

Testing the Cultural Differences of School Characteristics with Measurement Invariance

Ergül Demir¹

¹ Faculty of Educational Sciences Measurement and Evaluation Department, Ankara University, Ankara, Turkey
Correspondence: Ergül Demir, Faculty of Educational Sciences Measurement and Evaluation Department, Ankara University, Ankara, Turkey. E-mail: erguldemir@ankara.edu.tr

Received: April 8, 2016

Accepted: April 22, 2016

Online Published: April 26, 2016

doi:10.5539/jel.v5n2p337

URL: <http://dx.doi.org/10.5539/jel.v5n2p337>

Abstract

In this study, it was aimed to model the school characteristics in multivariate structure, and according to this model, aimed to test the invariance of this model across five randomly selected countries and economies from PISA 2012 sample. It is thought that significant differences across group in the context of school characteristics have the potential to explain the effectiveness of schools and educational systems. This study was conducted with correlational model as a basic research. Secondary level analyses were conducted on PISA 2012 School Questionnaire data. To construct “school characteristics model”, whole data from 65 participant countries and economies were considered. One country from each proficiency level and totally 5 countries were randomly selected for the research sample. These countries and economies are Shanghai-China, Korea, Ireland, Turkey and Uruguay. In this way sample was composed of totally 835 schools. Multi-group confirmatory factor analysis was used to test the invariance of school characteristics across countries. According to the results, Shanghai and Uruguay differed from each other and other countries. Across Korea, Ireland and Turkey, School characteristics provide strong invariance. These three cultures were more similar. Main result of this study is that school characteristics cannot be invariant across some cultural groups or sub-groups. In order to provide equal opportunity to all stakeholders of the educational system, and also provide school effectiveness, such kinds of differences are considered carefully.

Keywords: measurement invariance, multi-group confirmatory factor analysis, school characteristics, cultural differences, PISA 2012

1. Introduction

As the state populations become more diverse and international communications grow faster, the issue of culture and its effects becomes more important. Culture is one of the important key concepts in this context. As a general definition, culture is defined as “the values, norms, and traditions that affect how individuals of a particular group perceive, think, interact, behave, and make judgments about their world” (Chamberlain & Medeiros-Landurand, 1991). Beside the classical definitions, culture can be evaluated as a key in order to understanding the nature of the main differences across countries or some sub-groups, like gender subgroups, socioeconomic subgroups, minorities, educational sub-groups, etc. Time living in a culture, experience of interacting, increased endorsement of a main culture and decreased endorsement of one’s own culture could explain the differences across cultures (Bjornsdottir & Rule, 2016).

Culture is an important variable which has become widely considered in education. Many studies indicate that there are relationships between culture and individual learning style (Joy & Kolb, 2009), and between teachers cultural background and students cultural background (Hofstede, 1986), between school characteristics and immigration status (Rich, Ari, Amir, & Eliassy, 1996); between culture and student achievement (Hirchfeld & Brown, 2009; Husen, 1967; Kao & Thompson, 2003; Oakes, 1990), between culture and using computers (Li & Kirkup, 2007), culture and linguistic diverse of the students (Chamberlain, 2005). All these research provides some solutions to get educational systems better and more effective. Generally, findings of these researches indicate that there are significant differences across cultures and it is possible that these differences should be the source of inequalities of opportunities in education against some groups. For example Oakes (1990) emphasized that women and minorities, especially in inner cities, had lower educational opportunities and resources.

It is obvious that there are many latent dimensions of culture and it is difficult to define. For this reason, it is difficult to understand how and which cultural differences create problems in education (Gudykunst, 1997). There are some technical approaches in order to test the differences across cultures or some other groups. Some of these techniques are discussed under the concept of “measurement invariance”.

Measurement invariance is defined as equivalence of a psychological structure or model across groups. Initial definitions of this concept indicate that measurement invariance is one of the important psychometric characteristic of the measurement tools (Byrne, 2006; Hambleton, 1994; Jöreskog & Sörbom, 1996). It is stated that if groups are going to be compared each other in a specific characteristic, principally, this characteristic which is modelled with a psychological structure should be invariant across groups (Byrne, 2008; Milfont & Fischer, 2010; Wu, Li, & Zumbo, 2007). In this context, providing invariance indicates the generalizability of the psychological structures and that psychological structures aren't affected by differences of the groups. And also, bias which result from group differences are not significant (Meredith, 1993).

Technical and conceptual framework, and also process of measurement invariance, was determined by Jöreskog (1971). This technique was based on “variance-covariance structures model”, and it is widely known as “Multi-Group Confirmatory Factor Analysis (MG-CFA)” (Jöreskog, 1971; Little, 1997). This is a multivariate and includes hierarchical steps. According to the structure of the initial model, four or five steps can be followed forward steps model from 1 to 5 or backward steps model from 5 to 1, respectively. Testing invariance begin with an initial model. This model is multivariate and/or multilevel. And also, it is often about complex structures. For further steps, it is recommended that the initial model should be strong and high level of model-data fit (Meredith, 1993; Horn & McArdle, 1992).

