Between and Within Country Measurement Invariance Testing in a EU Comparative Research on School Dropout

3MC Conference, Chicago, July 26th 2016

Ward Nouwen
University of Antwerp (Belgium)

www.resl-eu.org

This project has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no 320223.
Introduction

- Doctoral Study within RESL.eu Project
  - Comparative study in 9 EU member states (BE, ES, PL, PT, NL, SE, UK, (AU & HU)
  - Financed by EU 7th Framework Program
  - Period: February 2013 – January 2018

- Data used for this paper:
  - Survey data from first wave of a comparative survey collected in secondary schools across 14 urban areas in 7 EU member states (N=19522)
School Engagement as a Predictor for Early School Leaving

- Low school engagement predicts ‘Early School Leaving’

Figure 1: estimated survival probability of (a) behavioural engagement and (b) emotional engagement by grade.

Source: Lamote et al., 2013; Based on Longitudinal Research in Flemish Secondary Education
School Engagement as a Multidimensional Concept

Fredricks et al. (2004) proposed a 3-dimensional concept:
- **Emotional** component: identification with ‘the school’/‘education’
- **Cognitive** component: self-regulated/strategic learning approach
- **Behavioural** component: participation in school-related activities
More recently scholars made distinction between (Appleton et al., 2008):

- **Behavioral engagement**: both positive (e.g. participation in extra-curricular activities) as well as negative (e.g. non-compliance)
- **Academic engagement**: more specific study related behavior like paying attention in class and putting time and effort in study work

Our data supports this using Confirmatory Factor Analysis (CFA)
Operationalisation of School Engagement concept based on Wang et al., 2011

Data were drawn from Maryland (US) Adolescent Development in Context Study (MADICS, 1998)

Article by Wang et al. (2011) showed measurement invariance across ethnic and gender groups
Exploratory Factor Analysis (EFA)

The EFA distinguished the same 6 first order factors.

Rotated Factor Matrix

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belonging</td>
<td>0.838</td>
<td>0.186</td>
<td>-0.073</td>
<td>0.114</td>
<td>0.054</td>
<td>0.037</td>
</tr>
<tr>
<td></td>
<td>0.736</td>
<td>0.2</td>
<td>-0.031</td>
<td>0.094</td>
<td>0.046</td>
<td>0.043</td>
</tr>
<tr>
<td></td>
<td>0.849</td>
<td>0.141</td>
<td>-0.059</td>
<td>0.047</td>
<td>0.034</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>0.169</td>
<td>0.728</td>
<td>0.003</td>
<td>0.05</td>
<td>0.083</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>0.107</td>
<td>0.63</td>
<td>-0.06</td>
<td>0.092</td>
<td>0.092</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>0.076</td>
<td>0.688</td>
<td>-0.017</td>
<td>0.09</td>
<td>0.078</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>-0.029</td>
<td>-0.102</td>
<td>0.779</td>
<td>0.023</td>
<td>-0.044</td>
<td>-0.051</td>
</tr>
<tr>
<td></td>
<td>-0.017</td>
<td>-0.106</td>
<td>0.925</td>
<td>0.027</td>
<td>-0.028</td>
<td>-0.038</td>
</tr>
<tr>
<td></td>
<td>-0.086</td>
<td>-0.08</td>
<td>0.592</td>
<td>0.005</td>
<td>-0.069</td>
<td>-0.117</td>
</tr>
<tr>
<td></td>
<td>0.052</td>
<td>0.041</td>
<td>0.01</td>
<td>0.68</td>
<td>0.051</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>0.037</td>
<td>0.022</td>
<td>0.022</td>
<td>0.79</td>
<td>0.1</td>
<td>0.103</td>
</tr>
<tr>
<td></td>
<td>0.077</td>
<td>-0.016</td>
<td>0.008</td>
<td>0.621</td>
<td>-0.036</td>
<td>-0.01</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.139</td>
<td>0.011</td>
<td>0.545</td>
<td>0.108</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>0.016</td>
<td>0.093</td>
<td>-0.041</td>
<td>0.092</td>
<td>0.597</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>0.043</td>
<td>0.063</td>
<td>-0.071</td>
<td>0.038</td>
<td>0.785</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>0.056</td>
<td>0.118</td>
<td>-0.027</td>
<td>0.051</td>
<td>0.701</td>
<td>0.119</td>
</tr>
<tr>
<td></td>
<td>0.047</td>
<td>0.112</td>
<td>-0.12</td>
<td>0.075</td>
<td>0.069</td>
<td>0.712</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
<td>0.136</td>
<td>-0.076</td>
<td>0.167</td>
<td>0.085</td>
<td>0.793</td>
</tr>
</tbody>
</table>

Extraction Method: Maximum Likelihood.
Rotation Method: Varimax with Kaiser Normalization.
a Rotation converged in 6 iterations.
CFA – 1st order factors

CFI = 0.971
RMSEA = 0.038
AIC = 3685.085
BCC = 3685.219
CFA – 2nd order factors (3)

CFI = 0.967
RMSEA = 0.040
AIC = 4185.041
BCC = 4185.164
CFA – 2nd order factors (3)

