

# *Testing for Cross-National Heterogeneity*

Jean-Paul Fox

University of Twente  
Department of Behavioural Science  
Enschede, Netherlands

October 14, 2009

# Overview

## *Measurement Invariance*

# Overview

*Measurement Invariance*

*Random Item Effects Approach*

# Overview

*Measurement Invariance*

*Random Item Effects Approach*

*International Comparisons of Student Achievement*

# *Measurement Invariance*

DIF:

- Unconditional methods (item x group interaction, Coffman, 1964; Echternacht, 1974)

# Measurement Invariance

DIF:

- Unconditional methods (item x group interaction, Coffman, 1964; Echternacht, 1974)
  1. Dependence on other items

# Measurement Invariance

DIF:

- Unconditional methods (item x group interaction, Coffman, 1964; Echternacht, 1974)
  1. Dependence on other items
  2. Dependence on the score distribution

# Measurement Invariance

DIF:

- Unconditional methods (item x group interaction, Coffman, 1964; Echternacht, 1974)
  1. Dependence on other items
  2. Dependence on the score distribution
- Conditional methods (Scheuneman, 1979; Lord, 1980; Mellenbergh, 1982)

# Measurement Invariance

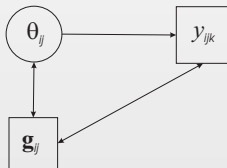
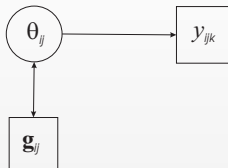
DIF:

- Unconditional methods (item x group interaction, Coffman, 1964; Echternacht, 1974)
  1. Dependence on other items
  2. Dependence on the score distribution
- Conditional methods (Scheuneman, 1979; Lord, 1980; Mellenbergh, 1982)

$$P(Y = 1 | \theta, g) = P(Y = 1 | \theta)$$

for all  $g, \theta$

# Graphical display of DIF



# *Conditional DIF Detection*

1. Specify item response model (likelihood model)

## *Conditional DIF Detection*

1. Specify item response model (likelihood model)
2. Specify distribution of the latent variable  $\theta$

## *Conditional DIF Detection*

1. Specify item response model (likelihood model)
2. Specify distribution of the latent variable  $\theta$
3. Estimate the ICCs in each group ( $\theta$  not on the same scale)

## *Conditional DIF Detection*

1. Specify item response model (likelihood model)
2. Specify distribution of the latent variable  $\theta$
3. Estimate the ICCs in each group ( $\theta$  not on the same scale)
4. Assess the degree of bias (Mellenbergh, Hambleton)

## *Conditional DIF Detection*

Distributional assumptions (Thissen et al, 1986; Reise et al, 1993)

1. Evaluate fit of item response model in each group

## *Conditional DIF Detection*

Distributional assumptions (Thissen et al, 1986; Reise et al, 1993)

1. Evaluate fit of item response model in each group
2. Different population distributions, same item curves, simultaneous analysis

## *Conditional DIF Detection*

Distributional assumptions (Thissen et al, 1986; Reise et al, 1993)

1. Evaluate fit of item response model in each group
2. Different population distributions, same item curves, simultaneous analysis
3. Full measurement invariance versus partial measurement invariance

## *Conditional DIF Detection*

Distributional assumptions (Thissen et al, 1986; Reise et al, 1993)

1. Evaluate fit of item response model in each group
2. Different population distributions, same item curves, simultaneous analysis
3. Full measurement invariance versus partial measurement invariance
4. Inspection of item parameters for DIF detection

# *Many groups, ...*

1. Specifying the population distribution

## *Many groups, ...*

1. Specifying the population distribution
2. Finding anchor items

## *Many groups, ...*

1. Specifying the population distribution
2. Finding anchor items
3. DIF in difficulty and discrimination parameters.  
(Tests for partial measurement invariance)

# Modeling Item Responses

Collection of responses  $Y$  of persons ( $i$ ) nested in groups ( $j$ ) who answer items ( $k$ ).

$$P(Y_{ijk} = 1 \mid \theta_{ij}, \tilde{a}_{kj}, \tilde{b}_{kj}) = \begin{cases} \frac{\exp(d(\tilde{a}_{kj}\theta_{ij} - \tilde{b}_{kj}))}{1 + \exp(d(\tilde{a}_{kj}\theta_{ij} - \tilde{b}_{kj}))} \\ \Phi(\tilde{a}_{kj}\theta_{ij} - \tilde{b}_{kj}) \end{cases}$$

# *International Item Parameters*

International item parameters:

$$(a_k, b_k)^t \sim \mathcal{N}(\boldsymbol{\mu}_\xi, \boldsymbol{\Sigma}_\xi)$$

# International Item Parameters

International item parameters:

$$(a_k, b_k)^t \sim \mathcal{N}(\boldsymbol{\mu}_\xi, \boldsymbol{\Sigma}_\xi)$$

Hyper priors:

$$\begin{aligned}\boldsymbol{\Sigma}_\xi &\sim \mathcal{IW}(\nu, \boldsymbol{\Sigma}_0) \\ \boldsymbol{\mu}_\xi \mid \boldsymbol{\Sigma}_\xi &\sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_\xi / K_0)\end{aligned}$$