Commonly, forward steps model form 1 to 5 is preferred to investigate the invariance. This way, if there is significant violation at a level, analysis can be stopped and made a decision about invariance level. Otherwise, steps of the analyses must be continued until defining the point of invariance level. For forward steps model, at first step, also known as full free model, all model parameters and path coefficients are set free for all groups. If there are significant model-data fit within each group, it means that “configural invariance” can be provided across groups. At second step, factor loadings or path coefficients between observed variables and latent variables are limited as constant. If there is no significant difference of model-data fit between first and second step, it means that “metric invariance” can be provided across groups. Configural and metric invariances are also known as “weak invariance”. At third step, also known as “strong invariance”, factor correlations, as well as factor loadings, are limited as constant. If there is no significant change in model-data fit, it means that “scalar invariance” can be provided. Finally, at fourth step, error variances of observed variables, as well as factor loadings and factor correlations, are limited as constant. If there is no significant difference, it means that “strict invariance” or “full invariance” can be provided across groups (Van de Schoot, Lugtig, & Hox, 2012; Widaman & Reise, 1997).

Initial studies on measurement invariances are very technical and on the theoretical baseline (Byrne, Shavelson, & Muthén, 1989; Cheung & Rensvold, 2002; Horn & McArdle, 1992; Jöreskog, 1971; Little, 1997; McArdle & Cattle, 1994; Meredith, 1993). Also, measurement invariance is used for providing validity evidence of a measurement tool (Hambleton, 1994; Gerber, Carvacho, & Gonzales, 2016; Grouzet, Otis, & Pelletier, 2006), testing measurement bias across sub-groups (Oakes, 1990), examining the gender differences or gender gap (Hirschfeld & Brown, 2009; Li & Kirkup, 2007; Oakes, 1990). In addition to these studies, measurement invariance is widely used for defining cultural differences in education and for understanding the nature of education (Glanville & Wildhagen, 2007; Joy & Kolb, 2009; Marsh, Hau, Artelt, Baumert, & Peschar, 2006; Stein, Lee, & Jones, 2006; Wu, Li, & Zumbo, 2007). Although school characteristics and educational environment impact on the effectiveness of the educational system (Heck & Marcoulides, 1996), relationship between culture and school characteristics with measurement invariance is less studied issue. As one of the quite rare example, Rich et al. (1996) investigated the school characteristics and impact of these characteristics on the immigrant students who recently immigrated to Israel from the former Soviet Union. They found that school characteristics delayed the integration across cultures. Appropriate schools were needed for schooling of immigrant students. As another example, Heck (1996) researched the invariance of administrative leaderships on school effectiveness across two schools from two different cultural settings. He observed that measurement invariance of the leaderships was not provided across cultural settings. According to this finding, he argued that there were some educational disadvantages for one of these two schools. These limited numbers of researches indicate that differences of school characteristics across cultural groups have the potential in order to understand and explain some important problems of educational systems.

In this study, it was aimed to construct the “school characteristics model” and to test the invariance of this model across five randomly selected countries and economies from PISA 2012 sample. It was investigated that whether this model provided invariance or not. And also, it was tried to define the significant differences across these cultural groups. It is thought that significant differences across group in the context of school characteristics have the potential to explain the effectiveness of schools and educational systems.

2. Methods

2.1 Research Design

This study was conducted with correlational model as a basic research. With basic researches, it is aimed to provide pure information about facts and phenomena. And, with correlational model, it is tried to explain the relationships between variables (Bryman, 2015; Slavin, 1992). In this study, it was aimed to explain the relations between school characteristics and cultural differences. And also, it was tried to determine the cultural differences in the context of school characteristics.

2.2 Research Sample

In this study, secondary level analyses were conducted on PISA 2012 School Questionnaire data (Note 1). This data set includes totally 18,139 schools from 65 different countries and economies. As an average, there are 279 schools for each. Mexico (8.1%) and Italy (6.6%) have the highest rates in the school samples. Macao-China (0.1%), Luxemburg (0.2%) and Liechtenstein (0.2%) have the lowest rates. For 52 out of 65 countries, this ratio is between 1% and 2%.

Data for all 65 countries and economies were considered in order to construct school characteristics model. After constructing this model, invariance of the model were testing across 5 countries and economies from all participants. For selecting countries, mathematical literacy proficiency levels were considered. In PISA 2012, mathematical literacy was the main domain for student assessment. For each domain, six proficiency levels are defined (OECD, 2013). In this study, it was decided that sample should include one country or economy from each level at least. According to the average scores of student achievements, there are no countries and economies at 6th level. Shanghai-China is only one economy placed at 5th level. Other levels include more than one country or economy. 4th level includes 4, 3rd level includes 28, 2nd level includes 19 and 1st level includes 13 countries and economies. One country from each level was selected randomly. With this way, totally 5 countries and economies were determined as sample of cultural group. Selected countries/economies and number of schools in these countries/economies are shown at the following table.

Table 1. Countries/economies in the sample and number of schools

Country/Economy	OECD Membership	Proficiency Level	Number of School
Shanghai-China	No	5	155
Korea	Yes	4	153
Ireland	Yes	3	183
Turkey	Yes	2	170
Uruguay	No	1	174
Toplam			835

As seen at Table 1, totally 835 schools were placed in the sample. Numbers of schools are close to each other. Two countries are the memberships of OECD.

2.3 Data Collection Tool

In this study, analyses were conducted on the PISA 2012 School Questionnaire data. School questionnaire is one of the instruments that providing information about background of the educational systems of the countries and economies. With this tool, some important information can be gathered about (1) the structure and organization of the school, (2) the student body and teachers, (3) the school’s resources, (4) the school’s instruction, curriculum and assessment, (5) the school climate, (6) the school’s policies and practices and optionally and (7) financial education at school. So, it includes 6 or 7 sections and 30 to 40 items. It was filled by school administrators (OECD, 2013).

