# Linear mixed models
# Part II

Marcello Gallucci
University of Milano-Bicocca

# GLM

When the assumptions are NOT met because the data, and thus the errors, have more complex structures, we generalize the GLM to the Linear Mixed Model

# Linear Mixed Model

## GLM

- Regression
- T-test
- ANOVA
- ANCOVA
- Moderation
- Mediation
- Path Analysis

## LMM

- Random coefficients models
- Random intercept regression models
- One-way ANOVA with random effects
- One-way ANCOVA with random effects
- Intercepts-and-slopes-as-outcomes models
- Multi-level models

# The mixed model

- We can now define a model with a regression for each cluster and the mean values of coefficients

**Overall model**

$$\hat{y}_{ij} = \bar{a} + a'_j + \bar{b} \cdot x_{ij} + b'_j \cdot x_{ij}$$

**Random coefficients**

**Fixed coefficient**

**A GLM which contains both fixed and random effects is called a Linear Mixed Model**

# The mixed model

- In practice, mixed models allow to estimate the kind of effects we can estimate with the GLM, but they allow the effects to vary across clusters.
- Effects that vary across clusters are called **random effects**
- Effects that do not vary (the ones that are the same across clusters) are said to be **fixed effects**

# Building a model

To build a model in a simple way, we need to answer very few questions:

- What is (are) the cluster variable(s)?
- What are the fixed effects?
- What are the random effects?

# Software

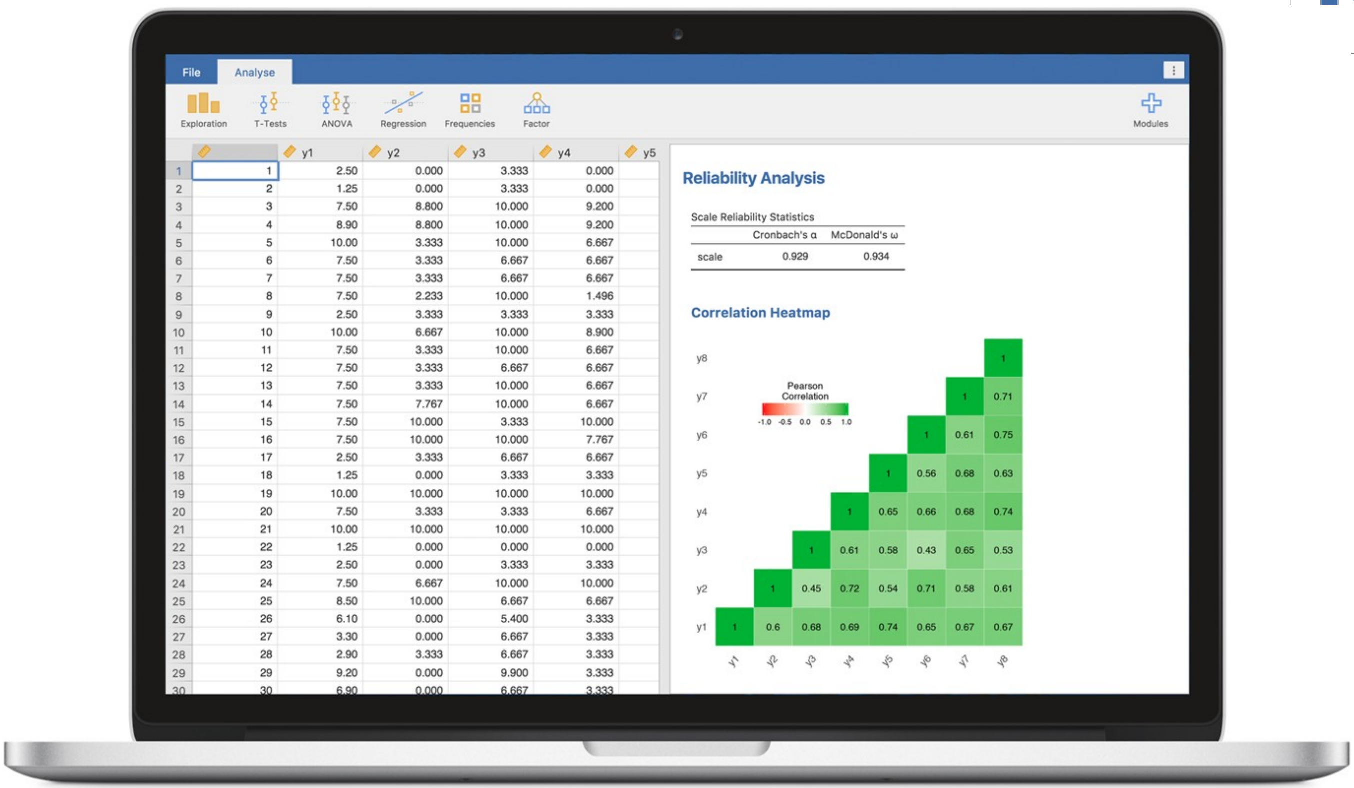
SPSS

R

jamovi Stats. Open. Now.

# Jamovi

www.jamovi.org




jamovi Stats. Open. Now.

# Repeated Measures Anova as a linear mixed model

# A repeated measures design

- Consider now a classical repeated measures design (within-subjects) the levels of the WS IV (5 different trials) are represented by different measures taken on the same person

trial

| | | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Participants | 1 | Y11 | Y21 | Y31 | Y41 | Y51 |
| | 2 | Y12 | Y22 | Y32 | Y42 | Y52 |
| | 3 | Y13 | Y23 | Y33 | Y43 | Y53 |
| | …. | | | | | |
| | N | Y1n | Y2n | Y3n | Y4n | Y5n |

# Standard file format

- As for many applications of the repeated-measure design, each level of the WS-factor is represented by a column in the file

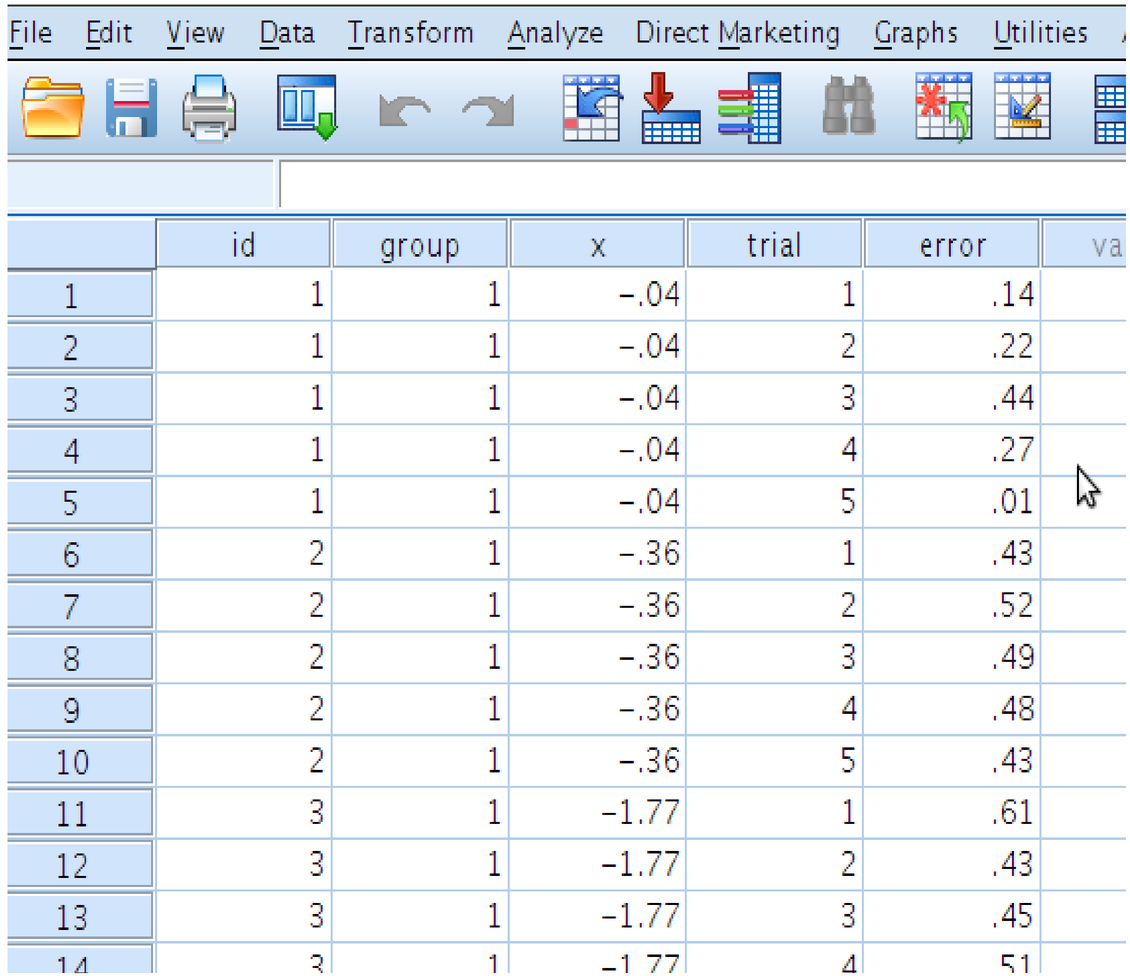
One participant, one row

| | group | err_t0 | err_t1 | err_t2 | err_t3 | err_t4 | x |
|---|---|---|---|---|---|---|---|
| 1 | 1 | .14 | .22 | .439 | .27 | .01 | -.04 |
| 2 | 1 | .43 | .52 | .492 | .48 | .43 | -.36 |
| 3 | 1 | .61 | .43 | .446 | .51 | .57 | -1.77 |
| 4 | 0 | .29 | .70 | 1.000 | .89 | .75 | 1.63 |
| 5 | 1 | .16 | .49 | .500 | .56 | .29 | -.32 |
| 6 | 0 | .70 | .36 | .573 | .57 | .69 | -1.16 |
| 7 | 0 | .35 | .51 | .572 | .46 | .77 | -.87 |
| 8 | 1 | .45 | .49 | .545 | .41 | .43 | -1.79 |
| 9 | 1 | .05 | .55 | .333 | .54 | .53 | 1.01 |
| 10 | 1 | .10 | .35 | .358 | .57 | .67 | .58 |
| 11 | 0 | .14 | .45 | .373 | .25 | .29 | -.88 |
| 12 | 0 | .04 | .74 | .541 | .53 | .35 | -.27 |
| 13 | 1 | .62 | .73 | .529 | .31 | .48 | 1.36 |
| 14 | 1 | .15 | .22 | .101 | .17 | .17 | -.32 |

# Long file format

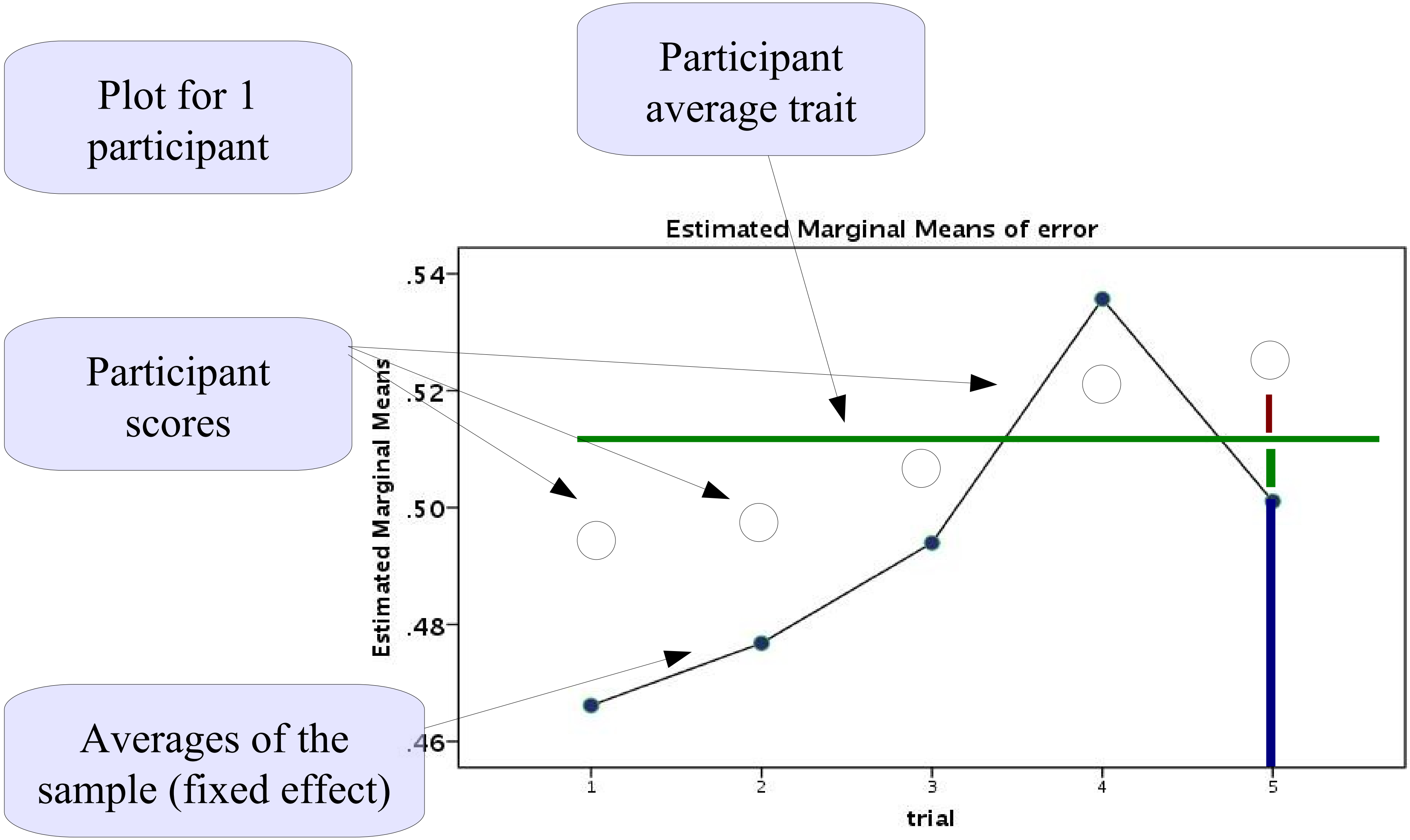
- For the mixed model we need to tabulate the data as if they came from a between-subject design

One measure, one row




| | id | group | x | trial | error |
|---|---|---|---|---|---|
| 1 | 1 | 1 | -.04 | 1 | .14 |
| 2 | 1 | 1 | -.04 | 2 | .22 |
| 3 | 1 | 1 | -.04 | 3 | .44 |
| 4 | 1 | 1 | -.04 | 4 | .27 |
| 5 | 1 | 1 | -.04 | 5 | .01 |
| 6 | 2 | 1 | -.36 | 1 | .43 |
| 7 | 2 | 1 | -.36 | 2 | .52 |
| 8 | 2 | 1 | -.36 | 3 | .49 |
| 9 | 2 | 1 | -.36 | 4 | .48 |
| 10 | 2 | 1 | -.36 | 5 | .43 |
| 11 | 3 | 1 | -1.77 | 1 | .61 |
| 12 | 3 | 1 | -1.77 | 2 | .43 |
| 13 | 3 | 1 | -1.77 | 3 | .45 |
| 14 | 3 | 1 | -1.77 | 4 | .51 |

# Participant scores

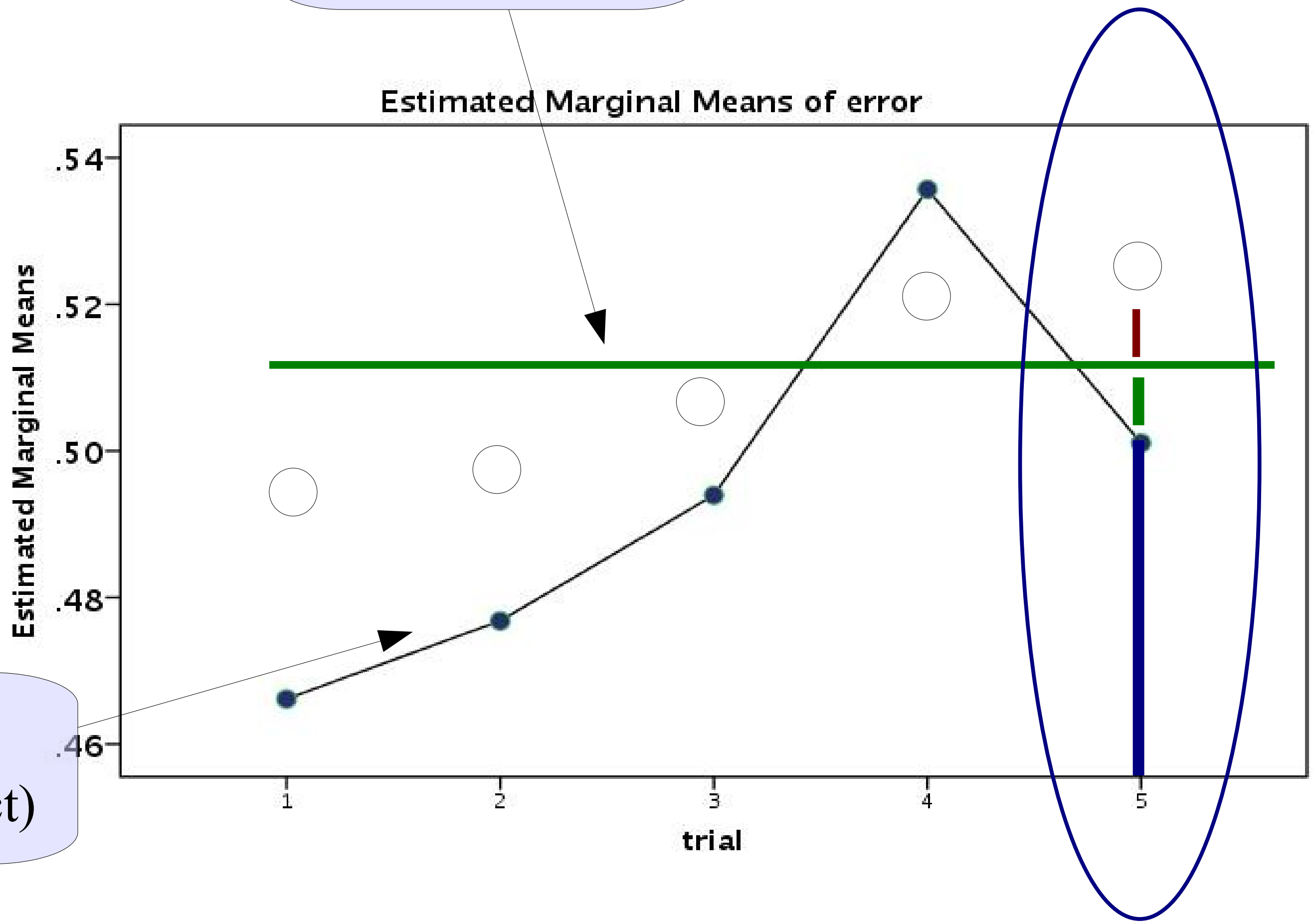


# Where does the score come from?

Plot for 1 participant

Participant average trait

Averages of the sample (fixed effect)

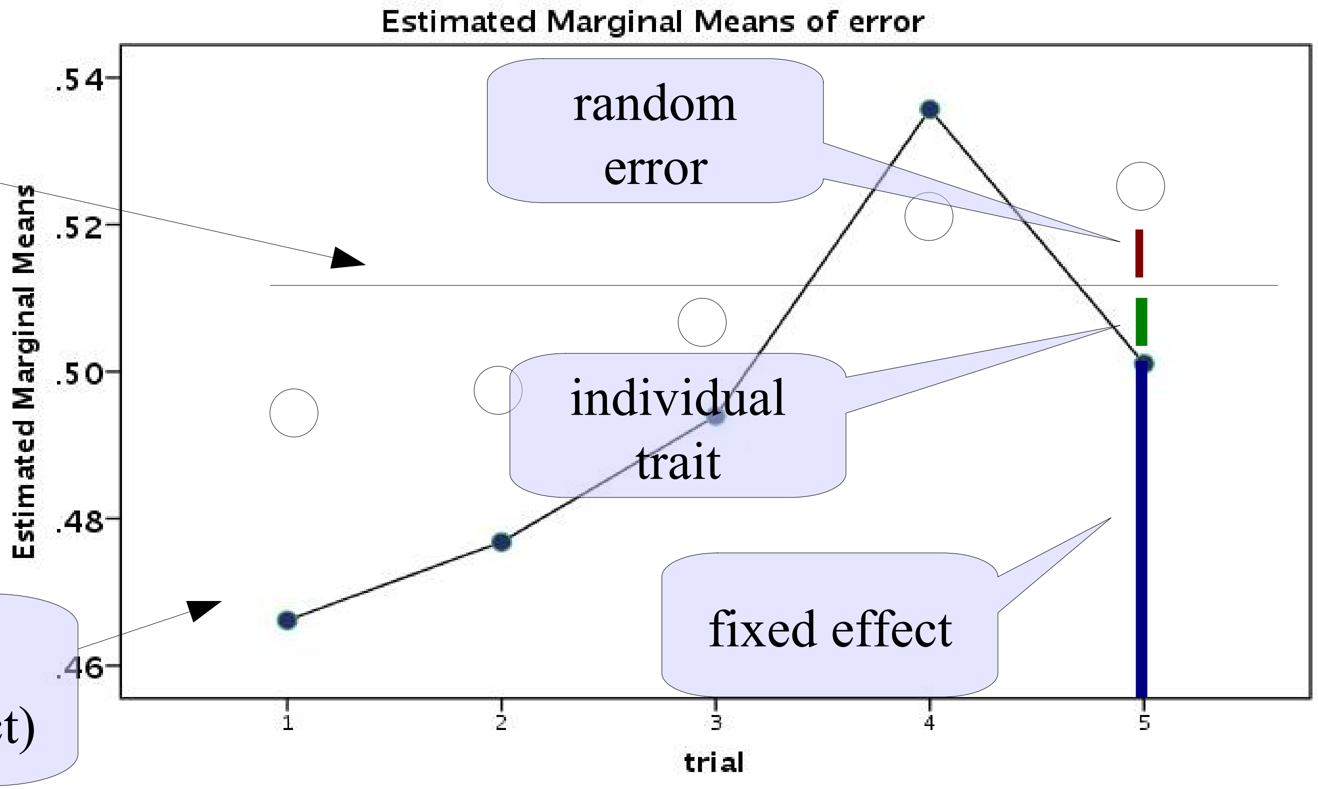


# Participant component

Plot for 1 participant

Participant individual trait

Averages of the sample (fixed effect)



