Using a Multilevel Structural Equation Modeling Approach to Explain Cross-Cultural Measurement Noninvariance

Eldad Davidov, University of Zurich
Hermann Dülmer, University of Cologne
Elmar Schlüter, University of Cologne
Peter Schmidt, University of Marburg

Overview

1. Introduction
2. What do we mean by invariance?
3. What to do when there is no invariance?
4. Using multilevel techniques to explain measurement noninvariance
5. Empirical illustration
6. Conclusions

1. Introduction

More and more ***studies test*** the ***invariance*** (equivalence) ***properties*** of scales across

- countries,
- cultural groups,
- time points,
- modes of data collection etc.

There are **two** typical scenarios:

- We analyze and find high levels of invariance and we are **happy**.
- Invariance across groups is absent and we are **disappointed** and question , whether it makes sense to continue comparisons.

1. Introduction

In this study we propose using

- **mulilevel structural equation modeling** (ML SEM)
- to **explain** noninvariance,
- **not to solve** it (diagnosis of the problem).

We will **exemplify** its use with an example from the European Social Survey (ESS).

2. What do we mean by invariance

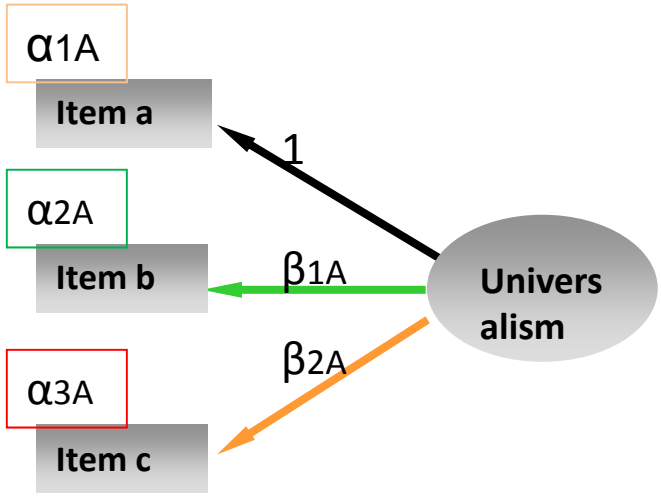
There are typically three important levels of testing measurement invariance:

- Configural Invariance,
- Metric Invariance,
- Scalar Invariance.

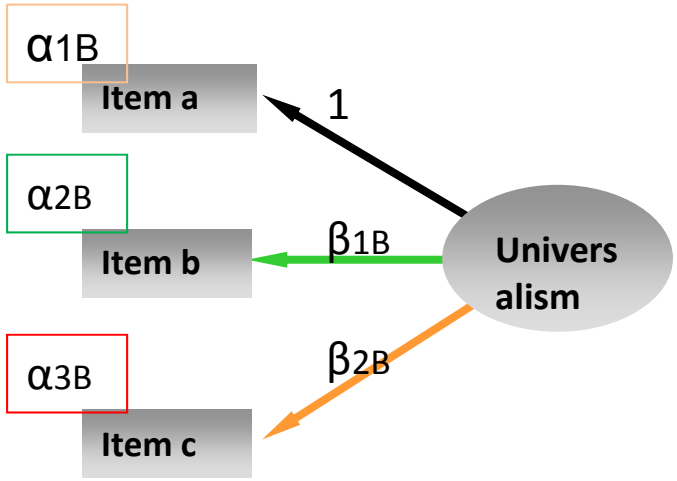
2. What do we mean by invariance

Measurement Invariance:

Group A (Culture, country, time point)



Group B (Culture, country, time point)



All levels of invariance are necessary for meaningful mean comparisons.

2. What do we mean by invariance

But:

Higher levels of invariance (i.e., scalar)

- are **seldom guaranteed**,
- **especially when many groups are compared.**

3. What to do when there is no invariance?

1) Resort to partial invariance

(Byrne/Shavelson/Muthén 1989; Steenkamp/Baumgartner 1998).

