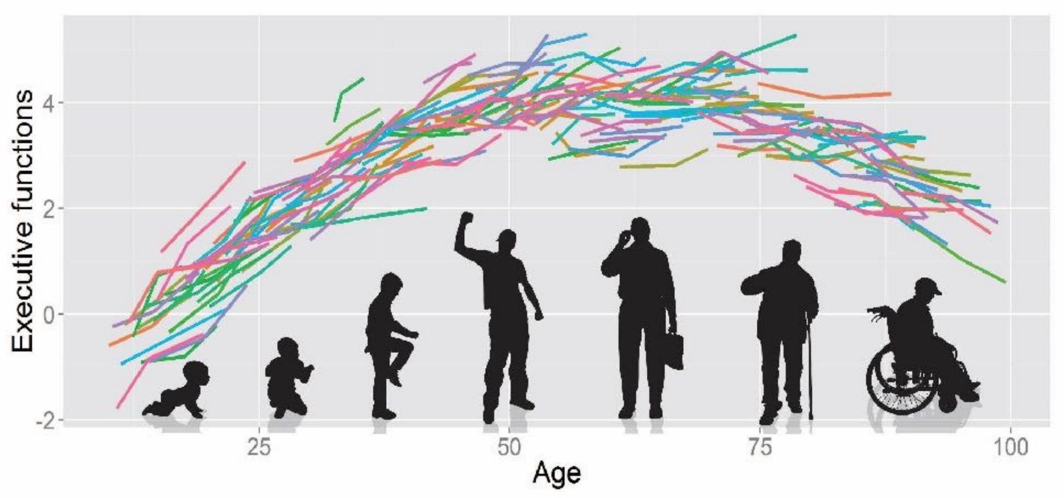
An introduction to Latent Growth Curve Modeling



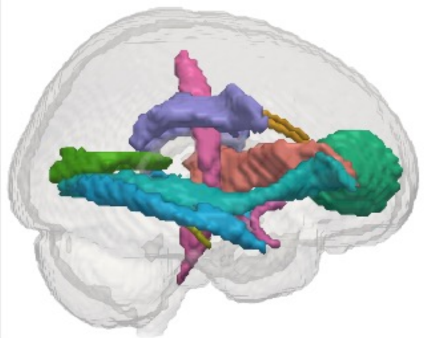
Rogier A. Kievit
Lifespan Cognitive Dynamics Group

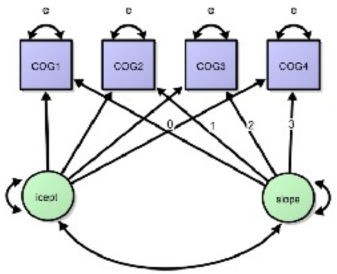
rogier.kievit@radboudumc.nl

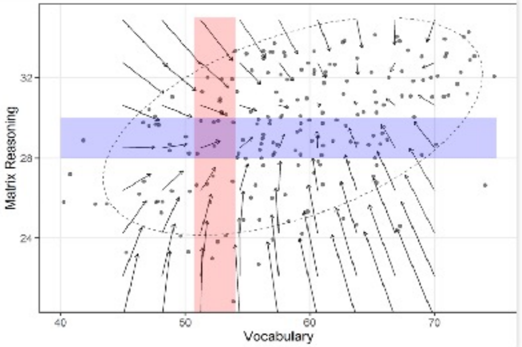
Lifespan Cognitive Dynamics Lab

Brains, Psychology and Psychometrics

LCD LAB | PEOPLE | RESEARCH INTERESTS | PUBLICATIONS | OPEN SCIENCE AND RESOURCES | JOINING THE LAB







How, and why, do cognitive abilities change across the lifespan?

(SEM) Model based, theory driven developmental cognitive neuroscience

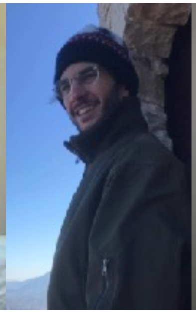
Rogier Kievit

Jessica Schaaf

Ilse Coolen

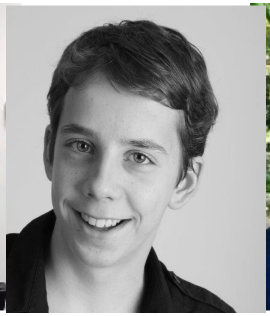
Jordy van Langen

Eleni Zimianiti



Michael Aristodemou

Lea Michel



Aran van Hout

Nick Judd

Emma Meeussen

Feline van Aagten

DONDERS INSTITUTE

Outline

- 1) Why study change?
- 2) Structural Equation Modeling – Getting started
- 3) Latent growth curve modeling
- 4) DIY

Magda Lazari

Léa Michel

Nick Judd

From snapshots to change

- Classic approach: ‘High resolution’ snapshot of brain structure, behavioural phenotypes
  - 7 Tesla, Connectom; better questionnaires, better tests
- But: Does not capture *change* or *development*
- To study development we should study *change*

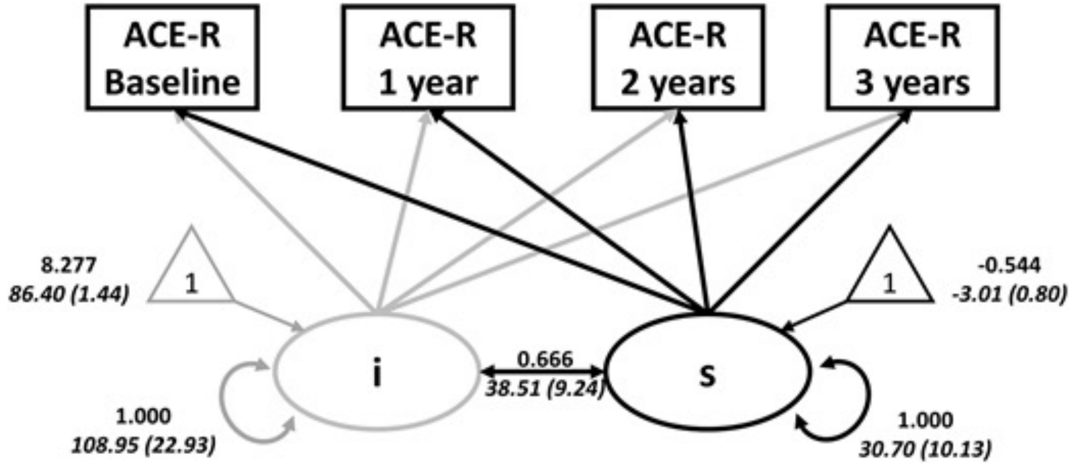
Why study change?

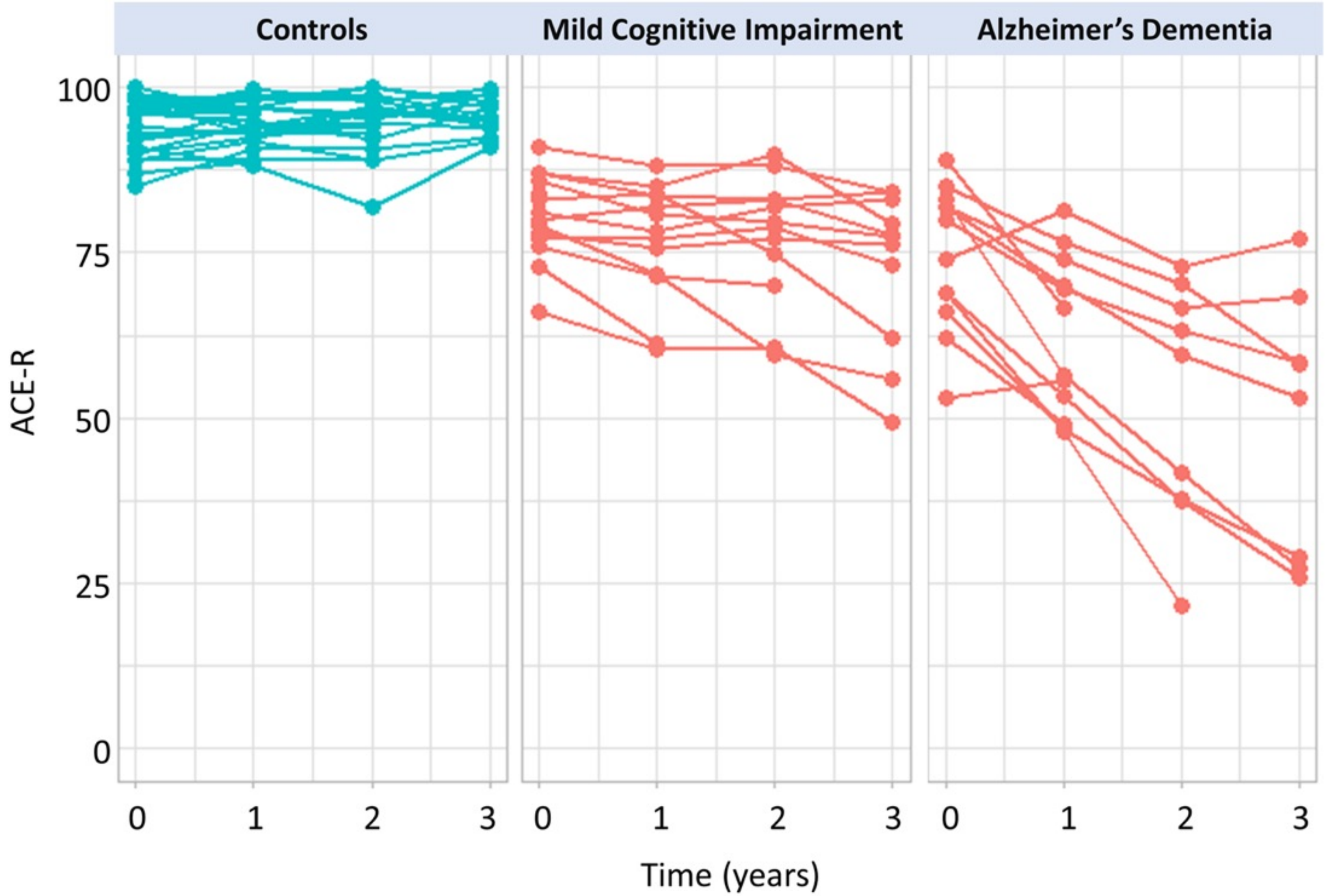
- 1) How do people change?
- 2) How do individuals differ in change?
- 3) How do changes co-occur?
- 4) What predicts – or is predicted by – change?

