Testing for measurement bias across large numbers of groups

Universiteit Utrecht

Suzanne Jak

Overview

- What is measurement invariance?
- Testing invariance across many groups
- Testing invariance in 3-level data
- Illustration
- Discussion

Universiteit Utrecht

Measurement invariance

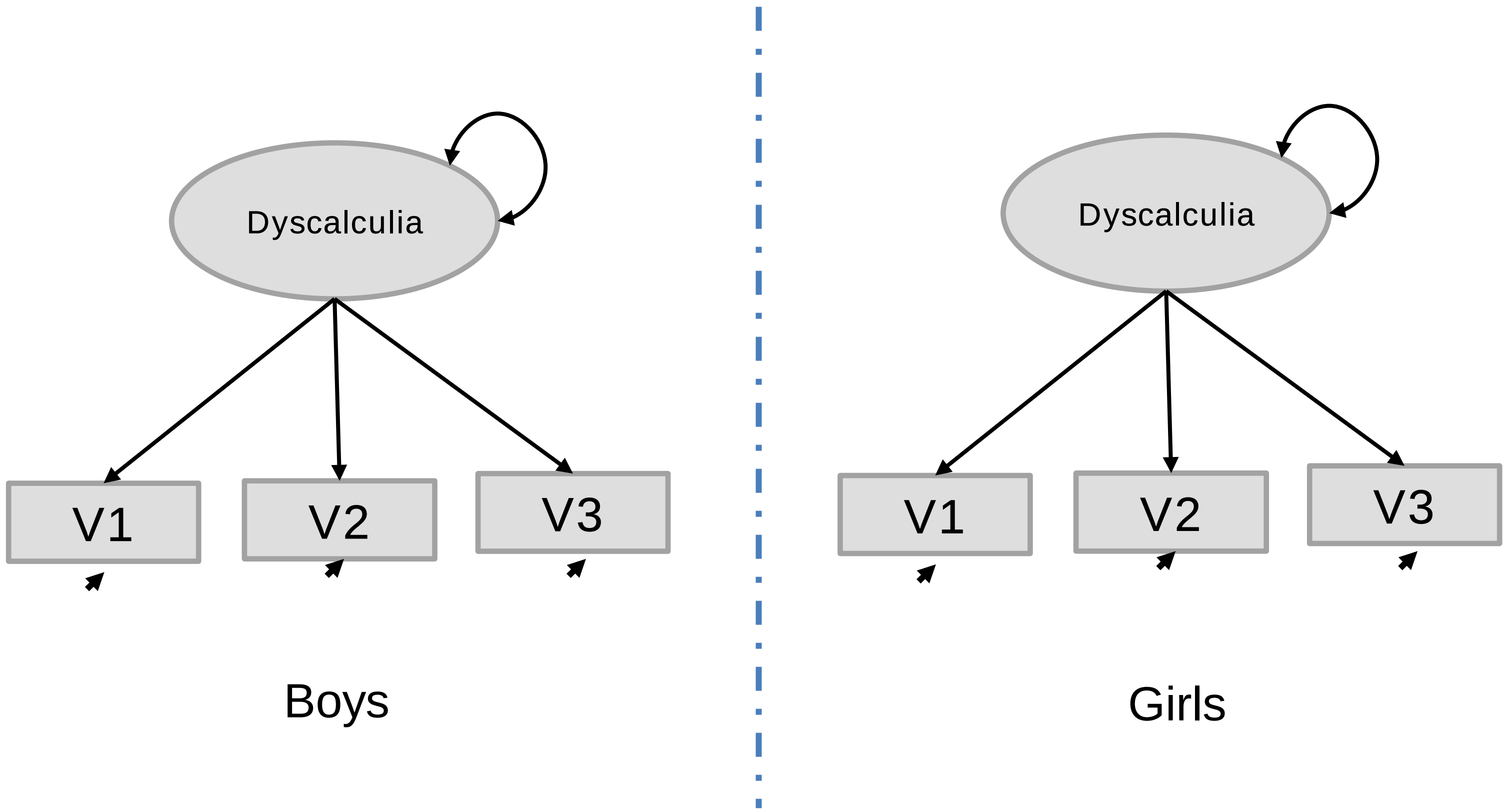
Measurement invariance

- Are observed differences attributable to differences in the trait that you wanted to measure?
- In context of factor analysis: Factorial invariance
- With a factor model as a measurement model, strong factorial invariance with respect to group implies equality of factor loadings and intercepts across groups (Meredith, 1994).
- Can be tested in a multigroup CFA model

Configural invariance

Equal factor structure across groups

$$\Sigma_g = \Lambda_g \Phi_g \Lambda_g' + \Theta_g$$
$$\mu_g = \nu_g + \Lambda_g \kappa_g$$