## National Item Parameters

$$\tilde{\xi}_{kj} = (\tilde{a}_{kj}, \tilde{b}_{kj})^t \sim \mathcal{N}((a_k, b_k)^t, \Sigma_{\tilde{\xi}})$$

## National Item Parameters

$$\tilde{\xi}_{kj} = (\tilde{a}_{kj}, \tilde{b}_{kj})^t \sim \mathcal{N}((a_k, b_k)^t, \Sigma_{\tilde{\xi}})$$

$$\text{Cov}(\tilde{\xi}_{kj}, \tilde{\xi}_{k'j'}) = \begin{cases} \Sigma_{\xi} + \Sigma_{\tilde{\xi}} & k = k', j = j' \\ \Sigma_{\xi} & k = k', j \neq j' \\ 0 & k \neq k'. \end{cases}$$

# Population Model

Ignore latent mean differences:

$$\theta_i \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$$

# Population Model

Ignore latent mean differences:

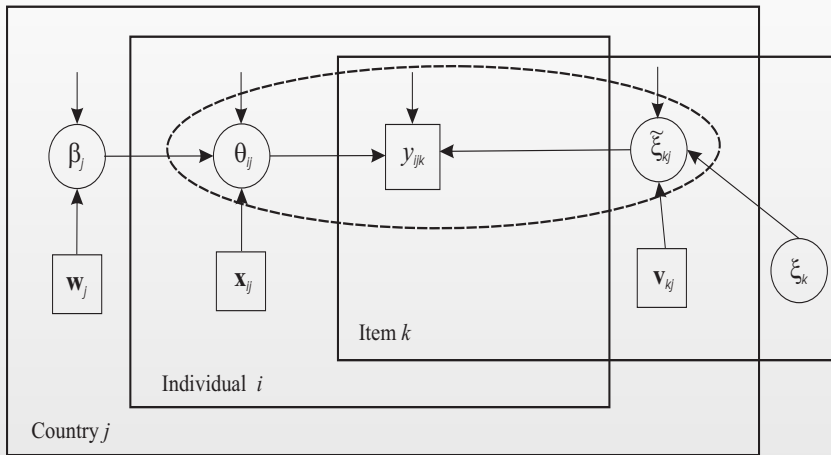
$$\theta_i \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$$

Multilevel model:

$$\theta_{ij} \sim \mathcal{N}(\mathbf{x}_{ij}^t \boldsymbol{\beta}_j, \sigma_j^2)$$

$$\boldsymbol{\beta}_j \sim \mathcal{N}(\mathbf{w}_j^t \boldsymbol{\gamma}, \tau^2)$$

# Graphical Display of Random Item Effects Model



# *International Comparisons of Student Achievement*

- Eight mathematic items from booklet one

# *International Comparisons of Student Achievement*

- Eight mathematic items from booklet one
- 9796 respondents from 40 participating countries

# International Comparisons of Student Achievement

- Eight mathematic items from booklet one
- 9796 respondents from 40 participating countries

$$\begin{aligned}P\left(Y_{ijk} \mid \theta_{ij}, \tilde{\xi}_k\right) &= \Phi\left(\tilde{a}_{kj}\theta_{ij} - \tilde{b}_{kj}\right) \\ \left(\tilde{a}_{kj}, \tilde{b}_{kj}\right)^t &= \left(a_k, b_k\right)^t + \left(\epsilon_{a_{kj}}, \epsilon_{b_{kj}}\right)^t \\ \theta_{ij} &= \beta_{0j} + e_{ij}\end{aligned}$$

# International Comparisons of Student Achievement

- Eight mathematic items from booklet one
- 9796 respondents from 40 participating countries

$$\begin{aligned}P\left(Y_{ijk} \mid \theta_{ij}, \tilde{\boldsymbol{\xi}}_k\right) &= \Phi\left(\tilde{a}_{kj}\theta_{ij} - \tilde{b}_{kj}\right) \\ \left(\tilde{a}_{kj}, \tilde{b}_{kj}\right)^t &= \left(a_k, b_k\right)^t + \left(\epsilon_{a_{kj}}, \epsilon_{b_{kj}}\right)^t \\ \theta_{ij} &= \beta_{0j} + e_{ij}\end{aligned}$$

$$e_{ij} \sim \mathcal{N}\left(0, \sigma_{\theta_j}^2\right)$$

$$\beta_{0j} \sim \mathcal{N}\left(\gamma_{00}, \tau^2\right)$$

$$\left(\epsilon_{a_{kj}}, \epsilon_{b_{kj}}\right)^t \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}_{\tilde{\boldsymbol{\xi}}}\right)$$