PISA 2012 Student Questionnaire data set includes separate variables for each sub-content. Also, there are index and continuous variables which are derived from these variables in the data set. These indexes are representing the sub-contents as standardized total scores. And, these are appropriate and available for further multilevel and advanced statistical analyses (OECD, 2014).

2.4 Data Analysis

In this study, “principal component analysis” and “confirmatory factor analysis” were used to construct the school characteristic model. After constructing the model, “Multi-Group Confirmatory Factor Analysis (MG-CFA)” was used to test measurement invariance of “school characteristics model” across cultural groups. Within the MG-CFA technique, measurement invariance was defined and considered four hierarchical steps: (1) configural invariance, (2) metric invariance, (3) scalar invariance and (4) strict invariance. First two steps were evaluated as weak invariance and third step was evaluated as strong invariance as well. Significance of χ^2 differences and difference of goodness of fit statistics between consecutive steps were considered to evaluate invariance. If χ^2 differences were statistically significant and differences of most of the goodness of fit statistics were higher than 0.01, then it was decided to violate the invariance (Jöreskog & Sörbom, 1996; Cheung & Rensvold, 2002).

As known, these techniques are multivariate. So, variables must be interval at least. For this reason, it was preferred to use the index variables and continuous variables in the PISA 2012 School Questionnaire data set. There are close to 300 variables defined in this data set. Most of them are categorical. On the other hand, some groups of continuous variables are defined with index or total standardized scores. According to these preliminary reviews, in this study, it was decided that 44 indexes and continuous variables were available for further analyses. Analyses for constructing the model and testing invariance were executed on these variables.

3. Results

3.1 Constructing the School Characteristics Model

As explained before, totally 44 index variables and continuous variables could be defined as available for constructing the initial model. First of all, these variables were checked for basic statistical assumptions; missing value, extremes, normality, linearity and multicollinearity. Variables providing these assumptions were preferred to include the initial regression model.

It was seen that there were possible missing data problem for most of these 44 variables. 27 out of 44 variables have missing values over 5%, 15 variables have over 10%, and 8 variables have over 15%. It is well-known that missing data leads to decrease the power of analysis, and it is possible source of bias. Although there are no strict criteria, especially for multivariate statistical techniques, it is recommended that there is no missing over 5% each variable. Also, if missing data mechanism is providing the complete or partial randomization, it is stated that some missing data methods like EM algorithms and multiple imputation should be used to handle with missing problem (Enders, 2010; Tabachnick & Fidell, 2013). In this context, it was decided that 17 out of 44 variables including missing values under 5% should be considered for further analysis. And, values were imputed to missing by using EM algorithms method, and completed data set were obtained.

On the completed data set, these 17 variables were checked for extreme. For this purpose, standardized Z scores were obtained for each variable. Observation units which had Z scores out of the range (-3, +3) were evaluated as extreme. And, totally 225 units were removed from the data set.

And then, these variables were checked for normality with graphical and descriptive methods. Except two variables, it was seen that 15 variables met normality assumption. For these 15 variables, mean and mod and median were close each other. Skewness and kurtosis indexes were between -1 and +1, and close to 0. Histograms and P-P Plots and Q-Q plots indicated that variables contributed normally. Two variables were removed from data set, and further analysis was continued with 15 variables.

Linearity was checked with Pearson correlation coefficient, and multicollinearity was checked with collinearity diagnostic statistics. It was seen that highest bivariate correlation between variables was 0.614. Other bivariate and partial correlations were under this value. Tolerances were over 0.30 and Variance Inflation Factors (VIF) were under 3.00 and Condition Indexes (CI) were under 15. These statistics indicated that there was no multicollinearity problem.

After controlling basic assumptions, principal component analysis was executed with 15 variables. Number of variables and observations were available for this analysis (KMO = 0.762, for Bartlett' test $\chi^2= 80503.354$, $df= 105$ and $p < 0.01$). It was seen that one variable (class size) had low extraction value (0.167) and factor loading

(under 0.27). Also, this variable had the same factor loading for more than one factor and there was an overlapping. So, it was decided that this variable should be removed from model. For other 14 variables, extractions were changing between 0.441 and 0.784. There were 4-factors structure and total variance explained by 4 factors was 64.10%. Respectively, variances of each factor were 21.74%, 20.30%, 12.89% and 9.17%. Also, bivariate correlations between factors are very low and close to 0 ($r_{12} = 0.10$, $r_{13} = 0.07$, $r_{14} = -0.26$, $r_{23} = -0.05$, $r_{24} = -0.06$ and $r_{34} = -0.17$). The distributions of the variables to the factors and factor loadings are shown with a pattern matrix at the following table.

Table 2. Distributions of the variables to the factors and factor loadings

	Factors			
	1	2	3	4
Teacher morale	.747			
Teacher focus	.725			
Teacher related factors affecting school climate	.709			
Student-related factors affecting school climate	.639			
Instructional Leadership		.882		
Framing and communicating the school's goals and curricular development		.838		
Teacher participation in leadership		.829		
Promoting instructional improvements and professional development		.799		
School autonomy			.824	
Index of school responsibility for curriculum and assessment			.809	
Teacher participation/autonomy			.803	
Quality of school educational resources				-.849
Quality of physical infrastructure				-.828
Shortage of teaching staff				.611

Note. Extraction Method: Principal Component Analysis.