2) Compare only a subset of countries (or other groups)

where invariance of the involved concepts does not hold

(e.g., Byrne/van de Vijver 2010).

3) Even refrain from cross-country comparisons

when invarince is not found even across a subset of countries.

4) Decrease the number of items and **delete those items** whose parameters are very different across groups. However, when this approach is applied, one has to address the **question whether** the **meaning of the concept has changed** after the item reduction (Bynre/van de Vijver 2010).

2. What to do when there is no invariance?

But:

None of these solutions explains

where the problem comes from!

4. Using mulilevel techniques to explain measurement noninvariance

5) Explain noninvariance by introducing **contextual variables** and using **multilevel analysis** (Schlüter/Meulemann 2009).

Noninvariance can be viewed as a **useful source of information on cross-cultural differences**
(e.g., Poortinga 1989; Schlüter/Meulemann 2009).

Although it was already referred to by some authors (e.g., Hox/de Leeuw/Brinkhuis 2010) and although the **technique is not new** (e.g., Muthén 1985)

to the best of our knowledge this possibility has **not yet been explicated** and **applied for** the goal of **explaining noninvariance systematically.**

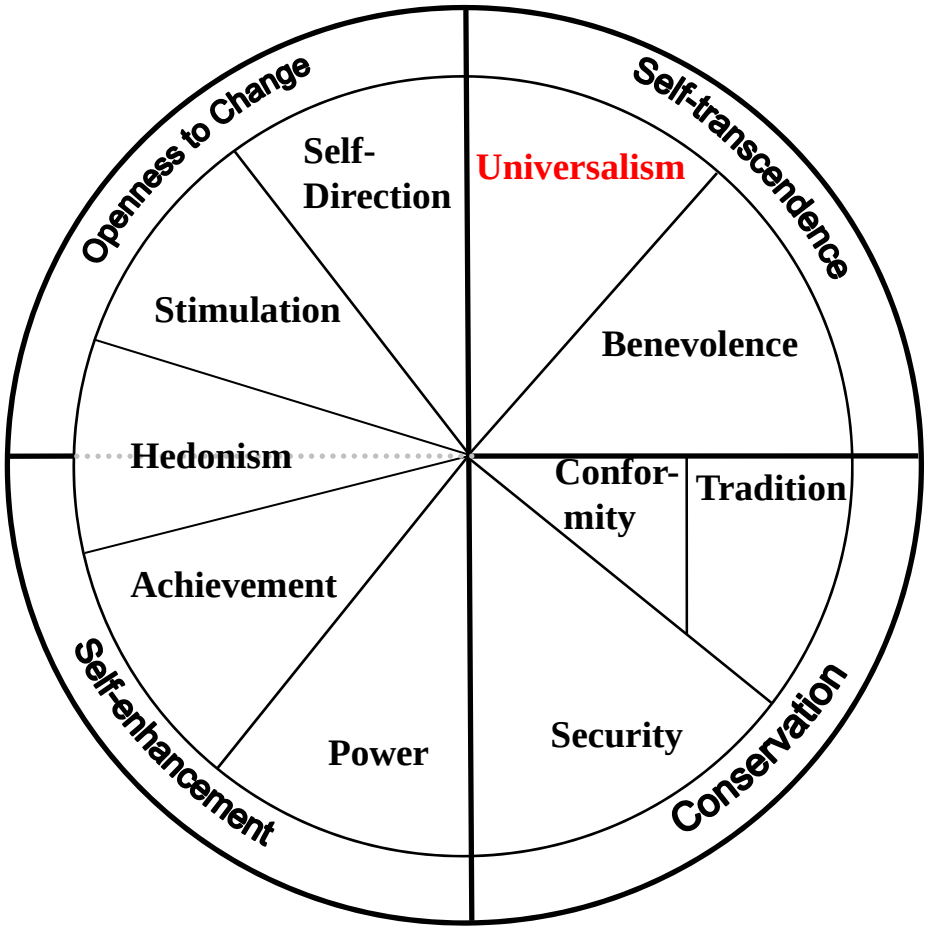
5. Empirical illustration

Theory of 10 Universal Values (Schwartz 1992)