# Solution

Thus, we should consider an extra residual term which represents participants individual characteristic. This term is the same within each participant

one participant
one trait

$$Y_{11} = a + b_1 \cdot T_1 + u_1 + e_{11}$$
$$Y_{21} = a + b_2 \cdot T_2 + u_1 + e_{21}$$
$$Y_{31} = a + b_3 \cdot T_3 + u_1 + e_{31}$$

Each score,
one residual

Average effects
of trials

$$Y_{1j} = a + b_1 \cdot T_1 + u_j + e_{1j}$$
$$Y_{2j} = a + b_2 \cdot T_2 + u_j + e_{2j}$$
$$Y_{3j} = a + b_3 \cdot T_3 + u_j + e_{3j}$$

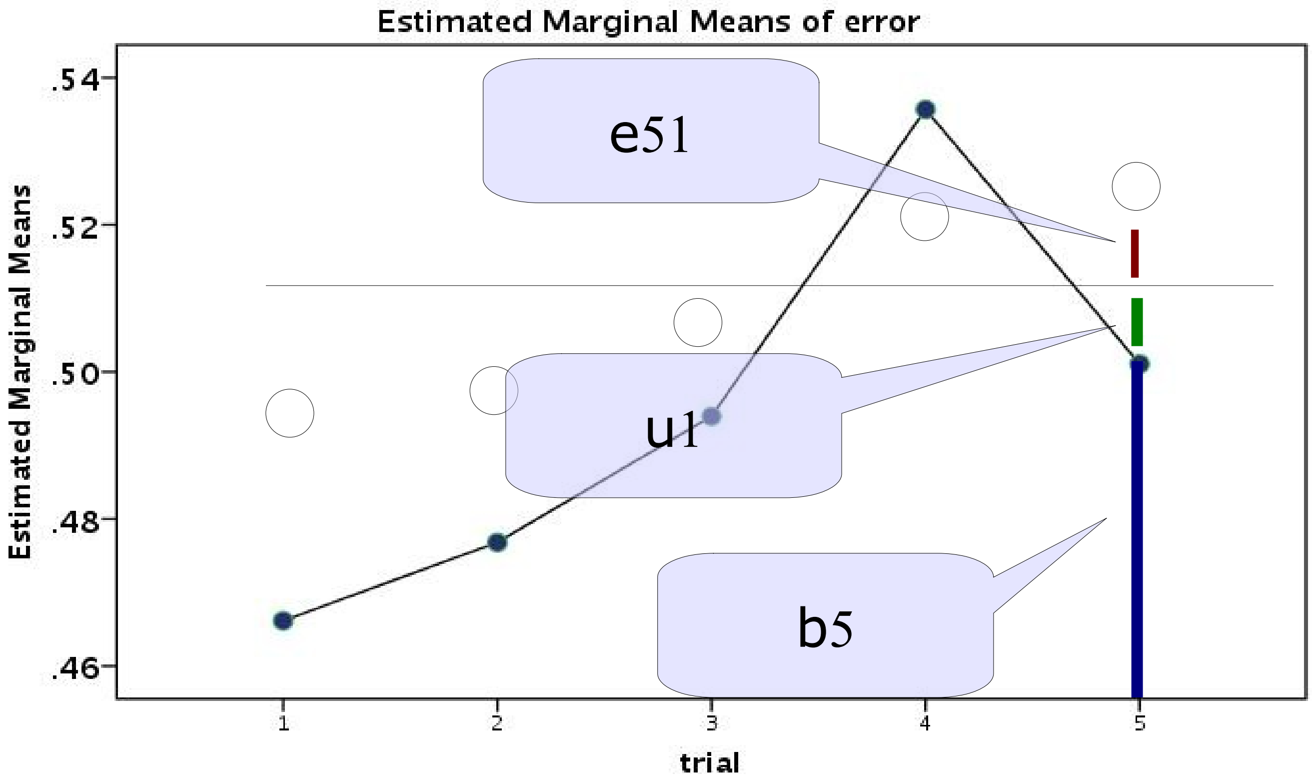
Each score,
one error

One participant
one trait

We assume the 5 trials are dummy coded

# Participant component

$$Y_{51} = a + b \cdot T_5 + u_1 + e_{51}$$



# Building the model

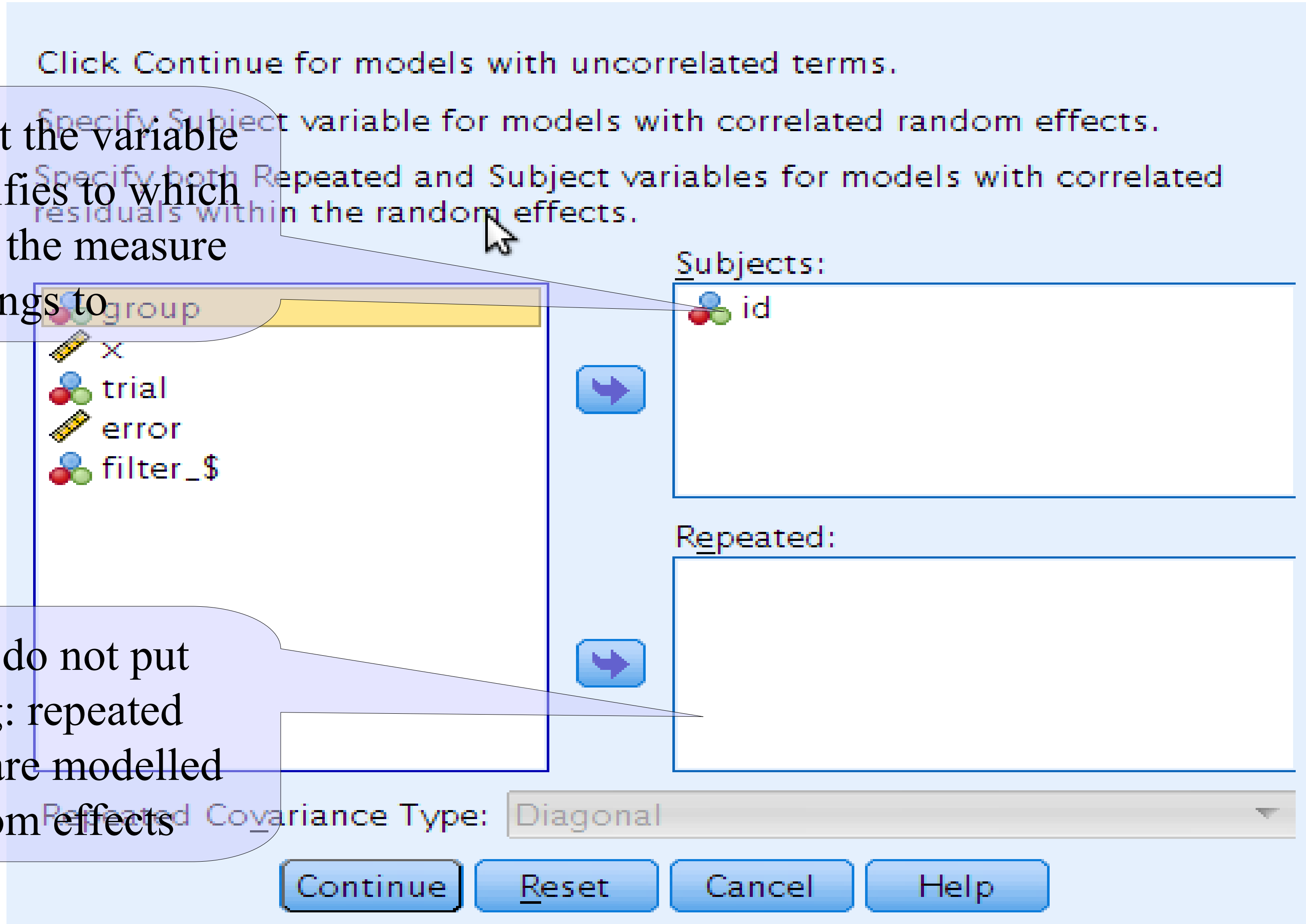
We translate this in the standard mixed model

$$Y_{ij} = a + b' \cdot T_i + u_j + e_{ij}$$

$$y_{ij} = \bar{a} + a_j + \bar{b} \cdot x_{ij} + e_{ij}$$

- Fixed effects? Intercept and trial effect
- Random effects? Intercepts
- Clusters? participants

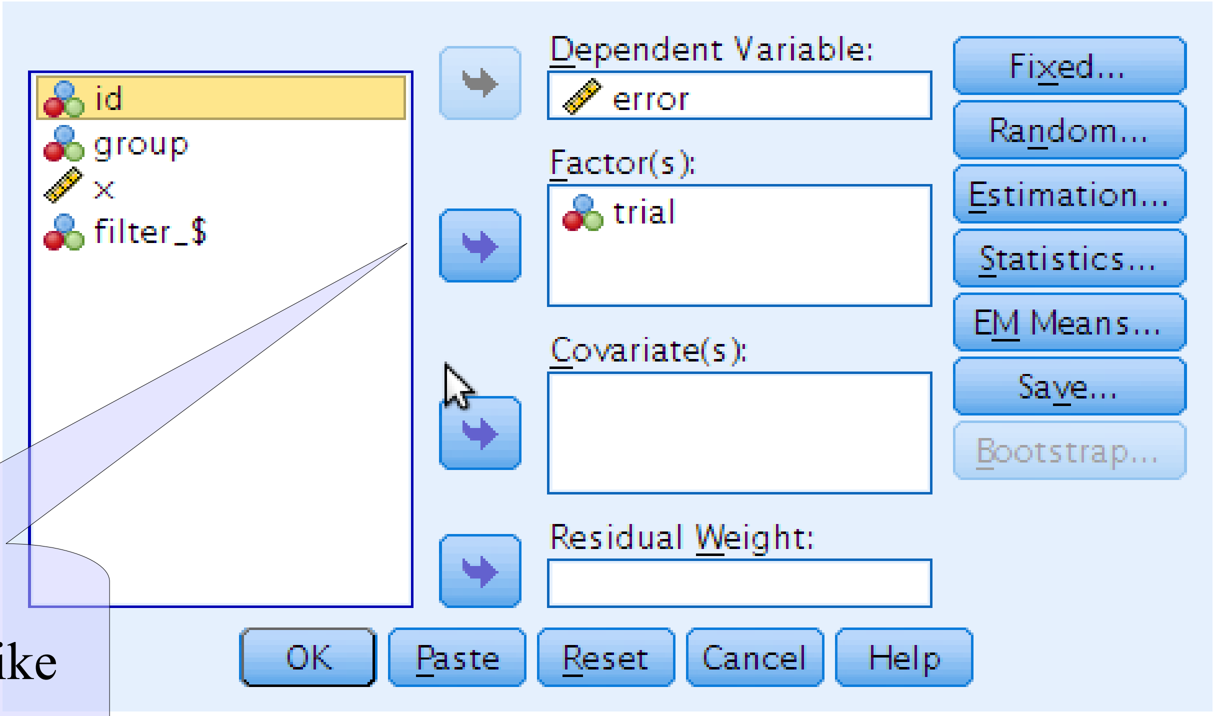
# SPSS: General mixed models




Here we put the variable which specifies to which participant the measure belongs to

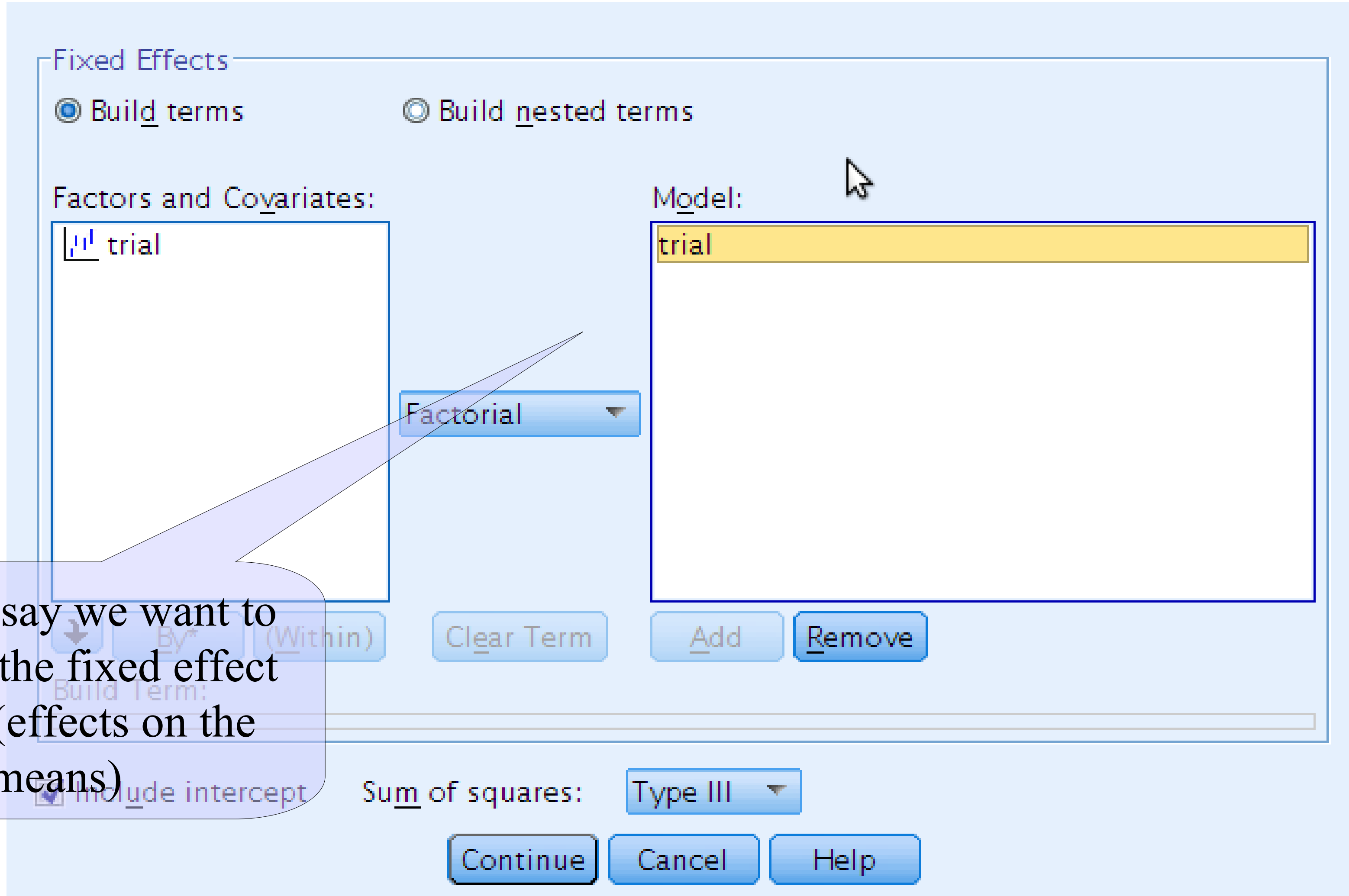
Here we do not put anything: repeated measures are modelled as random effects

# SPSS: General mixed models



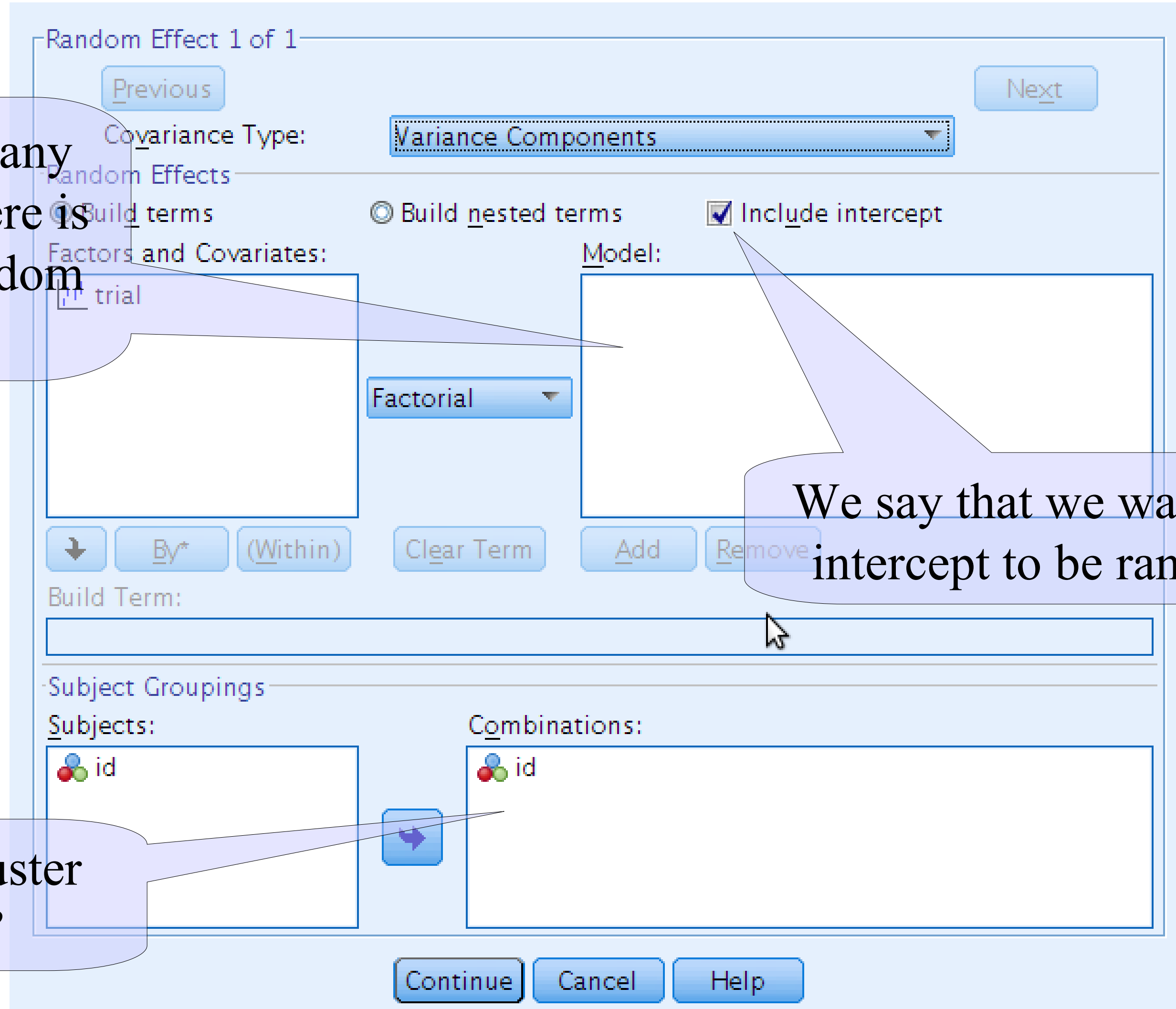

Independent and dependent variables like in GLM

# SPSS: General mixed models




Here we say we want to estimate the fixed effect of trial (effects on the means)

# SPSS: General mixed models




Here we do not put any variable, because there is no variable with random effects

We say that we want the intercept to be random

We say that the cluster variable is "id"

# SPSS: General mixed models

**Model Dimension[b]**

| | | Number of Levels | Covariance Structure | Number of Parameters | Subject Variables |
|---|---|---|---|---|---|
| Fixed Effects | Intercept | 1 | | 1 | |
| | trial | 5 | | 4 | |
| Random Effects | Intercept[a] | 1 | Variance Components | 1 | id |
| Residual | | | | 1 | |
| Total | | 7 | | 7 | |

a. As of version 11.5, the syntax rules for the RANDOM subcommand have changed. Your command syntax may yield results that differ from those produced by prior versions. If you are using version 11 syntax, please consult the current syntax reference guide for more information.

The model is as intended

# SPSS: General mixed models

## Fixed Effects

**Type III Tests of Fixed Effects[a]**

| Source | Numerator df | Denominator df | F | Sig. |
|---|---|---|---|---|
| Intercept | 1 | 199.000 | 3535.735 | .000 |
| trial | 4 | 796.000 | 4.724 | .001 |

a. Dependent Variable: error.