Rotation Method: Oblimin with Kaiser Normalization.

As seen at Table 2, first and second factors include four variables, and other factors include 3 variables. These factors were named respectively as “School Climate (SCCL)”, “School Leadership and Development (SCLD)”, “School Autonomy (SAUT)” and “School Resources (SCRS)”.

After defining the factors, confirmatory factor analysis was executed with this 4 factors and 14 variables structure. Some modifications were made to improve model-data fit. According to the results, it was seen that significant model with perfect model-data fit could be obtained. Some goodness of fit statistics are shown at the following table.

Table 3. Goodness of fit statistics for school characteristics model

	Statistics
Normal Theory Weighted Least Squares Chi-Square (X^2)	3376.64 (df = 66 and p < 0.05)
Root Mean Square Error of Approximation (RMSEA)	0.053
Root Mean Square Residual (RMR)	0.046
Standardized (RMR)	0.044
Goodness of Fit Index (GFI)	0.97
Adjusted Goodness of Fit Index (AGFI)	0.96
Normed Fit Index (NFI)	0.97
Non-Normed Fit Index (NNFI)	0.95
Comparative Fit Index (CFI)	0.97
Incremental Fit Index (IFI)	0.97
Relative Fit Index (RFI)	0.95

As seen at Table 3, school characteristics model with 4 factors and 14 variables show perfect model-data fit. X^2 is statistically significant. For error terms, RMSEA, RMR and SRMR are under 0.053. And, other model-fit indices are over 0.95. These results indicate that school characteristics model with 4 latent variables and 14 observed variables could be constructed significantly with high level of model-data fit. Conceptual diagram of this model is shown at following graphic. And also, standardized path coefficients are shown at following table.

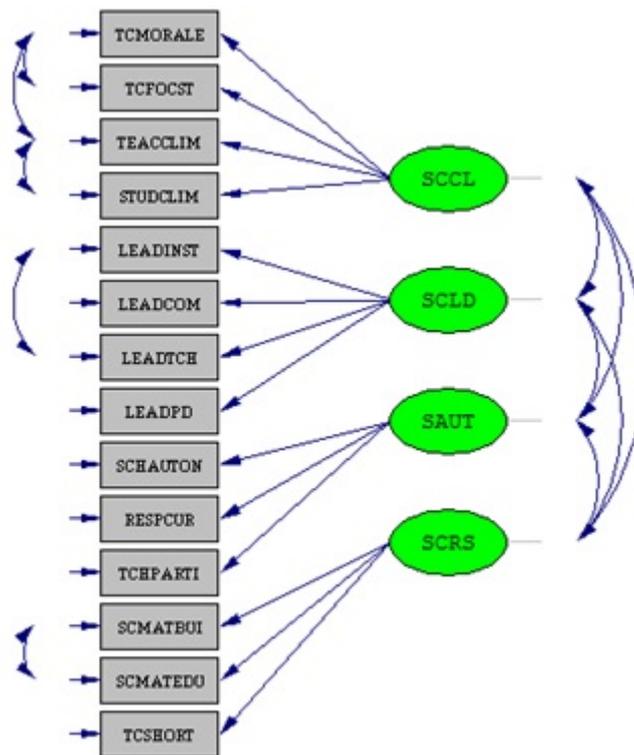


Figure 1. Conceptual diagram of school characteristics model

Table 4. Standardized path coefficients of school characteristics model

Error Terms	Observed Variables	Path Coefficients	Latent Variables
0.69	TCMORALE	0.55	
0.89	TCFOCST	0.33	
0.18	TEACCLIM	0.91	SCCL
0.63	STUDCLIM	0.60	
0.19	LEADINST	0.90	
0.39	LEADCOM	0.78	
0.36	LEADTCH	0.80	SCLD
0.55	LEADPD	0.67	
0.30	SCHAUTON	0.84	
0.57	RESPCUR	0.66	SAUT
0.63	TCHPARTI	0.61	
0.71	SCMATBUI	0.54	SCRS
0.65	SCMATEDU	0.59	
0.61	TCSHORT	-0.62	

* All statistics and paths are statistically significant at the 0.05 significance level.

As seen at Table 4, standardized error terms are positive and under 0.90. Except one, all variables are positive predictors. According to the path coefficients or factor loadings, best predictors of school characteristics are “Teacher Related Factors Affecting School Climate (TEACCLIM)” and “Instructional Leadership (LEADINST)”. One unit change of these variables leads 0.9 unit change of school characteristics. It is followed by “School Autonomy (SCHAUTON)” and “Teacher Participation in Leadership (LEADTCH)” respectively. One unit change of these variables leads 0.8 unit change for total structure.

3.2 Invariance of School Characteristics Model across Cultural Groups

Invariance of school characteristics model was tested across six different countries with four step model. For each step, goodness of fit statistics across groups and general were calculated. All these statistics are shown at Table 5.