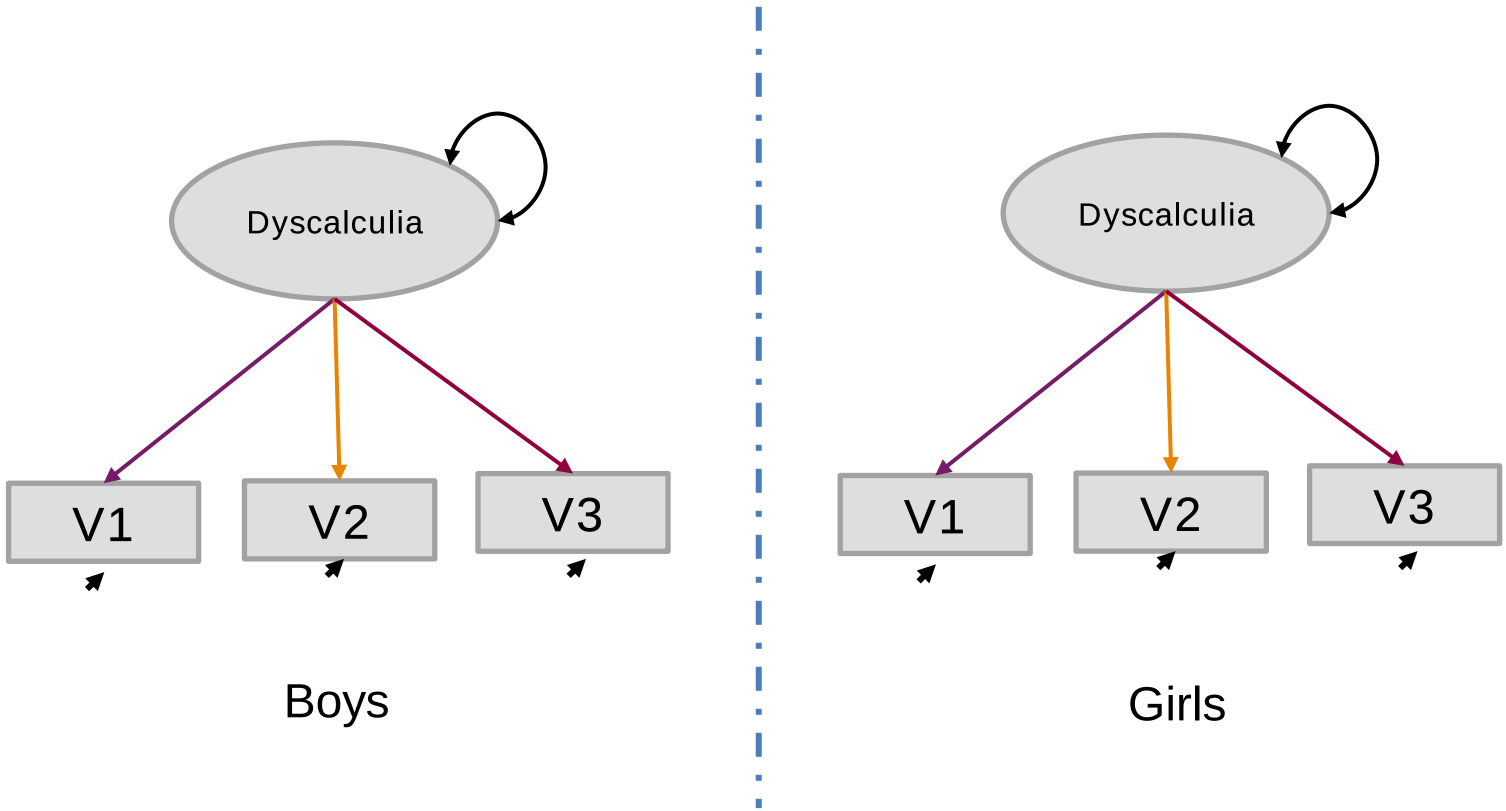


Weak factorial invariance

Equal factor loadings across groups

$$\Sigma_g = \Lambda \Phi_g \Lambda' + \Theta_g$$

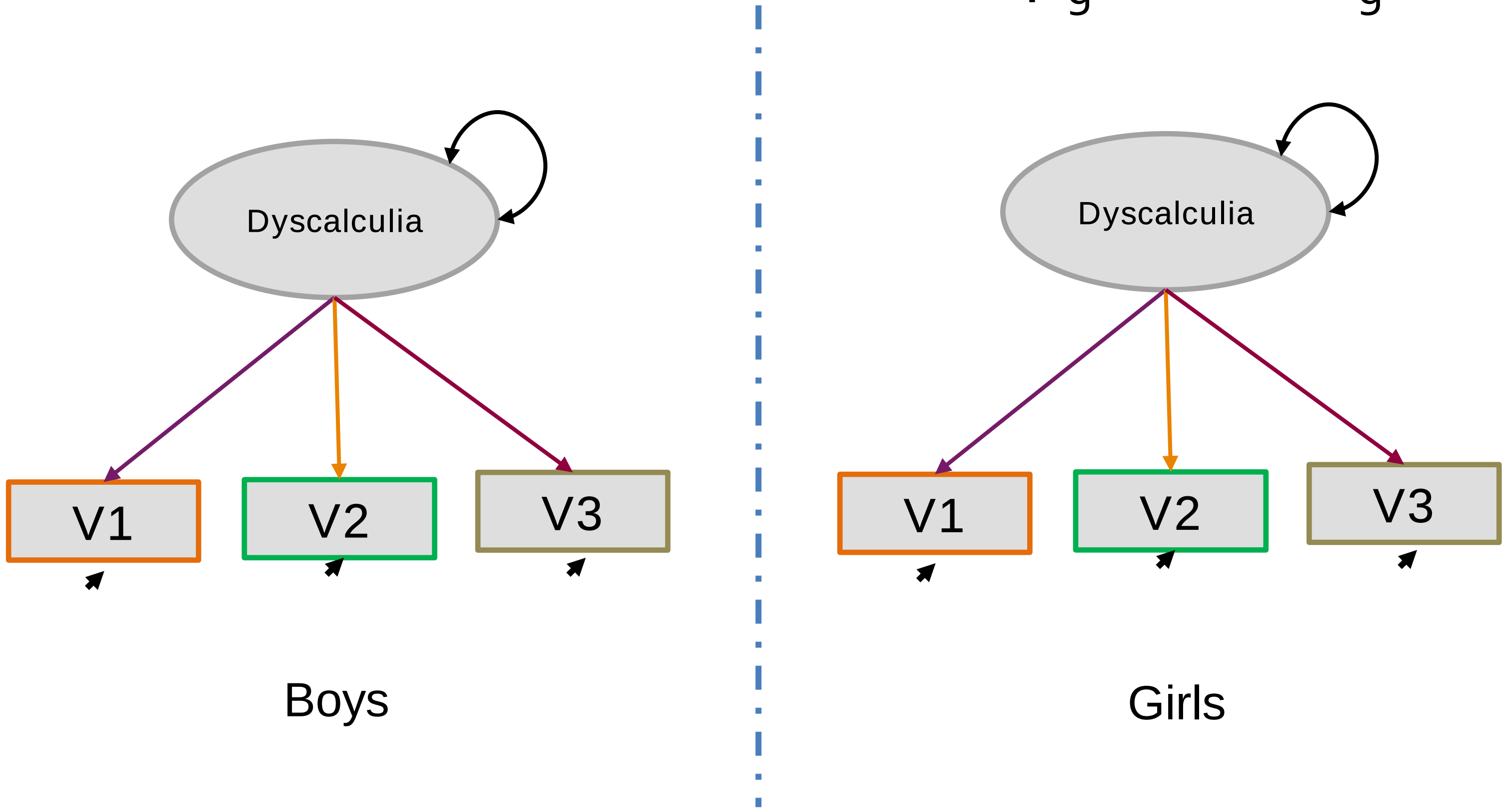
$$\mu_g = \nu_g + \Lambda \kappa_g$$



Strong factorial invariance

Equal factor loadings and intercepts across groups

$$\Sigma_g = \Lambda \Phi_g \Lambda' + \Theta_g$$
$$\mu_g = \nu + \Lambda \kappa_g$$



Strong factorial invariance across many groups

- Invariance across 156 school classes?

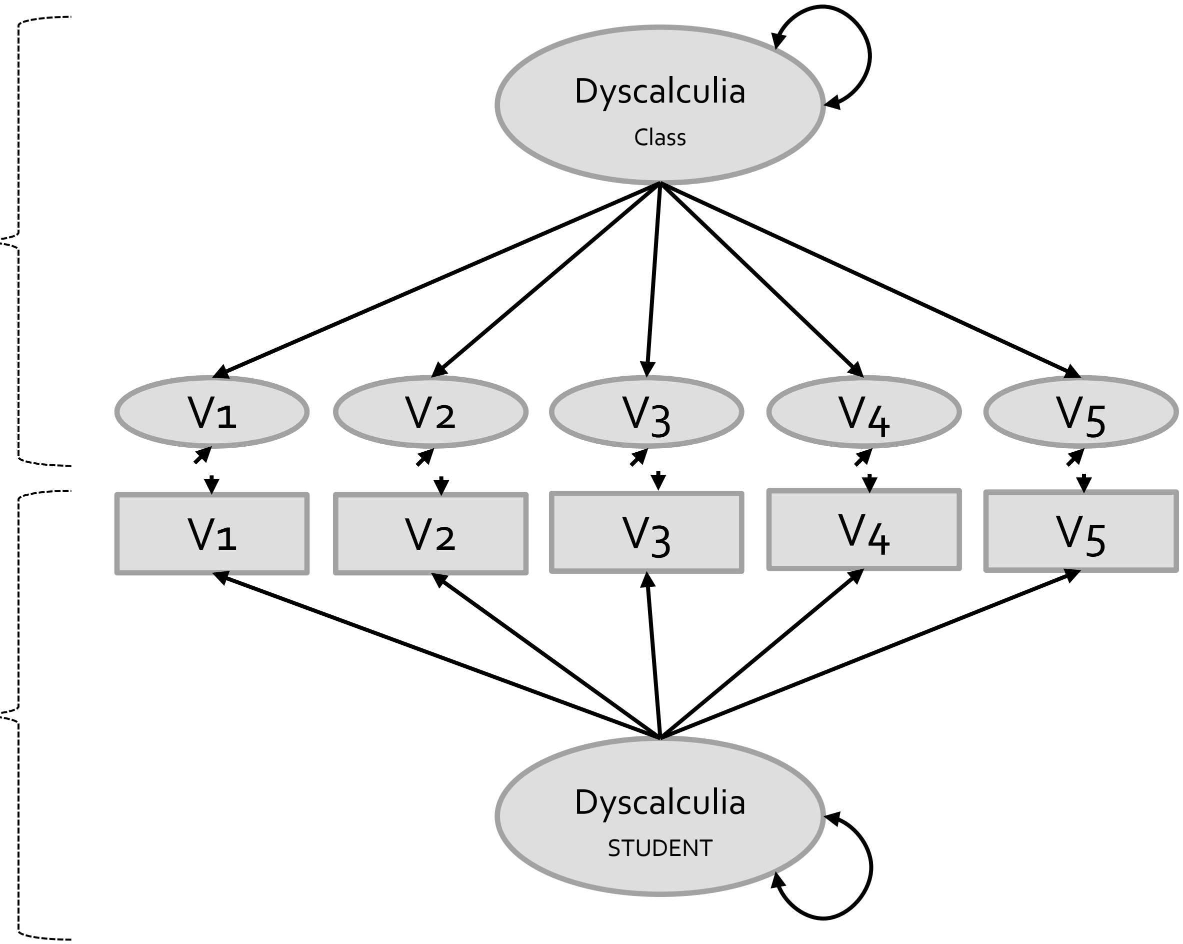
- Two-level model

- Test for cluster bias (Jak, Oort & Dolan, 2013)

Two-level SEM on Dyscalculia

- Level 2
- Between level
- Teacher level

- Level 1
- Within level
- Student level



Multilevel SEM

$\Sigma_{TOTAL} = \Sigma_{BETWEEN} + \Sigma_{WITHIN}$

Specify models for $\Sigma_{BETWEEN}$ and Σ_{WITHIN}

$$\Sigma_B = \Lambda_B \Phi_B \Lambda_B' + \Theta_B$$

$$\Sigma_W = \Lambda_W \Phi_W \Lambda_W' + \Theta_W$$

No cluster bias

- With equal factor loadings ($\Lambda_j = \Lambda$) and equal intercepts ($\tau_j = \tau$) for all clusters , the following model holds:

$$\Sigma_{WITHIN} = \Lambda\, \Phi_{WITHIN}\, \Lambda^t + \Theta_{WITHIN}$$

$$\Sigma_{BETWEEN} = \Lambda\, \Phi_{BETWEEN}\, \Lambda^t$$

Uniform cluster bias

- If there is uniform cluster bias ($\tau_j \neq \tau$), the following model holds:

$$\Sigma_{WITHIN} = \Lambda \, \Phi_{WITHIN} \, \Lambda^t + \Theta_{WITHIN}$$

$$\Sigma_{BETWEEN} = \Lambda \, \Phi_{BETWEEN} \, \Lambda^t + \Theta_{BETWEEN}$$

Test for cluster bias

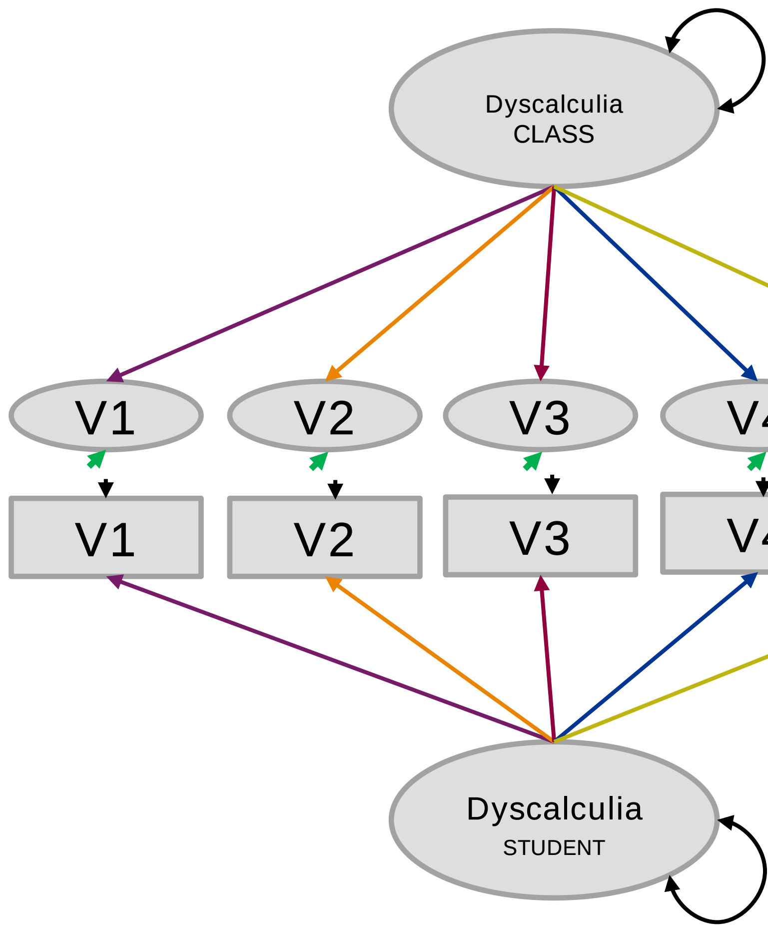
$\Sigma_{WITHIN} = \Lambda \, \Phi_W \, \Lambda' + \Theta_W$

$\Sigma_{BETWEEN} = \Lambda \, \Phi_B \, \Lambda'$

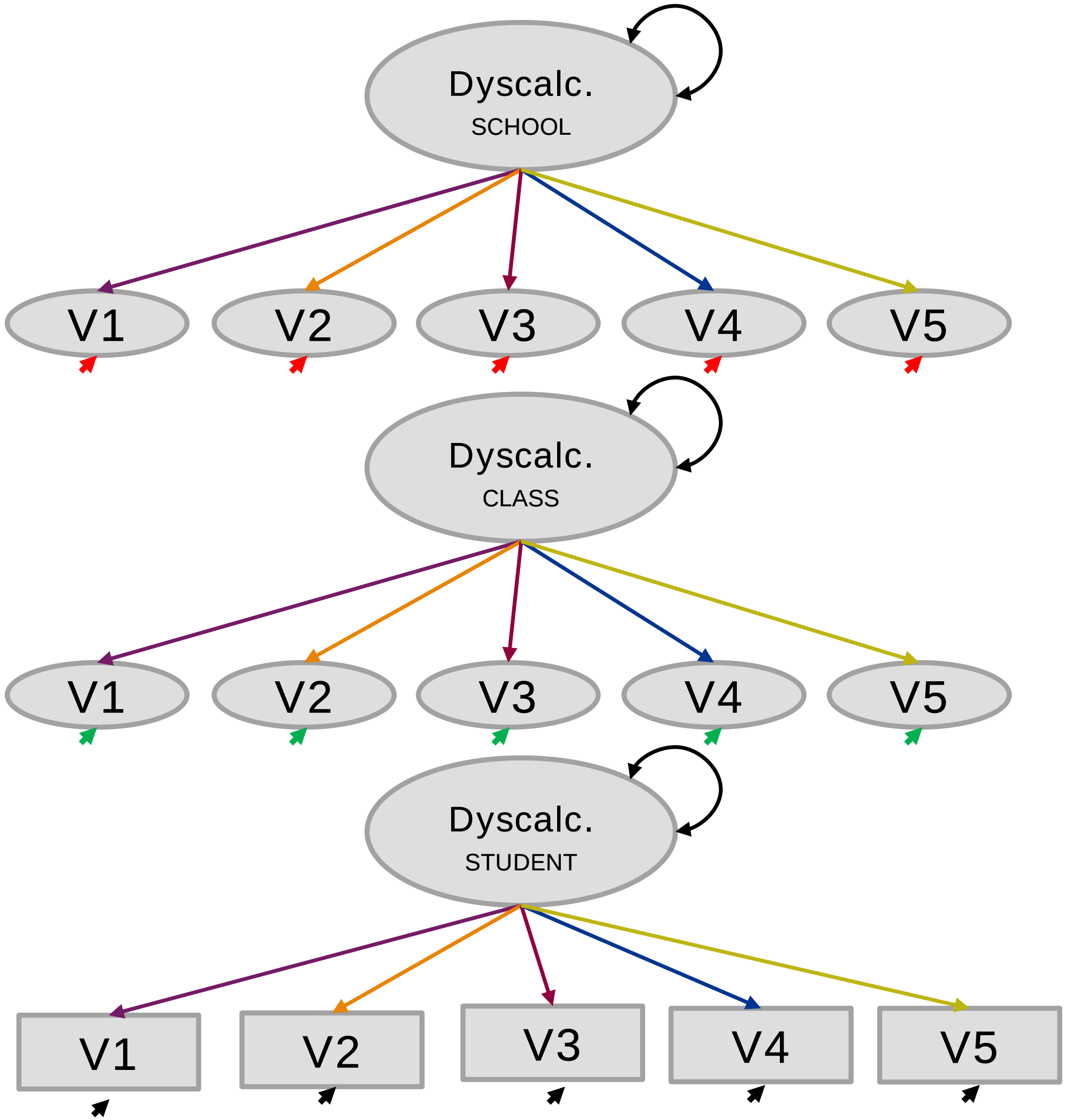
Set $\Lambda_{BETWEEN} = \Lambda_{WITHIN}$ and

