Relating measurement invariance, cross-level invariance, and multilevel reliability

Suzanne Jak and Terrence Jorgensen

University of Amsterdam

S.Jak@uva.nl

March 17, 2017
Overview

Introduction

How invariance between groups relates to between-level reliability

How invariance between groups relates to invariance across levels

How invariance across levels relates to reliability

Data illustration

Overview and discussion
Issues in applications of multilevel SEM

- Cross-level invariance restrictions
Issues in applications of multilevel SEM

- Cross-level invariance restrictions
- Negative (zero) residual variance at between-level
Issues in applications of multilevel SEM

- Cross-level invariance restrictions
- Negative (zero) residual variance at between-level
- High standardized factor loadings at between-level
Issues in applications of multilevel SEM

- Cross-level invariance restrictions
- Negative (zero) residual variance at between-level
- High standardized factor loadings at between-level
- Invariance across clusters
Strong factorial invariance across groups

Strong factorial invariance across groups implies equality of factor loadings and intercepts across groups (Meredith, 1994)

\[
\Sigma_g = \Lambda \Phi_g \Lambda^t + \Theta_g \\
\mu_g = \nu + \Lambda \kappa_g
\]
Two-level SEM

\[ \Sigma_{\text{TOTAL}} = \Sigma_{\text{WITHIN}} + \Sigma_{\text{BETWEEN}} \]
Two-level SEM

\[ \Sigma_{TOTAL} = \Sigma_{WITHIN} + \Sigma_{BETWEEN} \]

Specify models at the within- and between level:

\[ \Sigma_{WITHIN} = \Lambda_W \Phi_W \Lambda_W^t + \Theta_W \]

\[ \Sigma_{BETWEEN} = \Lambda_B \Phi_B \Lambda_B^t + \Theta_B \]
Two-level model and cross-level invariance

If the between-level factor represents the aggregate of the within-level factor, cross-level invariance of factor loadings is needed (Stapleton et al. 2016).
Two-level model and cross-level invariance

If the between-level factor represents the aggregate of the within-level factor, cross-level invariance of factor loadings is needed (Stapleton et al. 2016)
Composite reliability


\[\omega = \frac{(\sum_{i=1}^{k} \lambda_i)^2 \phi}{(\sum_{i=1}^{k} \lambda_i)^2 \phi + \sum_{i=1}^{k} \theta_i}\]
Composite reliability

  \[ \omega = \frac{(\sum_{i=1}^{k} \lambda_i)^2 \phi}{(\sum_{i=1}^{k} \lambda_i)^2 \phi + \sum_{i=1}^{k} \theta_i} \]

- Geldhof, Preacher & Zyphur (2014)
  \[ \omega_{\text{BETWEEN}} = \frac{(\sum_{i=1}^{k} \lambda_{ib})^2 \phi_b}{(\sum_{i=1}^{k} \lambda_{ib})^2 \phi_b + \sum_{i=1}^{k} \theta_{ib}} \]
How invariance between groups relates to between-level reliability

- Measurement invariance across groups implies that between-group differences cannot be due to other factors than those accounting for within-group differences (Lubke et al., 2003).
How invariance between groups relates to between-level reliability

- Measurement invariance across groups implies that between-group differences cannot be due to other factors than those accounting for within-group differences (Lubke et al., 2003).

- The measurement of between-group differences is only reliable if differences in observed scores across groups reflect differences in common factors across groups.
How invariance between groups relates to invariance across levels
How invariance between groups relates to invariance across levels

With equal factor loadings ($\Lambda_g = \Lambda$), and equal intercepts ($\nu_g = \nu$) for all groups (clusters), the following model holds:
How invariance between groups relates to invariance across levels

With equal factor loadings ($\Lambda_g = \Lambda$), and equal intercepts ($\nu_g = \nu$) for all groups (clusters), the following model holds:

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

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

How invariance between groups relates to invariance across levels

With equal factor loadings ($\Lambda_g = \Lambda$), but different intercepts ($\nu_g = \nu_g$) across groups (clusters), the following model holds:
How invariance between groups relates to invariance across levels

With equal factor loadings ($\Lambda_g = \Lambda$), but different intercepts ($\nu_g = \nu_g$) across groups (clusters), the following model holds:

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

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

Test for strong factorial invariance across clusters

\[
\Sigma_{\text{WITHIN}} = \Lambda \Phi_{\text{WITHIN}} \Lambda^t + \Theta_{\text{WITHIN}}
\]

\[
\Sigma_{\text{BETWEEN}} = \Lambda \Phi_{\text{BETWEEN}} \Lambda^t
\]