Table 5. Goodness of fit statistics for invariance of school characteristics model across countries

Invariance Steps	Country/ Economy	Group Goodness of Fit Statistics			Model Goodness of Fit Statistics				χ^2	df	p
		GFI	RMR	χ^2 Contribution (%)	NFI	IFI	CFI	RMSEA			
Configural Invariance	Shanghai	0.92	0.075	19.35							
	Korea	0.93	0.060	15.70							
	Ireland	0.91	0.061	22.33	0.90	0.96	0.96	0.061	532.35	330	0.0
	Turkey	0.93	0.052	18.89							
	Uruguay	0.90	0.088	23.73							
Metric Invariance	Shanghai	0.86	0.180	23.32							
	Korea	0.91	0.090	12.95							
	Ireland	0.90	0.090	19.26	0.86	0.93	0.93	0.077	766.82	386	0.0
	Turkey	0.91	0.076	17.38							

	Uruguay	0.84	0.150	27.08							
	Shanghai	0.86	0.190	22.80							
	Korea	0.91	0.10	12.67							
Scalar Invariance	Ireland	0.89	0.10	18.90	0.85	0.92	0.92	0.079	837.14	410	0.0
	Turkey	0.90	0.087	17.07							
	Uruguay	0.81	0.18	28.56							
	Shanghai	0.78	0.28	19.23							
	Korea	0.82	0.13	15.95							
Strict Invariance	Ireland	0.84	0.16	19.20	0.74	0.81	0.81	0.12	1502.2	466	0.0
	Turkey	0.83	0.12	20.79							
	Uruguay	0.74	0.20	24.82							

As seen at Table 5, at first step of invariance, configural invariance are provided each culture. Model-data fit indices are at acceptable range. RMSEA and RMR are between 0.05 and 0.10. GFI and NFI are between 0.90 and 0.96. On the other hand IFI and CFI indices are over 0.95. Also, X^2 is statistically significant ($p < 0.05$) and X^2/df ratio is 1.613. These three values indicate to the perfect model-data fit. As a result, it is seen that school characteristics model are statistically significant and provides high model-data fit for each country and economy.

At second step, it is seen that metric invariance are not provided across cultures. Differences between indices are higher than 0.01 ($\Delta NFI = 0.04$, $\Delta IFI = 0.03$, $\Delta CFI = 0.03$, $\Delta RMSEA = 0.016$). Also, differences of X^2 is significant ($\Delta X^2 = 234.47$, $df = 56$ and $p < 0.10$). Because metric invariance cannot be provided, it is obvious that scalar and strict invariances cannot be provided. Similar to the previous one, for other steps, it can be seen that differences between model-data fit indices were over 0.01, and also differences of X^2 is not significant. To support this, changes of the group model-data fit indices for each country or economy indicate that school characteristics are not invariant in itself. This finding indicates that there are significant differences across cultures in the context of the parameters of the model.

As further and deeper analyses, measurement invariance was tested for all possible pairwise of cultures in order to understand the reason and resource of the violation. Totally 10 pairwise comparisons were executed separately. According to the results of the comparisons;

- For Shanghai and Korea, metric invariance and further steps cannot be provided. For first and second steps, differences between indices are higher than 0.01 ($\Delta NFI = 0.03$, $\Delta IFI = 0.03$, $\Delta CFI = 0.04$, $\Delta RMSEA = 0.01$). Also, differences of X^2 is significant ($\Delta X^2 = 54.81$, $df = 14$ and $p < 0.005$). There are just configural and weak invariance between Shanghai and Korea in School Characteristics Model.

- For Shanghai and Ireland, metric invariance and further steps cannot be provided. For first and second steps, differences between indices are higher than 0.01 ($\Delta NFI = 0.03$, $\Delta IFI = 0.03$, $\Delta CFI = 0.04$, $\Delta RMSEA = 0.011$). Also, differences of X^2 is significant ($\Delta X^2 = 60.76$, $df = 14$ and $p < 0.005$). There are just configural and weak invariance between Shanghai and Ireland in School Characteristics Model.

- For Shanghai and Turkey, metric invariance and further steps cannot be provided. For first and second steps, differences between indices are higher than 0.01 ($\Delta NFI = 0.03$, $\Delta IFI = 0.02$, $\Delta CFI = 0.03$, $\Delta RMSEA = 0.017$). Also, differences of X^2 is significant ($\Delta X^2 = 68.49$, $df = 14$ and $p < 0.005$). There are just configural and weak invariance between Shanghai and Turkey in School Characteristics Model.

- For Shanghai and Uruguay, metric invariance and further steps cannot be provided. For first and second steps, differences between indices are higher than 0.01 ($\Delta NFI = 0.04$, $\Delta IFI = 0.03$, $\Delta CFI = 0.03$, $\Delta RMSEA = 0.014$). Also, differences of X^2 is significant ($\Delta X^2 = 74.98$, $df = 14$ and $p < 0.005$). There are just configural and weak invariance between Shanghai and Uruguay in School Characteristics Model.

- For Korea and Ireland, metric invariance and scalar invariance can be provided. For first and second steps, differences between indices are lower than 0.01 ($\Delta NFI = 0.00$, $\Delta IFI = 0.00$, $\Delta CFI = 0.00$, $\Delta RMSEA = 0.002$). Also, differences of X^2 is not significant ($\Delta X^2 = 16.36$, $df = 14$ and $p > 0.10$). For second and third steps, differences between indices are lower than 0.01 ($\Delta NFI = 0.01$, $\Delta IFI = 0.00$, $\Delta CFI = 0.00$, $\Delta RMSEA = 0.001$).