whether $\Theta_{BETWEEN}$ equals zero.

Testing for cluster bias in dyscalculia scale



Strong factorial invariance in 3-level data



Universiteit Utrecht

Illustration

—

Illustration

- Dyscalculia
  - Learning difficulty specific to mathematics learning (Butterworth, 2008, Devine et al., 2013).
  - Deficits in understanding basic concepts such as quantity conservation and reversibility, despite otherwise typically developing mental abilities (Kosc, 1974; Gross-Tsur, Manor & Shalev, 1996).
  - Affect between 1.3 and 10 % of the population, which is equivalent to the prevalence of dyslexia

Instrument

- Dyscalculia (NDS, Nederlandse Dyscalculie Screener, Milikowski & Vermeire, 2013)

- 8 subtests

- E.g

Test 1 - Largest number. Cross-out the largest number of two numbers that are smaller than 10, for example:
2 or 1
6 or 8

Instrument

Test 4 - Number series. Fill in the number between two other numbers, for example:
20 _ 22
20 _ 18

Test 6 - Number line until 10. Which number should be on this position?
1___________-_______10

Test 7 - Fast subtraction. Subtract a number from a given number, for example:

Subtract 2 from:
8 6 3 9

Data

- Scores from a dyscalculia screening instrument
- 4527 students from 156 school classes from 50 schools
- Mean age 11.42 years (SD = 1.27)
- 49 % boys

Analysis

- M*plus* 7 , full information maximum likelihood estimation
- Step 1: Evaluate ICC’s and significance of variance at Level 2 and Level 3
- Step 2: Find a measurement model at Level 1 (saturated Level 2 and Level 3 model)
- Step 3: Test cluster bias at Level 2
- Step 4: Test cluster bias at Level 3

Results

- Step 1: Evaluate ICC's and significance of variance at Level 2 and Level 3
  - ICC at class level varies between 0.19 (Test 4) and 0.43 (Test 8)
  - ICC at school level varies between 0.004 (Test 5) and 0.02 (Test 8)
  - Significant variance at class level but not at the school level

Results

- Step 2: Find a measurement model at Level 1
  - Saturated Level 3 model was not possible
  - Alternative: Type = twolevel complex
  - Fit of a one-factormodel:
    - $\chi^2(20) = 304.51, p < .05, \mathrm{RMSEA} = .056, \mathrm{CFI} = .93$
      - Large modification index between Test 1 and Test 2

Results

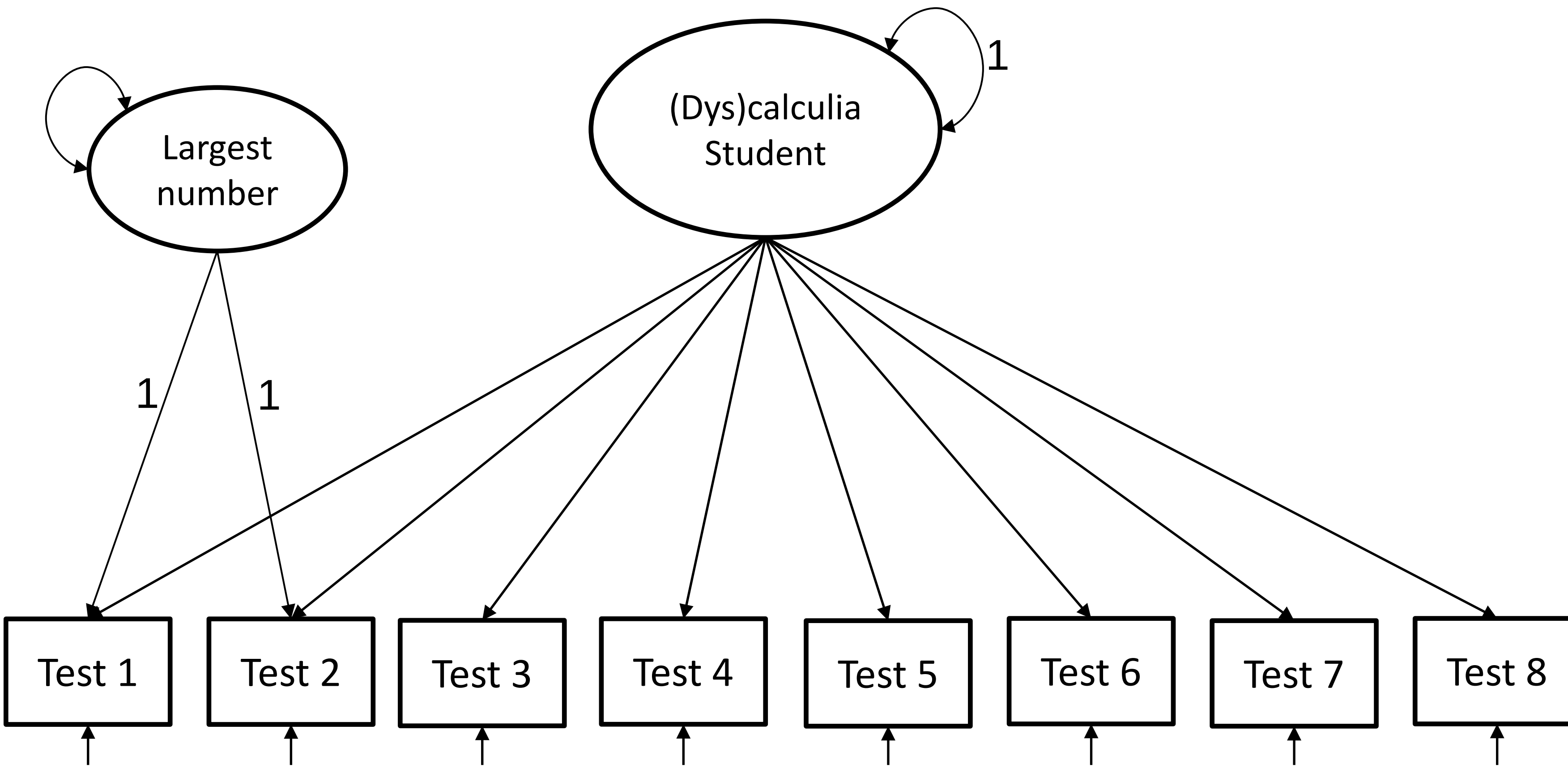
- Step 2: Find a measurement model at Level 1
  - Fit of a one-factor model with residual covariance between Test 1 and Test 2:

    $\chi^2(19) = 135.69$, $p < .05$, RMSEA = .037, CFI = .97

  - But, we cannot model residual covariances at Level 2 and Level 3 when testing for cluster bias.

Results

- Measurement model at Level 1



Results

- Step 3: Test cluster bias at Level 2

Model	df	χ^2	RMSEA	CFI
1. Baseline model (equal factor loadings)	71	731.95	.045	.96
2. Strong factorial invariance at Level 2	79	2821.77	.088	.84
Difference between model 1 and 2	8	2089.82		

Results

- Step 3: Test cluster bias at Level 3
  - Three-level model, equal factor loadings across levels

Model	df	χ^2	RMSEA	CFI
1. Baseline model (equal factor loadings)	71	731.95	.045	.96
3. Strong factorial invariance at Level 3	79	738.45	.043	.96
Difference between model 1 and 3	8	6.50		

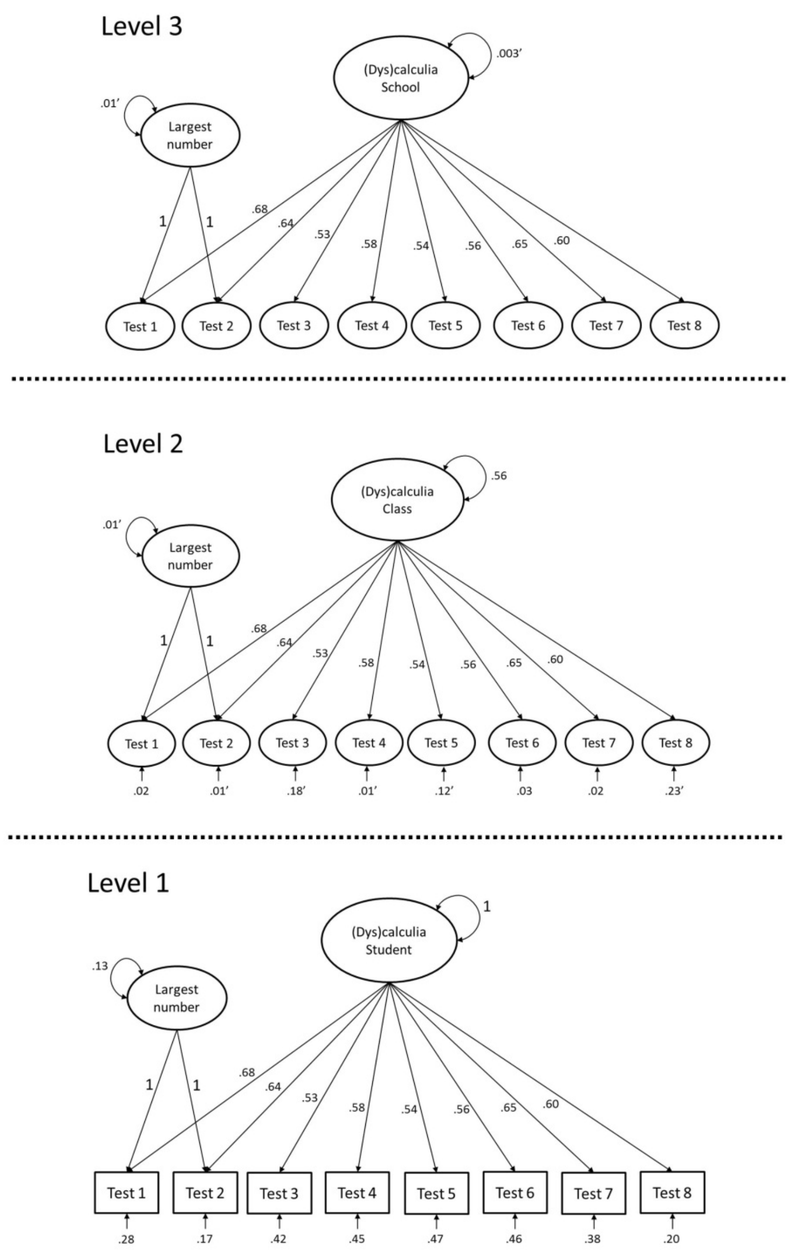
Final model

Significant cluster bias at Level 2 for Test 1, Test 6 and Test 7

Size of the bias:
Proportion of variance due to cluster bias

Proportion of bias in Test 1 wrt Level 2 variance:
$.02 / (.01 + .68^2 * .56 + .02) = .071$