- Focus of our Analysis:
  Universalism

- Universalism
  is measured by 3 Items:
  - Equality
  - Understanding Different People
  - Environment Protection

Data Source: ESS 2004-5



5. Empirical illustration

Test of invariance of universalism across 25 ESS countries by using multiple group confirmatory factor analysis (**MGCFA**)

Results:

- **Configural** and **metric invariance** were confirmed by the data:
  - Factor loadings are equal across countries.
- **Scalar invariance** was **rejected:**
  - Indicator intercepts are different. People use the scale differently.

More specifically:

- The **intercept of the indicator Environment** is particularly different across countries.

5. Empirical illustration

Theoretical Expectations (Inglehart 1990, 1997):

H1: The value of Universalism is more important in Postmodern advanced industrial countries than in less developed Modern countries.

- Intergenerational value change from **Materialist** to **Postmaterialist** value priorities in advanced Modern societies (Inglehart 1990):
  - Postmaterialist needs (belonging, self-expression) become more important than more fundamental Materialist needs (physical/economical security).
- Shift from Materialist to Postmaterialist value priorities is a key indicator for a broader syndrome of **Postmodernization** (Inglehart 1997):
  - Postmodernization includes a shift from an emphasis on self-denying achievement orientations toward individual choice of lifestyles, greater **tolerance for ethic, cultural and sexual diversity**, and an increasing emphasis on **protection of environment** and **quality of life** in general.

5. Empirical illustration

Theoretical Expectations (Inglehart 1990, 1997):

H2: Environment protection is perceived more important in less developed countries than in Postmodern advanced industrial countries.

- Inglehart (1997: 242) also states that in less developed countries where air/water pollution are far worse than in advanced industrial societies, environment protection is less a quality of life issue but much more a matter of physical health.

5. Empirical illustration

Theoretical Expectations (Inglehart 1990, 1997):

Thus:

People may ‘react’ differently to the environment item, because of differences in the level of human development (**HDI**) between the countries.