## Covariance Parameters

**Estimates of Covariance Parameters[a]**

| Parameter | | Estimate | Std. Error |
|---|---|---|---|
| Residual | | .030204 | .001514 |
| Intercept [subject = id] | Variance | .007804 | .001421 |

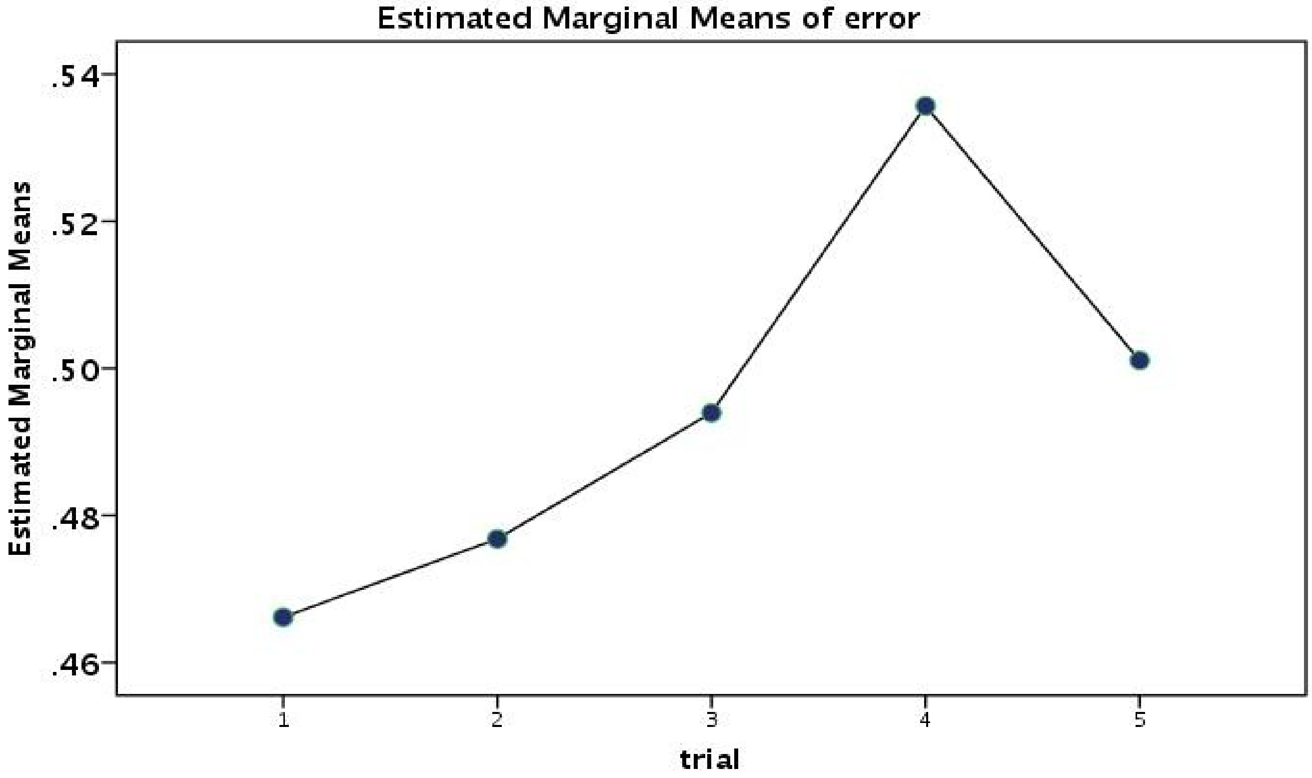
a. Dependent Variable: error.

# Interpreting the effects

- As in GLM (Anova). We interpret the main effect looking at the means

Fixed effects

Means of the 5 trials



# Dependency of scores

We can quantify the dependency of scores within clusters (participants) by computing the intra-class correlation

$$ICR = \frac{\sigma_a}{\sigma_a + \sigma}$$

**Estimates of Covariance Parameters[a]**

| Parameter | | Estimate | Std. Error |
|---|---|---|---|
| Residual | | .030204 | .001514 |
| Intercept [subject = id] | Variance | .007804 | .001421 |

a. Dependent Variable: error.

$\sigma$

$\sigma_a$

# Dependency of scores

We can quantify the dependency of scores within clusters (participants) by computing the intra-class correlation

$$ICR = \frac{.0078}{.0078 + .0302} = .205$$

**Estimates of Covariance Parameters[a]**

| Parameter | | Estimate | Std. Error |
|---|---|---|---|
| Residual | | .030204 | .001514 |
| Intercept [subject = id] | Variance | .007804 | .001421 |

a. Dependent Variable: error.

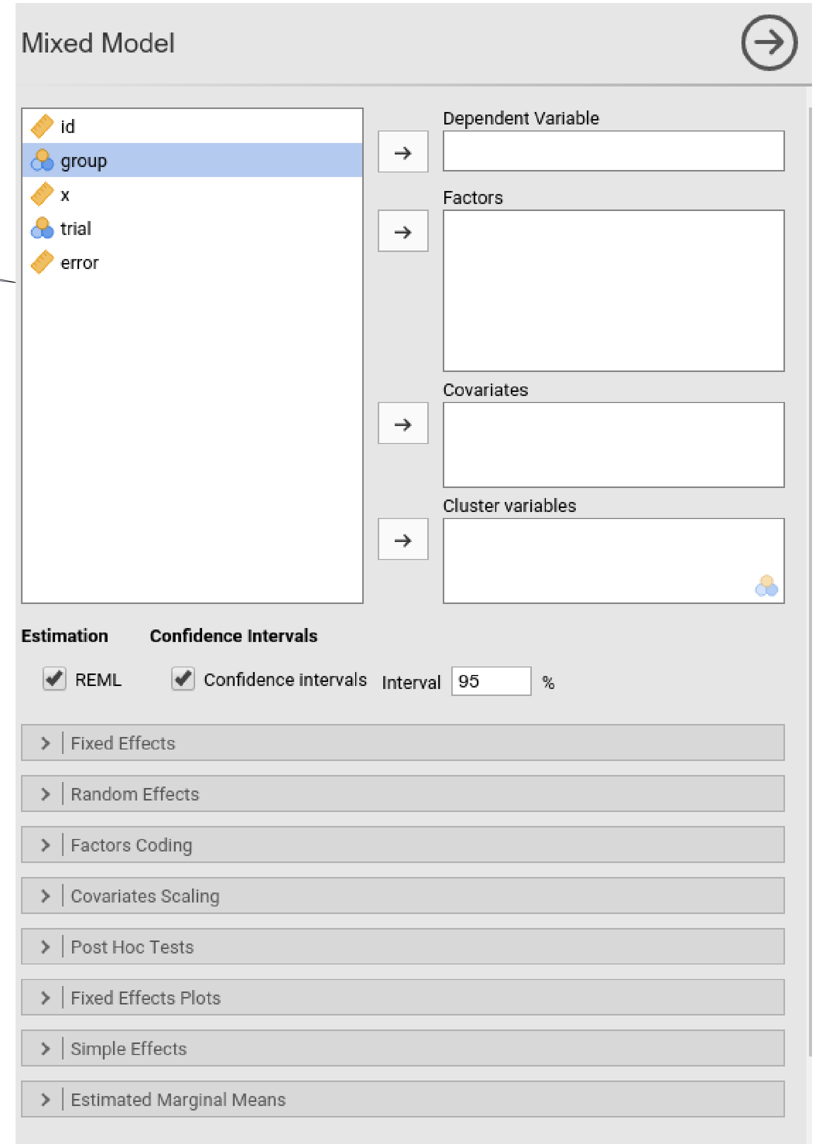
$\sigma$ (Residual)

$\sigma_a$ (Intercept variance)

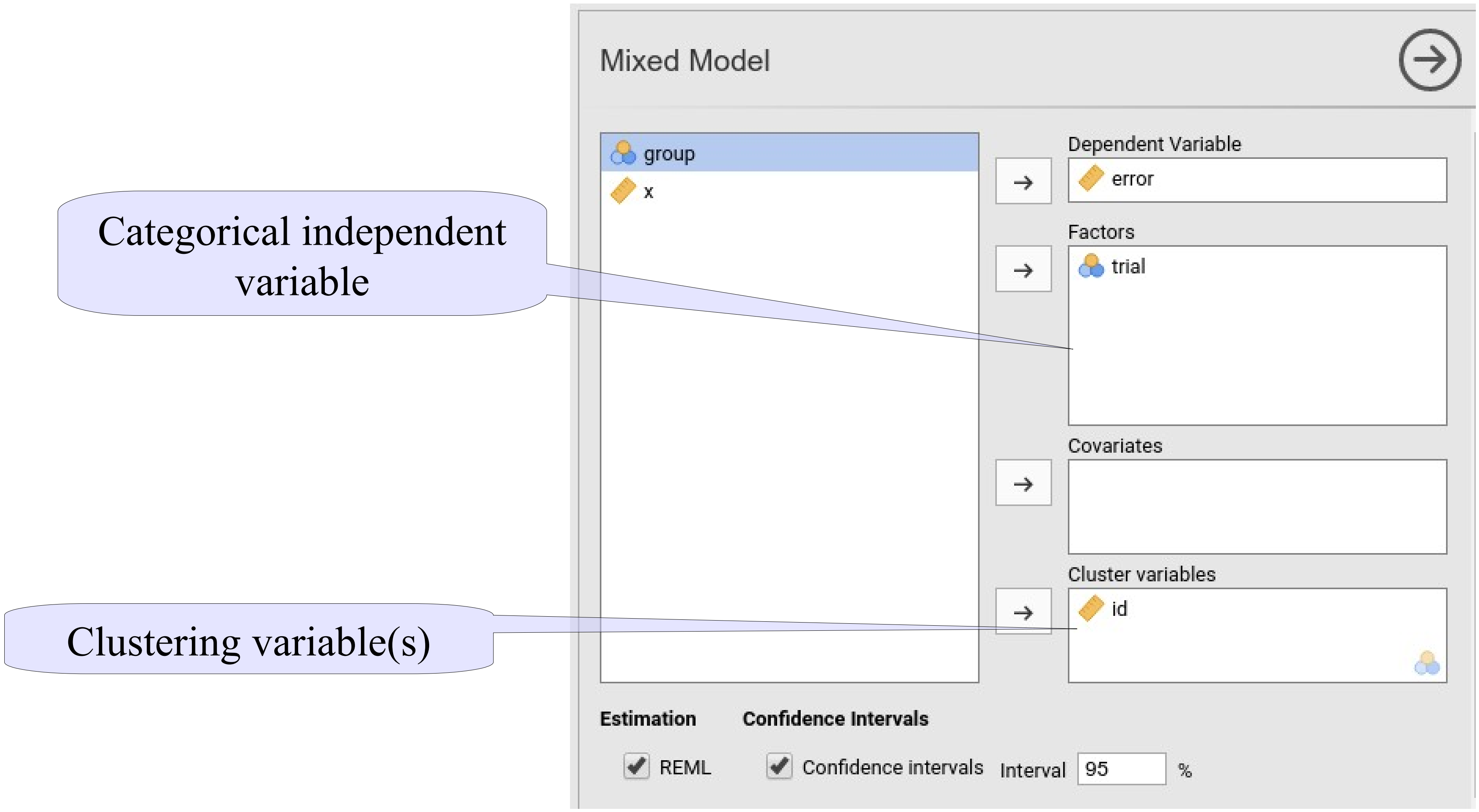
# GAMLj: mixed models

Variables

Options



# GAMLj: mixed models

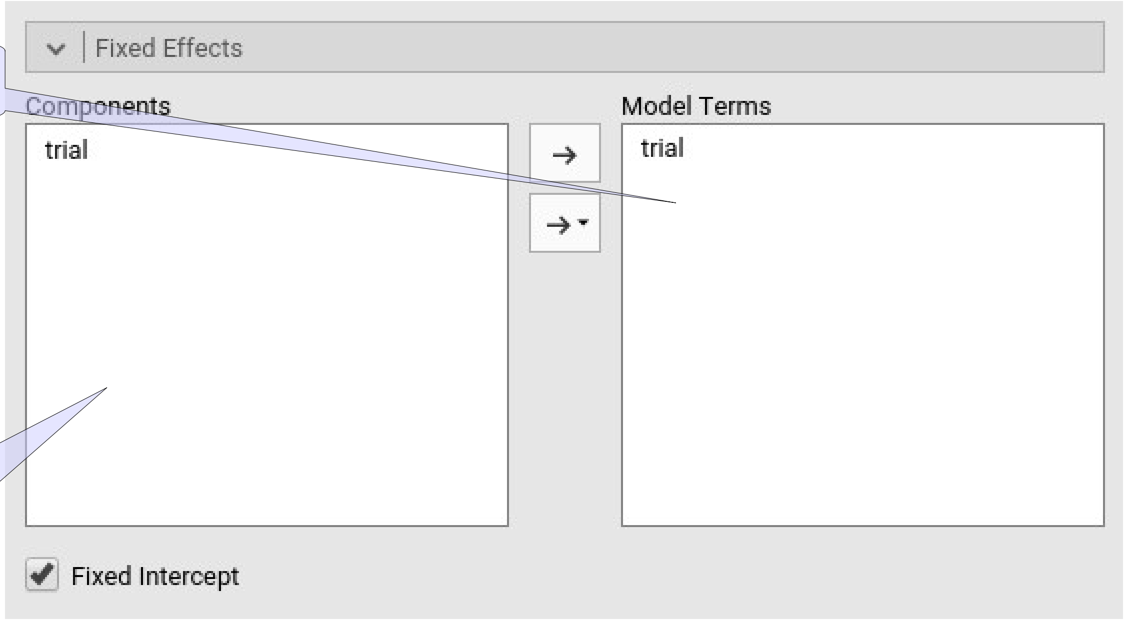


# GAMLj: random coefficients

Random intercepts


All possible random coefficients

# GAMLj: fixed coefficients




Fixed effect

All possible fixed coefficients

# GAMLj: Results: model

Model Info

| | Info |
|---|---|
| Estimate | Linear mixed model fit by REML |
| Call | error ~ 1 + (1 \| id) + trial |
| AIC | −463.8270 |
| R-squared Marginal | 0.0148 |
| R-squared Conditional | 0.2171 |

R-squared

**R-squared Conditional: How much variance can the fixed and random effects together explain of the overall variance**

**R-squared Marginal: How much variance can the fixed effects alone explain of the overall variance**

# GAMLj: Results: random

Variance of intercepts

Random Components

| Groups | Name | SD | Variance |
|---|---|---|---|
| id | (Intercept) | 0.0883 | 0.00780 |
| Residual | | 0.1738 | 0.03020 |

*Note.* Numer of Obs: 1000 , groups: id , 200

**As long as the variance is non-zero, we are fine**

# GAMLj: Results: fixed

F-tests

Fixed Effect ANOVA

| | F | Num df | Den df | p |
|---|---|---|---|---|
| trial | 4.72 | 4 | 796 | < .001 |

*Note.* Satterthwaite method for degrees of freedom

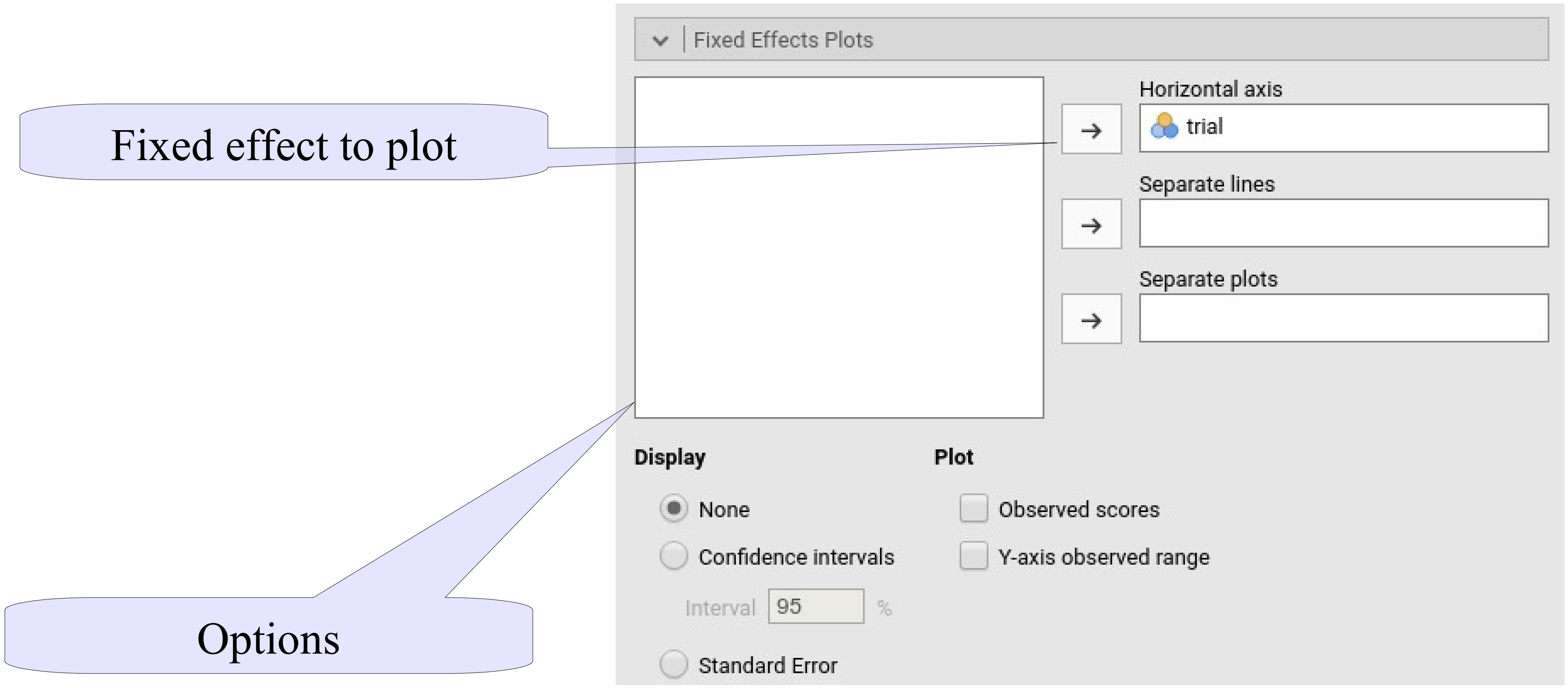
Coefficients

Fixed Effects Parameter Estimates

| | | | | 95% Confidence Interval | | | | |
|---|---|---|---|---|---|---|---|---|
| Effect | Contrast | Estimate | SE | Lower | Upper | df | t | p |
| (Intercept) | Intercept | 0.49474 | 0.00832 | 0.4784 | 0.51104 | 199 | 59.4620 | < .001 |
| trial1 | 2 - ( 1, 2, 3, 4, 5 ) | −0.01791 | 0.01099 | −0.0395 | 0.00363 | 796 | −1.6296 | 0.104 |
| trial2 | 3 - ( 1, 2, 3, 4, 5 ) | −7.92e−4 | 0.01099 | −0.0223 | 0.02075 | 796 | −0.0720 | 0.943 |
| trial3 | 4 - ( 1, 2, 3, 4, 5 ) | 0.04094 | 0.01099 | 0.0194 | 0.06248 | 796 | 3.7246 | < .001 |
| trial4 | 5 - ( 1, 2, 3, 4, 5 ) | 0.00634 | 0.01099 | −0.0152 | 0.02788 | 796 | 0.5764 | 0.564 |

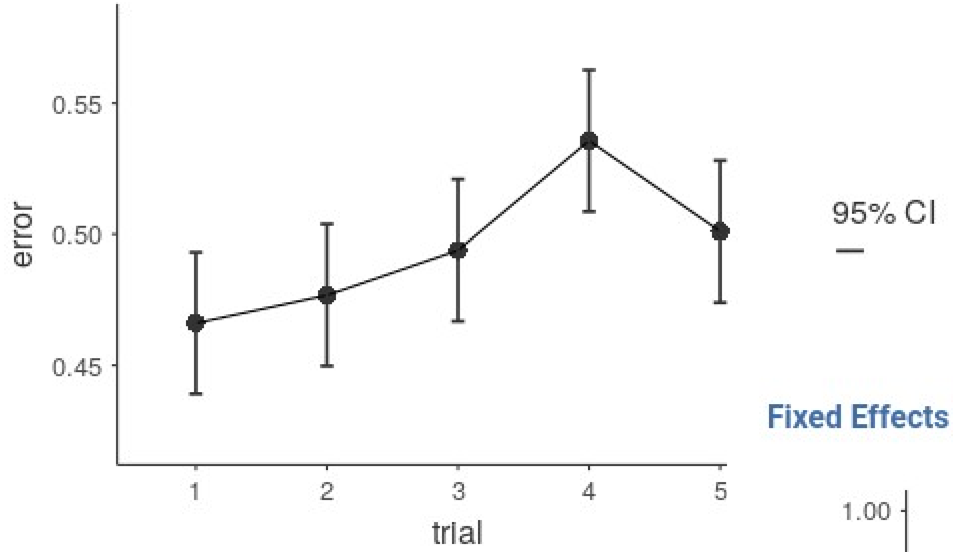
Contrasts used to cast the categorical IV

# GAMLj: plot

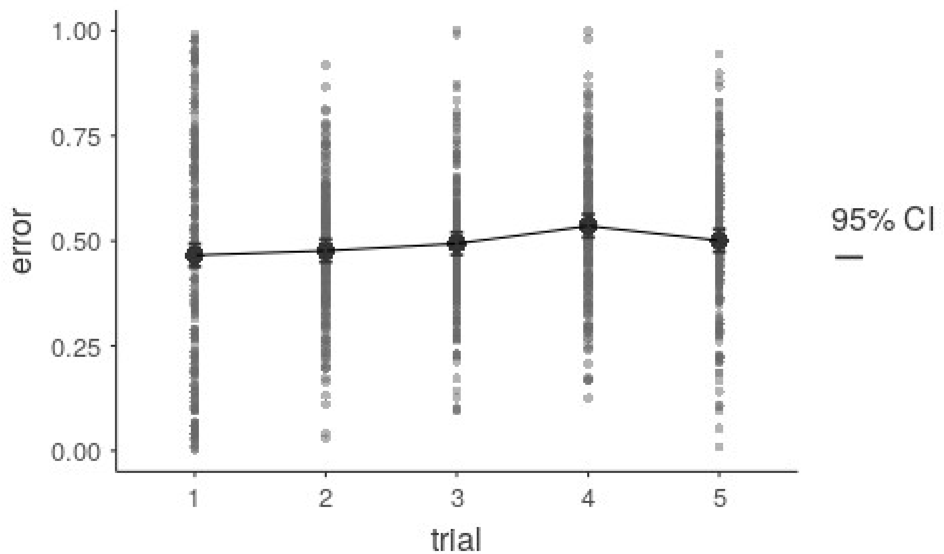

Fixed effect to plot
Options
Fixed Effects Plots
Horizontal axis
trial
Separate lines
Separate plots
Display
None
Confidence intervals
Interval
95
%
Standard Error
Plot
Observed scores
Y-axis observed range

# GAMLj: plot

Fixed Effects Plots




Fixed Effects Plots



# GAMLj: post-hoc

- As in GLM (Anova), sometimes we want to compares conditions using post-hoc tests. GAMLj allows for Bonferroni and Holm (more liberal) p-value adjustement

# GAMLj: post-hoc

- The interpretation follows as for any standard ANOVA

## Post Hoc Tests

Post Hoc Comparisons - trial

| Comparison | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| trial | | trial | Difference | SE | t | df | $p_{bonferroni}$ | $p_{holm}$ |
| 1 | - | 2 | −0.01066 | 0.0174 | −0.613 | 796 | 1.000 | 1.000 |
| | - | 3 | −0.02778 | 0.0174 | −1.598 | 796 | 1.000 | 0.552 |
| | - | 4 | −0.06951 | 0.0174 | −4.000 | 796 | < .001 | < .001 |
| | - | 5 | −0.03491 | 0.0174 | −2.009 | 796 | 0.449 | 0.314 |
| 2 | - | 3 | −0.01712 | 0.0174 | −0.985 | 796 | 1.000 | 0.975 |
| | - | 4 | −0.05885 | 0.0174 | −3.386 | 796 | 0.007 | 0.007 |
| | - | 5 | −0.02425 | 0.0174 | −1.395 | 796 | 1.000 | 0.653 |
| 3 | - | 4 | −0.04173 | 0.0174 | −2.401 | 796 | 0.166 | 0.133 |
| | - | 5 | −0.00713 | 0.0174 | −0.410 | 796 | 1.000 | 1.000 |
| 4 | - | 5 | 0.03460 | 0.0174 | 1.991 | 796 | 0.468 | 0.314 |

# Between and Repeated Measures Anova

# linear mixed model

# Standard design

- There are two groups - a Control group and a Treatment group, measured at 4 times. These times are labeled as 1 (pretest), 2 (one month posttest), 3 (3 months follow-up), and 4 (6 months follow-up).
- The dependent variable is a depression score (e.g. Beck Depression Inventory) and the treatment is drug versus no drug. If the drug worked about as well for all subjects the slopes would be comparable and negative across time. For the control group we would expect some subjects to get better on their own and some to stay depressed, which would lead to differences in slope for that group (*)

*) https://www.uvm.edu/~dhowell/StatPages/More_Stuff/Mixed-Models-Repeated/Mixed-Models-for-Repeated-Measures1.html

# Standard design

- There are two groups - a Control group and a Treatment group, measured at 4 times. These times are labeled as 1 (pretest), 2 (one month posttest), 3 (3 months follow-up), and 4 (6 months follow-up).