Also, differences of X^2 is not significant ($\Delta X^2 = 8.36$, $df = 6$ and $p > 0.10$). But, strict invariance cannot be provided. For third and fourth steps, differences between indices are higher than 0.01 ($\Delta NFI = 0.03$, $\Delta IFI = 0.02$, $\Delta CFI = 0.03$, $\Delta RMSEA = 0.017$). Also, differences of X^2 is significant ($\Delta X^2 = 94.95$, $df = 14$ and $p < 0.005$). As a result, there are scalar and strong invariance between Korea and Ireland in School Characteristics Model.

- For Korea and Turkey, metric invariance and scalar invariance can be provided. For first and second steps, differences between indices are lower than 0.01 ($\Delta NFI = 0.00$, $\Delta IFI = 0.01$, $\Delta CFI = 0.00$, $\Delta RMSEA = 0.004$). Also, differences of X^2 is not significant ($\Delta X^2 = 10.95$, $df = 14$ and $p > 0.10$). For second and third steps, differences between indices are lower than 0.01 ($\Delta NFI = 0.01$, $\Delta IFI = 0.01$, $\Delta CFI = 0.00$, $\Delta RMSEA = 0.000$). Also, differences of X^2 is not significant ($\Delta X^2 = 6.56$, $df = 6$ and $p > 0.10$). But, strict invariance cannot be provided. For third and fourth steps, differences between indices are higher than 0.01 ($\Delta NFI = 0.08$, $\Delta IFI = 0.08$, $\Delta CFI = 0.08$, $\Delta RMSEA = 0.039$). Also, differences of X^2 is significant ($\Delta X^2 = 148.12$, $df = 14$ and $p < 0.005$). As a result, there are scalar and strong invariance between Korea and Turkey in School Characteristics Model.

- For Korea and Uruguay, metric invariance and further steps cannot be provided. For first and second steps, differences between indices are higher than 0.01 ($\Delta NFI = 0.01$, $\Delta IFI = 0.01$, $\Delta CFI = 0.01$, $\Delta RMSEA = 0.01$). Also, differences of X^2 is significant ($\Delta X^2 = 57.97$, $df = 14$ and $p < 0.005$). There are just configural and week invariance between Korea and Uruguay in School Characteristics Model.

- For Ireland and Turkey, metric invariance and scalar invariance can be provided. For first and second steps, differences between indices are lower than 0.01 ($\Delta NFI = 0.01$, $\Delta IFI = 0.00$, $\Delta CFI = 0.00$, $\Delta RMSEA = 0.001$). Also, differences of X^2 is not significant ($\Delta X^2 = 21.89$, $df = 14$ and $p > 0.05$). For second and third steps, differences between indices are lower than 0.01 ($\Delta NFI = 0.00$, $\Delta IFI = 0.00$, $\Delta CFI = 0.00$, $\Delta RMSEA = 0.000$). Also, differences of X^2 is not significant ($\Delta X^2 = 9.14$, $df = 6$ and $p > 0.10$). But, strict invariance cannot be provided. For third and fourth steps, differences between indices are higher than 0.01 ($\Delta NFI = 0.07$, $\Delta IFI = 0.07$, $\Delta CFI = 0.07$, $\Delta RMSEA = 0.018$). Also, differences of X^2 is significant ($\Delta X^2 = 96.04$, $df = 14$ and $p < 0.005$). As a result, there are scalar and strong invariance between Ireland and Turkey in School Characteristics Model.

- For Ireland and Uruguay, metric invariance and further steps cannot be provided. For first and second steps, differences between indices are higher than 0.01 ($\Delta NFI = 0.03$, $\Delta IFI = 0.03$, $\Delta CFI = 0.03$, $\Delta RMSEA = 0.014$). Also, differences of X^2 is significant ($\Delta X^2 = 83.56$, $df = 14$ and $p < 0.005$). There are just configural and week invariance between Ireland and Uruguay in School Characteristics Model.

- For Turkey and Uruguay, metric invariance and further steps cannot be provided. For first and second steps, differences between indices are higher than 0.01 ($\Delta NFI = 0.03$, $\Delta IFI = 0.04$, $\Delta CFI = 0.03$, $\Delta RMSEA = 0.021$). Also, differences of X^2 is significant ($\Delta X^2 = 103.87$, $df = 14$ and $p < 0.005$). There are just configural and week invariance between Turkey and Uruguay in School Characteristics Model.

In summary, invariance levels across cultures are shown at the following table.

Table 6. Results of the pairwise comparisons for invariance of school characteristics model

	Shanghai	Korea	Ireland	Turkey
Korea	Week (Configural)			
Ireland	Week (Configural)	Strong (Scalar)		
Turkey	Week (Configural)	Strong (Scalar)	Strong (Scalar)	
Uruguay	Week (Configural)	Week (Configural)	Week (Configural)	Week (Configural)

As seen at Table 6, there are week invariances between Shanghai and other countries and between Uruguay and other countries. Except Shanghai and Uruguay, there are not strict but strong invariances between other three countries. Given that it is difficult to provide the strict invariance in education, it can be claimed that school characteristics remain invariant across these cultures.