Baltes, P. B., & Nesselroade, J. R. (Eds.). (1979). *Longitudinal research in the study of behavior and development*. Academic Press.

Example: Modeling decline in dementia

- Group comparison: Do groups differ in their mean linear slope?
- Groups differed significantly in their slopes
- Differences in slopes are predicted by tau burden

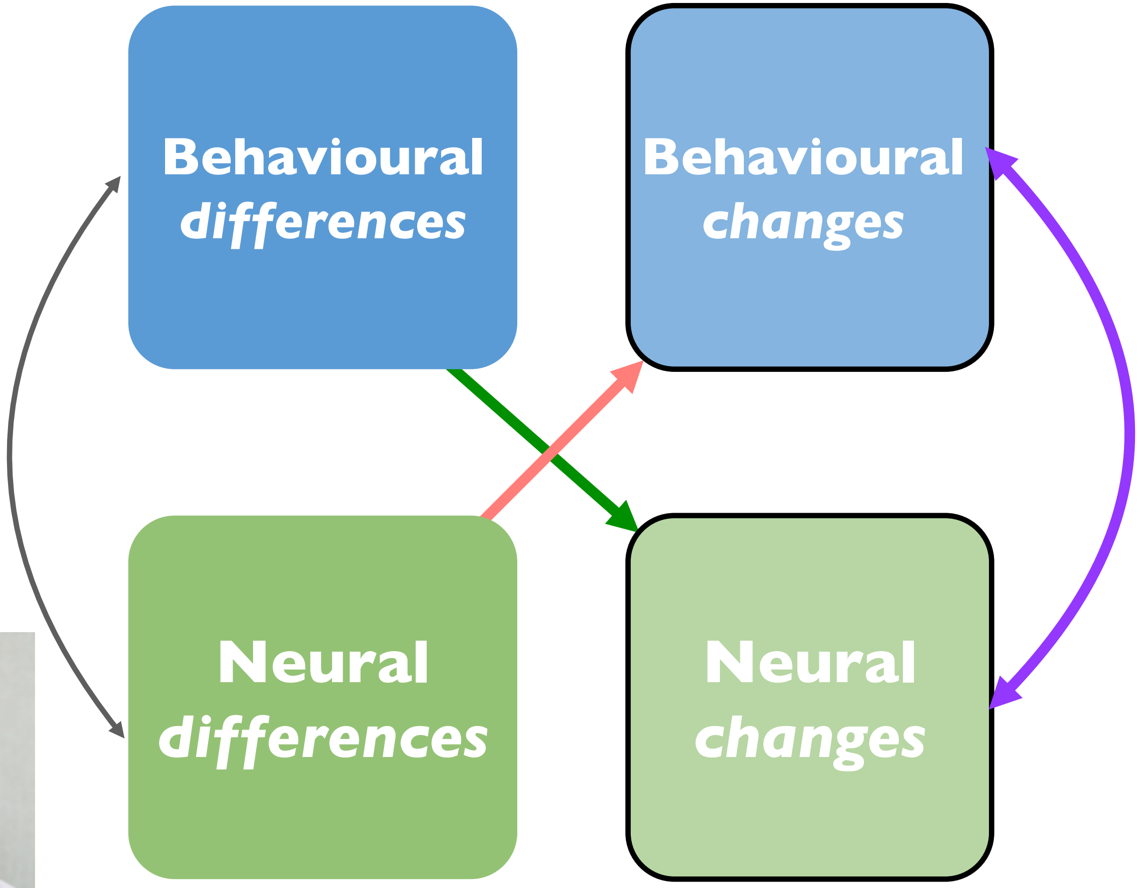




Malpetti, M., Kievit, R. A., Passamonti, L., Jones, P. S., Tsvetanov, K. A., Rittman, T., ... & Rowe, J. B. (2020). Microglial activation and tau burden predict cognitive decline in Alzheimer's disease. Brain, 143(5), 1588-1602.

How do behavioural and neural phenotypes co-develop in childhood and early adolescence?

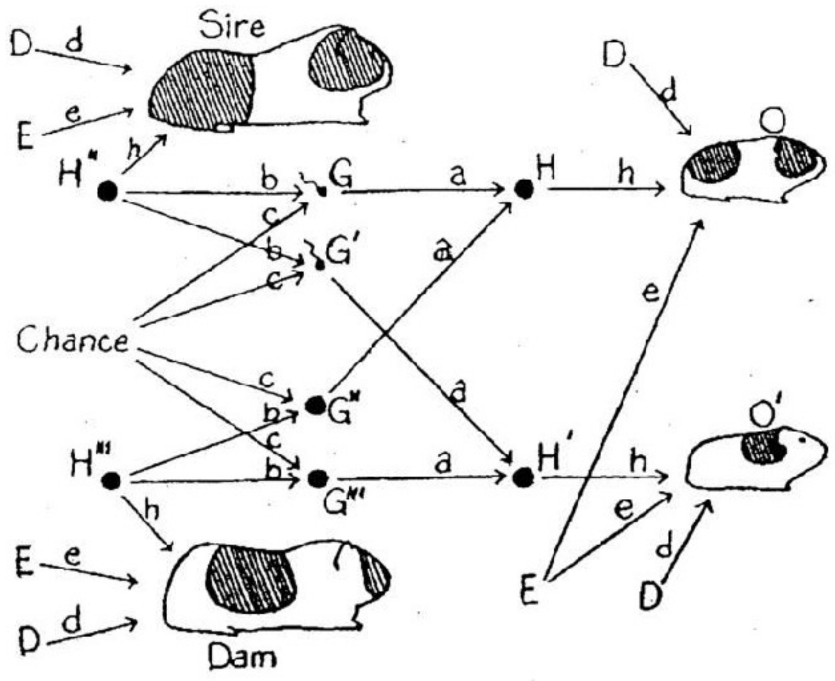
- 99% of Developmental Cognitive Neuroscience: *Age heterogeneous differences*
- Increase focus on three (empirically) underrepresented questions
- 1) Do changes in brain structure and behavioral characteristics go hand in hand?
- 2) Does brain structure drive the emergence of behavioural characteristics?
- 3) Do behavioural differences drive neural change?

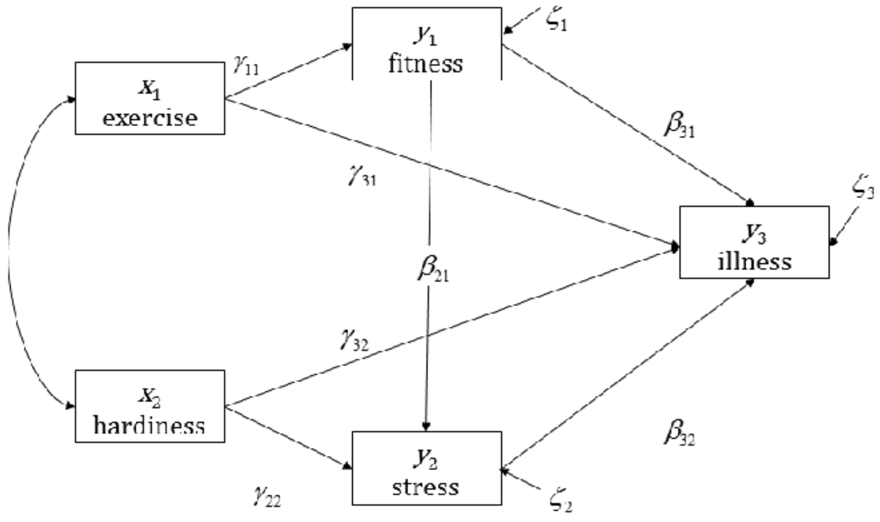


Kievit, R., & Simpson-Kent, I. L. (2019). It's About Time: Towards a Longitudinal Cognitive Neuroscience of Intelligence. https://psyarxiv.com/n2yg7/

What is structural equation modeling (SEM)?

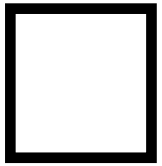
- A combination of two tools
  - Path analysis
    - Sewall Wright (1921)
    - Unique: path can be the cause and consequence at the same time
    - Estimated simultaneously
- Latent variable analysis
    - Relating measured variables to hypothesized constructs



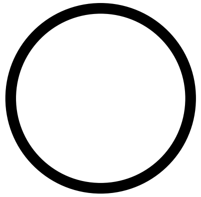


Graphical representation

- One-to-one mapping between graphical representation and matrix algebra
If you can draw your hypothesis, you can (likely) fit it as a model



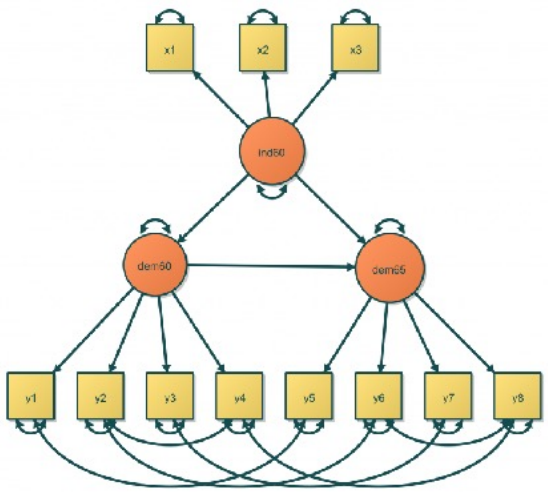
Observed variable

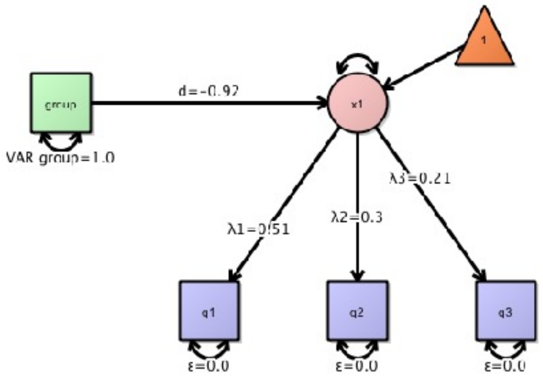


Latent variable

Directed relationship
(e.g. factor loadings, regressions)

Undirected relationship
(e.g. (co)variance, error)





Model fitting in SEM

- 1. Overall model fit: chi square test
- 2. Model fit: Comparative fit indices
- 3. Model fit: Model misfit
- 4. Comparative model fit
- 5. Parameter estimates

Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures

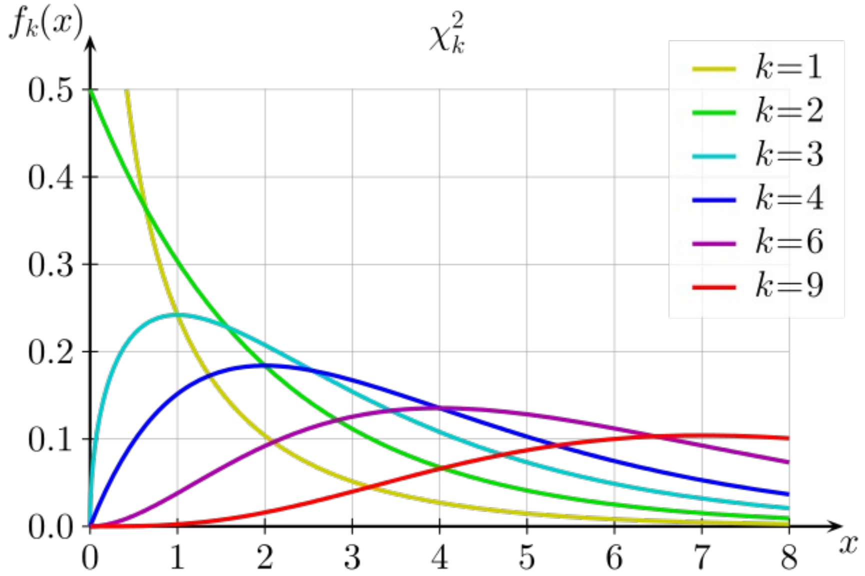
Karin Schermelleh-Engel[1] and Helfried Moosbrugger

Goethe University, Frankfurt

Hans Müller

University of Erfurt

1. Chi square test



- Basic principle: Test of perfect model fit
- We have a proposed model
- IF that model generated our data…
- THEN we know the expected misfit of the data to this model
  - Chi square, with *k* degrees of freedom
- If the data deviates significantly more than expected…
- Then the hypothesis of perfect fit is rejected

Different..

P<0.05

Standard stats

NOVEMBER 3, 2012

New version SPSS will include 'celebratory fireworks' for significant results

An official press release has confirmed that the newes
equipped with 'performance-rewarding features'. The
popular data-analysis package will light up with song,
whenever a statistical test is significant. 'We want to p
line with the day-to-day experiences of researchers. V
the publish, and the relief that is felt by many when th
appear in the results table.'

The level of significance will determine the abundance
value is below 0.05, researchers will automatically hea
cheerful tone', according to a company spokesman.'

SEM

2. Model fit: Comparative fit indices

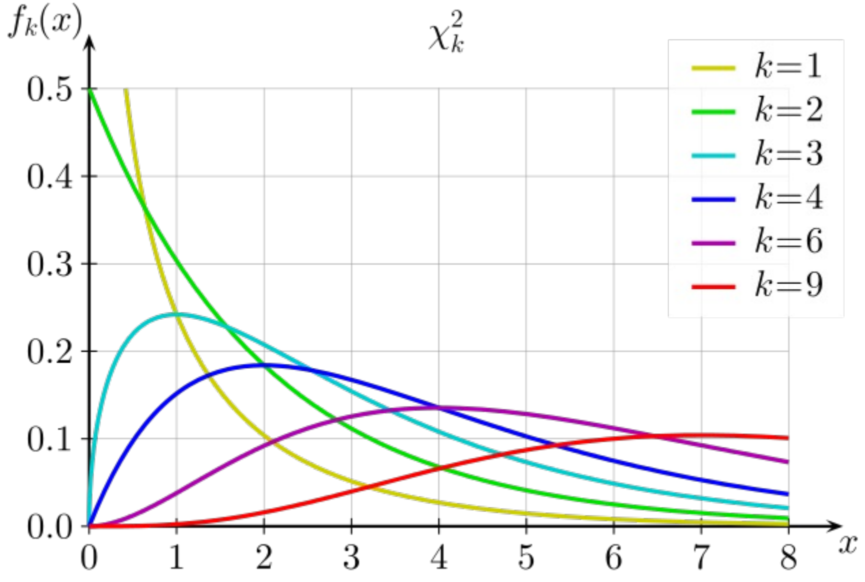
- Basic idea: How much better does my model perform than a baseline, reference (e.g. independence) model?
- Conceptually similar to (adjusted) r-squared: given my model complexity, how well do I explain the data
- Range: 0-1
- Cutoff: >0.95-0.97
- Examples:
  - CFI, TLI, NNFI, GFI

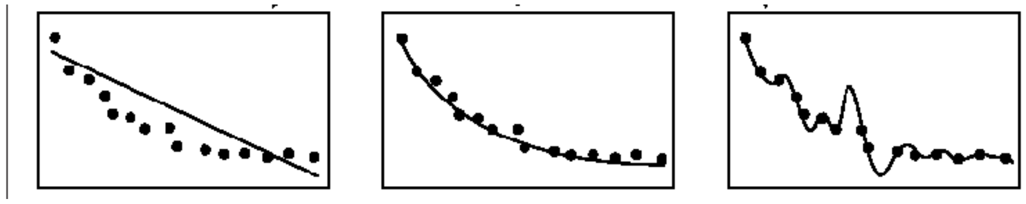
3. Model fit: Model misfit indices

- Basic idea: How *wrong* is my fitted model?
- Conceptually: what is the *average wrongness*
- Examples:
  - RMSEA, SRMR
- Range: ~0-1
- Cutoffs: <0.08-0.05

4. Model Comparison A: chi square

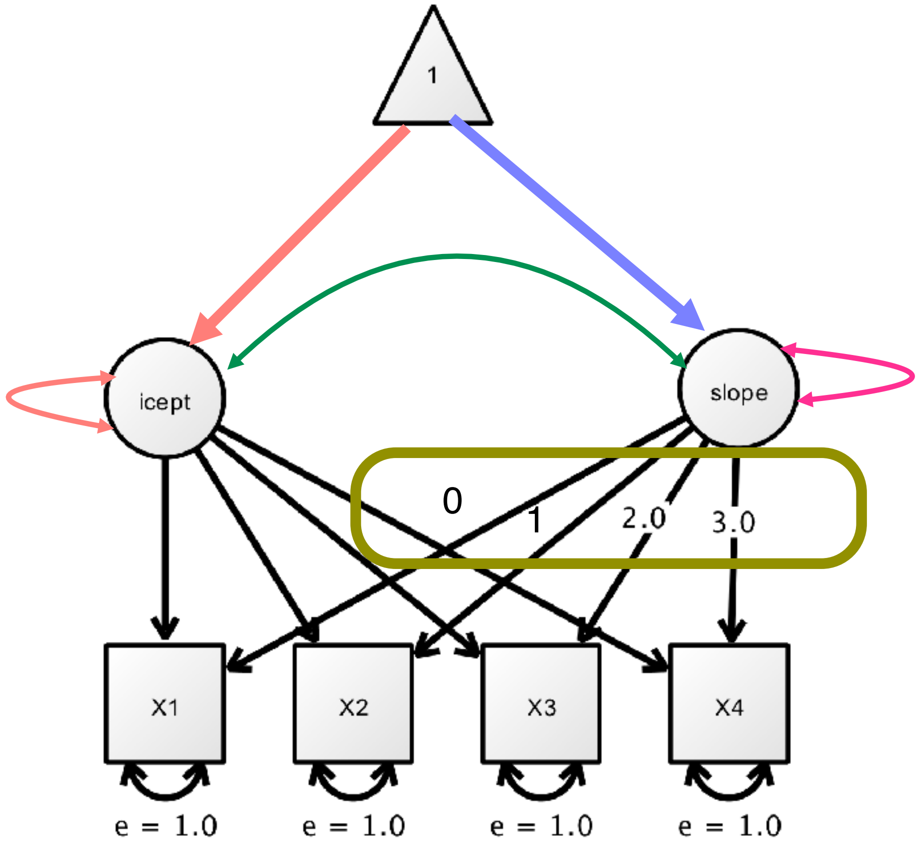
- The *difference* in model fit is *also* chi-square distributed
- Only works if models are ***nested***
  - *You can turn model 1 into model 2 by simply freeing parameters*
- Conceptually:
  - Does the more complicated model fit significantly better than would be expected just by making *any model* more (needlessly) complicated
- Can also use information criteria (AIC/BIC)





A particular type of SEM: *Latent Growth Curve Models*

- Flexible framework to capture:
- Individual differences at baseline
- The shape of change
- Mean change over time (slope)
- Individual differences in change over time (slope variance)
- The association between present and future (slope-intercept covariance)



Kievit et al., (2018). Developmental cognitive neuroscience using latent change score models: A tutorial and applications. Developmental cognitive neuroscience, 33, 99-117.

Duncan & Duncan (2004). An introduction to latent curve modelling. Behaviour Therapy, 35, 333-363.

McNeish, D. & Matta, T. (2017). Differentiating Between Mixed Effects and Latent Curve Approaches to Growth Modeling.

Fitting a latent growth curve in Lavaan

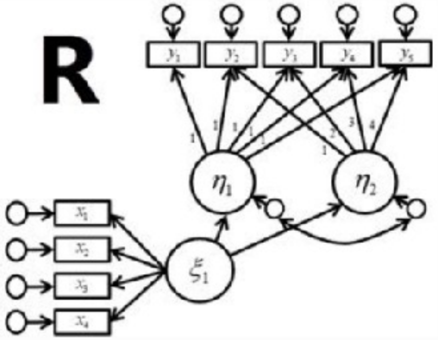
- Build up a 3 wave example model

- As *graphical model*

- As *syntax*

Do you have any materials that demonstrate how to estimate structural equation models using lavaan in R?

April 24, 2019



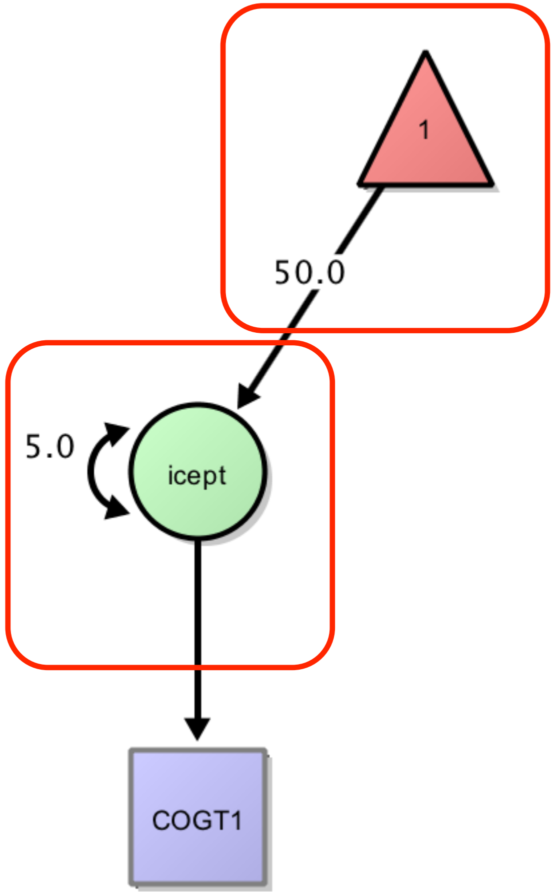
This is a question we often hear, particularly from students and junior researchers who don't have access to sometimes expensive commercial software for fitting structural equation models. It is possible to estimate a wide array of SEMs, ranging from simple path models to fully latent SEMs to growth curve models and beyond, using the lavaan package within R. For those who may be interested, we have developed detailed demonstrations of how to estimate a broad class of SEMs using lavaan and these are now freely available for download.