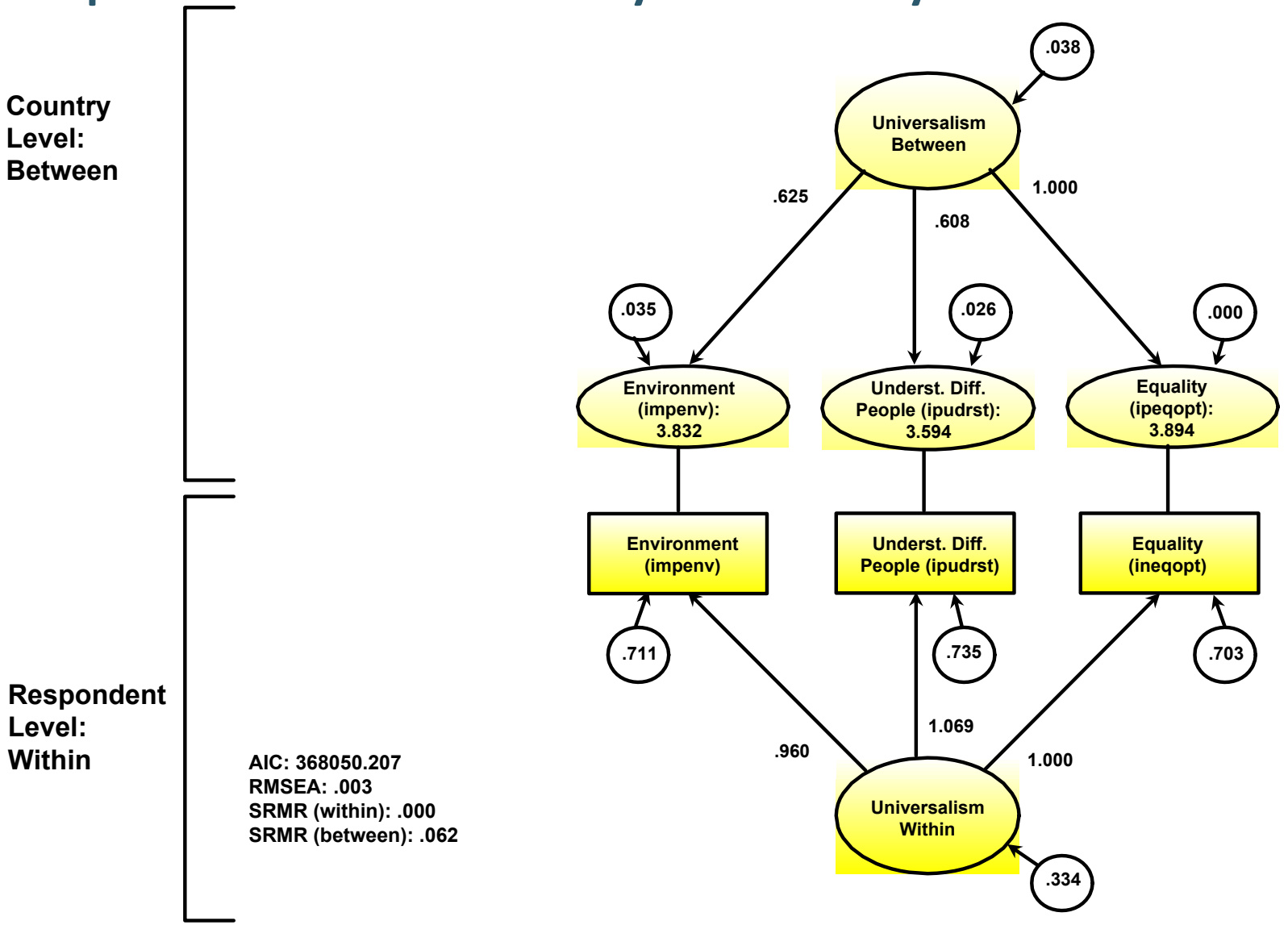
Human Development Index (HDI 2004):

Geometric Mean of

- Standard of Living: Logarithm of the GDP pc in PPP $US
- Educational Attainment: Index for Adult Literaty Rate and Gross Enrolment Rate in Primary/Secondary/Tertiary Schools
- Health: Life Expectation at Birth