Set factor loadings equal across levels and test \( \Theta_{\text{BETWEEN}} = 0 \)

Jak, Oort & Dolan (2013, 2014)
Test for cluster bias
How invariance across levels relates to reliability

$\omega_{BETWEEN} = \frac{(\sum_{i=1}^{k} \lambda_{ib})^2 \phi_{b}}{(\sum_{i=1}^{k} \lambda_{ib})^2 \phi_{b} + \sum_{i=1}^{k} \theta_{ib}}$
How invariance across levels relates to reliability

\[ \omega_{\text{BETWEEN}} = \frac{\left( \sum_{i=1}^{k} \lambda_{ib} \right)^2 \phi_b}{\left( \sum_{i=1}^{k} \lambda_{ib} \right)^2 \phi_b + \sum_{i=1}^{k} \theta_{ib}} \]

- If factor loadings are equal across levels, differences in reliability across levels are only a function of differences in factor variances and residual variances.
- If residual variance is zero, reliability will be perfect.
Illustration Wellbeing

- Six items (three positive, three negative)
- Data from European Social Survey (round 2012)
- Responses from 54,673 respondents from 29 countries
Illustration Wellbeing

- Six items (three positive, three negative)
- Data from European Social Survey (round 2012)
- Responses from 54,673 respondents from 29 countries

- Test measurement across countries with multigroup model
- Test measurement across countries with the test for cluster bias
- Evaluate level-specific composite reliability
Measurement model on data of all countries

$\chi^2(7) = 1352.81$, RMSEA = .059 90%CI [.057 ; .062], CFI = 0.99
## Multigroup measurement invariance analyses

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$\chi^2$</th>
<th>RMSEA</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Configural invariance</td>
<td>203</td>
<td>1742.85</td>
<td>.063</td>
<td>.99</td>
</tr>
<tr>
<td>2. Weak factorial invariance</td>
<td>343</td>
<td>3168.430</td>
<td>.066</td>
<td>.97</td>
</tr>
<tr>
<td>3. Strong factorial invariance</td>
<td>455</td>
<td>12471.47</td>
<td>.118</td>
<td>.88</td>
</tr>
</tbody>
</table>

![Chi-squared vs df graph](image)
Multigroup measurement invariance analyses

<table>
<thead>
<tr>
<th>Item</th>
<th># MI&gt;50</th>
<th># MI&gt;100</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRHPPY</td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td>ENJLF</td>
<td>13</td>
<td>5</td>
</tr>
<tr>
<td>FLTPCFL</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>FLTDPR</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>FLTSD</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>FLTANX</td>
<td>18</td>
<td>14</td>
</tr>
</tbody>
</table>
Two-level analysis

- Step 1: Two-level model without across-level constraints
- Step 2: Two-level model with equal factor loadings across levels
- Step 3: Two-level model with equal factor loadings across levels and zero residual variance at between-level
Two-level invariance analyses

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$\chi^2$</th>
<th>RMSEA</th>
<th>CFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Two-level CFA</td>
<td>14</td>
<td>516.69</td>
<td>.026</td>
<td>.98</td>
</tr>
<tr>
<td>2. Cross-level invariance</td>
<td>19</td>
<td>619.52</td>
<td>.024</td>
<td>.97</td>
</tr>
<tr>
<td>3. Strong factorial invariance</td>
<td>25</td>
<td>6880.93</td>
<td>.071</td>
<td>.68</td>
</tr>
</tbody>
</table>

![Graph showing the relationship between df and Chi-squared values](image)
## Two-level invariance analyses

<table>
<thead>
<tr>
<th>Item</th>
<th>MI</th>
<th>$\Delta \chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRHPPY</td>
<td>8895</td>
<td>661.02</td>
</tr>
<tr>
<td>ENJLF</td>
<td>28999</td>
<td>1229.30</td>
</tr>
<tr>
<td>FLTPCFL</td>
<td>40919</td>
<td>1410.16</td>
</tr>
<tr>
<td>FLTDPR</td>
<td>36531</td>
<td>1380.28</td>
</tr>
<tr>
<td>FLTSD</td>
<td>8491</td>
<td>641.51</td>
</tr>
<tr>
<td>FLTANX</td>
<td>147722</td>
<td>2868.18</td>
</tr>
</tbody>
</table>
Two-level invariance analyses