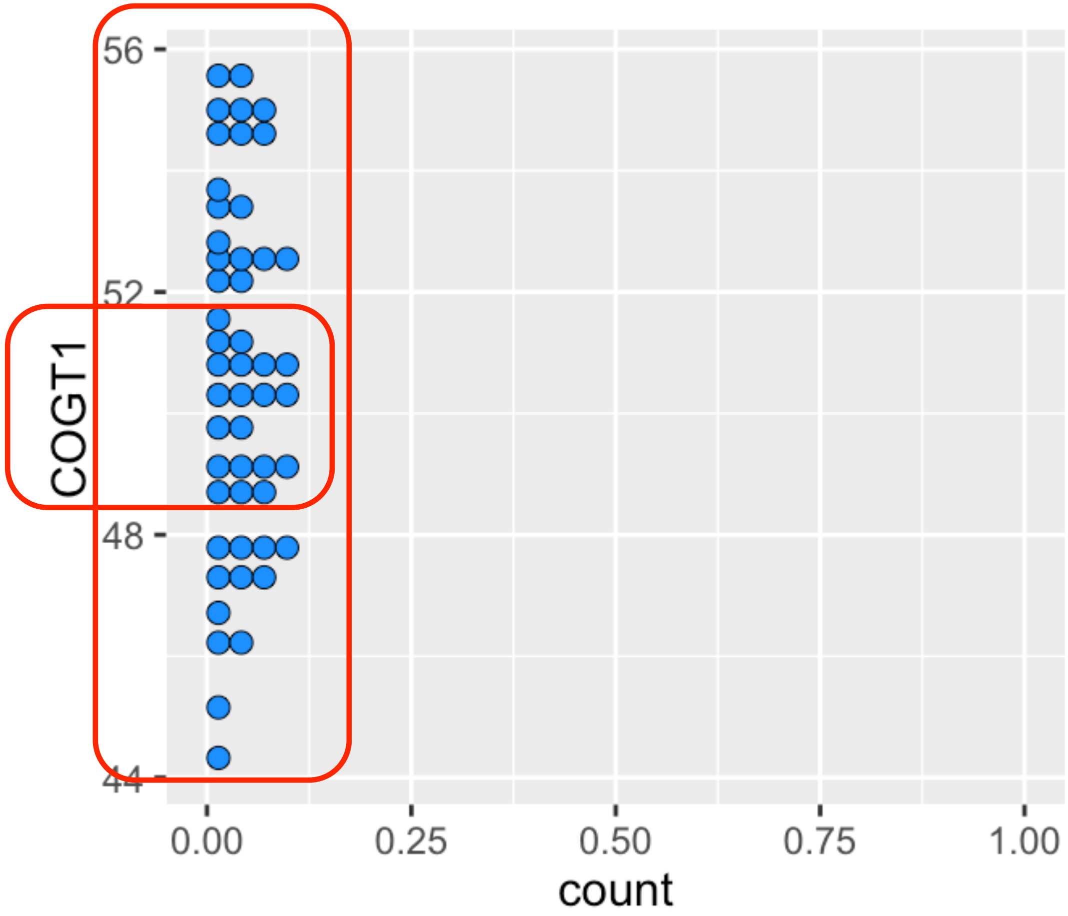
Summary of Formulae Types

formula type	operator	mnemonic
latent variable definition	`=~`	is measured by
regression	`~`	is regressed on
(residual) (co)variance	`~~`	is correlated with
intercept	`~ 1`	intercept

Time 1

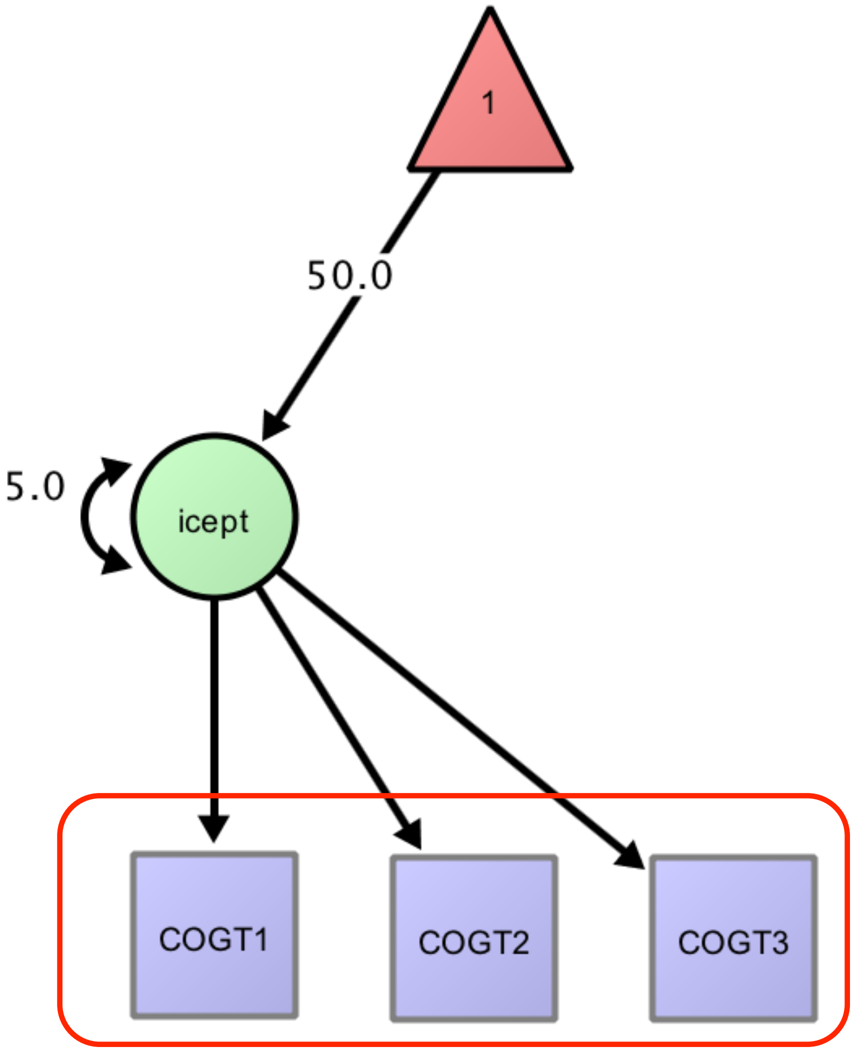
```
'
i=~1*COGT1  #intercept
i~1         #mean
i~~i        #variance
'
```