**96 observations**
**24 subjects**

## Contingency Tables

Contingency Tables

| | group | | |
|---|---|---|---|
| time | 1 | 2 | Total |
| 0 | 12 | 12 | 24 |
| 1 | 12 | 12 | 24 |
| 3 | 12 | 12 | 24 |
| 6 | 12 | 12 | 24 |
| Total | 48 | 48 | 96 |

*) https://www.uvm.edu/~dhowell/StatPages/More_Stuff/Mixed-Models-Repeated/Mixed-Models-for-Repeated-Measures1.html

# Standard design: data

- Data are in the long format

One subject 4 rows


| | subj | time | group | dv |
|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 296 |
| 2 | 1 | 1 | 1 | 175 |
| 3 | 1 | 3 | 1 | 187 |
| 4 | 1 | 6 | 1 | 192 |
| 5 | 2 | 0 | 1 | 376 |
| 6 | 2 | 1 | 1 | 329 |
| 7 | 2 | 3 | 1 | 236 |
| 8 | 2 | 6 | 1 | 76 |
| 9 | 3 | 0 | 1 | 309 |
| 10 | 3 | 1 | 1 | 238 |
| 11 | 3 | 3 | 1 | 150 |
| 12 | 3 | 6 | 1 | 123 |
| 13 | 4 | 0 | 1 | 222 |
| 14 | 4 | 1 | 1 | 60 |
| 15 | 4 | 3 | 1 | 82 |
| 16 | 4 | 6 | 1 | 85 |
| 17 | 5 | 0 | 1 | 150 |
| 18 | 5 | 1 | 1 | 271 |
| 19 | 5 | 3 | 1 | 250 |

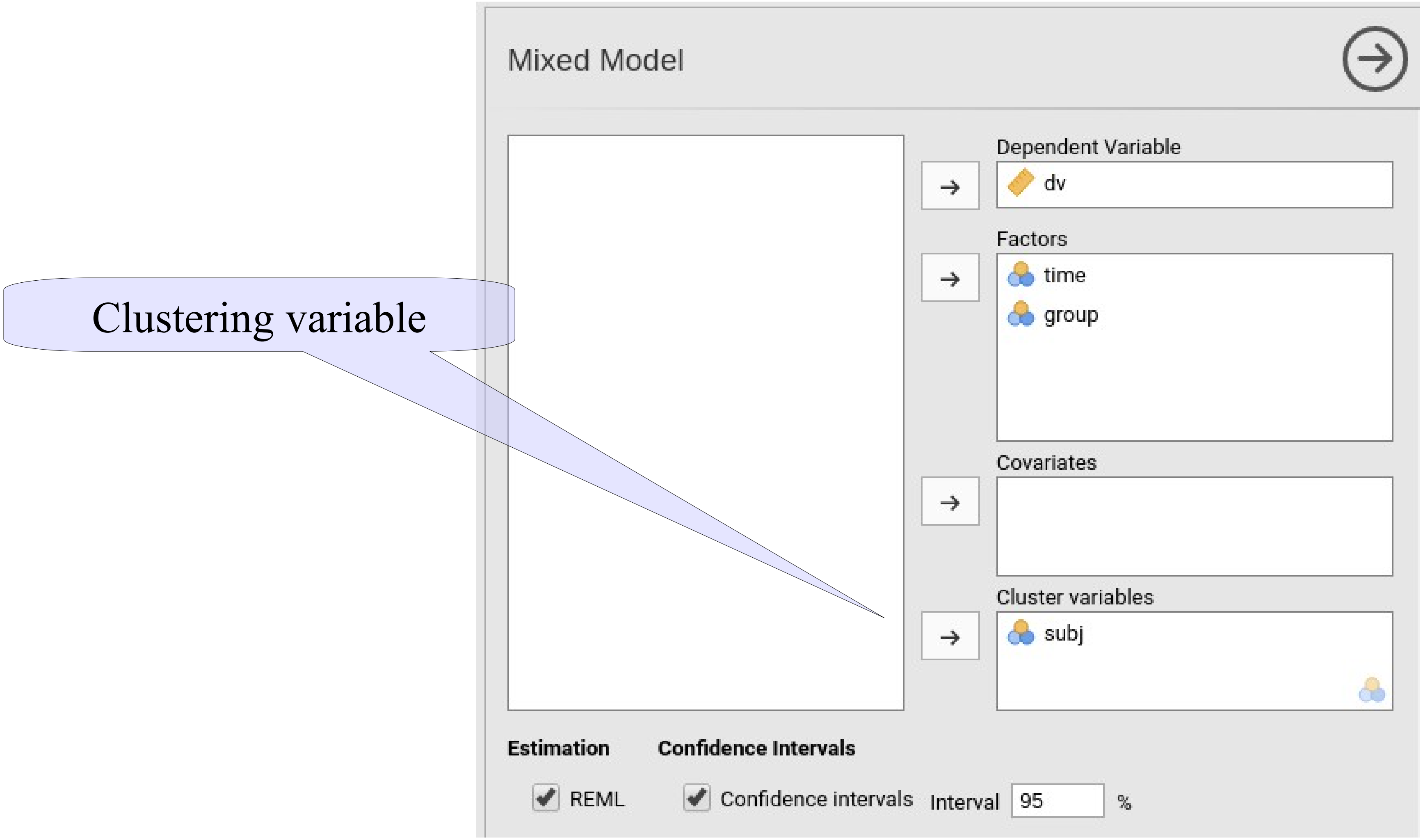
# Mixed model

We can translate this in a standard mixed model

- Fixed effects? Intercept and group,time, and interaction effect
- Random effects? Intercepts
- Clusters? subjects

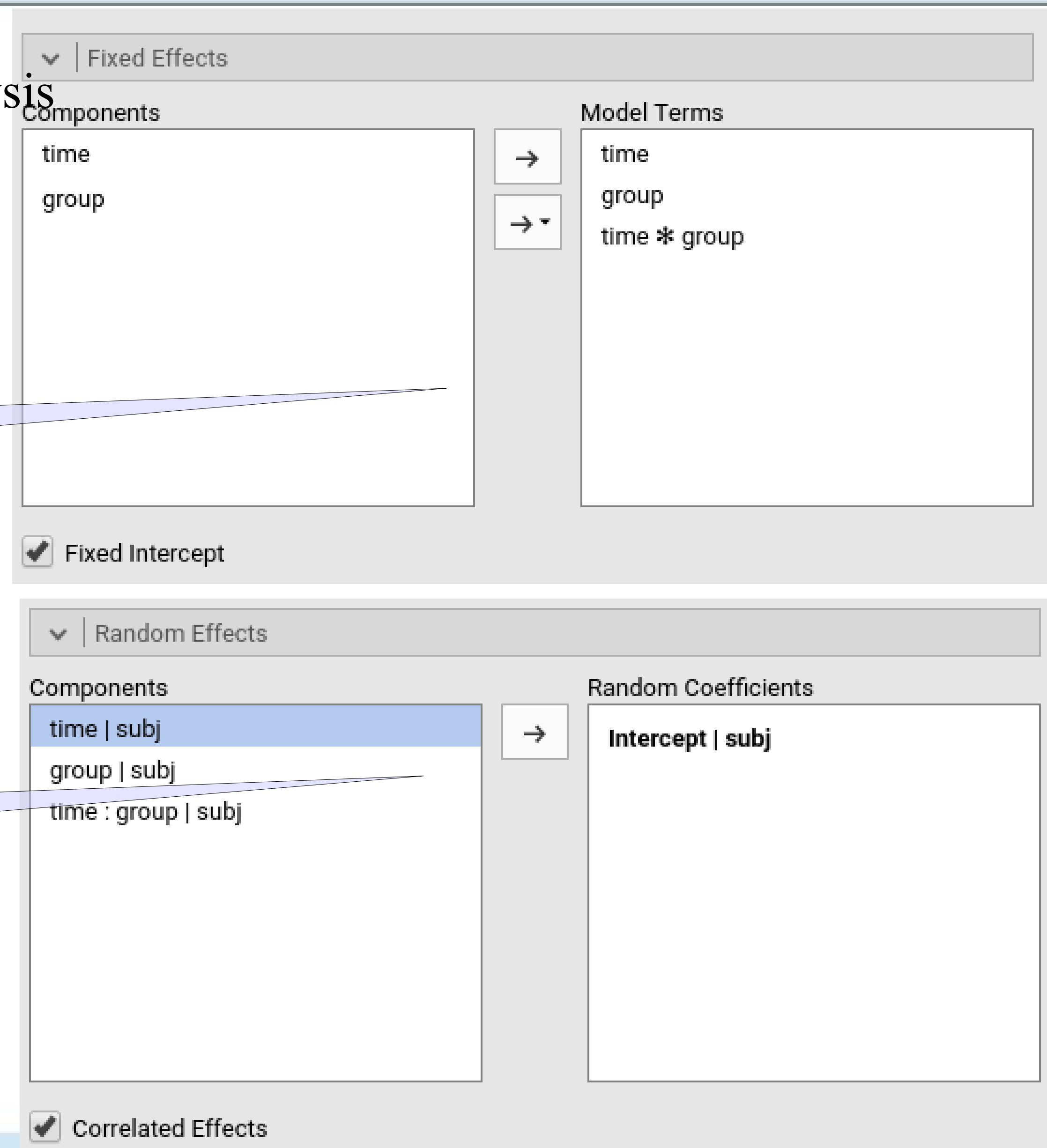
# Variables

- Definition of the analysis



# Model

- Definition of the analysis




Fixed effects

Random effects

# Results

- Interpretation of results

Model

Random effects

## Mixed Model

Model Info

| | Info |
|---|---|
| Estimate | Linear mixed model fit by REML |
| Call | dv ~ 1 + (1 \| subj) + time + group + time:group |
| AIC | 1011.895 |
| R-squared Marginal | 0.554 |
| R-squared Conditional | 0.768 |

Random Components

| Groups | Name | SD | Variance |
|---|---|---|---|
| subj | (Intercept) | 50.4 | 2539 |
| Residual | | 52.5 | 2761 |

*Note.* Numer of Obs: 96 , groups: subj , 24

# Results

- Interpretation of results

Fixed F-tests

Fixed Effect ANOVA

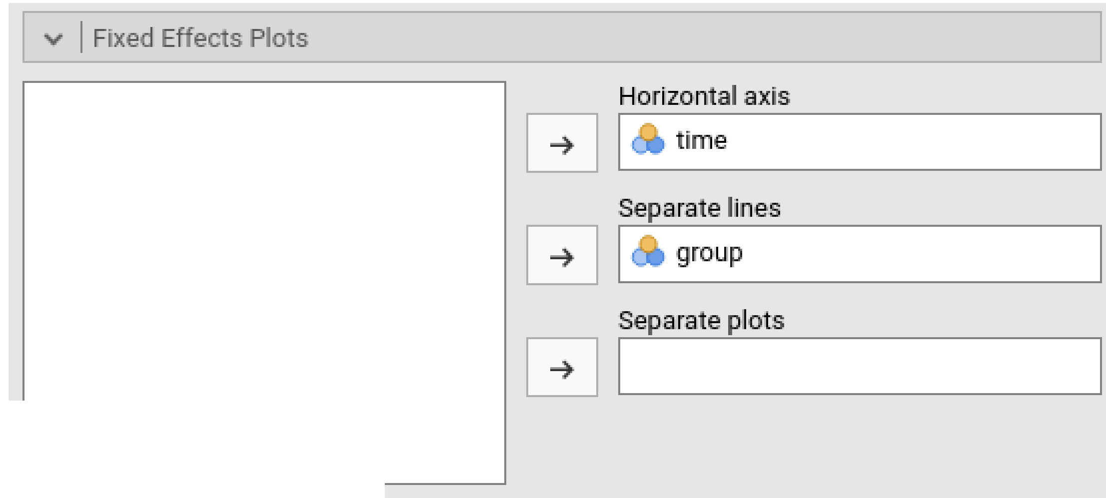
| | F | Num df | Den df | p |
|---|---|---|---|---|
| time | 45.14 | 3 | 66.0 | < .001 |
| group | 13.71 | 1 | 22.0 | 0.001 |
| time:group | 9.01 | 3 | 66.0 | < .001 |

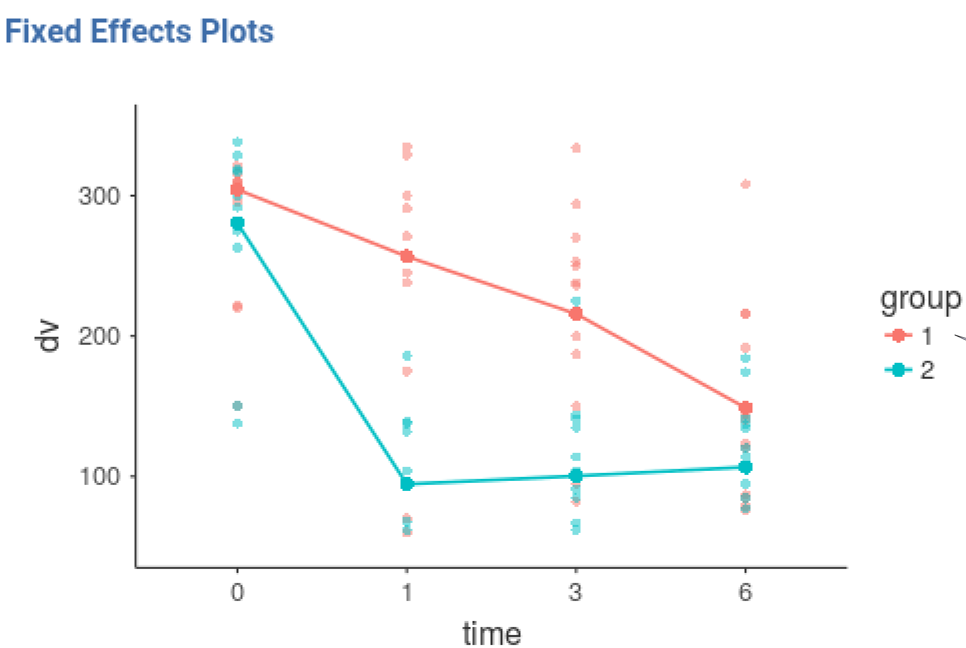
*Note.* Satterthwaite method for degrees of freedom

- For the moment we ignore the coefficients of the parameter estimates

# Results: plot

- Interpretation of results




Red is the control group

# Probing the results

- We can probe the interaction (and the pattern of means) in different ways (all available in GAMLj):
- Simple effects: Test if the effects of time is there (and how strong it is) for different groups
- Trend analysis: Checking the polynomial trend for time in general and for different groups
- Post-hoc test: not nice, but doable

# Simple effect analysis

# Simple Effects

- Simple effects are effects of one variable evaluated at one level of the other variable (like simple slopes for continuous variables)

Is the effect of B for A1 different from zero?

Is there an effect here?

| | A1 | A2 | A3 | |
|---|---|---|---|---|
| B1 | E | | | |
| B2 | E | | | |
| B3 | E | | | |
| Totals | | | | |

# Simple Effects

- Simple effects are effects of one variable evaluated at one level of the other variable (like simple slopes for continuous variables)

Is the effect of B for A2 different from zero?

Is there an effect here?

| | A1 | A2 | A3 | T |
|---|---|---|---|---|
| B1 | | E | | |
| B2 | | E | | |
| B3 | | E | | |
| Totals | | | | |

# Simple Effects

- Simple effects are effects of one variable evaluated at one level of the other variable (like simple slopes for continuous variables)

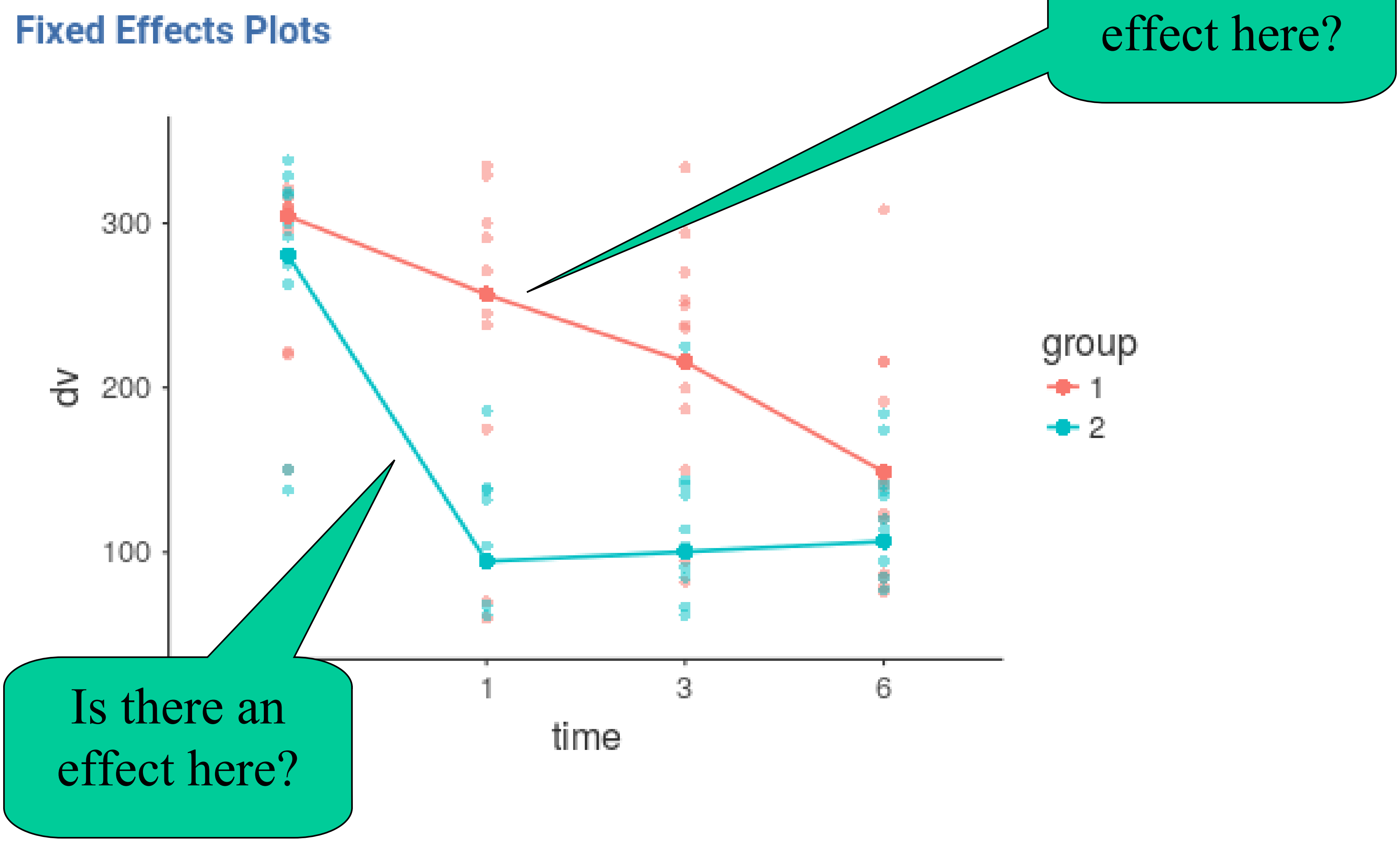
Is the effect of B for A3 different from zero?

| | A1 | A2 | A3 | Totals |
|---|---|---|---|---|
| B1 | | | E | |
| B2 | | | E | |
| B3 | | | E | |
| Totals | | | | |

Is there an effect here?