# PISA 2003: Exploring Cross-National Item Variation

Item	Invariance		Non-Invariance						
	Model $\mathcal{M}_0$		Model $\mathcal{M}_1$			Model $\mathcal{M}_2$			
	Mean	SD	Mean	SD	$\sigma_{a_k}$	Mean	SD	$\sigma_{a_k}$	$p_0(a_k)$
<b>Discrimination parameter</b>									
1	.80	.02	.72	.03	.07	.81	.03	.08	.86
2	1.06	.03	1.07	.06	.28	1.09	.05	.23	.99
3	.72	.02	.61	.02	.06	.71	.02	.06	.56
4	.68	.02	.61	.02	.08	.69	.03	.11	.97
5	.55	.02	.53	.02	.09	.58	.02	.11	.99
6	.36	.02	.34	.03	.10	.40	.04	.16	.99
7	.69	.02	.60	.03	.07	.69	.03	.09	.92
8	.66	.02	.63	.03	.09	.68	.03	.11	.97

# PISA 2003: Exploring Cross-National Item Variation

Item	Invariance		Non-Invariance						
	Model $\mathcal{M}_0$		Model $\mathcal{M}_1$			Model $\mathcal{M}_2$			
	Mean	SD	Mean	SD	$\sigma_{b_k}$	Mean	SD	$\sigma_{b_k}$	$p_0(b_k)$
<b>Difficulty parameter</b>									
1	-.58	.01	-.58	.06	.38	-.59	.03	.13	.99
2	.18	.01	.17	.07	.45	.19	.03	.11	.96
3	-.04	.01	-.04	.06	.37	-.03	.03	.10	.99
4	-.35	.01	-.37	.05	.34	-.35	.02	.10	.97
5	-.01	.01	-.02	.04	.24	-.02	.02	.08	.72
6	-1.51	.02	-1.54	.04	.24	-1.59	.02	.10	.90
7	-.78	.01	-.78	.05	.34	-.78	.02	.09	.95
8	-.94	.01	-.96	.04	.26	-.95	.02	.08	.76

# PISA 2003: Exploring Cross-National Item Variation

	Invariance		Non-Invariance			
	Model $\mathcal{M}_0$		Model $\mathcal{M}_1$		Model $\mathcal{M}_2$	
<b>Structural Part</b>						
<i>Fixed</i>						
$\gamma_{00}$	.010	.083	.000	-	.734	.083
<i>Random</i>						
$\sigma_{\theta}^2$	.791	.013	1.000	-	.791	.014
$\tau_{00}^2$	.269	.063			.270	.063
<b>Information Criteria</b>						
-2log L		95727.5		102322.1		94681.0
DIC ( $p_D$ )		100395.1(4667)				98960.2(4279)

# PISA 2003: Exploring Cross-National Item Variation

Item	Invariance		Non-Invariance			
	Model $\mathcal{M}_3$		Model $\mathcal{M}_4$			
	Mean	SD	Mean	SD	$\sigma_{a_k}$	$p_0(a_k)$
<b>Discrimination parameter</b>						
1	.80	.02	.73	.04	.08	.43
2	1.06	.03	1.02	.12	.12	.82
3	.72	.02	.63	.03	.07	.17
4	.68	.02	.62	.03	.07	.29
5	.55	.02	.52	.03	.07	.39
6	.36	.02	.32	.06	.05	.09
7	.69	.02	.62	.03	.07	.21
8	.66	.02	.61	.04	.07	.25

# PISA 2003: Exploring Cross-National Item Variation

Item	Invariance		Non-Invariance			
	Model $\mathcal{M}_3$		Model $\mathcal{M}_4$			
	Mean	SD	Mean	SD	$\sigma_{b_k}$	$p_0(b_k)$
<b>Difficulty parameter</b>						
1	-.58	.02	-.59	.03	.16	.99
2	.18	.02	.19	.09	.17	.99
3	-.04	.01	-.00	.04	.12	.99
4	-.35	.01	-.37	.03	.13	.99
5	-.01	.01	-.02	.03	.11	.98
6	-1.51	.02	-1.52	.09	.13	.99
7	-.78	.02	-.79	.03	.12	.99
8	-.94	.02	-.96	.04	.09	.93

# PISA 2003: Exploring Cross-National Item Variation

	Invariance		Non-Invariance	
	Model $\mathcal{M}_3$		Model $\mathcal{M}_4$	
<b>Structural Part</b>				
<i>Fixed</i>				
$\gamma_{00}$	.010	.073	.496	.051
<i>Random</i>				
$\bar{\sigma}_\theta^2$	.797	.023	.370	.011
$\tau_{00}^2$	.218	.056	.104	.025
<b>Information Criteria</b>				
-2log L	94331.2		87277.4	
DIC ( $p_D$ )	99984.3(5653.1)		91593.5(4316.1)	

## *Discussion*

- A common measurement scale

## *Discussion*

- A common measurement scale
- Explain cross-national heterogeneity

## *Discussion*

- A common measurement scale
- Explain cross-national heterogeneity
- Bayes factor for nested models (common likelihood)