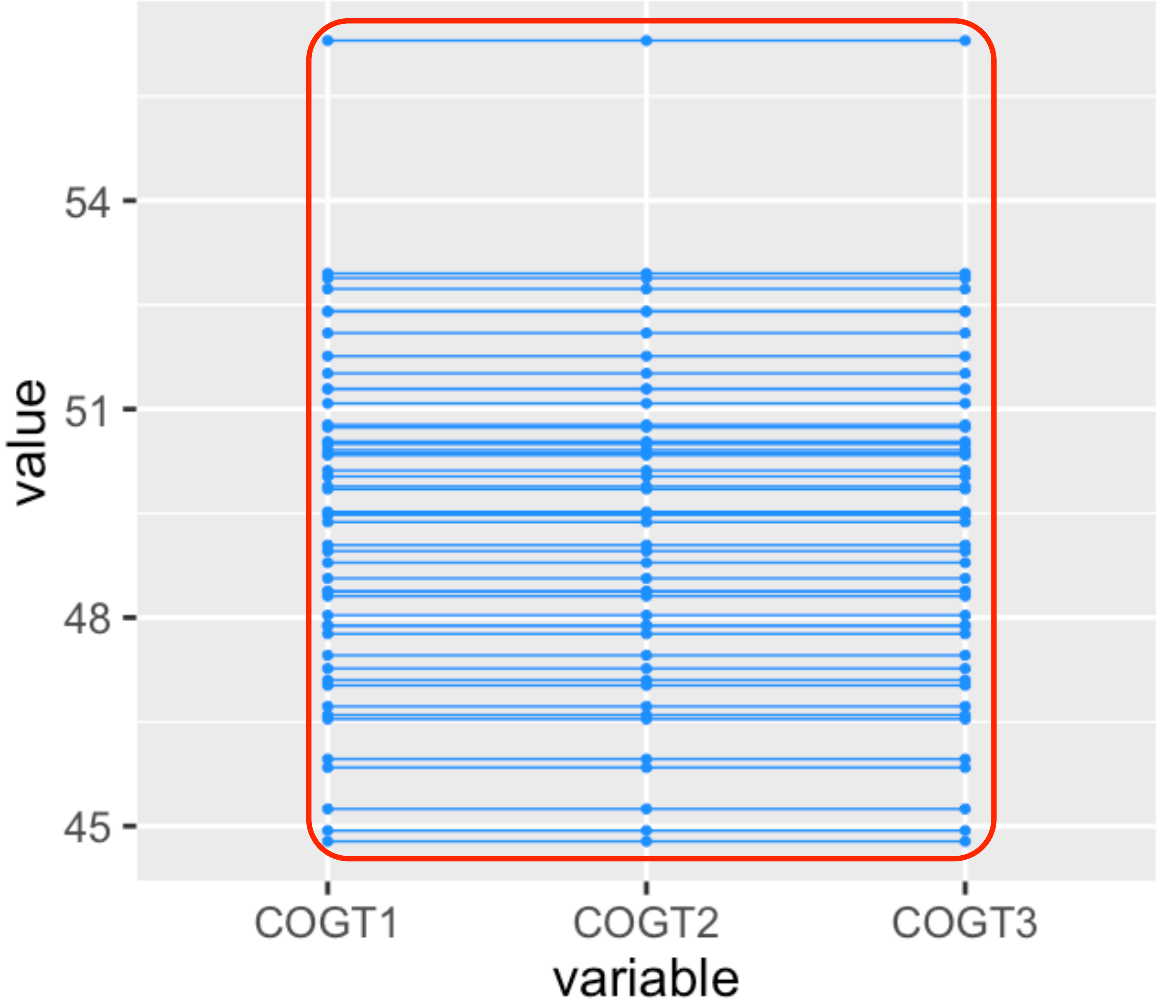




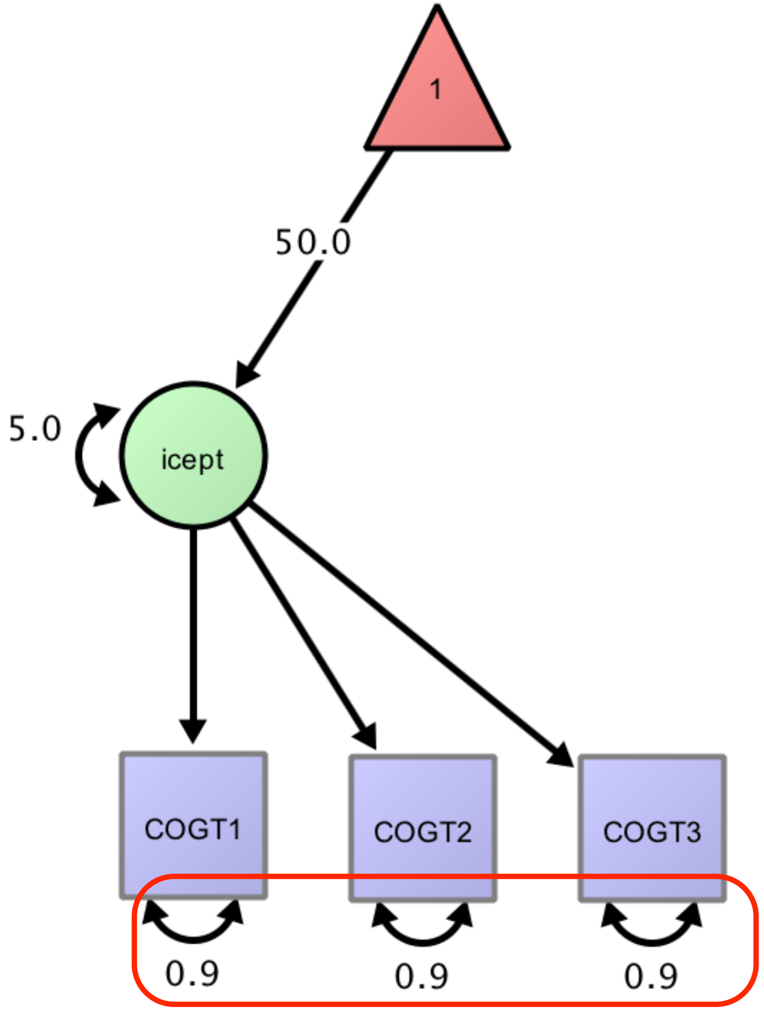
Repeated measures

```
'
i=~1*COGT1+1*COGT2+1*COGT3
i~1
i~~i
'
```