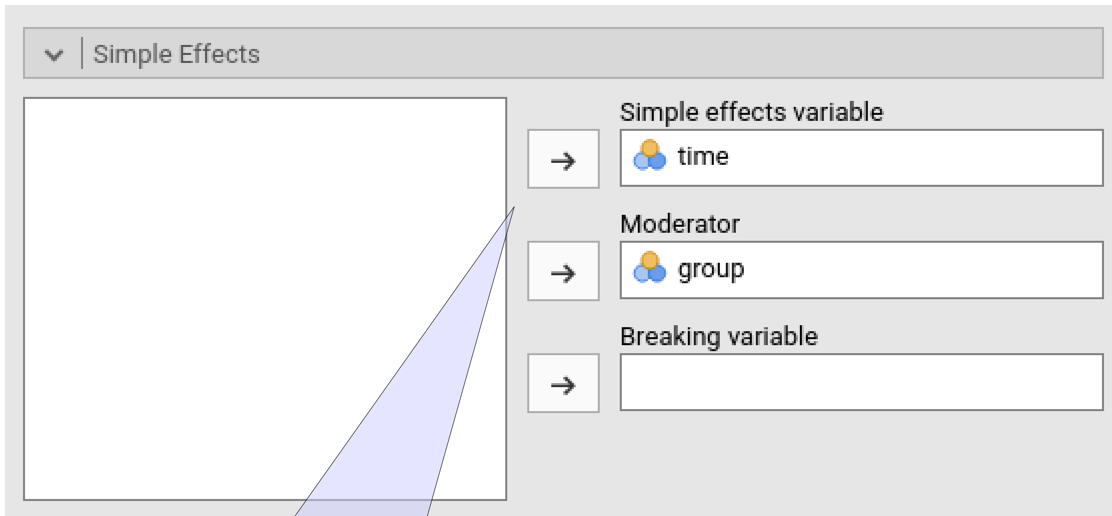
# Simple effects

- Interpretation of results



# Simple effects

- We should declare which is the variable we want the effect for and which is the moderator




Effects of time for different groups

# Simple effects

- We can say that the treatment works for both groups, although in a different way (recall the interaction)

### Simple Effects ANOVA

Simple effects of time

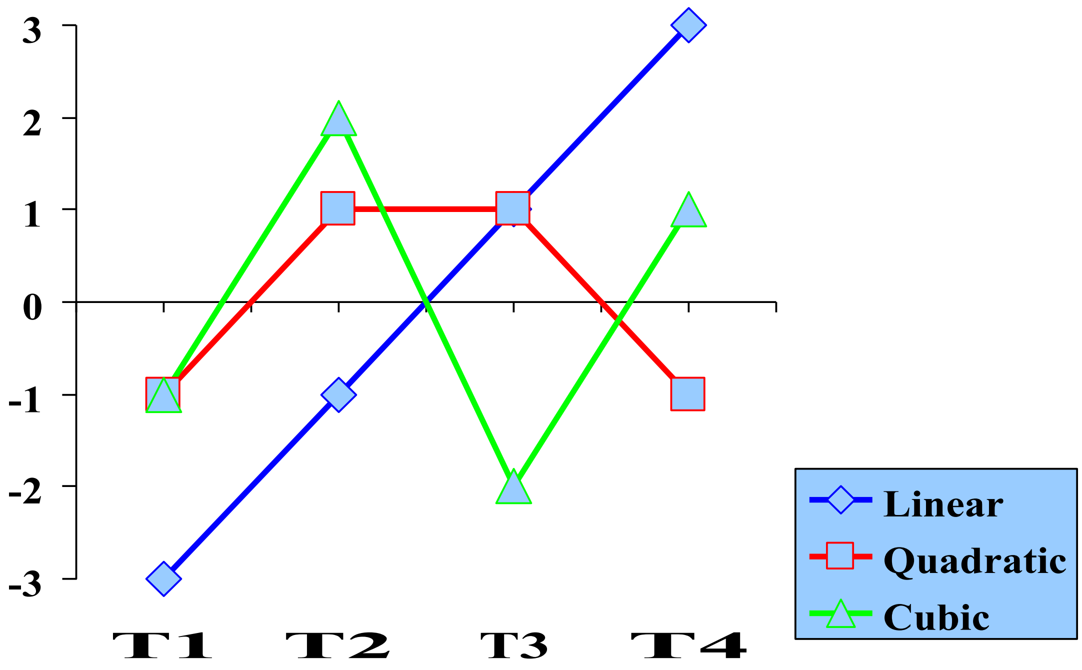
| Effect | Moderator Levels | df Num | df Den | F | p |
|---|---|---|---|---|---|
| time | group at 1 | 3.00 | 66.0 | 18.9 | < .001 |
| time | group at 2 | 3.00 | 66.0 | 35.3 | < .001 |

In both groups there is an affect of time

# Trend analysis

# Polynomial Contrasts

Trend analysis is based on Polynomial contrasts: each contrast features weights which follow well-known shapes (polynomial functions)




$$L = \begin{pmatrix} -3 & -1 & 1 & 3 \end{pmatrix}$$

$$Q = \begin{pmatrix} -1 & 1 & 1 & -1 \end{pmatrix}$$

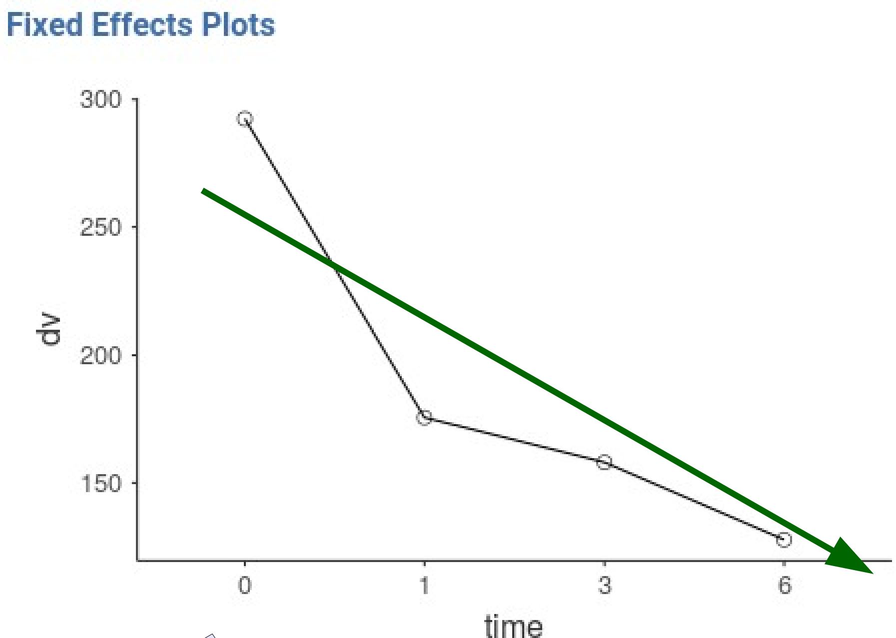
$$C = \begin{pmatrix} -1 & 2 & -2 & 1 \end{pmatrix}$$

# Trend analysis

- It is useful to test what kind of trend is present in the pattern of means
- It can be applied to any ordered categorical variables
- It is often used (and SPSS gives it by default) in repeated measures analysis
- One can estimate K-1 trends (linear, quadratic, cubic etc), where K is the number of means (conditions)

# Trend analysis

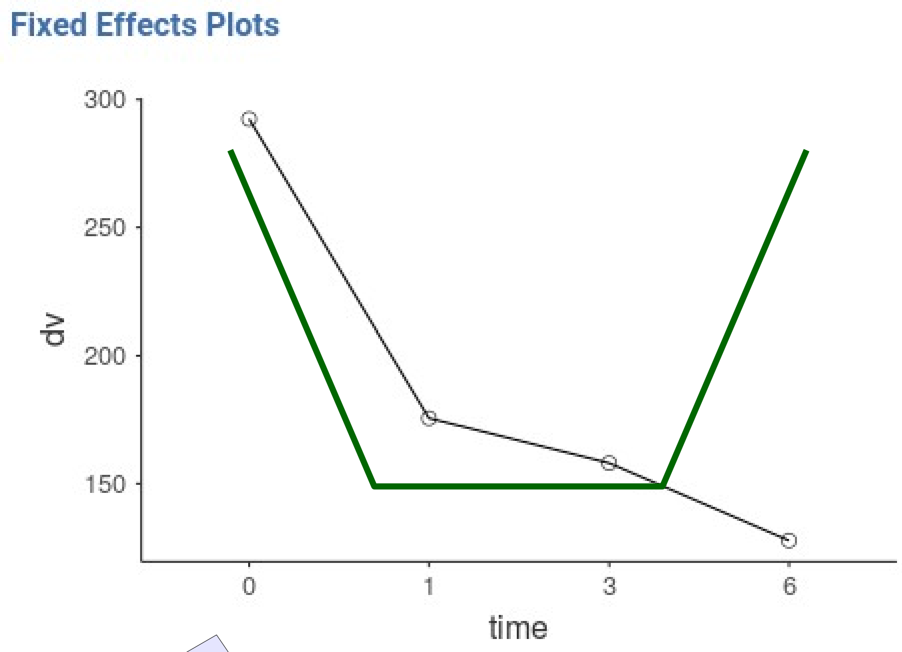
- Each trend (linear, quadratic, etc) tests a particular shape of the mean pattern




Linear: on average means go down (or up, not flat)

# Trend analysis

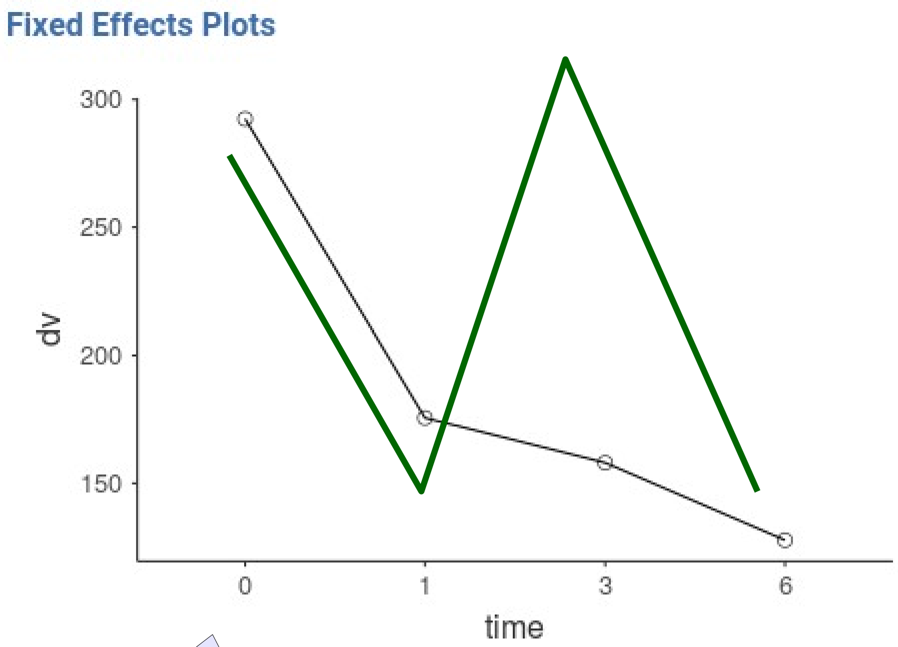
- Each trend (linear, quadratic, etc) tests a particular shape of the mean pattern




Quadratic: on average means go down and then up

# Trend analysis

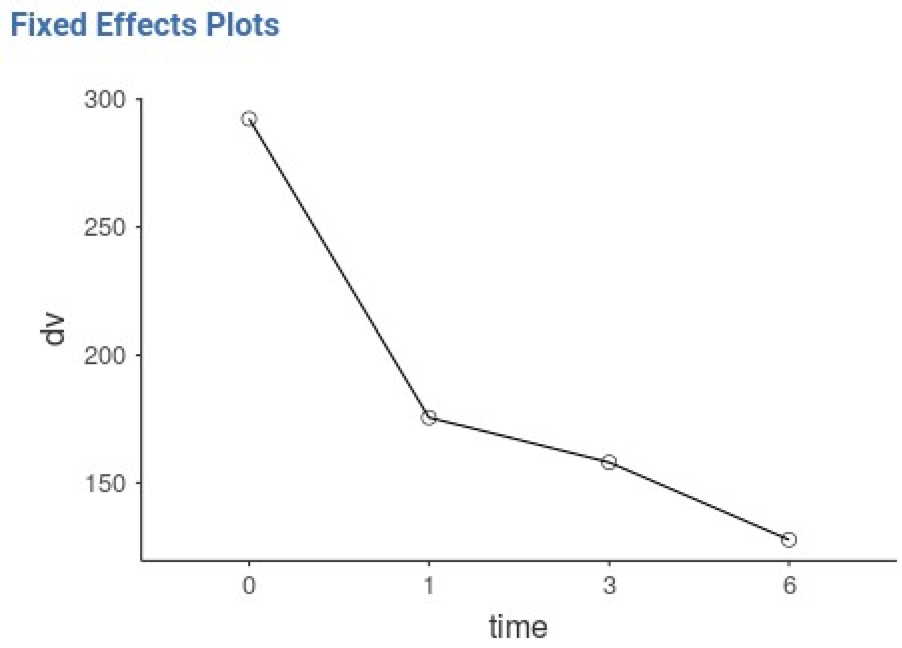
- Each trend (linear, quadratic, etc) tests a particular shape of the mean pattern




Cubic: on average means fluctuate

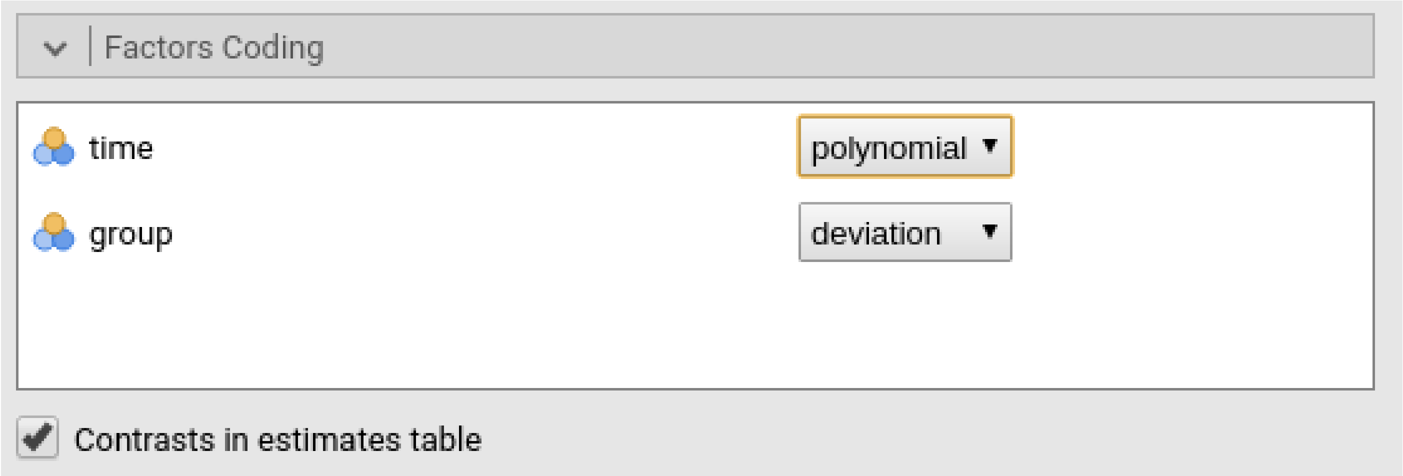
# Trend analysis

- Each significant trend justifying interpreting a particular characteristic of the mean pattern



# GAMLj:Trend analysis

- First, we should code the categorical variable “time” as a polynomial contrast




- We can leave “group” as deviation (default) which means “centered contrasts”

# GAMLj:Trend analysis

- Second, look at the parameter estimates

Contrast labels

Contrast average effects

Fixed Effects Parameter Estimates

| Effect | Contrast | Estimate | SE | 95% Confidence Interval Lower | 95% Confidence Interval Upper | df | t | p |
|---|---|---|---|---|---|---|---|---|
| (Intercept) | Intercept | 188.437 | 11.6 | 165.7 | 211.17 | 22.0 | 16.2444 | < .001 |
| time1 | linear | −114.356 | 10.7 | −135.4 | −93.34 | 65.9 | −10.6626 | < .001 |
| time2 | quadratic | 43.250 | 10.7 | 22.2 | 64.27 | 65.9 | 4.0326 | < .001 |
| time3 | cubic | −25.044 | 10.7 | −46.1 | −4.02 | 65.9 | −2.3351 | 0.023 |
| group1 | 2 - ( 1, 2 ) | −42.958 | 11.6 | −65.7 | −20.22 | 22.0 | −3.7033 | 0.001 |
| time1 ✻ group1 | linear ✻ 2 - ( 1, 2 ) | −0.894 | 10.7 | −21.9 | 20.13 | 65.9 | −0.0834 | 0.934 |
| time2 ✻ group1 | quadratic ✻ 2 - ( 1, 2 ) | 52.875 | 10.7 | 31.9 | 73.90 | 65.9 | 4.9301 | < .001 |
| time3 ✻ group1 | cubic ✻ 2 - ( 1, 2 ) | −17.721 | 10.7 | −38.7 | 3.30 | 65.9 | −1.6523 | 0.103 |

Contrast interaction with group

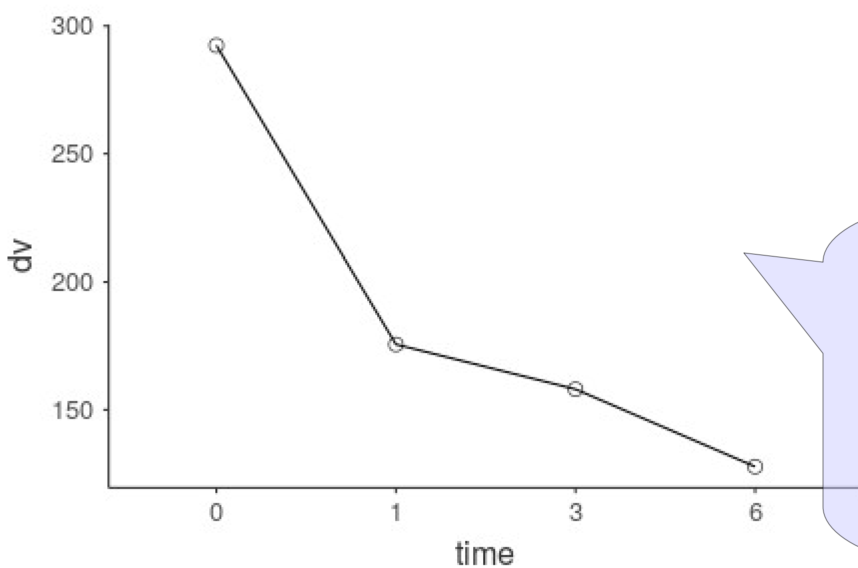
# GAMLj:Trend analysis

- Average effects of the contrasts

Fixed Effects Parameter Estimates

| Effect | Contrast | Estimate | SE | 95% Confidence Interval Lower | 95% Confidence Interval Upper | df | t | p |
|---|---|---|---|---|---|---|---|---|
| (Intercept) | Intercept | 188.437 | 11.6 | 165.7 | 211.17 | 22.0 | 16.2444 | < .001 |
| time1 | linear | -114.356 | 10.7 | -135.4 | -93.34 | 65.9 | -10.6626 | < .001 |
| time2 | quadratic | 43.250 | 10.7 | 22.2 | 64.27 | 65.9 | 4.0326 | < .001 |
| time3 | cubic | -25.044 | 10.7 | -46.1 | -4.02 | 65.9 | -2.3351 | 0.023 |
| | | | | 5.7 | -20.22 | 22.0 | -3.7033 | 0.001 |
| | | | | 1.9 | 20.13 | 65.9 | -0.0834 | 0.934 |
| | | | | 1.9 | 73.90 | 65.9 | 4.9301 | < .001 |
| | | | | 8.7 | 3.30 | 65.9 | -1.6523 | 0.103 |

**Fixed Effects Plots**




The pattern (on average) shows all three trends:

1. it goes down (linear)
2. it tend to go down and then up
3. if fluctuates a bit

# GAMLj:Trend analysis

- Trend analysis by group

Fixed Effects Parameter Estimates

| Effect | Contrast | Estimate | SE | 95% Confidence Interval Lower | 95% Confidence Interval Upper | df | t | p |
|---|---|---|---|---|---|---|---|---|
| (Intercept) | Intercept | 188.437 | 11.6 | 165.7 | 211.17 | 22.0 | 16.2444 | < .001 |
| time1 | linear | −114.356 | 10.7 | −135.4 | −93.34 | 65.9 | −10.6626 | < .001 |
| time2 | quadratic | 43.250 | 10.7 | 22.2 | 64.27 | 65.9 | 4.0326 | < .001 |
| time3 | cubic | −25.044 | 10.7 | −46.1 | −4.02 | 65.9 | −2.3351 | 0.023 |
| group1 | 2 - ( 1, 2 ) | −42.958 | 11.6 | −65.7 | −20.22 | 22.0 | −3.7033 | 0.001 |
| time1 ✻ group1 | linear ✻ 2 - ( 1, 2 ) | −0.894 | 10.7 | −21.9 | 20.13 | 65.9 | −0.0834 | 0.934 |
| time2 ✻ group1 | quadratic ✻ 2 - ( 1, 2 ) | 52.875 | 10.7 | 31.9 | 73.90 | 65.9 | 4.9301 | < .001 |
| time3 ✻ group1 | cubic ✻ 2 - ( 1, 2 ) | −17.721 | 10.7 | −38.7 | 3.30 | 65.9 | −1.6523 | 0.103 |