Data illustration

Unstandardized

Positive wellbeing

ICC positive wellbeing factor: \( \frac{0.06}{1 + 0.06} = 0.057 \)

Negative wellbeing

ICC negative wellbeing factor: \( \frac{0.13}{1 + 0.13} = 0.117 \)

Between countries

Within countries
Two-level invariance analyses

ICC positive wellbeing factor: \( \frac{0.060}{1 + 0.060} = 0.057 \)

ICC negative wellbeing factor: \( \frac{0.133}{1 + 0.133} = 0.117 \)
Two-level invariance analyses

Standardized

Between countries

Within countries
Two-level invariance analyses

Standardized residual variance: Proportion of variance across countries caused by biasing factors
Reliability at within- and between-level

\[ \omega_{WITHIN} = \frac{(\sum_{i=1}^{k} \lambda_i)^2 \phi_w}{(\sum_{i=1}^{k} \lambda_i)^2 \phi_w + \sum_{i=1}^{k} \theta_{iw}} \]

\[ \omega_{BETWEEN} = \frac{(\sum_{i=1}^{k} \lambda_i)^2 \phi_b}{(\sum_{i=1}^{k} \lambda_i)^2 \phi_b + \sum_{i=1}^{k} \theta_{ib}} \]

<table>
<thead>
<tr>
<th>Factor</th>
<th>Within</th>
<th>Between</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive wellbeing</td>
<td>0.78</td>
<td>0.87</td>
</tr>
<tr>
<td>Negative wellbeing</td>
<td>0.73</td>
<td>0.85</td>
</tr>
</tbody>
</table>
Reliability at within- and between-level

\[ \omega_{\text{WITHIN}} = \frac{(\sum_{i=1}^{k} \lambda_i)^2 \phi_w}{(\sum_{i=1}^{k} \lambda_i)^2 \phi_w + \sum_{i=1}^{k} \theta_{iw}} \]

\[ \omega_{\text{BETWEEN}} = \frac{(\sum_{i=1}^{k} \lambda_i)^2 \phi_b}{(\sum_{i=1}^{k} \lambda_i)^2 \phi_b + \sum_{i=1}^{k} \theta_{ib}} \]

<table>
<thead>
<tr>
<th>Factor</th>
<th>Within</th>
<th>Between</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive wellbeing</td>
<td>0.78</td>
<td>0.87</td>
</tr>
<tr>
<td>Negative wellbeing</td>
<td>0.73</td>
<td>0.85</td>
</tr>
</tbody>
</table>
Reliability at within- and between-level

Data illustration
Reliability at within- and between-level

Items with highest standardized factor loadings contribute most to level-specific reliability (and are least biased across countries)
### Overview

<table>
<thead>
<tr>
<th>Condition</th>
<th>Multigroup model</th>
<th>Two-level model</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configural</td>
<td><strong>pat(Λ_g) = pat(Λ)</strong></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
## Overview

<table>
<thead>
<tr>
<th>Condition</th>
<th>Multigroup model</th>
<th>Two-level model</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configural</td>
<td>$\text{pat}(\Lambda_g) = \text{pat}(\Lambda)$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Weak</td>
<td>$\Lambda_g = \Lambda$</td>
<td>$\Lambda_W = \Lambda_B$</td>
<td>-</td>
</tr>
</tbody>
</table>
### Overview

<table>
<thead>
<tr>
<th>Condition</th>
<th>Multigroup model</th>
<th>Two-level model</th>
<th>Reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Configural</td>
<td>(\text{pat}(\Lambda_g) = \text{pat}(\Lambda))</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>Weak</td>
<td>(\Lambda_g = \Lambda)</td>
<td>(\Lambda_W = \Lambda_B)</td>
<td>(-)</td>
</tr>
<tr>
<td>Strong</td>
<td>(\Lambda_g = \Lambda, \nu_g = \nu)</td>
<td>(\Lambda_W = \Lambda_B, \Theta_B = 0)</td>
<td>(\omega_B = 1)</td>
</tr>
</tbody>
</table>
Conclusion

- Cross-level invariance is a reasonable assumption
Conclusion

- Cross-level invariance is a reasonable assumption
- Absence of residual variance at the between-level is desirable
Conclusion

- Cross-level invariance is a reasonable assumption
- Absence of residual variance at the between-level is desirable
- Measurement invariance across clusters implies perfect reliability at the between-level
Thank you for your attention