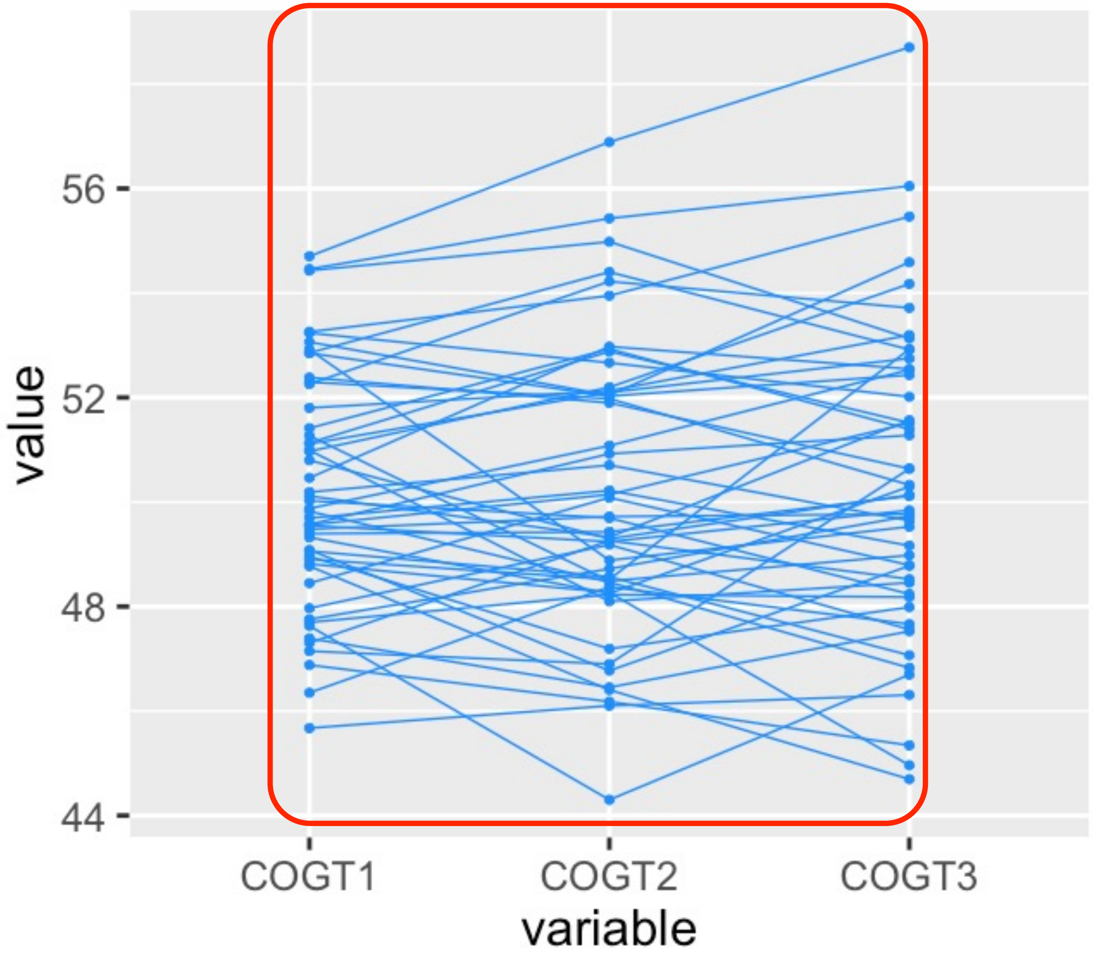




Measurement error

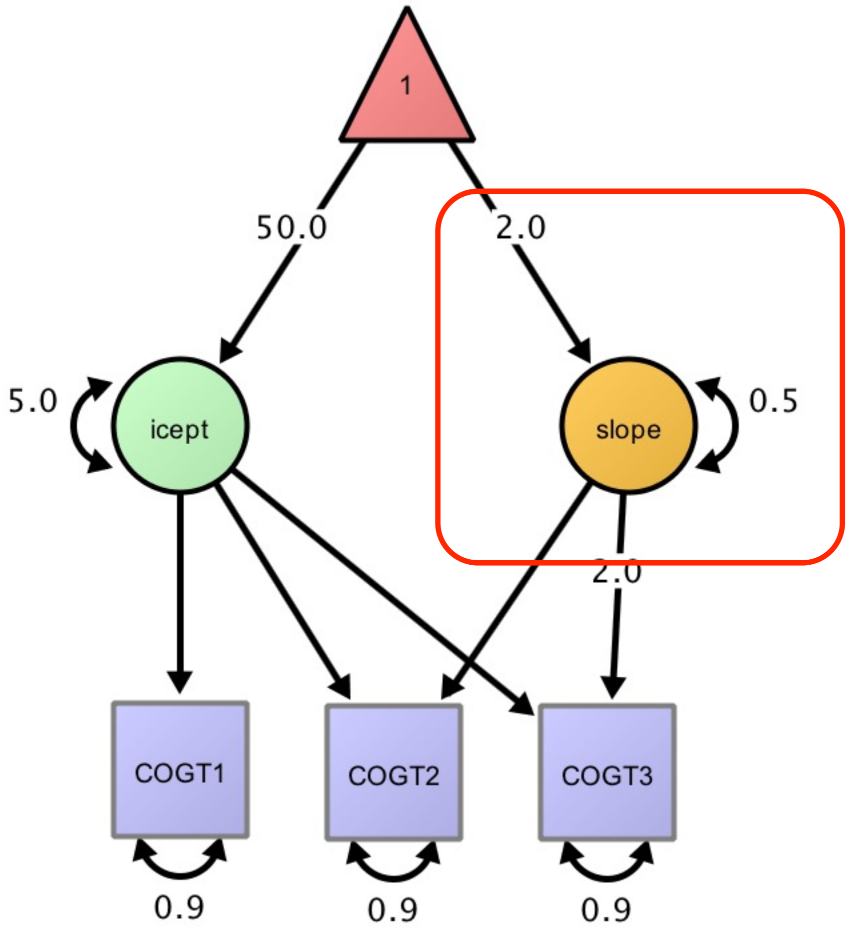


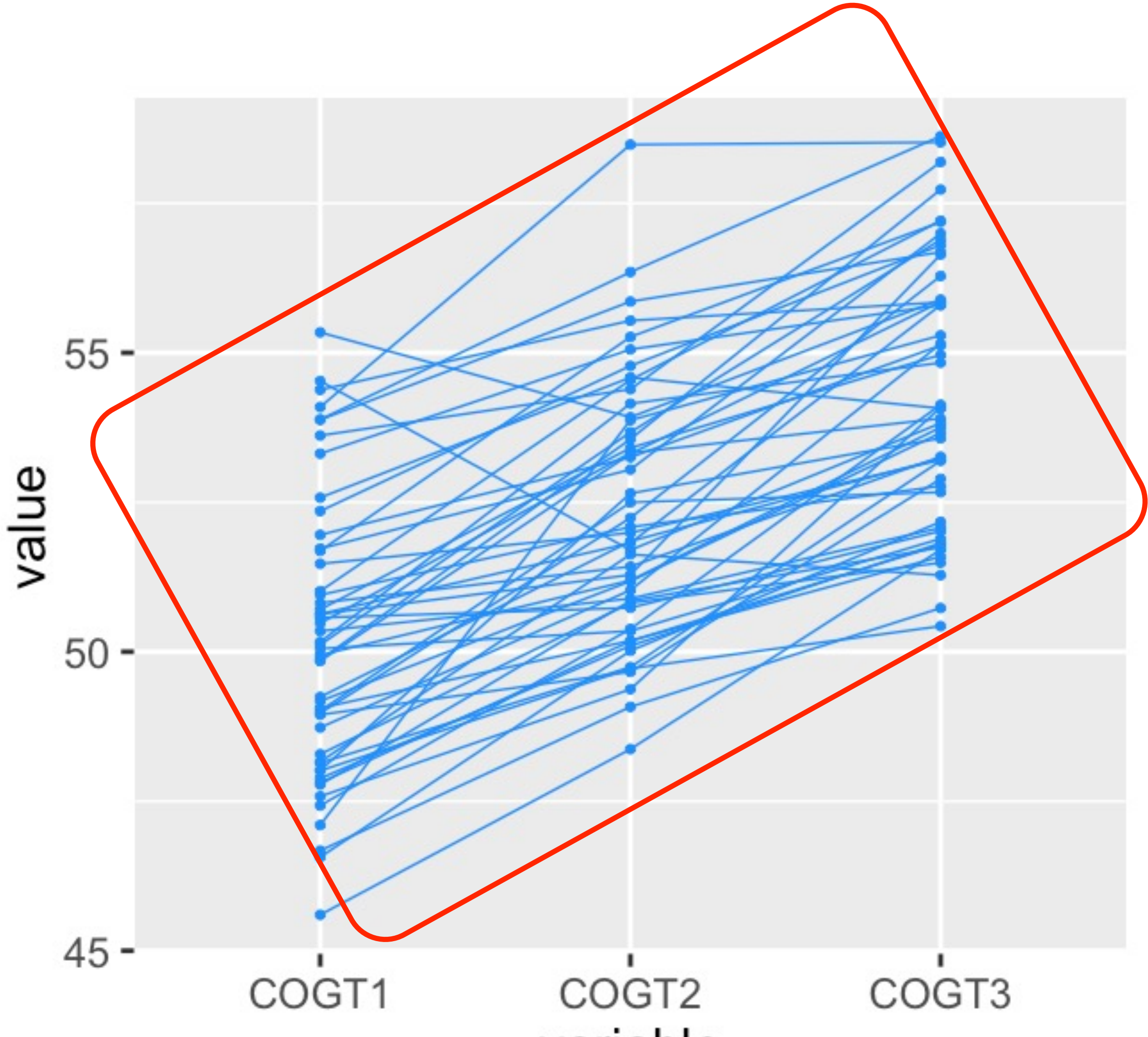

```
'
i=~1*COGT1+1*COGT2+1*COGT3
i~1
i~~i
COGT1~~ervar*COGT1
COGT2~~ervar*COGT2
COGT3~~ervar*COGT3
'
```



Adding a slope

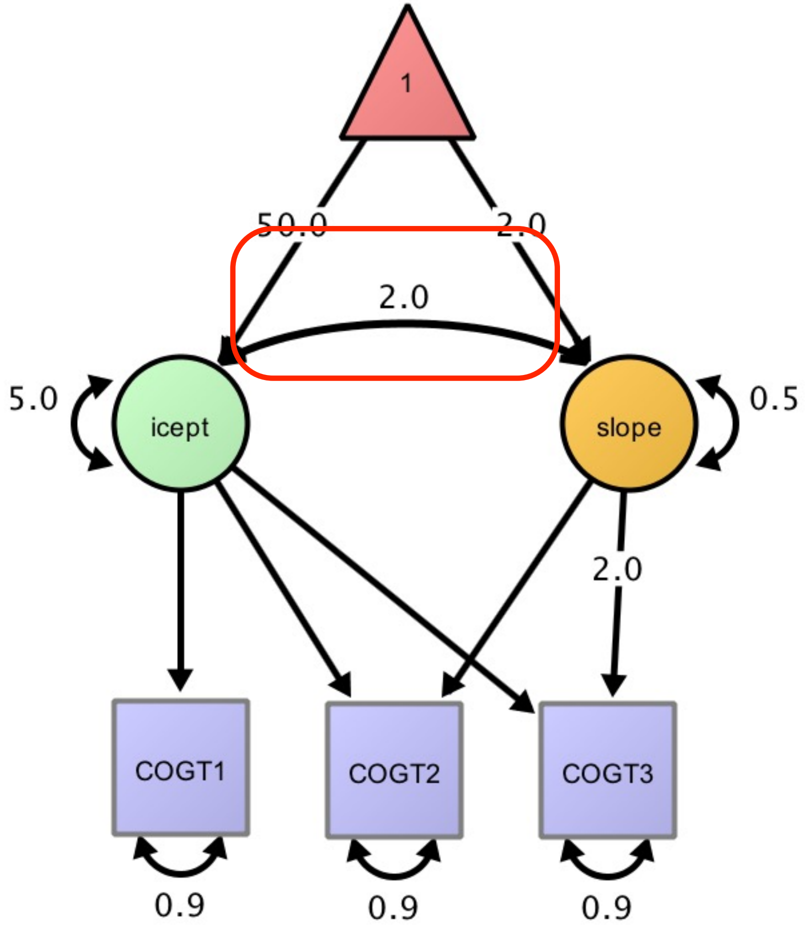
```
'
i=~1*COGT1+1*COGT2+1*COGT3
s=~0*COGT1+1*COGT2+2*COGT3
s~~s
'
```

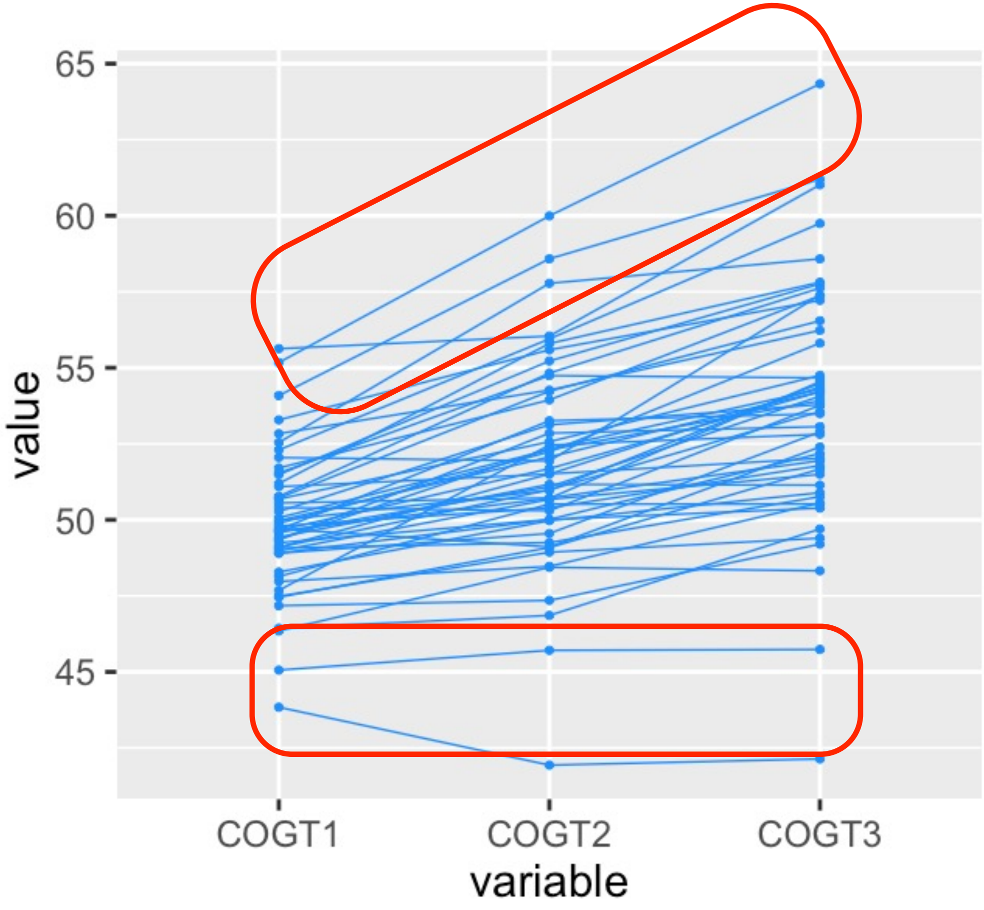




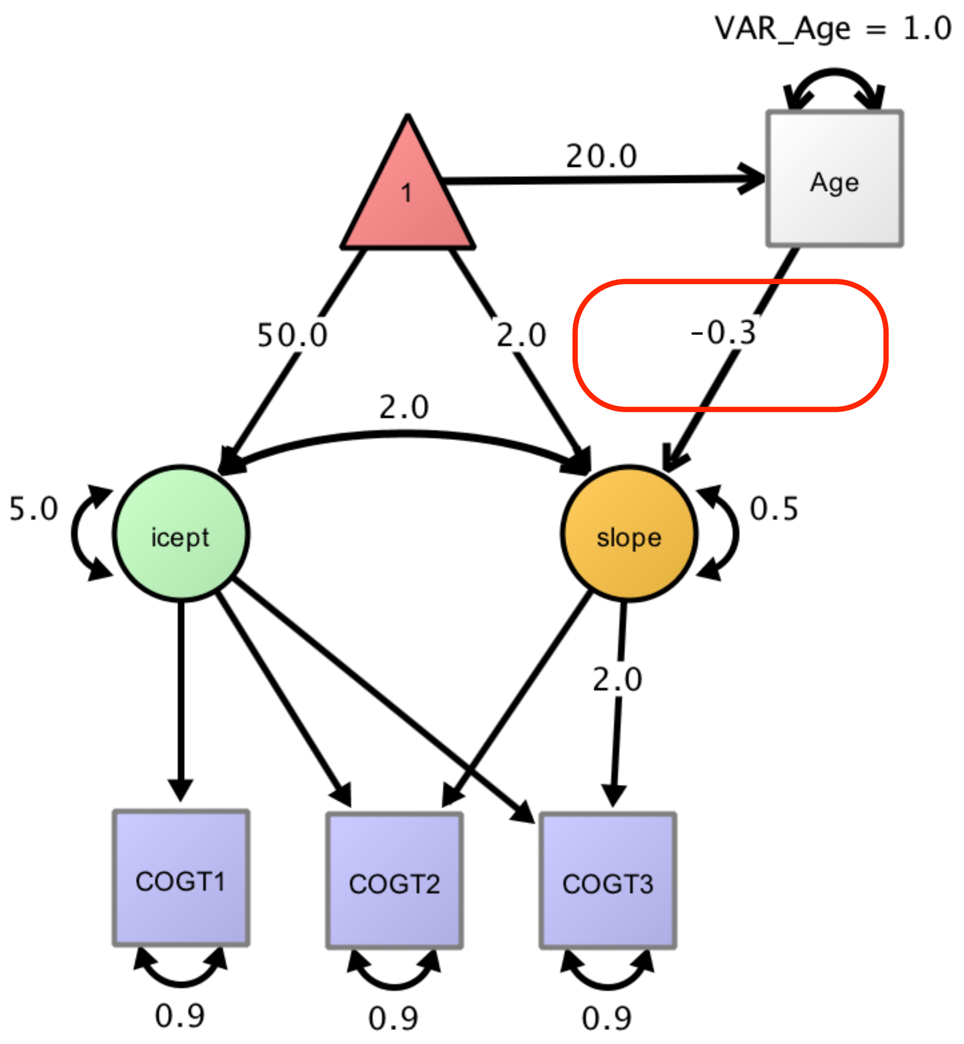
Adding latent covariance

```
'
i=~1*COGT1+1*COGT2+1*COGT3
s=~0*COGT1+1*COGT2+2*COGT3
s~~i
'
```