Those tell us if the trend is different between the two groups:
Linear: no
Quadratic: yes
Cubib: no

Both groups decreases
Group 2 curve is stronger
They both fluctuates a bit

# GAMLj:Trend analysis

- Trend analysis by group

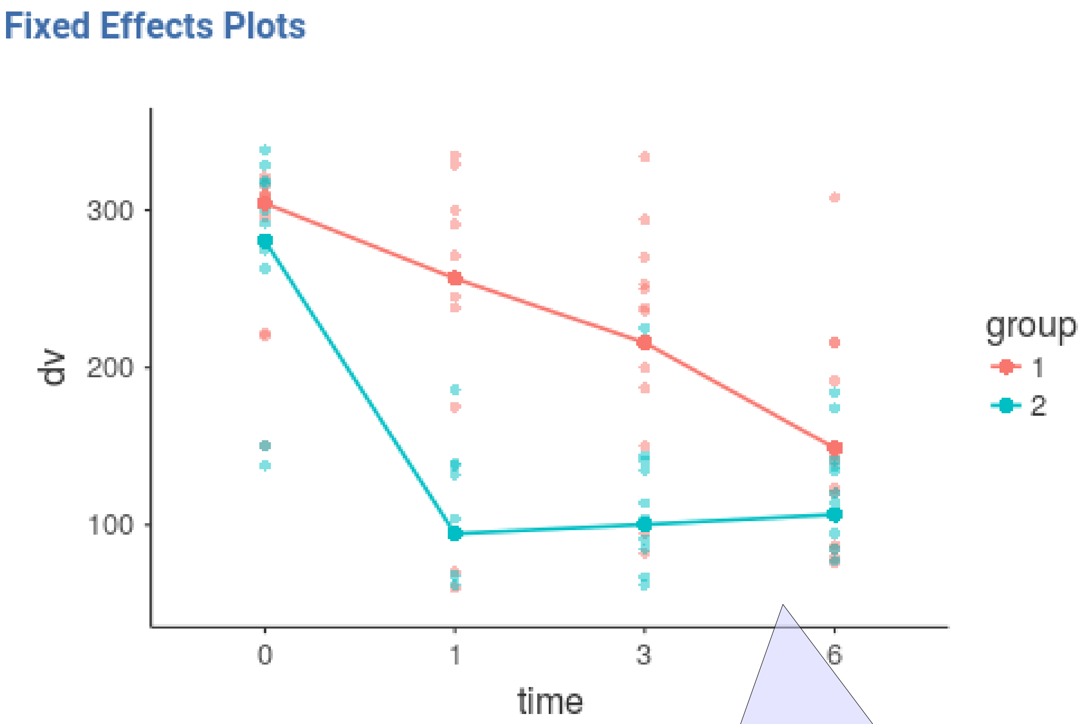
Fixed Effects Parameter Estimates

| Effect | Contrast | Estimate | SE | 95% Confidence Interval | | df | t | p |
|---|---|---|---|---|---|---|---|---|
| | | | | Lower | Upper | | | |
| (Intercept) | Intercept | 188.437 | 11.6 | 165.7 | 211.17 | 22.0 | 16.2444 | < .001 |
| time1 | linear | -114.356 | 10.7 | -135.4 | -93.34 | 65.9 | -10.6626 | < .001 |
| time2 | quadratic | 43.250 | 10.7 | 22.2 | 64.27 | 65.9 | 4.0326 | < .001 |
| time3 | cubic | -25.044 | 10.7 | -46.1 | -4.02 | 65.9 | -2.3351 | 0.023 |
| group1 | 2 - ( 1, 2 ) | -42.958 | 11.6 | -65.7 | -20.22 | 22.0 | -3.7033 | 0.001 |
| time1 ✻ group1 | linear ✻ 2 - ( 1, 2 ) | -0.894 | 10.7 | -21.9 | 20.13 | 65.9 | -0.0834 | 0.934 |
| time2 ✻ group1 | quadratic ✻ 2 - ( 1, 2 ) | 52.875 | 10.7 | 31.9 | 73.90 | 65.9 | 4.9301 | < .001 |
| time3 ✻ group1 | cubic ✻ 2 - ( 1, 2 ) | -17.721 | 10.7 | -38.7 | 3.30 | 65.9 | -1.6523 | 0.103 |

Those tell us if the trend is
different between the two groups:
Linear: no
Quadratic: yes
Cubic: mild

Both groups decreases
One group has a stronger curve
They both fluctuates a bit

# GAMLj:Trend analysis




Those tell us if the trend is different between the two groups:
Linear: no
Quadratic: yes
Cubic: mild

Both groups decreases
Group 2 curve is stronger
The fluctuation is similar

# GAMLj:Trend analysis

- Simple effects trend analysis: We can now look at the parameters of the simple effects analysis

Time1: linear
Time2: quadratic
Time3: cubic

**Simple Effects Parameters**

Simple effects of time

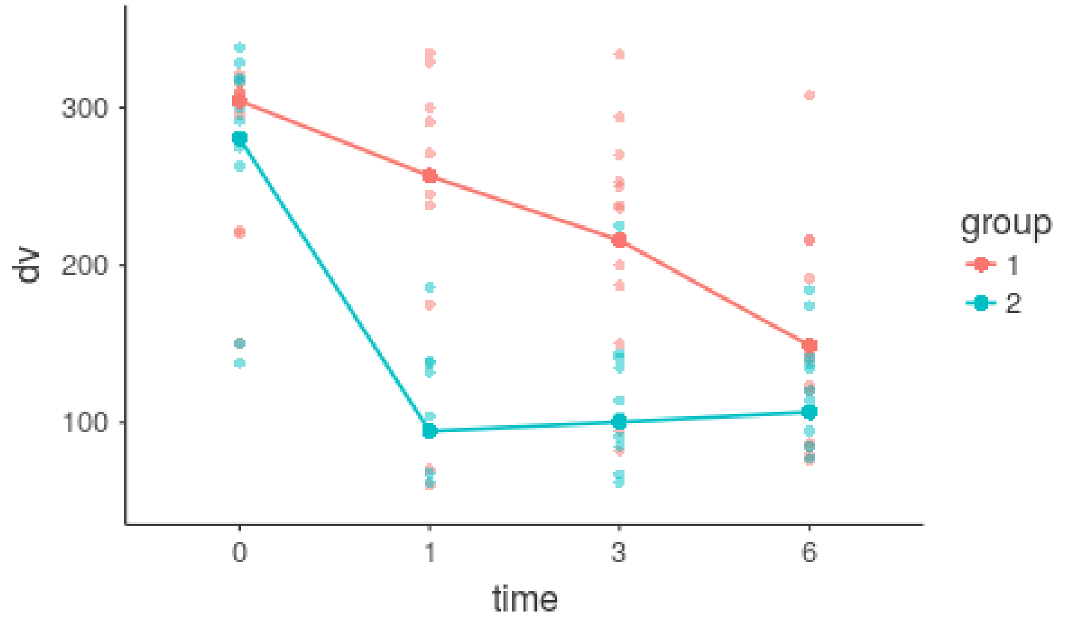
| Effect | Moderator Level | Estimate | SE | t | p |
|---|---|---|---|---|---|
| time1 | group at 1 | -113.46 | 15.2 | -7.481 | < .001 |
| time2 | group at 1 | -9.63 | 15.2 | -0.635 | 0.528 |
| time3 | group at 1 | -7.32 | 15.2 | -0.483 | 0.631 |
| time1 | group at 2 | -115.25 | 15.2 | -7.599 | < .001 |
| time2 | group at 2 | 96.13 | 15.2 | 6.338 | < .001 |
| time3 | group at 2 | -42.76 | 15.2 | -2.820 | 0.006 |

- In group 1 there's only a linear trend
- In group 2 all three trend are there

# GAMLj:Trend analysis

- Simple effects trend analysis: We can now interpret the parameters of the simple effects analysis

Fixed Effects Plots



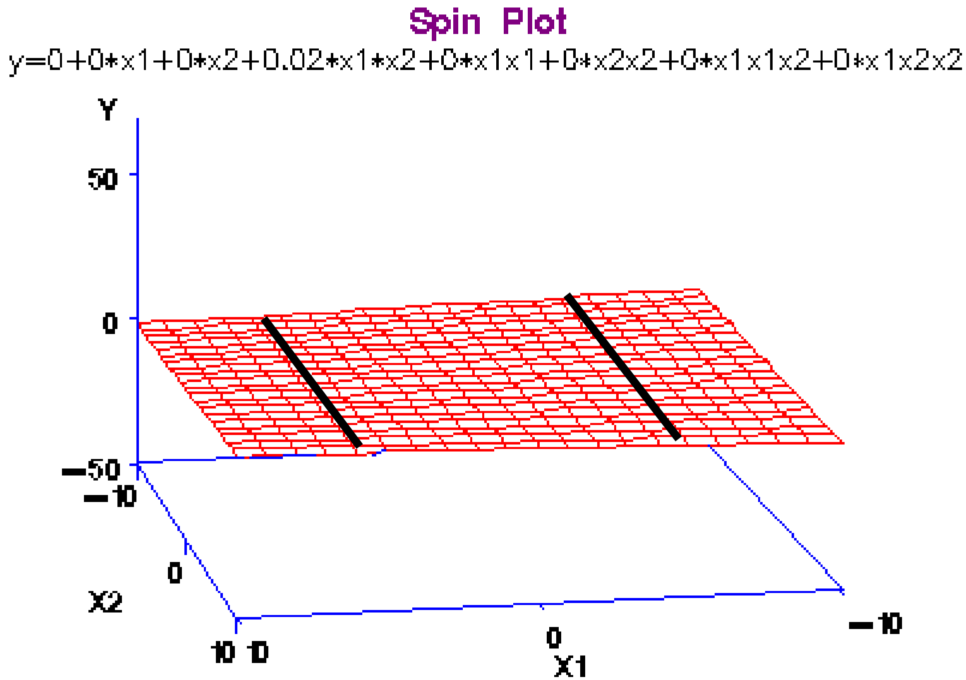

- In group 1 there's only a linear trend
- In group 2 all three trends are there

# Interactions between continuous variables

# Two continuous variables

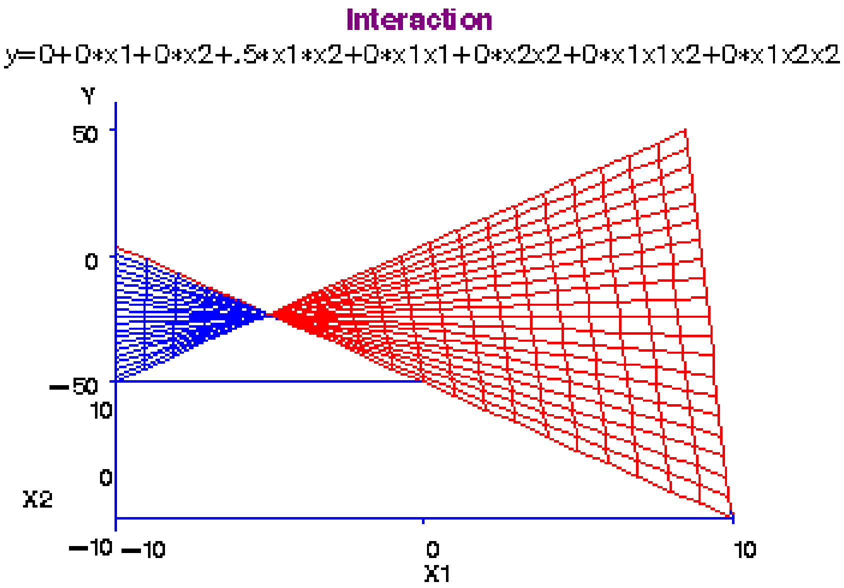
- In the multiple regression we have seen, lines are parallels, making a flat surface

- The effect of one IV is constant (the same) for each level of the other IV



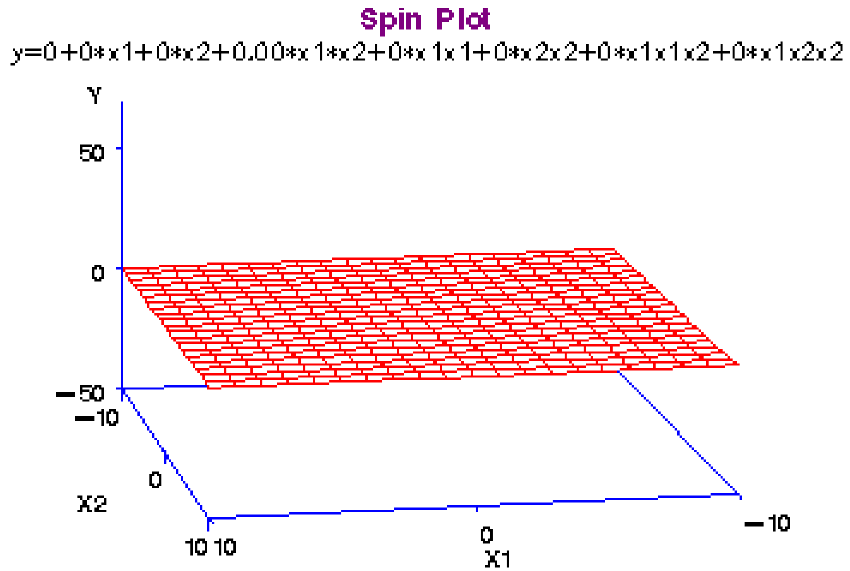
# Interactions lines

- Interaction: Lines are **not** parallel
- The effect of one IV is different for each level of the other IV



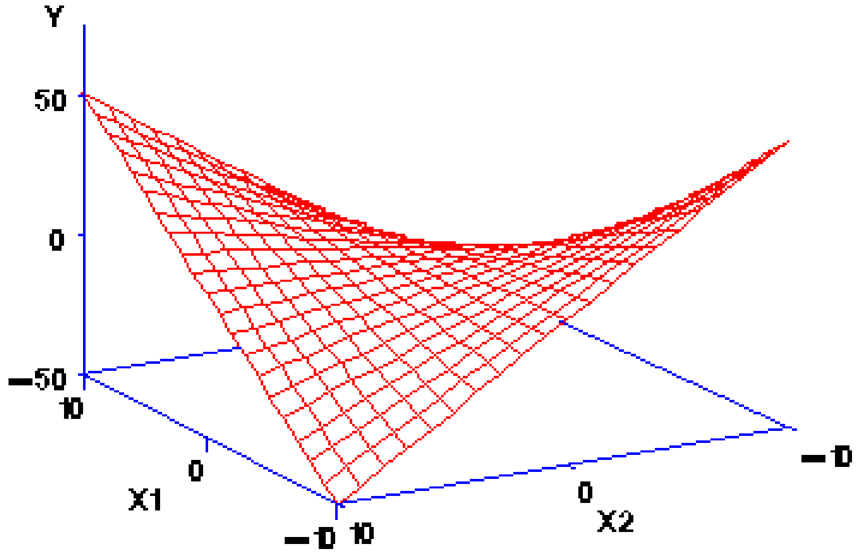
# Interactions line

- The bigger the interaction, the less parallel the lines: Bigger difference in the slopes



# Interactions line

- The bigger the interaction, the less parallel the lines: Bigger difference in the slopes



# Multiplicative effect

- The interaction effect is captured in the regression by a multiplicative term

The product of the two independent variables

$$\hat{y}_i = a + b_1 \cdot x_1 + b_2 \cdot x_2 + b_{\text{int}} x_1 x_2$$

The coefficient of $x_1$ is changing as $x_2$ changes

$$\hat{y}_i = a + (b_1 + b_{\text{int}} x_2) \cdot x_1 + b_2 \cdot x_2$$

The effect of one IV changes at different levels of the other IV

# Conditional effect

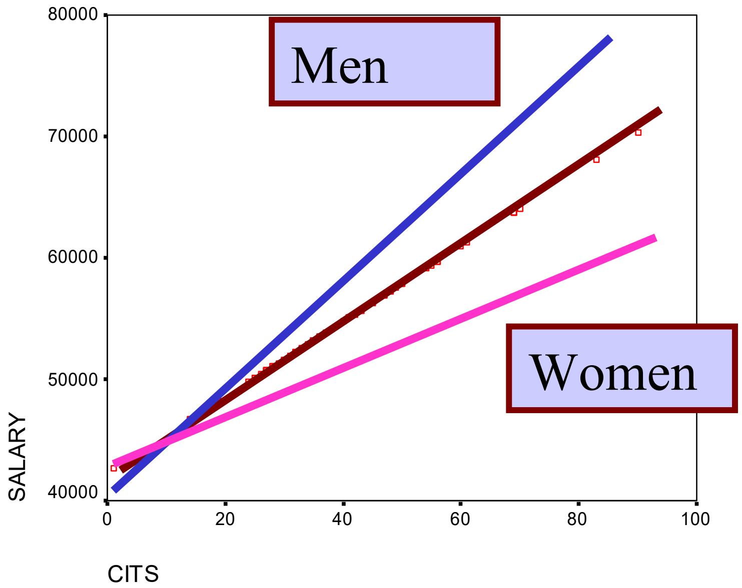
- We say that the effect of one IV is conditional to the level of the other IV

For Women (0) the slope is different

$$\hat{y}_i = a + (b_2 + b_{int} 0) \cdot x_2 + b_1 \cdot 0$$

…than for Men (1)

$$\hat{y}_i = a + (b_2 + b_{int} 1) \cdot x_2 + b_1 \cdot 1$$



# Conditional vs linear effect

- A linear effect (when no interaction is present) tells you how much change there is in the DV when you change the IV

$$\hat{y}_i = a + b_1 \cdot x_1 + b_2 \cdot x_2 + b_{int} x_1 x_2$$

Change in the DV

- An interaction effect (the B of the product term) tells you how much change there is **in the effect** of one IV on the DV when you change the other IV

$$\hat{y}_i = a + \left( b_1 + b_{int} x_2 \right) \cdot x_1 + b_2 \cdot x_2$$

Change in the effect

Change in the DV

# Terminology

- When there is an interaction term in the equation, one refers to the linear effect (the ones that are not interactions) as the first-order effect

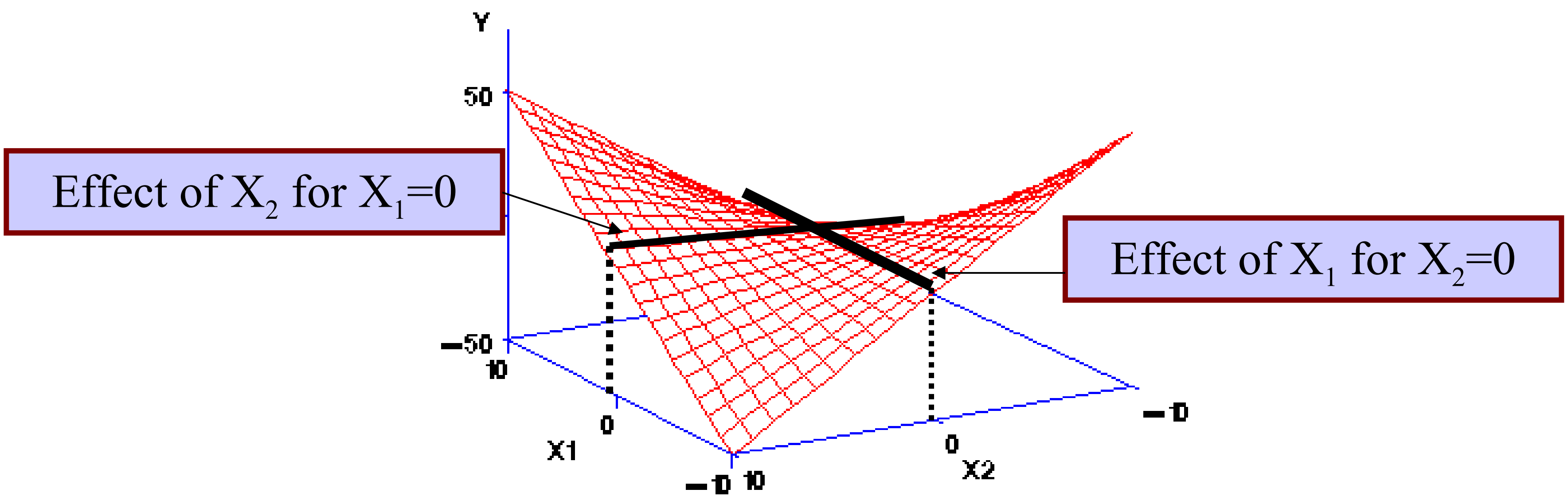
$$\hat{y}_i = a + b_1 \cdot x_1 + b_2 \cdot x_2 + b_{\text{int}} x_1 x_2$$