It is clear that Shanghai and Uruguay differ from other three countries in school characteristics. Indeed, best predictors of school characteristics are “teacher related factors affecting school climate” and “student-related factors affecting school climate” in Shanghai and “school autonomy” and “teacher participation in leadership” in Uruguay. On the other hand, predictive levels to the school characteristics and order of predictors are close to each other in Korea, Ireland and Turkey. In these countries, best predictors are “teacher morale” and “instructional leadership”. On the other hand, lowest predictors are “teacher morale” and “teacher focus” in

Shanghai, while “index of school responsibility for curriculum and assessment” and “teacher focus” in other countries. In addition, correlations between factors or latent variables are higher in Shanghai and Uruguay than the other countries in general.

Furthermore, if paid attention, it can be seen that Shanghai is located at the highest proficiency level and Uruguay is located at the lowest proficiency level. This finding indicates the possibility that school characteristics can be correlated with academic performance of the students.

4. Discussion

One of the important results is that it is difficult to construct the psychological multivariate model like school characteristics. In this study, over 300 school variables were examined, 44 continuous variables were defined, but school characteristics could be structured with just 14 variables. Gudykunst (1997) emphasized that modelling the cultural characteristics was difficult in education, because these characteristics are often latent and very complex.

In this study, school characteristics could be modelled in a multivariate structure with significant and high model-data fit. This model constructed with all PISA 2012 school data. Meredith (1993) emphasized that initial model was important for such kind of multivariate analysis. He recommended that strong initial models should be structured with high level of model-data fit for further analyses.

According to the “School Characteristics Model”, best predictors of school characteristics are “teacher related factors affecting school climate” and “instructional leadership”. It is followed by “school autonomy” and “teacher participation in leadership” respectively. These findings indicate the impact of the teachers and leaderships on effectiveness of the school. These findings are supported by some other researches. For example, Heck (1996) observed that the cultural settings and school leaderships were highly correlated each other. Also, he explained the differences of school effectiveness with cultural differences. In another study, Hofstede (1986) and Chamberlain (2005) emphasized the teacher role on education.

As another results, it was seen that school characteristics were different across cultures at some invariance steps. For Shanghai and Uruguay, school characteristics are completely different from other countries and each other. Just configural and weak invariance can be provided between Shanghai and Uruguay and the other countries. On the other hand, for Korea, Ireland and Turkey, these differences are lower. Not strict, but strong invariance can be provided across these countries. So, Korea, Ireland and Turkey are more homogenous cultural groups in the context of school characteristics. Although not quite similar, Wu et al. (2007) observed the differences across countries on students’ mathematical achievement according to the TIMSS 1999 results of six countries; New Zealand, Canada, USA, Taiwan, Korea and Japan. They found that, although strict invariance could be provided within each country, just configural and weak invariance could be provided across cultures. In another study, Rich et al. (1996) found that school characteristics were not invariant across some cultures and this situation led to delay the integration across cultures. In the similar way, Oakes (1990) observed that some cultural sub-groups, especially minorities, could not benefit educational resources and opportunities equally.

A remarkable detail is that Shanghai is located at the highest proficiency level and Uruguay is located at the lowest proficiency level. This finding indicates the possibility that school characteristics can be correlated with academic performance of the students. Also, it is possible that school characteristics can be predictors of students’ academic achievement.

Main result of this study is that school characteristics cannot be invariant across some cultural groups or sub-groups. In order to provide equal opportunity to all stakeholders of the educational system, and also provide school effectiveness, such kinds of differences are considered carefully. For further researches, it is recommended that school characteristics should be modelled with multilevel by using secondary level latent variables, like school achievement. Also, it is thought that these relations can be studied on the different sample.

References

- Bjornsdottir, R. T., & Rule, N. O. (2016). On the relationship between acculturation and intercultural understanding: Insight from the Reading the Mind in the Eyes test. *International Journal of Intercultural Relations*, 52, 39-48. <http://dx.doi.org/10.1016/j.ijintrel.2016.03.003>
- Bryman, A. (2015). *Social research methods* (5th ed.). UK: Oxford university press.
- Byrne, B. M. (2006). *Structural equation with EQS: Basic concepts, applications, and programming* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associations, Inc.

- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, 20(4), 872-882.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105(3), 456-466. <http://dx.doi.org/10.1037/0033-2909.105.3.456>
- Chamberlain, P., & Medeiros-Landurand, P. (1991). Practical considerations for the assessment of LEP students with special needs. In E. V. Hamayan, & J. S. Damico (Eds.), *Limiting bias in the assessment of bilingual students* (pp. 122-156). Austin, TX: PRO-ED.
- Chamberlain, S. P. (2005). Recognizing and responding to cultural differences in the education of culturally and linguistically diverse learners. *Intervention in School and Clinic*, 40(4), 195-211. <http://dx.doi.org/10.1177/10534512050400040101>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. http://dx.doi.org/10.1207/S15328007SEM0902_5
- Enders, C. K. (2010). *Applied missing data analysis*. New York: The Guilford Press.
- Gerber, M. M., Carvacho, H., & Gonzales, R. (2016). Development and validation of a scale of support for violence in the context of intergroup conflict (SVIC): The case of violence perpetrated by Mapuche people and the police in Chile. *International Journal of Intercultural Relations*, 51, 61-68. <http://dx.doi.org/10.1016/j.ijintrel.2016.01.004>
- Glanville, J. L., & Wildhagen, T. (2007). The measurement of school engagement: Assessing dimensionality and measurement invariance across race and ethnicity. *Educational and Psychological Measurement*, 67(6), 1019-1041. <http://dx.doi.org/10.1177/0013164406299126>
- Grouzet, F. M. E., Otis, N., & Pelletier, L. G. (2006). Longitudinal cross-gender factorial invariance of the academic motivation scale. *Structural Equation Modeling: A Multidisciplinary Journal*, 13(1), 73-98. http://dx.doi.org/10.1207/s15328007sem1301_4
- Gudykunst, W. B. (1997). Cultural variability in communication an introduction. *Communication Research*, 24(4), 327-348. <http://dx.doi.org/10.1177/009365097024004001>
- Hambleton, R. K. (1994). Guidelines for adapting educational and psychological tests: A progress report. *European Journal of Psychological Assessment*, 10(3), 229-244.
- Heck, R. H. (1996). Leadership and culture. *Journal of Educational Administration*, 34(5), 74-97. <http://dx.doi.org/10.1108/09578239610148287>
- Heck, R. H., & Marcoulides, G. A. (1996). School culture and performance: Testing the invariance of an organizational model. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, 7(1), 76-95. <http://dx.doi.org/10.1080/0924345960070104>
- Hirschfeld, G. H. F., & Brown, G. T. L. (2009). Students' conceptions of assessment factorial and structural invariance of the SCoA across sex, age, and ethnicity. *European Journal of Psychological Assessment*, 25(1), 30-38. <http://dx.doi.org/10.1027/1015-5759.25.1.30>
- Hofstede, G. (1986). Cultural differences in teaching and learning. *International Journal of Intercultural Relations*, 10(3), 301-320. [http://dx.doi.org/10.1016/0147-1767\(86\)90015-5](http://dx.doi.org/10.1016/0147-1767(86)90015-5)
- Horn, J. L., & McArdle, J. J. (1992). A practical and theoretical guide to measurement equivalence in aging research. *Experimental Aging Research*, 18(3), 117-144. <http://dx.doi.org/10.1080/03610739208253916>
- Husen, T. (Ed.). (1967). *International study of achievement in mathematics: A comparison of twelve countries* (Vol. 1, 2). New York: John Wiley.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36(4), 409-426. <http://dx.doi.org/10.1007/BF02291366>
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software International.
- Joy, S., & Kolb, D. A. (2009). Are there cultural differences in learning style? *International Journal of Intercultural Relations*, 33(1), 69-85. <http://dx.doi.org/10.1016/j.ijintrel.2008.11.002>

- Kao, G., & Thompson, J. S. (2003). Racial and Ethnic Stratification in Educational Achievement and Attainment. *Annual Review of Sociology*, 29, 417-442. <http://dx.doi.org/10.1146/annurev.soc.29.010202.100019>
- Li, N., & Kirkup, G. (2007). Gender and cultural differences in Internet use: A study of China and the UK. *Computers & Education*, 48(2), 301-317. <http://dx.doi.org/10.1016/j.compedu.2005.01.007>
- Little, T. D. (1997). Mean And Covariance Structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research*, 32(1), 53-76. http://dx.doi.org/10.1207/s15327906mbr3201_3
- Marsh, H., Hau, K. T., Artelt, C., Baumert, J., & Peschar, J. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing*, 6(4), 311-360. http://dx.doi.org/10.1207/s15327574ijt0604_1
- McArdle, J. J., & Cattell, R. B. (1994). Structural equation models of factorial invariance in parallel proportional profiles and oblique confactor problems. *Multivariate Behavioral Research*, 29(1), 63-113. http://dx.doi.org/10.1207/s15327906mbr2901_3
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525-543. <http://dx.doi.org/10.1007/BF02294825>
- Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3(1), 11-121.
- Oakes, J. (1990). *Multiplying inequalities: The effect of race, social class and tracking on opportunities to learn mathematics and science*. Santa Monica, CA: The Rand Corporation.
- OECD. (2013). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. PISA, OECD Publishing. <http://dx.doi.org/10.1787/9789264190511-en>
- OECD. (2014). *PISA 2012 technical report*. PISA, OECD Publishing. Retrieved April 8, 2016, from <http://www.oecd.org/pisa/pisaproducts/PISA-2012-technical-report-final.pdf>
- Rich, Y., Ari, R. B., Amir, Y., & Eliassy, L. (1996). Effectiveness of schools with a mixed student body of natives and immigrants. *International Journal of Intercultural Relations*, 20(3-4), 323-339. [http://dx.doi.org/10.1016/0147-1767\(96\)00022-3](http://dx.doi.org/10.1016/0147-1767(96)00022-3)
- Slavin, R. E. (1992). *Research methods in education* (2nd ed.). Needham Heights, MA: Allyn & Bacon.
- Stein, J. A., Lee, J. W., & Jones, P. S. (2006). Assessing crosscultural differences through use of multiple-group invariance analyses. *Journal of Personality Assessment*, 87(3), 249-258. http://dx.doi.org/10.1207/s15327752jpa8703_05
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Boston: Pearson.
- Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurement invariance. *European Journal of Developmental Psychology*, 9(4), 486-492. <http://dx.doi.org/10.1080/17405629.2012.686740>
- Widaman, K. F., & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. *The Science of Prevention: Methodological Advances from Alcohol and Substance Abuse Research*, 281-324. <http://dx.doi.org/10.1037/10222-009>
- Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment Research & Evaluation*, 12(3), 1-26.

Note

Note 1. Retrieved from <https://www.oecd.org/pisa/pisaproducts/pisa2012database-downloadabledata.htm>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).