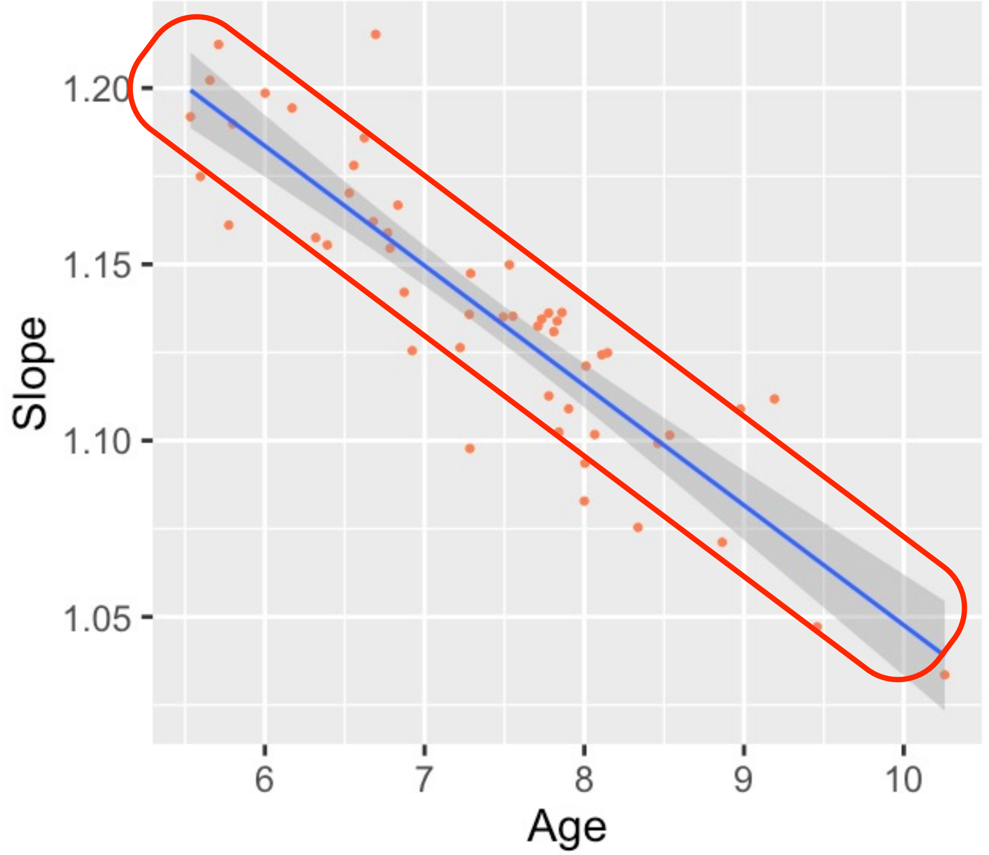




Adding predictors

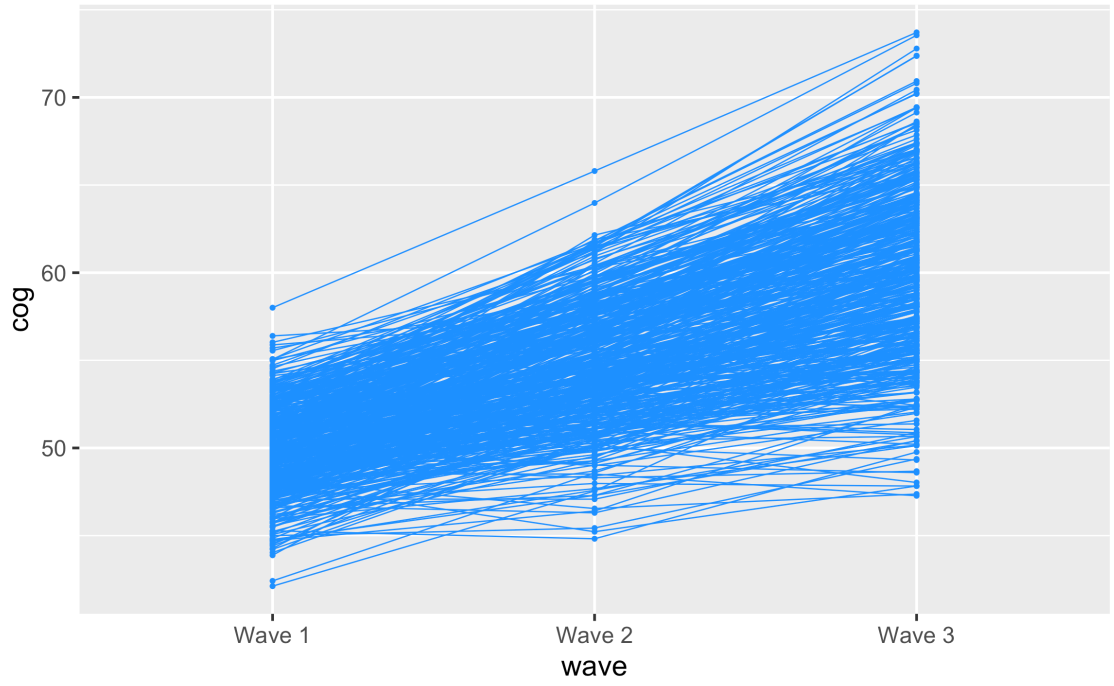



```
'
i=~1*COGT1+1*COGT2+1*COGT3
s=~0*COGT1+1*COGT2+2*COGT3
s~~i
s~Age
'
```



What does this look like in Lavaan?

```
> head(simdat_3waves)
     COGT1    COGT2    COGT3      age
1 53.58734 59.41773 65.45417 14.58179
2 54.64161 58.73914 62.56442 14.13963
3 49.94976 54.86229 63.12696 15.71563
4 53.69024 60.73795 67.02106 17.97347
5 47.70542 55.70050 62.30530 15.96064
6 51.01630 59.79515 65.29828 17.96462
```




```
growth_model_example<-
'
i=~1*COGT1+1*COGT2+1*COGT3
s=~0*COGT1+1*COGT2+2*COGT3
s~~i
s~age
'

#model fitting command
fit_growth_model_example <- growth(growth_model_example, data=simdat_3waves,missing='fiml')
# Output results
summary(fit_growth_model_example, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE,ci=T)
```

-The name of my model is 'growth_model_example'
-The data file is called 'simdat_3waves'

-The model has:
-an intercept (i)
-a linear slope (s)
-age predicting the rate of change.
-If data is missing I want to use 'FIML' (i.e. include all data)

- summary(fit_growth_model_example, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE,ci=T)

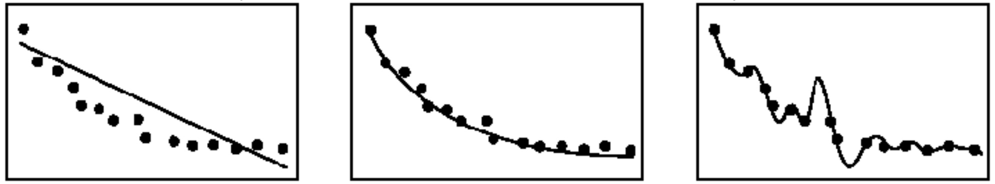
Step 1: Inspect model fit

```
> summary(fit_growth_model_example, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE,ci=T)
lavaan 0.6-18 ended normally after 123 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                         9

  Number of observations                           500
  Number of missing patterns                         1

Model Test User Model:

  Test statistic                                 2.561
  Degrees of freedom                                 3
  P-value (Chi-square)                           0.464

Model Test Baseline Model:

  Test statistic                              1421.790
  Degrees of freedom                                 6
  P-value                                        0.000

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                    1.000
  Tucker-Lewis Index (TLI)                       1.001

  Robust Comparative Fit Index (CFI)             1.000
  Robust Tucker-Lewis Index (TLI)                1.001

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)              -3397.408
  Loglikelihood unrestricted model (H1)      -3396.127

  Akaike (AIC)                                6812.816
  Bayesian (BIC)                              6850.747
  Sample-size adjusted Bayesian (SABIC)       6822.180

Root Mean Square Error of Approximation:

  RMSEA                                          0.000
  90 Percent confidence interval - lower         0.000
  90 Percent confidence interval - upper         0.071
  P-value H_0: RMSEA <= 0.050                    0.826
  P-value H_0: RMSEA >= 0.080                    0.025
```

Step 2: interpret parameters

```
Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper   Std.lv  Std.all
  i =~
    COGT1             1.000                               1.000    1.000    2.259    0.947
    COGT2             1.000                               1.000    1.000    2.259    0.584
    COGT3             1.000                               1.000    1.000    2.259    0.396
  s =~
    COGT1             0.000                               0.000    0.000    0.000    0.000
    COGT2             1.000                               1.000    1.000    2.193    0.567
    COGT3             2.000                               2.000    2.000    4.386    0.769

Regressions:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper   Std.lv  Std.all
  s ~
    age              -0.243    0.094   -2.583    0.010   -0.427   -0.059   -0.111   -0.113

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper   Std.lv  Std.all
  i ~~
   .s                 1.901    0.288    6.612    0.000    1.338    2.465    0.386    0.386

Intercepts:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper   Std.lv  Std.all
    i                50.037    0.106  471.113    0.000   49.828   50.245   22.146   22.146
   .s                 2.644    1.411    1.874    0.061   -0.121    5.408    1.206    1.206

Variances:
                   Estimate  Std.Err  z-value  P(>|z|) ci.lower ci.upper   Std.lv  Std.all
   .COGT1             0.592    0.267    2.219    0.026    0.069    1.115    0.592    0.104
   .COGT2             1.246    0.222    5.618    0.000    0.811    1.681    1.246    0.083
   .COGT3             0.609    0.633    0.962    0.336   -0.632    1.849    0.609    0.019
    i                 5.105    0.417   12.238    0.000    4.287    5.922    1.000    1.000
   .s                 4.747    0.381   12.464    0.000    4.000    5.493    0.987    0.987

R-Square:
                   Estimate
    COGT1             0.896
    COGT2             0.917
    COGT3             0.981
    s                 0.013
```