CFI = 0.967
RMSEA = 0.040
AIC = 4185.041
BCC = 4185.164
CFA – 2nd order factors (2)

In line with theoretical developments that distinguish behavioural from academic engagement

CFI = 0.968
RMSEA = 0.039
AIC = 3961.886
BCC = 3962.011

Reducing Early School Leaving in Europe
Between and Within Country Measurement Invariance (MI) Testing

- Using multi-group CFA to test for MI
  - Between countries: data from 7 EU member states
  - Within country: between educational tracks in Belgium
    - Testing for **configural invariance**
      - No equality constraints between groups = Baseline Model 1
    - Testing for **metric invariance**:
      - Testing for first-order factor loading invariance (Model 2)
      - Testing for second-order factor loading invariance (Model 3)
    - Testing for **scalar invariance**:
      - Testing for intercept of observed variables invariance (Model 4)
      - Testing for means of first-order latent factorial invariance (Model 5)
…by comparing nested models

“…the difference in the Satorra–Bentler scaled chi-square statistic is sensitive to large sample sizes”

(Wang, et al., 2011)

→ rely on guidelines who suggested that
“a difference of larger than .01 in the CFI indicates a meaningful difference in model fit for testing measurement invariance.”

(Cheung & Rensvold, 2002; Chen et al, 2005)
Between County MI: Testing for configural invariance (Model 1)

- Multi-group CFA with no equality constraints (=baseline model; CFI = 0.959)
  2nd order factor for cognitive engagement under pressure for Belgian and Polish data due to low factor loadings of self-regulated learning (‘regulation’)

- This unconstrained multi-group model serves as a baseline model against which we evaluated the model fits of successively more restrictive models (models 2 → 5).
Between county MI: testing for 1st and 2nd order Metric and Scalar Invariance

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
<th>CFI</th>
<th>ΔCFI (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Unconstrained</td>
<td>0.959</td>
<td>-</td>
</tr>
<tr>
<td>Model 2</td>
<td>1st order factor loadings</td>
<td>0.951</td>
<td>0.008</td>
</tr>
<tr>
<td>Model 3</td>
<td>2nd order factor loadings</td>
<td>0.947</td>
<td>0.012</td>
</tr>
<tr>
<td>Model 4</td>
<td>Intercepts of obs. variables</td>
<td>0.864</td>
<td>0.095</td>
</tr>
<tr>
<td>Model 5</td>
<td>1st order factor means</td>
<td>0.864</td>
<td>0.095</td>
</tr>
</tbody>
</table>

(*) Represents difference to unconstrained model (Model 1)
Educational tracking in Flanders:

- socio-ethnic school segregation (Wouters & Groenez, 2014)
- academic culture among staff and students in general versus VET schools (Van Houtte, 2004 & 2006)

Source: Eurydice, 2014
Within County MI: Testing for configural invariance (Model 1)

- Multi-group CFA with no equality constraints (baseline model; CFI = 0.947)
  - The unconstrained multi-group baseline model for the different tracks in the Flemish data is inferior to the multi-group model for different countries using the EU level pooled data (CFI=0.959). Mostly related to the issue of the 2nd order factor for cognitive engagement (see supra).

- Again, the unconstrained multi-group model serves as a baseline model against which we evaluated the fits of successively more restrictive models (models 2 → 5).
Within county MI: testing for 1st and 2nd order Metric and Scalar Invariance

<table>
<thead>
<tr>
<th>Model</th>
<th>Metric Invariance</th>
<th>CFI</th>
<th>ΔCFI (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Unconstrained</td>
<td>0.947-</td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>1st order factor loadings</td>
<td>0.946</td>
<td>0.001</td>
</tr>
<tr>
<td>Model 3</td>
<td>2nd order factor loadings</td>
<td>0.945</td>
<td>0.002</td>
</tr>
<tr>
<td>Model 4</td>
<td>Intercepts of obs. variables</td>
<td>0.925</td>
<td>0.022</td>
</tr>
<tr>
<td>Model 5</td>
<td>1st order factor means</td>
<td>0.925</td>
<td>0.022</td>
</tr>
</tbody>
</table>

(*) Represents difference to unconstrained model (Model 1)
Conclusion

- Between country MI (7 EU countries):
  - Weak but acceptable metric equivalence for 1st (and 2nd) order factor loadings
  - Weak and unacceptable scalar invariance for both intercepts of observed variables as well as 1st order factor means

- Within country MI (3 Flemish educational tracks):
  - Strong metric invariance between tracks
  - Weak and unacceptable scalar invariance
Discussion

What to do now? ➔ I’m here to learn…

“Davidov et al. (2012) have introduced a multilevel structural equation modelling (MLSEM) approach that can be used to interpret deviations from scalar equivalence substantively by modelling how cross-national differences in item intercepts are linked to contextual variables.”

- Not possible because of low number of groups (7 countries/ 3 tracks)

Questions for discussion:

- What are the implications of the scalar non-equivalence between tracks for using MLSEM with students clustered by schools?
- Include type of school (e.g. provision of tracks, socio-ethnic student composition, shared school culture) as contextual variables at the school level?
References