First order effects

# First-order effects with interaction

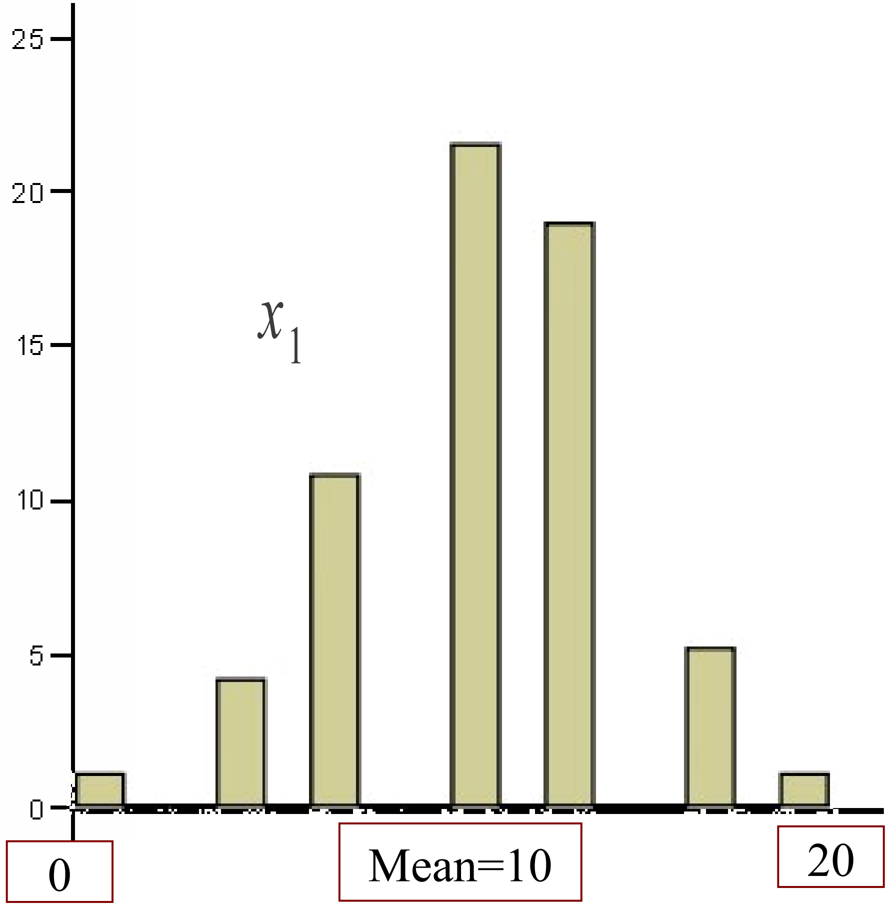
- When the interaction is in the regression, the first order effects become the effect of the IV while keeping the other IV's constant to zero

$$\hat{y}_i = a + b_1 \cdot x_1 + b_2 \cdot 0 + b_{\text{int}} x_1 0 = a + b_1 \cdot x_1$$



# Making zero meaningful

- We can always make zero a meaningful value by centering the variables before computing the product term:



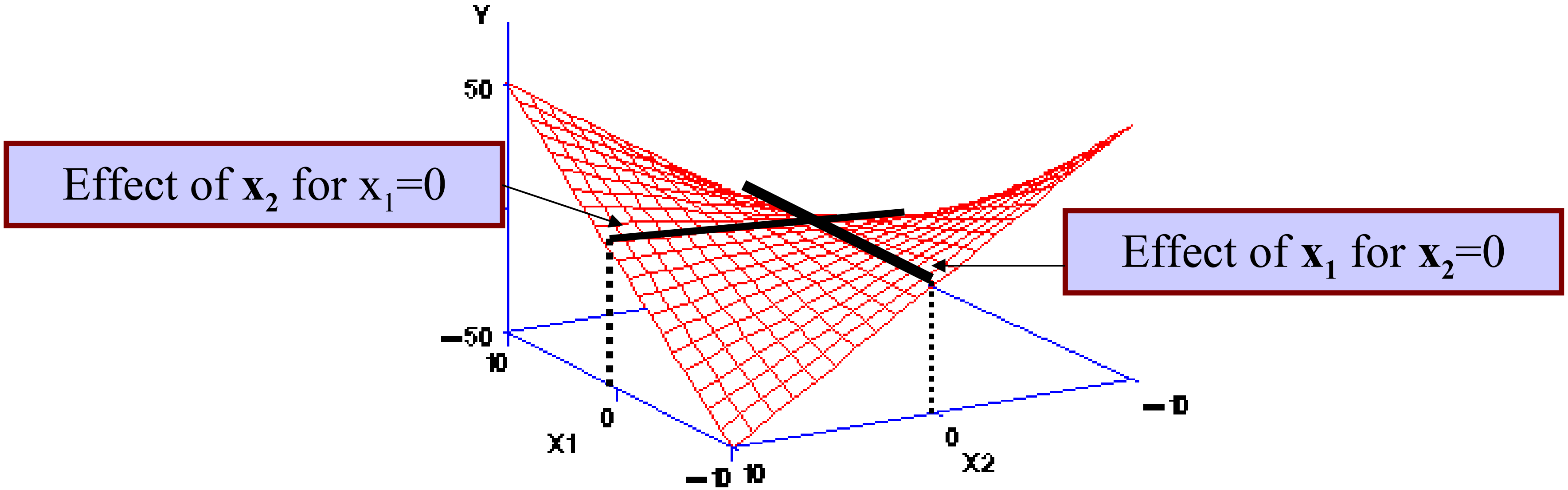

For each participant, compute a new variable as the old minus the average

$$c = x_1 - \bar{x}_1$$

The new variable has mean=0

# Centering

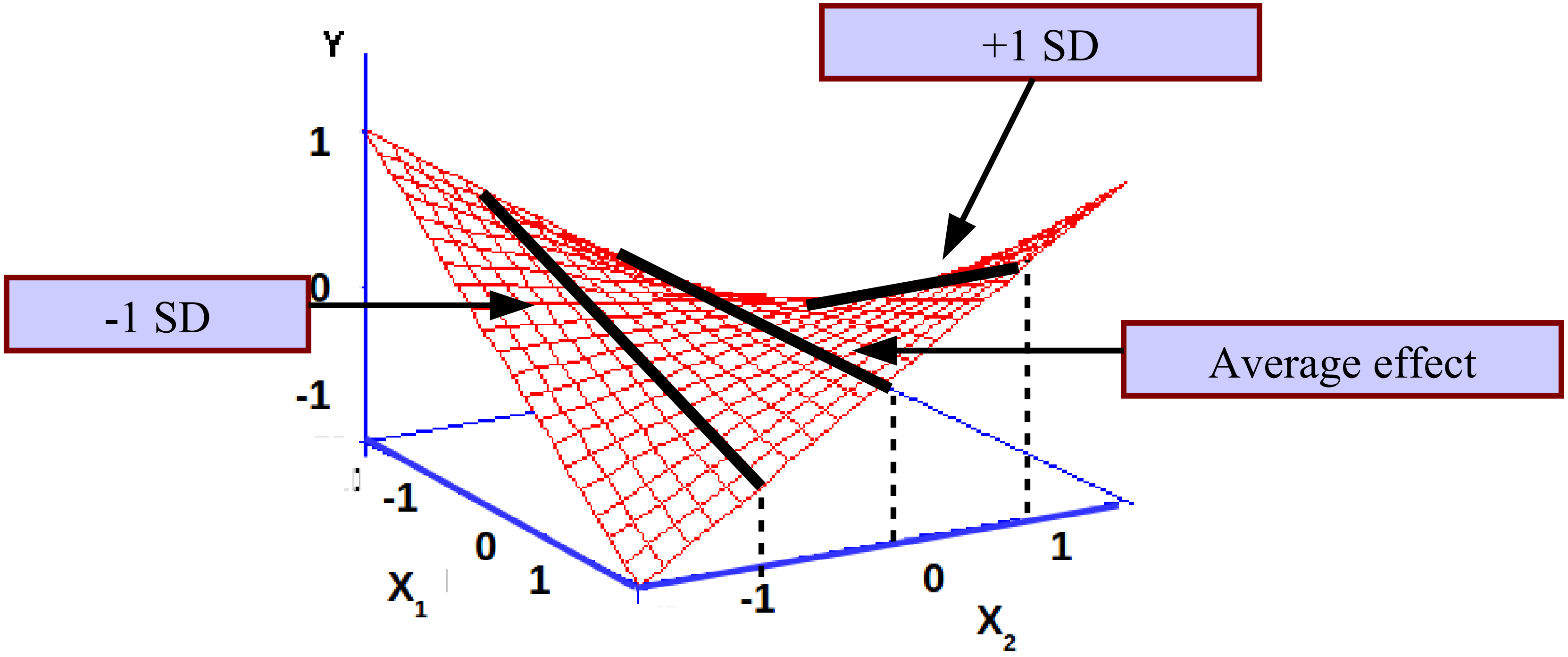
- The first-order effects computed on centered variables represent the average effect (the one in the middle) of the IV, across all levels of the other IV



# Simple slope analysis

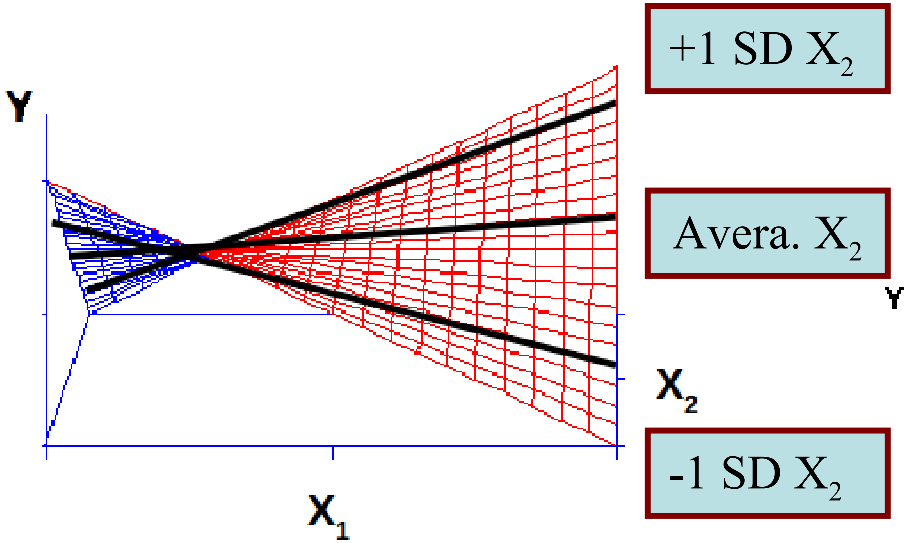
- We can study the interaction by evaluating the effect of one independent variables for low (-1 SD), average (Mean), and high (+1 SD) levels of the moderator

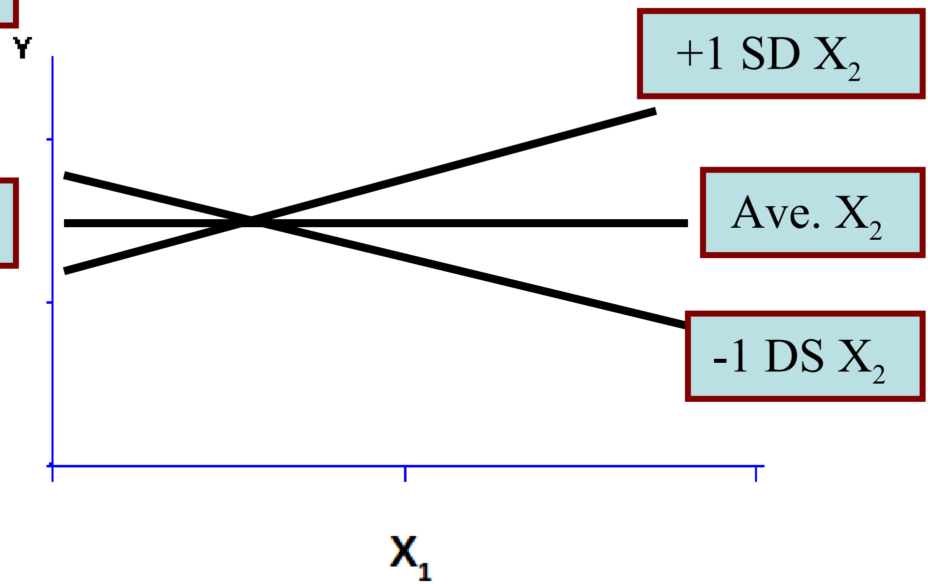
- We pick three lines out of many in the regression plane, and plot them



# Simple slope analysis

- We represent them in two dimensions



# Example

- 50 different school classes were assessed on students reading ability and self-efficacy. In each class, the teacher was assessed as well for her/his self-efficacy.

**1182 subjects**
**50 school clasess**

### Frequencies

Frequencies of schoolclass

| Levels | Counts | % of Total | Cumulative % |
|---|---|---|---|
| 1 | 24 | 2.0 % | 2.0 % |
| 2 | 23 | 1.9 % | 4.0 % |
| 3 | 24 | 2.0 % | 6.0 % |
| 4 | 24 | 2.0 % | 8.0 % |
| 5 | 23 | 1.9 % | 10.0 % |
| 6 | 25 | 2.1 % | 12.1 % |
| 7 | 22 | 1.9 % | 14.0 % |
| 8 | 25 | 2.1 % | 16.1 % |
| 9 | 23 | 1.9 % | 18.0 % |
| 10 | 24 | 2.0 % | 20.1 % |
| 11 | 24 | 2.0 % | 22.1 % |
| 12 | 23 | 1.9 % | 24.0 % |
| 13 | 23 | 1.9 % | 26.0 % |
| 14 | 23 | 1.9 % | 27.9 % |
| 15 | 24 | 2.0 % | 29.9 % |
| 16 | 24 | 2.0 % | 32.0 % |
| 17 | 25 | 2.1 % | 34.1 % |
| 18 | 24 | 2.0 % | 36.1 % |
| 19 | 25 | 2.1 % | 38.2 % |
| 20 | 24 | 2.0 % | 40.3 % |
| 21 | 24 | 2.0 % | 42.3 % |
| 22 | 23 | 1.9 % | 44.2 % |
| 23 | 23 | 1.9 % | 46.2 % |
| 24 | 23 | 1.9 % | 48.1 % |
| 25 | 25 | 2.1 % | 50.3 % |

# Example

- Efficacy and reading ability varies from participant to participant, whereas teaching efficacy varies from class to class, but not within each class

| | efficacy | read | teacheffic | schoolcla... |
|---|---|---|---|---|
| 17 | 57 | 8 | 22 | 1 |
| 18 | 41 | 13 | 22 | 1 |
| 19 | 43 | 16 | 22 | 1 |
| 20 | 47 | 17 | 22 | 1 |
| 21 | 58 | 13 | 22 | 1 |
| 22 | 41 | 8 | 22 | 1 |
| 23 | 42 | 14 | 22 | 1 |
| 24 | 30 | 11 | 22 | 1 |
| 25 | 54 | 16 | 21 | 2 |
| 26 | 44 | 11 | 21 | 2 |
| 27 | 75 | 17 | 21 | 2 |
| 28 | 48 | 12 | 21 | 2 |
| 29 | 50 | 11 | 21 | 2 |
| 30 | 61 | 20 | 21 | 2 |
| 31 | 31 | 15 | 21 | 2 |
| 32 | 46 | 9 | 21 | 2 |
| 33 | 61 | 20 | 21 | 2 |

# Example

- We wish to estimate the effect of reading ability to participants self-efficacy, the effect of teacher efficacy and the interaction between reading ability and teacher efficacy

$$\hat{SE} = a + b_1 REA + b_2 TE + b_3 TE \cdot REA$$

- We want to use a mixed model to take into the account the school class clustering effect

# Mixed model

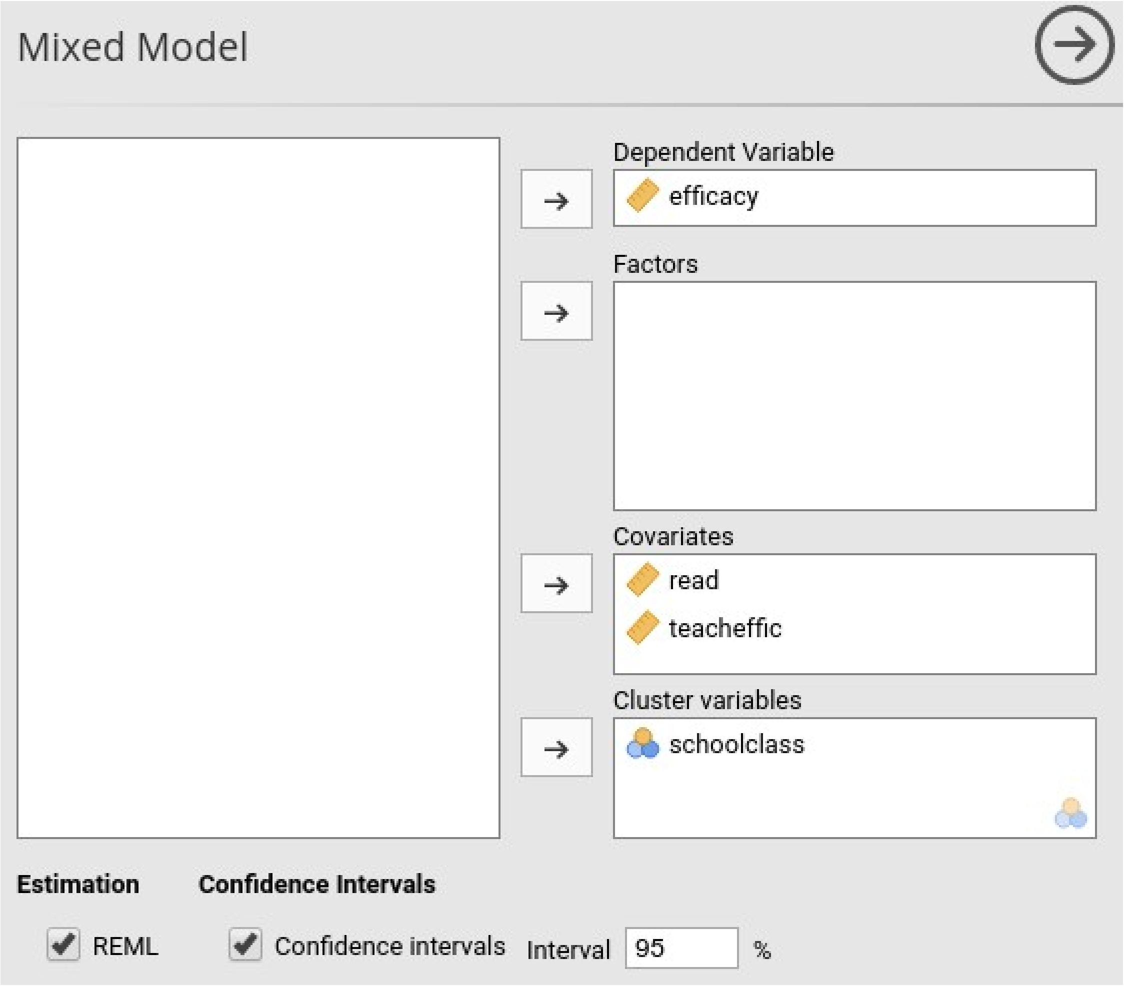
We can translate this in a standard mixed model

$$\hat{SE} = a + b_1 \, REA + b_2 \, TE + b_3 \, TE \cdot REA$$

- Fixed effects? Intercept and read ,teacher, and interaction effect
- Random effects? Intercepts read effect
- Clusters? School class

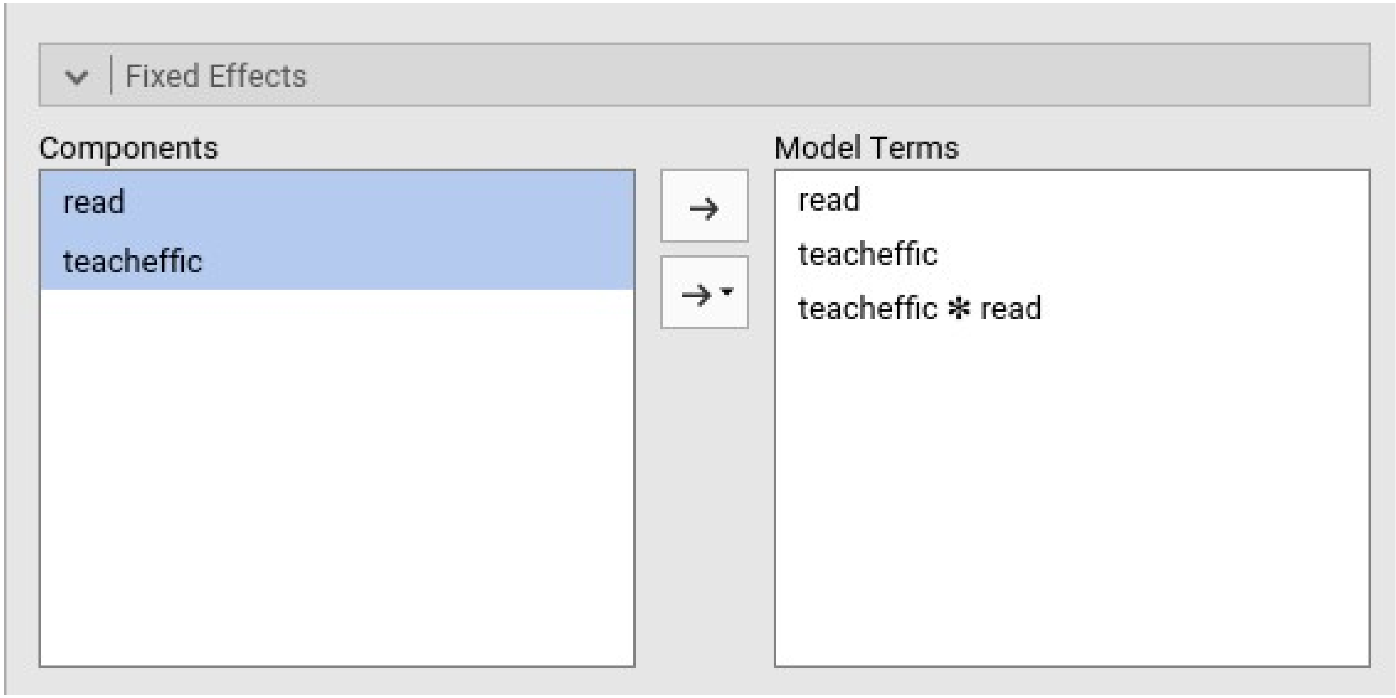
# Example jamovi

- First we define the variables in the model and their role



# Example: fixed effects

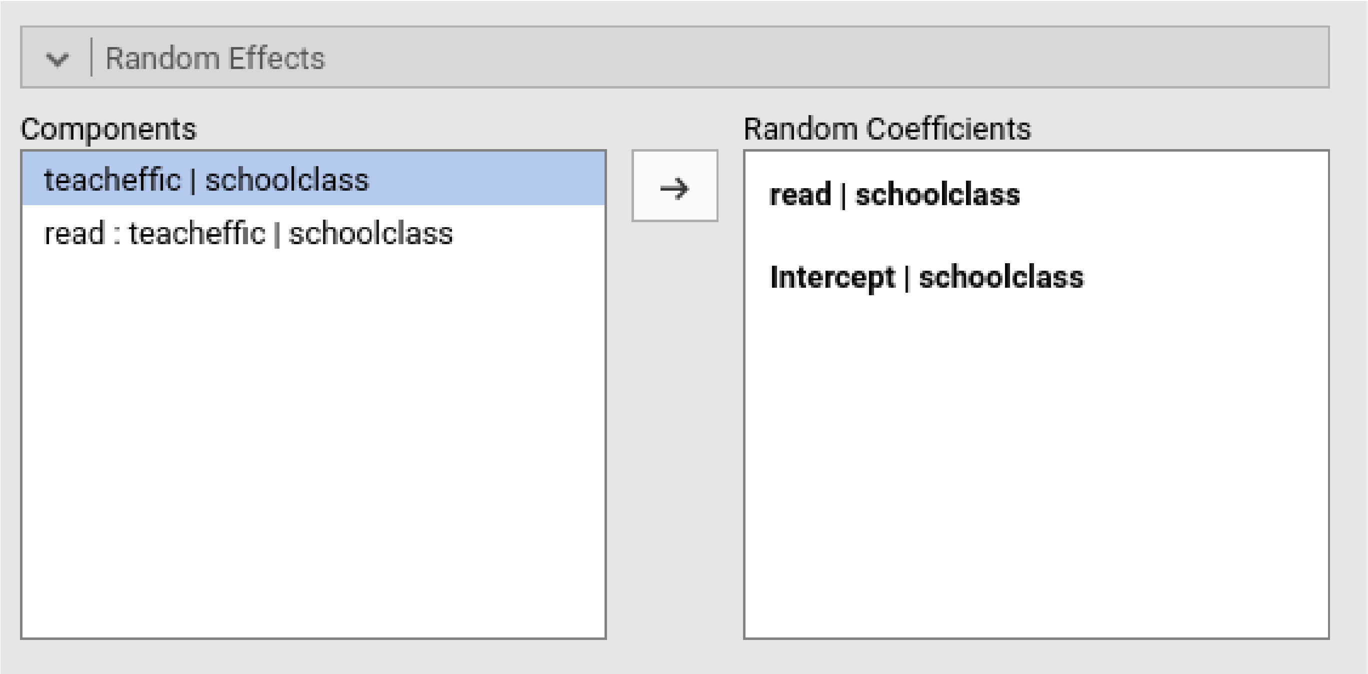
- We define the fixed effects in the model