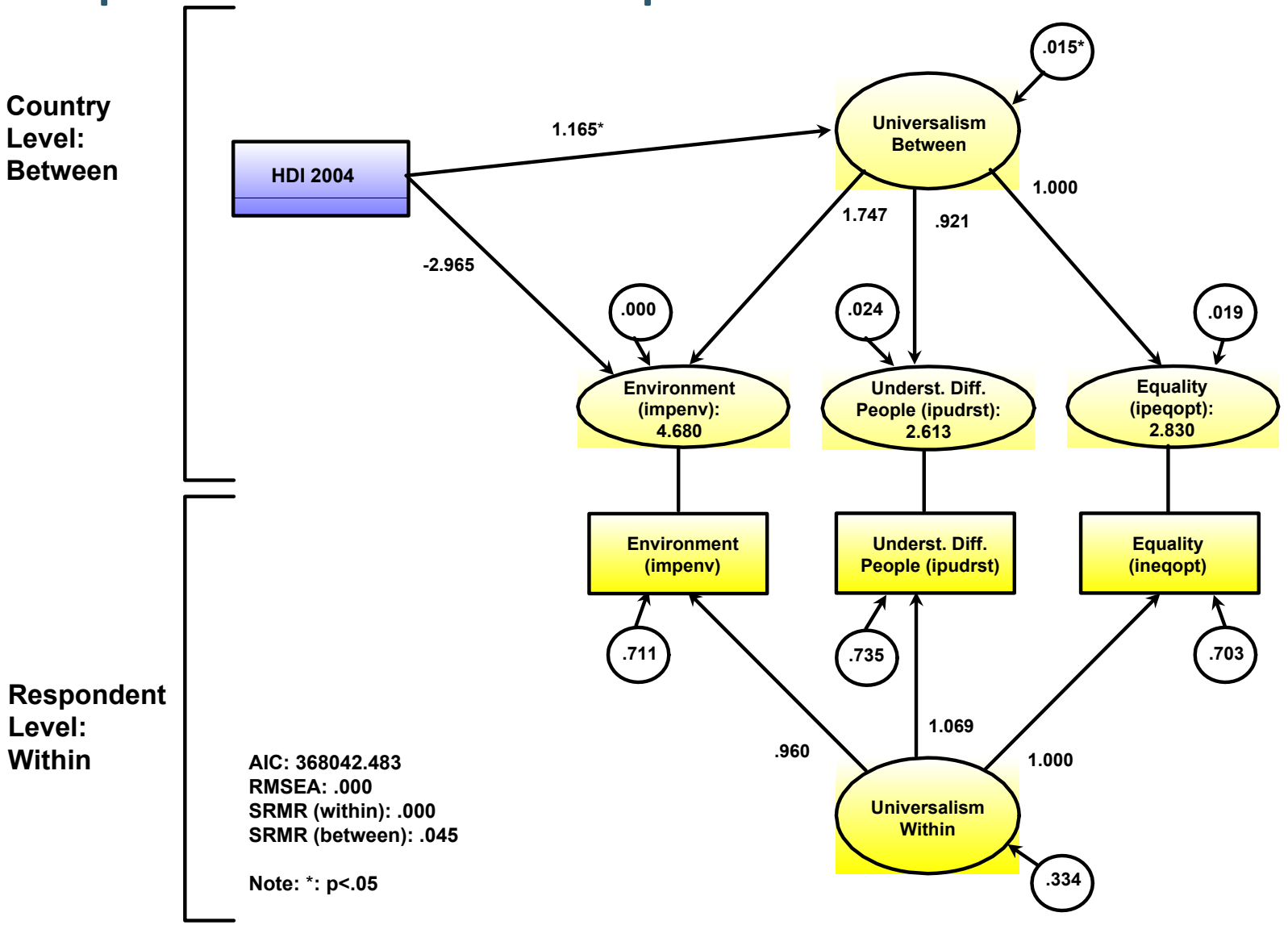
5. Empirical illustration

1st step: Two-Level Confirmatory Factor Analysis



5. Empirical illustration

2nd step: Two-Level Structural Equation Model



6. Conclusion

Our empirical results

confirm our hypotheses concerning the impact of a country‘s level of economic/technical development (HDI)

- on the value of Universalism (H1)
- on the citizens attitude toward environment protection (H2, intercept noninvariance)

The level of human development explains

why scalar invariance was absent for the item Environment.

6. Conclusion

- ML SEM is a rather straightforward technique to **explain variation** in an item‘s intercept.

- Although ML SEM does **not** provide **a resolution** to the problem of noninvariance, it is a useful tool for the **diagnosis of the causes**.

- Diagnosis and systematic identification of the sources of item bias are an **essential step before measurement scales can be improved for cross-cultural analysis.**

Thank you very much

for

your attention!!!

Descriptive Findings

Variances, Covariances, and ICC for the Indicators of Universalism

		Pooled Sample Correlations and Covariances				Within and Between Countries Correlations and Covariances		
		1	2	3		1	2	3
					within			
1	Equality (ipeqopt)	1.079	**.344**	**.324**		1.037	**.332**	**.312**
2	Understanding Diff. People (ipudrst)	.385	1.162	**.334**		.357	1.117	**.321**
3	Environment (impenv)	.347	.372	1.066		.321	.343	1.019
					between			
1	Equality (ipeqopt)					.038	**.591**	**.547**
2	Understanding Diff. People (ipudrst)					.023	.040	**.477**
3	Environment (impenv)					.024	.021	.049
	Means	**3.890**	**3.589**	**3.845**		**3.894**	**3.593**	**3.832**
	ICC					.036	.035	.046