Proportion of bias in Test 1 wrt total variance:
$.02 / (.01 + .68^2 * .56 + .02) +$
$(.01 + .68^2 * .003) +$
$(.13 + .68^2 + .28) = .019$

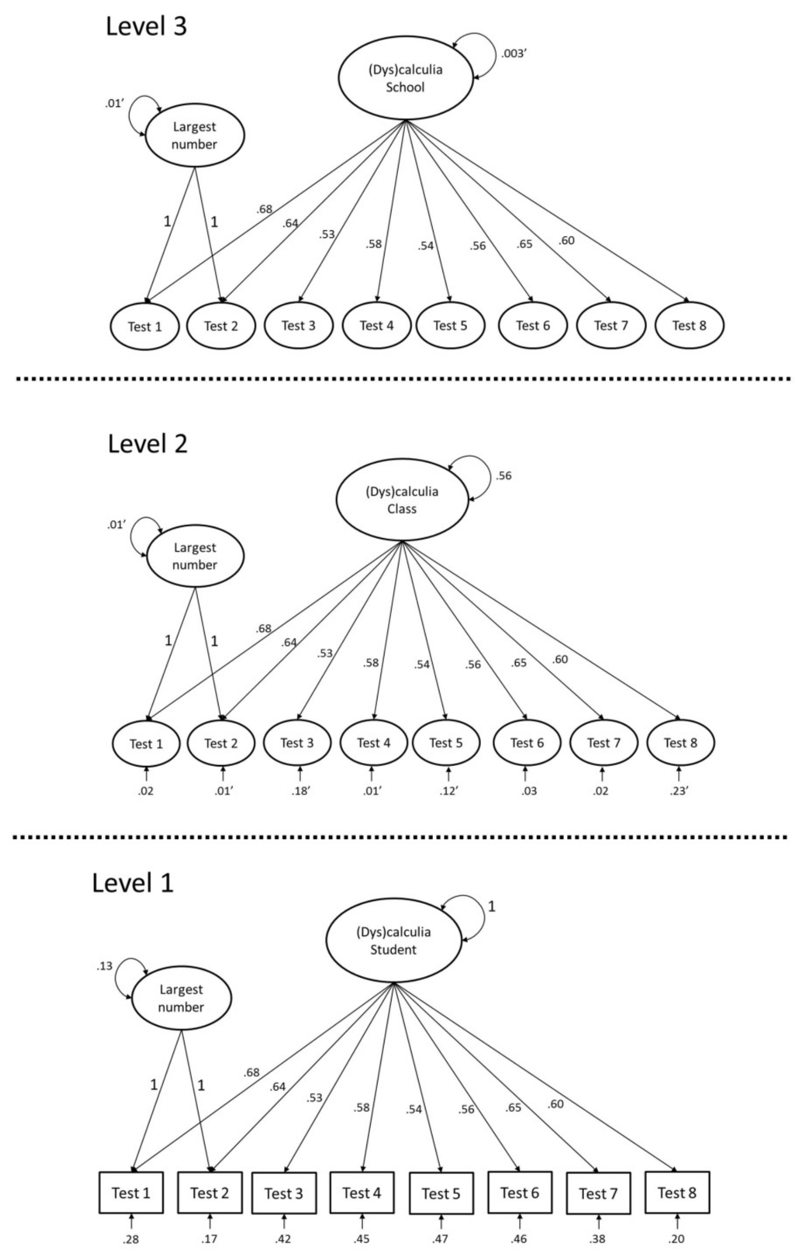


Final model

Significant cluster bias at Level 2 for Test 1, Test 6 and Test 7

Table 2 | Proportions of variance caused by biasing variables at Level 2.

Test	Proportion bias Level 2	Proportion bias Total
Test 1	0.071	0.019
Test 6	0.146	0.031
Test 7	0.090	0.020



Results

- Interpretation of the bias
- Test 1: Cross out largest number

- Test 6: Number line

- Test 7: Fast subtraction

Results

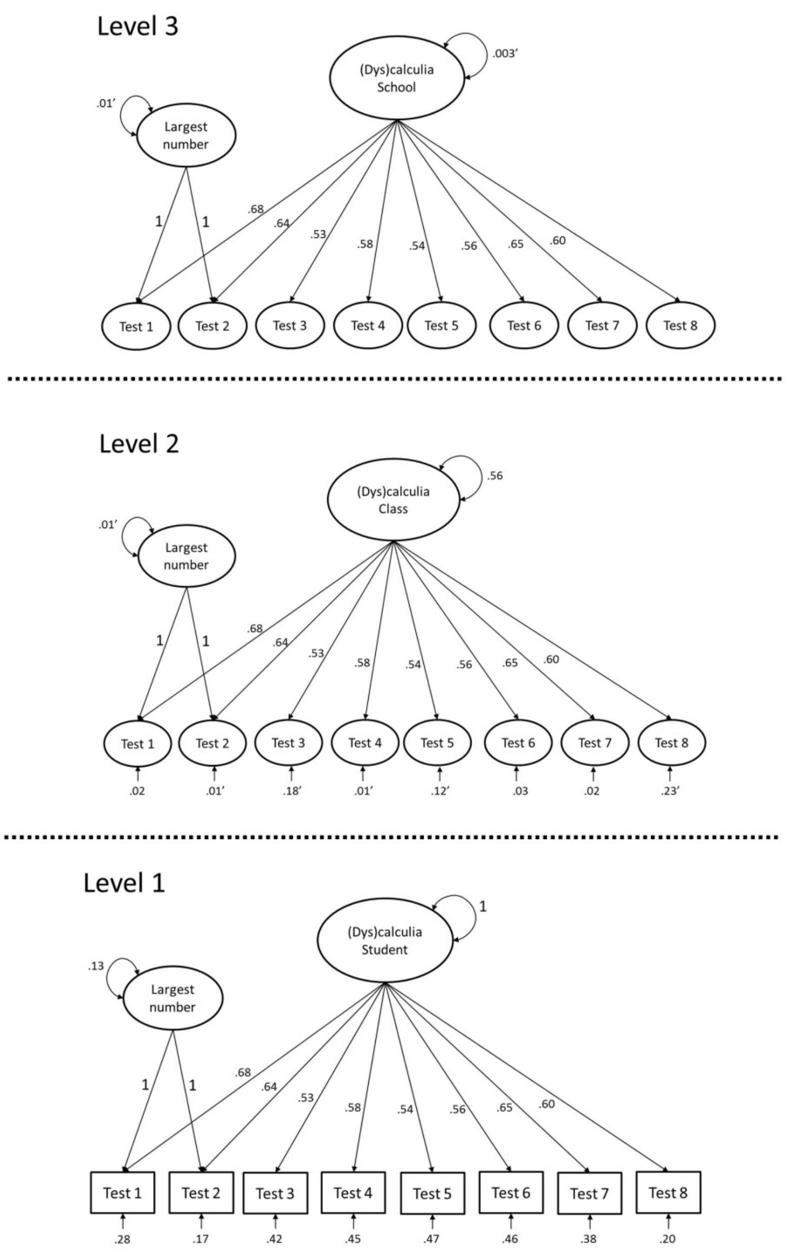
- ICC of factor (Mehta & Neale, 2005; Kim, Kwok & Yoon, 2012)

