + -1*gpa1 + 0*gpa2 + 1*gpa3 +
2*gpa4'
```
- Our intercept mean is now 2.81 and this is from the third timepoint of measurement (it was 2.60 from T1)
 - Our slope mean remains unchanged

Latent Variables:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
i =~						
gpa0	1.000				0.239	0.746
gpa1	1.000				0.239	0.730
gpa2	1.000				0.239	0.710
gpa3	1.000				0.239	0.682
gpa4	1.000				0.239	0.670
s =~						
gpa0	-2.000				-0.154	-0.482
gpa1	-1.000				-0.077	-0.236
gpa2	0.000				0.000	0.000
gpa3	1.000				0.077	0.220
gpa4	2.000				0.154	0.432

Covariances:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
i ~~						
s	0.012	0.002	6.534	0.000	0.638	0.638

Intercepts:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
i	2.810	0.018	155.038	0.000	11.756	11.756
s	0.104	0.006	16.444	0.000	1.351	1.351

Variances:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
.gpa0	0.069	0.009	7.263	0.000	0.069	0.670
.gpa1	0.068	0.008	8.628	0.000	0.068	0.631
.gpa2	0.056	0.006	9.195	0.000	0.056	0.496
.gpa3	0.036	0.005	7.997	0.000	0.036	0.294
.gpa4	-0.001	0.005	-0.121	0.903	-0.001	-0.004
i	0.057	0.007	8.633	0.000	1.000	1.000
s	0.006	0.001	5.880	0.000	1.000	1.000

Modify and Extend

- Lets consider some modifications of that basic SEM growth model
- Second, lets now look for “sex” differences

```
model_sex <- '
# Growth factors
i =~ 1*gpa0 + 1*gpa1 + 1*gpa2 + 1*gpa3 + 1*gpa4
s =~ 0*gpa0 + 1*gpa1 + 2*gpa2 + 3*gpa3 + 4*gpa4

# Regressions: sex predicts growth factors
i ~ sex
s ~ sex

# Variances/covariances (optional if already familiar)
i ~~ i
s ~~ s
i ~~ s

gpa0 ~~ gpa0
gpa1 ~~ gpa1
gpa2 ~~ gpa2
gpa3 ~~ gpa3
gpa4 ~~ gpa4'
```

```
fit_sex <- growth(model_sex, data = widedat)
summary(fit_sex, fit.measures = TRUE, standardized = TRUE)
```

• Fit indices:

- CFI: 0.998
- RMSEA: 0.019
- SRMR: 0.049

Regressions:

	Estimate	Std.Err	z-value	P(> z)	Std.lv	Std.all
i ~ sex	0.073	0.037	1.994	0.046	0.397	0.198
s ~ sex	0.031	0.013	2.508	0.012	0.407	0.203

- There’s an initial significant gender difference, and one that widens over time
 - Aside: Note the peril of labelling a binary variable by its concept not the category “coded high” – we have no idea the direction of the sex difference here as a result

Suggested follow-ups:



[Quantitative
Methods in
Education Group](#)

1. Expand your SEM toolkit:
 - Model multiple processes: Parallel growth models (e.g., GPA & motivation)
 - Explore Growth Mixture Models: Subgroups with distinct trajectories
 - Add time-varying covariates or nonlinear growth (e.g., quadratic)
 - Try multigroup growth models: Do growth patterns differ across groups?
2. Learn more R tools:
 - semTools: Fit indices, modification indices, invariance testing
 - blavaan: Bayesian latent growth models
 - OpenMx, brms, nlme: Advanced or alternative SEM frameworks
 - Explore plotting with semPlot or tidySEM
3. Recommended learning resources:
 - Hox et al. (2017) – Multilevel Analysis
 - Grimm, Ram & Estabrook (2016) – Growth Modeling: SEM and MLM Approaches
 - lavaan tutorials: lavaan.org
 - Rstudio help with packages and commend syntax: Any LLM