Note: Bold entries in the upper diagonal are the correlations, entries in the diagonal are variances, and entries in the lower diagonal are covariances; the total sample includes 43,779 respondents from 26 countries

Two-Level CFA and Two-Level SEM for Universalism

Respondent Level (n = 43,779) Country Level (n = 26)	Model 1: Two-Level CFA			Model 2 Multilevel SEM		
AIC	**368050.207**			**368042.483**		
SRMR Within						
SRMR Between	**.062**			**.045**		
	b	z		b	z	
Confirmatory Factor Analysis						
Intercept Level 2						
Equality (ipeqopt)	3.894	100.478**		2.830	4.968**	
Underst. Diff. People (ipudrst)	3.594	90.518**		2.613	4.795**	
Environment (impenv)	3.832	87.229**		4.680	6.384**	
Factor Loadings	b	z	beta	b	z	beta
Equality (ipeqopt)	1.000	-.-	1.000	1.000	-.-	.707
Underst. Diff. People (ipudrst)	.608	3.666**	.592	.921	3.197**	.637
Environment (impenv)	.625	3.277**	.549	1.747	4.599**	1.090
Equality (ipeqopt)	1.000	-.-	.568	1.000	-.-	.568
Underst. Diff. People (ipudrst)	1.069	57.275**	.585	1.069	57.275**	.585
Environment (impenv)	.960	58.203**	.550	.960	58.202**	.550
Regression						
Environment (impenv) on HDI 2004				**-2.965**	**-3.757****	**-.712**
Universalism (betw.) on HDI 2004				**1.165**	**1.871***	**.449**

Note: * $p \leq 0.05$; ** $p \leq 0.01$; Estimator: Full Maximum Likelihood (ML);
Estimates for level 2 parameters are indented to the right in the first column. Variances/Residual Variances tested one-tailed. Since we formulated hypotheses for the impact of the HDI on environment and universalism (between), the significance level of both b-coefficients are based on a one-tailed test. The beta-coefficients are fully standardized.

Two-Level CFA and Two-Level SEM for Universalism

Respondent Level (n = 43,779) Country Level (n = 26)	Model 1: Two-Level CFA		Model 2 Multilevel SEM	
AIC	**368050.207**		**368042.483**	
SRMR Within				
SRMR Between	**.062**		**.045**	
Residual Variances	Variance	z	Variance	
Equality (ipeqopt)	-.-	-.-	.019	3.478**
Underst. Diff. People (ipudrst)	.026	3.504**	.024	3.477**
Environment (impenv)	.035	3.511**	-.-	-.-
Equality (ipeqopt)	.703	93.393**	.703	93.391**
Underst. Diff. People (ipudrst)	.735	88.396**	.735	88.395**
Environment (impenv)	.711	98.524**	.711	98.525**
Variance	Variance	z	Variance	z
Latent Factor (Universalism betw.)	.038	3.542**		
Latent Factor (Universalism within)	.334	42.894**		
Variance Comp./Residual Var. Level 2				
Intercept Level 2: Universalism (betw.)			.015	1.943*
Level 1: Universalism (within)			.334	42.894**

Note: * $p \leq 0.05$; ** $p \leq 0.01$; Estimator: Full Maximum Likelihood (ML);
Estimates for level 2 parameters are indented to the right in the first column. Variances/Residual Variances tested one-tailed. Since we formulated hypotheses for the impact of the HDI on environment and universalism (between), the significance level of both b-coefficients are based on a one-tailed test.