Main effects and interactions

# Example: random effects

- We define the random effects in the model




Main  that can be computed within each school class

# Results: model recap

- R-squared measures

Model Info

| | Info |
|---|---|
| Estimate | Linear mixed model fit by REML |
| Call | efficacy ~ 1 + (read + 1 \| schoolclass) + read + teacheffic + teacheffic:read |
| AIC | 8788.131 |
| R-squared Marginal | 0.119 |
| R-squared Conditional | 0.551 |

# Results: model fixed effects

- F-tests and p-values: we interpret them as any regression with interaction

Fixed Effect ANOVA

| | F | Num df | Den df | p |
|---|---|---|---|---|
| read | 23.84 | 1 | 47.0 | < .001 |
| teacheffic | 1.73 | 1 | 47.9 | 0.194 |
| read:teacheffic | 9.05 | 1 | 45.5 | 0.004 |

*Note.* Satterthwaite method for degrees of freedom

# Results: model fixed effects

- B coefficients and p-values: To interpret the linear effects we should know the meaning zero of the independent variable: jamovi **centers** the independent variable by default

B coefficients

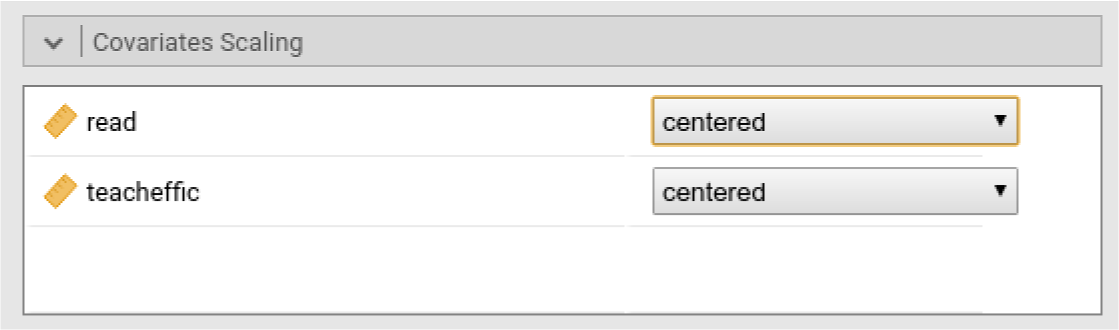
Fixed Effects Parameter Estimates

| Effect | Contrast | Estimate | SE | 95% Confidence Interval Lower | 95% Confidence Interval Upper | df | t | p |
|---|---|---|---|---|---|---|---|---|
| (Intercept) | Intercept | 49.8077 | 1.0620 | 47.7262 | 51.889 | 48.1 | 46.90 | < .001 |
| read | read | 0.8262 | 0.1692 | 0.4945 | 1.158 | 47.0 | 4.88 | < .001 |
| teacheffic | teacheffic | 0.2239 | 0.1701 | -0.1095 | 0.557 | 47.9 | 1.32 | 0.194 |
| read ✻ teacheffic | read ✻ teacheffic | 0.0809 | 0.0269 | 0.0282 | 0.134 | 45.5 | 3.01 | 0.004 |

Linear effects are average effects

# Centering IV

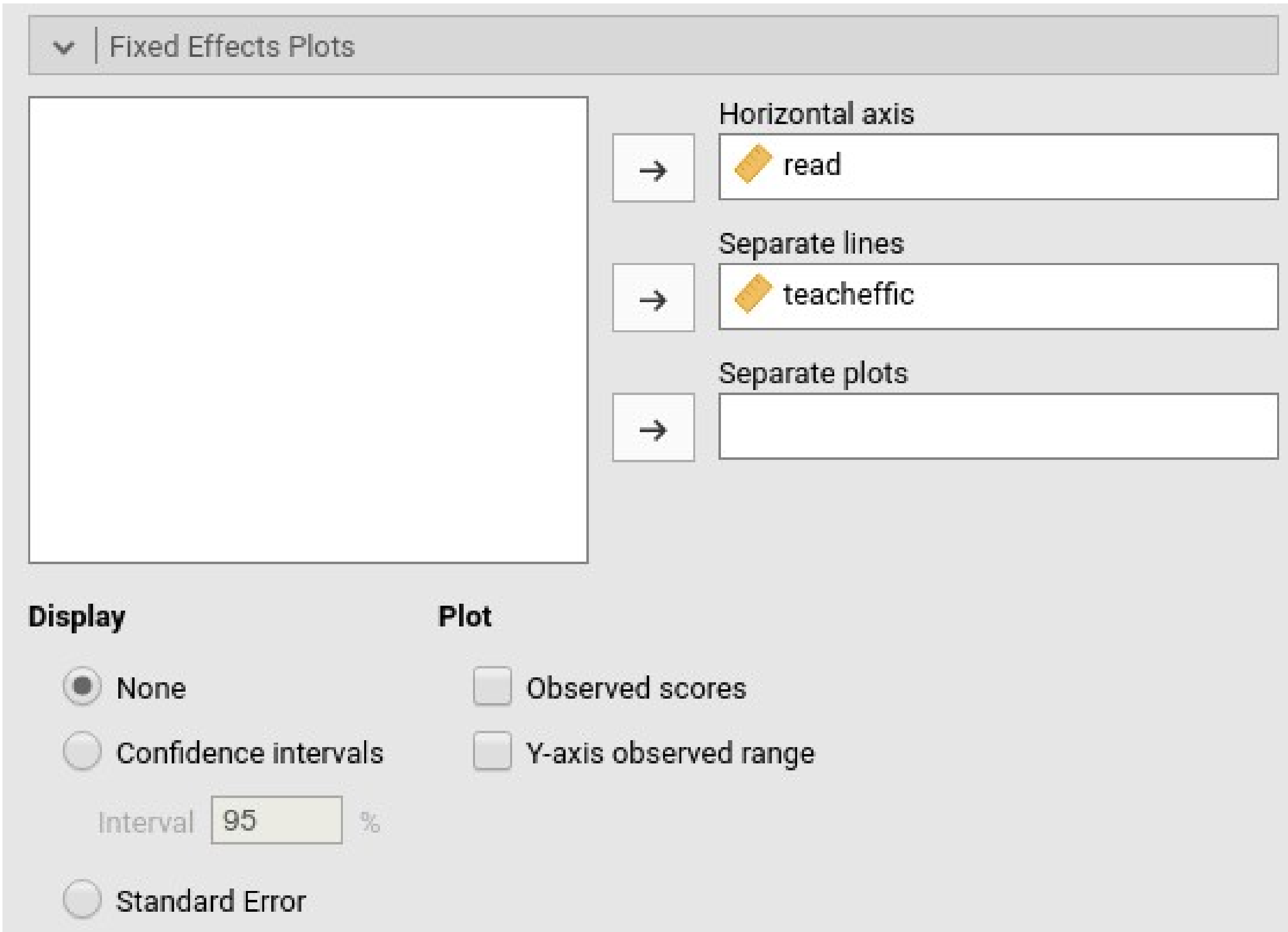
- Jamovi by default centers the IVs to their means, but different options are available




- Centered: centered using total sample mean
- Cluster-based centered: centered using each cluster mean
- Standardized: using mean and standard deviation of the total sample
- Cluster-based Standardized: using means and standard deviations of each cluster

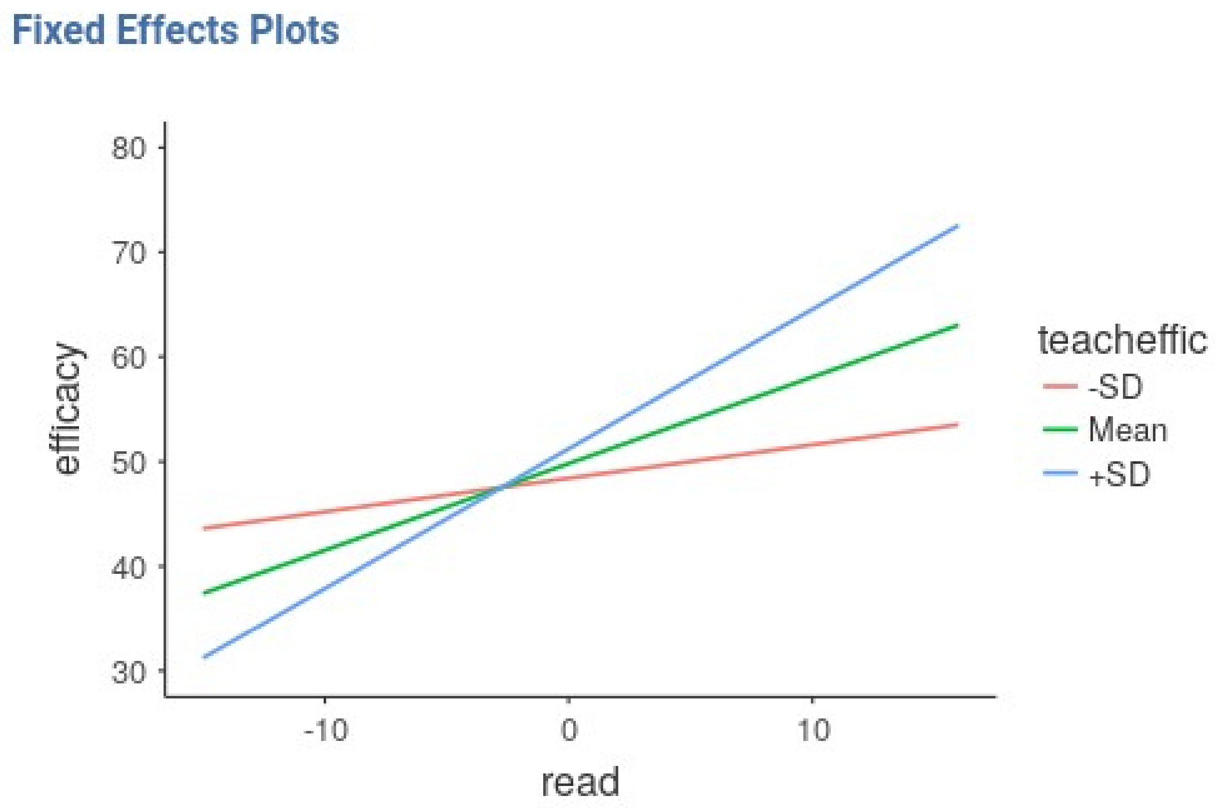
# Simple slope analysis

- Estimating the effect of one independent variable (read) at different levels of the moderator (teachefficacy) and **make a plot**



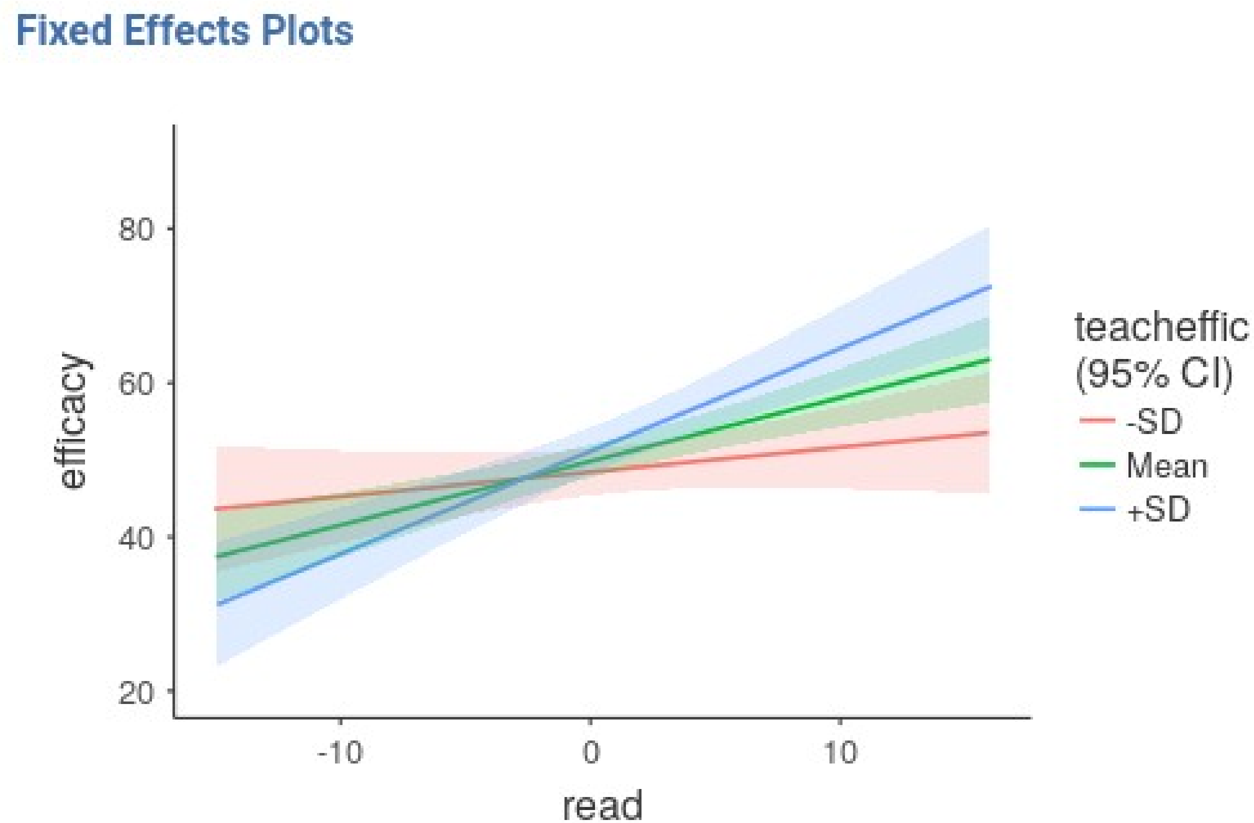
# Simple slope analysis

- Estimating the effect of one independent variable (read) at different levels of the moderator (teachefficacy) and **make a plot**



# Simple slope analysis

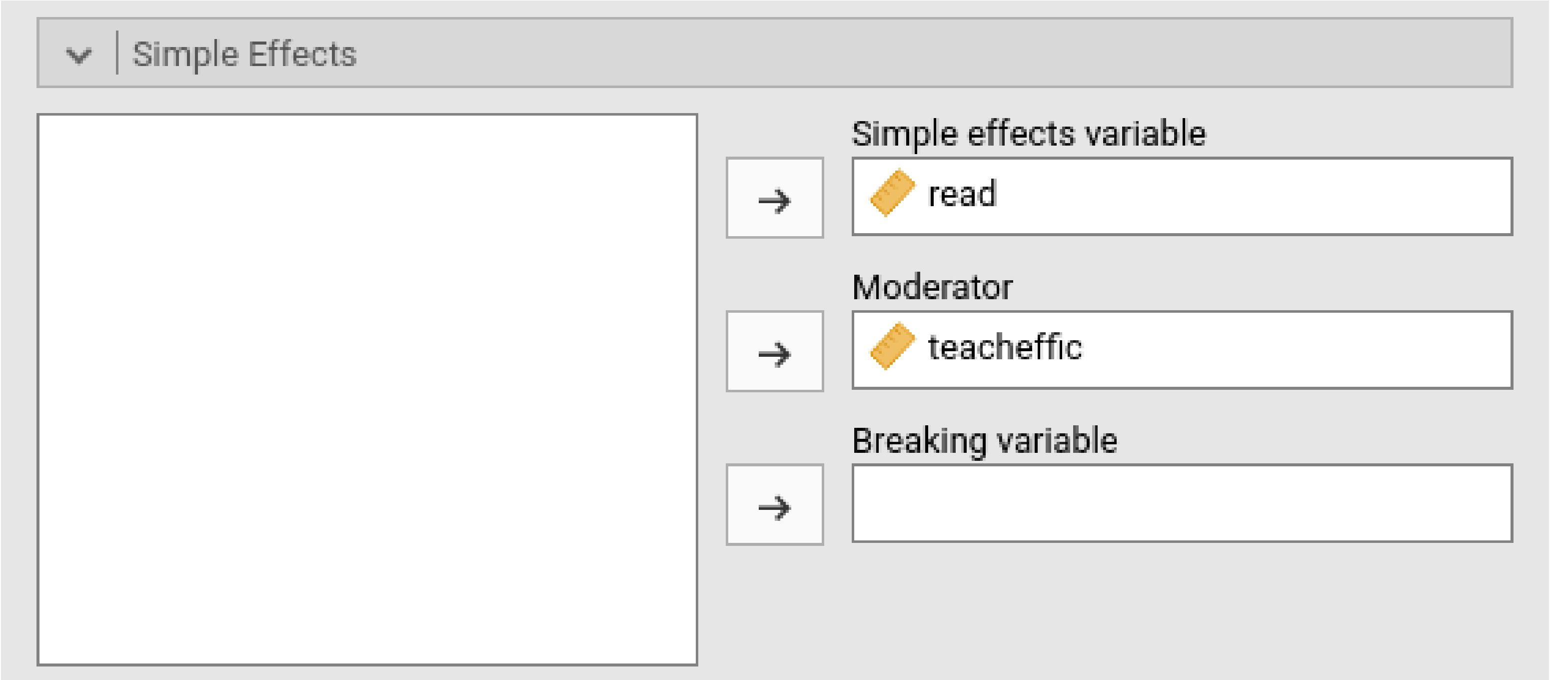
- One can add confidence bands: confidence intervals for continuous predicted values




- At the moment, the moderator is set to +1SD, mean, -1SD. More options will be added in the future

# Simple slope analysis

- Estimating the effect of one independent variable (read) at different levels of the moderator (teachefficacy) and **test the effects**



# Simple slope analysis

- Estimating the effect of one independent variable (read) at different levels of the moderator (teachefficacy) and **test the effects**

## Simple Effects ANOVA

Simple effects of read

| Effect | Moderator Levels | df Num | df Den | F | p |
|---|---|---|---|---|---|
| read | teacheffic at -6.26 | 1.00 | 46.7 | 1.78 | 0.188 |
| read | teacheffic at 0 | 1.00 | 47.9 | 23.83 | < .001 |
| read | teacheffic at 6.26 | 1.00 | 47.5 | 31.27 | < .001 |

- At the moment, the moderator is set to +1SD, mean, -1SD. More options will be added in the future

# Simple slope analysis

- Estimating the effect of one independent variable (read) at different levels of the moderator (teachefficacy) and **test the effects**

## Simple Effects Parameters

B coefficients

Simple effects of read

| Effect | Moderator Level | Estimate | SE | t | p |
|---|---|---|---|---|---|
| read | teacheffic at -6.26 | 0.320 | 0.239 | 1.34 | 0.188 |
| read | teacheffic at 0 | 0.826 | 0.169 | 4.88 | < .001 |
| read | teacheffic at 6.26 | 1.333 | 0.238 | 5.59 | < .001 |

- At the moment, the moderator is set to +1SD, mean, -1SD. More options will be added in the future

# Questions

- How many clusters, how many scores within cluster
- Convergences
- Multiple classifications
  - Subjects by items design

# The end