ICC_{class}:

$.56 / (1 + .56 + .003) = .358$

ICC_{school}:

$.003 / (1 + .56 + .003) = .002$



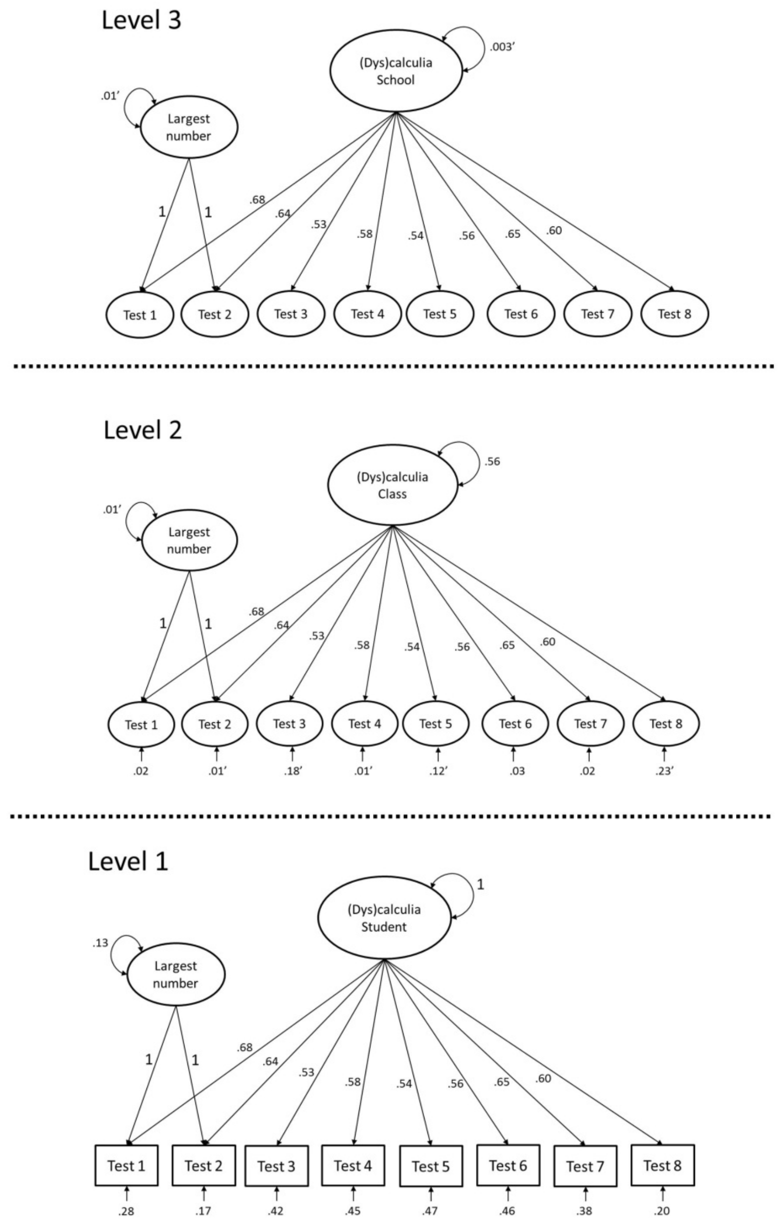
Syntax

```
Analysis:   TYPE = THREELEVEL;


Model:%WITHIN%
        calc by v1* v2 - v8 (Lambda1-Lambda8);
        calc@1;
        number by v1@1 v2@1;
        number;
        calc with number@0;
        v1-v8;

        %BETWEEN classcluster%
        L2calc by v1* v2 - v8 (Lambda1-Lambda8);
        L2calc;
        L2number by v1@1 v2@1;
        L2number;
        L2calc with L2number@0;
        v1-v8;

        %BETWEEN schoolcluster%
        L3calc by v1* v2 - v8 (Lambda1-Lambda8);
        L3calc;
        L3number by v1@1 v2@1;
        L3number;
        L3calc with L3number@0;
        v1-v8@0;
```



Discussion

- Downside of this approach
  - Cannot differentiate between strong and weak invariance
- Alternative approaches with 2-level data
  - Bayesian SEM (Muthén & Asparouhov , 2013)
  - Bayesian IRT (Verhagen & Fox, 2012)
- Alternative approaches with 3-level data

Thank you for listening!

- Questions?