Two-Level CFA

Respondent Level (within): Country Level (between):

$$y_{ijk} = \alpha_{jk} + \lambda_{Wk} \cdot \eta_{Wij} + \varepsilon_{Wijk} \qquad \alpha_{jk} = \upsilon_{k} + \lambda_{Bk} \cdot \eta_{Bj} + \varepsilon_{Bjk}$$

where

- y_{ijk} refers to the observed value of respondent *i* of country *j* on indicator variable *k*,
- α_{jk} refers to the intercept of country *j* on indicator variable *k*,
- υ_{k} refers to the intercept (usually called grand mean in multilevel analysis) of indicator variable *k*,
- η_{Wij} refers to the score of respondent *i* of country *j* on the within-level latent η_{W} ,
- η_{Bj} refers to the score of country *j* on the between-level latent variable η_{B} ,
- λ_{Wk} refers to the within-level factor loading λ_{W} of indicator variable *k*,
- λ_{Bk} refers to the between-level factor loading λ_{B} of indicator variable *k*,
- ε_{Wijk} refers to the within-level error term ε_{W} for respondent *i* of country *j* on indicator variable *k*, and
- ε_{Bjk} refers to the between-level error term ε_{B} (usually called random term in multilevel analysis) for country *j* on indicator variable *k*.

Two-Level CFA

$$y_{ijk} = \upsilon_k + \lambda_{Bk} \cdot \eta_{Bj} + \varepsilon_{Bjk} \quad + \quad \lambda_{Wk} \cdot \eta_{Wij} + \varepsilon_{Wijk}$$

where

- y_{ijk} refers to the observed value of respondent *i* of country *j* on indicator variable *k*,
- η_{Bj} refers to the score of country *j* on the between level latent variable η_B ,
- η_{Wij} refers to the score of respondent *i* of country *j* on the within level latent η_W ,
- υ_k refers to the intercept (usually called grand mean in multilevel analysis) of indicator variable *k*,
- λ_{Bk} refers to the between level factor loading λ_B of indicator variable *k*,
- λ_{Wk} refers to the within level factor loading λ_W of indicator variable *k*,
- ε_{Bjk} refers to the between level error term ε_B (usually called random term or u_{jk}-term in multilevel analysis) for country *j* on indicator variable *k*, and
- ε_{Wijk} refers to the within level error term ε_W for respondent *i* of country *j* on indicator variable *k*.

One-Level CFA vs. Two-Level CFA for Universalism

Respondent Level (n = 43,779) Country Level (n = 26)	Model 1a: One-Level CFA			Model 1b: Two-Level CFA		
AIC	**372209.497**			**368050.207**		
SRMR Within						
SRMR Between				**.062**		
	b	z		b	z	
Intercept Level 1						
Intercept Level 2						
Equality (ipeqopt)				3.894	100.478**	
Underst. Diff. People (ipudrst)				3.594	90.518**	
Environment (impenv)				3.832	87.229**	
Equality (ipeqopt)	3.890	783.571**				
Underst. Diff. People (ipeqopt)	3.589	696.455**				
Environment (impenv)	3.845	779.247**				
Factor Loadings	b	z	beta	b	z	beta
Equality (ipeqopt)				1.000	-.-	1.000
Underst. Diff. People (ipudrst)				.608	3.666**	.592
Environment (impenv)				.625	3.277**	.549
Equality (ipeqopt)	1.000	-.-	.578	1.000	-.-	.568
Underst. Diff. People (ipudrst)	1.069	60.117**	.596	1.069	57.275**	.585
Environment (impenv)	.964	60.997**	.561	.960	58.203**	.550