Modeling the shape of change

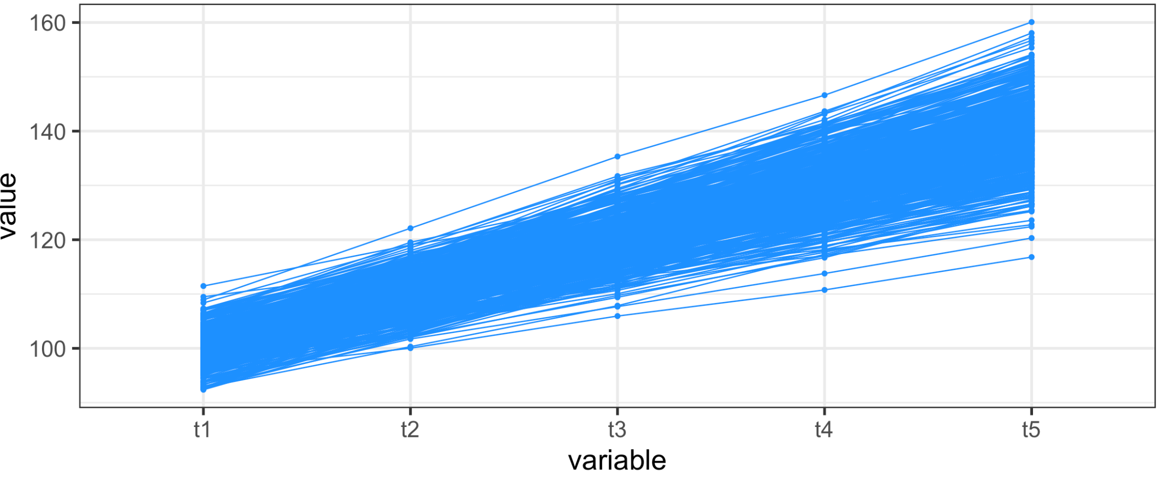


- Think through plausible candidates (given waves)
- Compare model fit using
  - Chi-square: Does the more complex model fit significantly better than expected just by chance
  - AIC/BIC: Does the model fit outweigh the penalty on the number of parameters
- 3 candidates:
  - Linear
  - Quadratic (=linear+quadratic)
  - Free ('basis' model)
- Balance explanatory power (fit to data) and parsimony (# parameters)

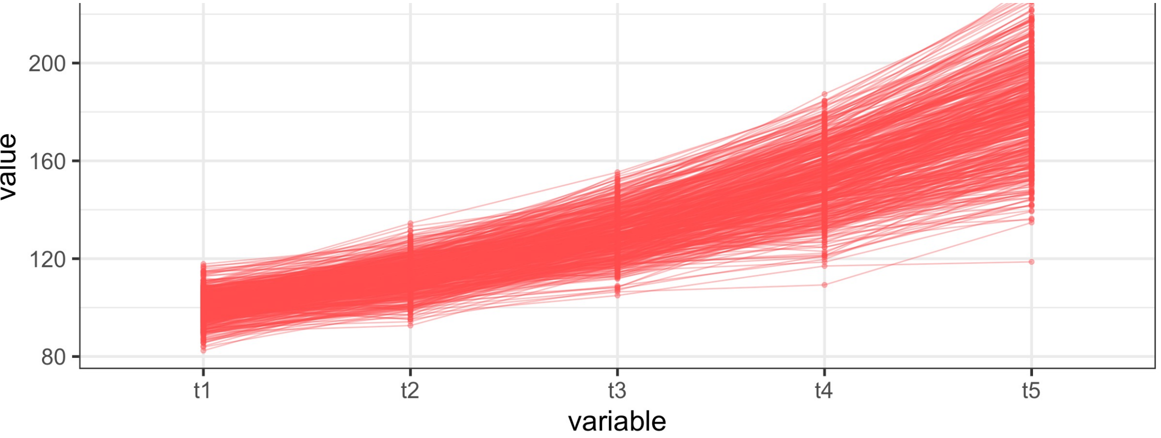
s_lin =~ 0*t1 + 1*t2 + 2*t3 + 3*t4+4*t5

s_lin =~ 0*t1 + 1*t2 + 2*t3 + 3*t4+4*t5
s_quad =~ 0*t1 + 1*t2 + 4*t3 + 9*t4 +16*t5

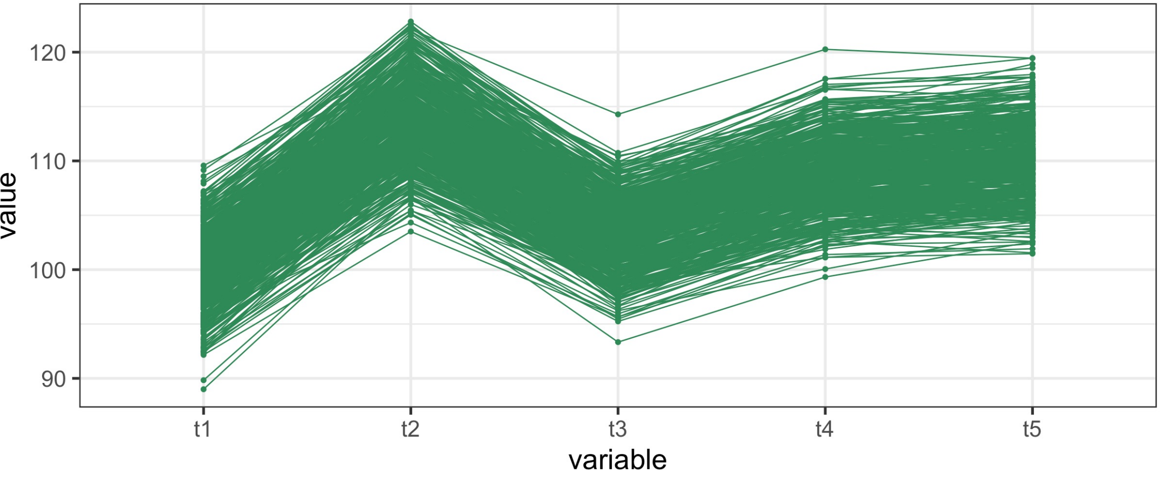
s_basis =~ 0*t1 + **t2 + t3 + t4+**1*t5




```
Chi-Squared Difference Test

                Df   AIC   BIC  Chisq Chisq diff RMSEA Df diff Pr(>Chisq)
fit_model_quad  10 10699 10741 6.4161
fit_model_basis 11 10697 10735 6.5470    0.13094     0       1     0.7175
fit_model_lin   14 10693 10718 8.5030    1.95596     0       3     0.5816
>
```




```
> anova(fit_model_lin,fit_model_quad,fit_model_basis)

Chi-Squared Difference Test

                Df   AIC   BIC     Chisq Chisq diff   RMSEA Df diff Pr(>Chisq)
fit_model_quad  10 16869 16911    4.4421
fit_model_basis 11 16894 16932   31.1694      26.73 0.22684       1  2.343e-07 ***
fit_model_lin   14 18296 18322 1439.4909    1408.32 0.96793       3  < 2.2e-16 ***
---
```



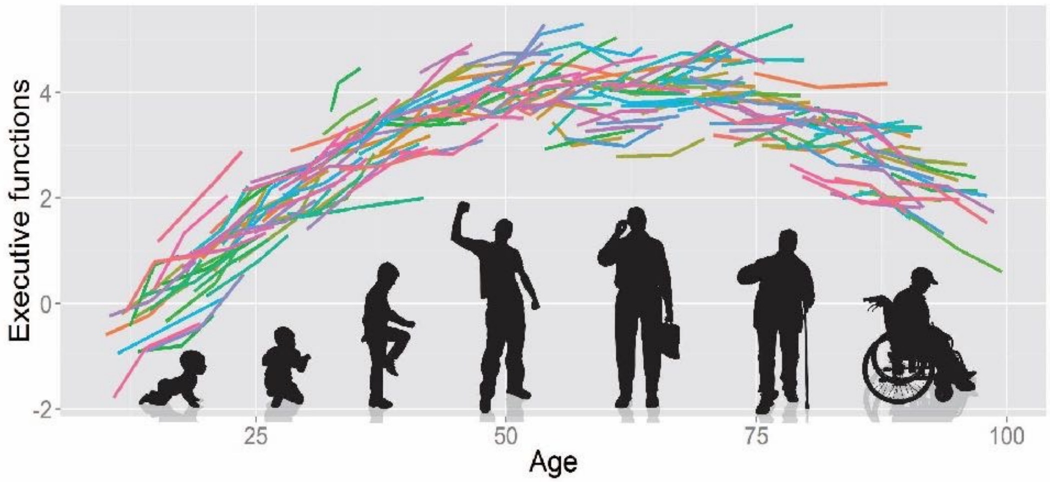

```
Chi-Squared Difference Test

                Df     AIC     BIC    Chisq Chisq diff  RMSEA Df diff Pr(>Chisq)
fit_model_quad  10 13331.3 13373.4 3615.028
fit_model_basis 11  9741.5  9779.4   27.225    -3587.8 0.0000       1          1
fit_model_lin   14 14721.6 14746.9 5013.336     4986.1 1.8227       3     <2e-16 ***
```

Léa Michel

Magda Lazari

Nick Judd

An introduction to Latent Growth Curve Modeling



Rogier A. Kievit
Lifespan Cognitive Dynamics Group

rogier.kievit@radboudumc.nl

Simulating data in Lavaan

```
simdat<-'
#basic growth model
i=~1*COGT1+1*COGT2+1*COGT3
s=~0*COGT1+1*COGT2+2*COGT3
#intercepts (mean)
i~50*1
s~2*1
#variances
s~~4*s
i~~5*i

#covariance
i~~2*s

#predictors
s~-0.2*age

#covariate properties
age~15*1
age~~1*age

'

simdat_3waves <- simulateData(simdat, sample.nobs = 500)
colMeans(simdat_3waves)
```