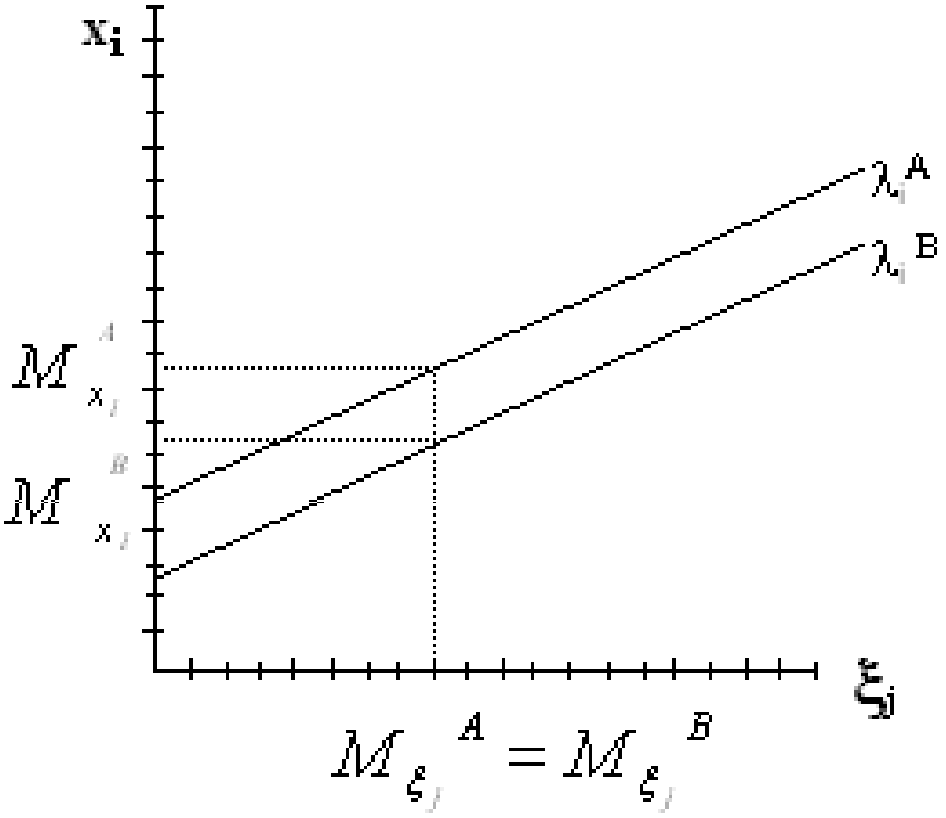
* $p \leq 0.05$; ** $p \leq 0.01$; Estimator: Full Maximum Likelihood (ML); the beta-coefficients are fully standardized; the residual variance of equality at level 2 turned out to be insignificant and has been fixed for that reason

One-Level CFA vs. Two-Level CFA for Universalism

Respondent Level (n = 43,779) Country Level (n = 26)	Model 1a: One-Level CFA		Model 1b: Two-Level CFA	
AIC	**372209.497**		**368050.207**	
SRMR Within				
SRMR Between			**.062**	
Residual Variances	Variance	z	Variance	z
Equality (ipeqopt)			-.-	-.-
Underst. Diff. People (ipudrst)			.026	3.504**
Environment (ipeqopt)			.035	3.511**
Equality (ipeqopt)	.718	93.533**	.703	93.393**
Underst. Diff. People (ipudrst)	.750	88.392**	.735	88.396**
Environment (ipeqopt)	.730	98.486**	.711	98.524**
Variance			Variance	z
Latent Factor (Universalism betw.)			.038	3.542**
Latent Factor (Universalism within)	.360	44.711**	.334	42.894**

* $p \leq 0.05$; ** $p \leq 0.01$; Estimator: Full Maximum Likelihood (ML); the residual variance of equality at level 2 turned out to be insignificant and has been fixed for that reason

Intercepts and Slopes:
Equal Slopes and Unequal Intercepts



Using mulilevel technique to explain measurement invariance

- The basic idea behind a two-level confirmatory factor analysis is to **decompose the variability** of the indicator variables **into** a respondent-level **“within” variability** and a country-level (group level) **“between” variability**.

- It **allows**, in a subsequent step, to **account for differences** in the **parameters between groups** by **including contextual variables**.

- In this way, the **multilevel CFA** (cf. Muthén 1994; Hox 2002) is extended to a **multilevel SEM** (cf. Muthén 1994; Selig et al. 2008), which allows the explanation of measurement noninvariance by contextual variables.