Quantifying teacher Professional Community in 36 countries – a test for measurement invariance using the Multiple-Group Confirmatory Factor Analysis (MGCFA) method

Lomos Cătălina [1]

Abstract: This study presents the results of a test of the professional community latent concept for measurement invariance in 36 countries and more than 58000 teachers, using the International Civic and Citizenship education Study (ICCS 2009). The Multiple-Group Confirmatory Factor Analysis was used for this analysis. Teachers become part of a professional community when they agree on a common school visions, engage in reflective dialogues and collaborative practices, and feel responsible for school improvement and student learning. The study discusses the implications of having (or not having) an invariant measure of the concept for quantifying and comparing the teacher professional community practices in the different countries involved. We establish that the latent concept of professional community can be meaningfully discussed in 35 countries, excluding Liechtenstein. In addition, in 34 countries all the items that represent the five specific dimensions are strongly related to the latent concept of professional community, except for the reflective dialogue item in Switzerland. As regards scalar invariance, we identify many noninvariant intercepts in different countries, especially for items measuring the dimensions of the deprivatisation of practice and collaborative activity, allowing to compare the professional community mean practices in 23 countries only, within partial scalar invariance.

Keywords: The Latent Concept of Professional Community; Multiple-Group Confirmatory Factor Analysis (MGCFA); Cross-Cultural approach

1. Introduction

Professional communities have been studied separately in individual countries, but the large-scale international studies and reports, such as the Trends in International Mathematics and Science Study (TIMSS) or Progress in International Reading Literacy (PIRLS), take also the comparative approach on the school processes that relate to student performance (e.g. Isac, da Costa, Araújo, Soto Calvo, & Albergaria-Almeida, 2015; OECD, 2014). However, before we compare the scores of such latent school processes across countries, the issue of measurement

[1] Dr. Cătălina Lomos, Researcher
LISER (Luxembourg Institute of Socio-Economic Research), Esch-sur-Alzette, Luxembourg
catalina.lomos@liser.lu

equivalence needs to be approached, valid comparisons being conditioned by invariant measurements. More specifically, we can only perform valid comparisons of latent constructs after we make sure that the scores we are comparing are only dependent on the latent concept of interest and not on measurement noninvariance (Billiet & Welkenhuysen-Gybels, 2004).

2. Theoretical background

The concept of professional community and its measurement

Considering that teachers' collaborative work within schools encompasses many activities and facets, authors such as Bryk, Camburn, and Louis, (1999) or Louis, Marks, and Kruse, (1996) concluded that the latent concept of professional community is a multidimensional one, covering different aspects of teacher interaction. Regarding the dimensions of the latent concept of professional community, Kruse, Louis, and Bryk, (1995) ''designated five interconnected variables that describe what they called genuine professional communities in such a broad manner that they can be applied to diverse settings'' (Toole & Louis, 2002, p. 249). These five dimensions used to define one overarching latent concept have been identified by these and by other authors to be *reflective dialogue, deprivatisation of practice, collaborative activity, shared sense of purpose* and *a collective focus or responsibility for student learning*.

Measurement equivalence, a necessary test for comparisons across countries

There are different stages in establishing measurement equivalence, going from less restrictive to more restrictive models, in a bottom-up approach (van der Veld & Saris, 2011). A first stage is 'configural invariance, which means that "the latent concept can be meaningfully discussed in all countries, but it does not guarantee any cross-cultural score comparability" (Meuleman & Billiet, 2012, p.9). In technical terms, configural invariance checks whether the items measure the latent concept they are intended to measure (van der Veld & Saris, 2011) and whose confirmation allows one to say that we have the same factors and the same structure in all groups. The second stage is 'metric invariance', whose confirmation allows to evaluate and compare the presence and size of the relationships of the concept of interest with other concepts of interest across groups. In technical terms, at this higher stage the factor loadings are restrained to be the same in all groups. If full metric invariance does not hold, partial metric invariance will show us in which countries one or more items is hardly measuring the concept of professional community. And finally, in order to be able to reliably compare the presence of professional community between countries, we need to establish 'scalar invariance. 'Scalar invariance' allows to compare and to test the difference between the average degree of participation in professional community across the countries. It is a strong equivalence level, which implies that both the loadings and the intercept parameters are equal in all groups. Full scalar invariance will establish not only that the concept can be meaningfully discussed in all countries, that the items are strongly related to the latent concept they are measuring in all countries, but also that the level of participation in each professional community

dimension is the same for all countries, given their general level of participation in professional community. If full scalar invariance does not hold, partial scalar invariance will show us in which country and for which item, more or less initial perceived participation is manifested, given the general level of participation in professional community across all countries.

3. Method

Data and variables

In the ICCS 2009 teacher questionnaire there are five items, common to all teachers in all modules, which refer to dimensions of the latent concept of professional community. Namely, the teachers were asked how many teachers in their schools, with reference to the current school year:

- *Support good discipline throughout the school even with students not belonging to their own class or classes?* (*Collective Responsibility*/CR)
- *Work collaboratively with one another in devising teaching activities?* (*Reflective Dialogue*/RD)
- *Take on tasks and responsibilities in addition to teaching (tutoring, school projects, etc.)?* (*Deprivatisation of Practice*/DP)
- *Actively take part in <school development/improvement activities>?* (*Shared sense of Practice*/SP)
- *Cooperate in defining and drafting the <school development plan>?* (*Collaborative Activity*/CA)

These items, presented in the order they appear in the original questionnaire, refer to teachers' practices embedded into the five dimensions of the overarching latent concept of professional community. The five items were measured using a four-point Likert scale, going from "all or nearly all" to "none or hardly any". For the analysis, all indicators were inverted in order to interpret the high-numerical values as an indication of high presence of professional community practices. After listwise deleting around 3.5% of the missing completely at random data across all items and all countries, the analysis was performed on 36[2] countries and more than 58000 8th grade teachers (sample size per country is presented in Table 6 of this paper).

Figure 1 shows the model that will be tested in order to establish the measurement invariance of the latent concept of professional community across countries.

[2] The countries in the European module are Austria (AUT), Belgium (Flemish) (BFL), Bulgaria (BGR), Cyprus (CYP), Czech Republic (CZE), Denmark (DNK), England (ENG), Estonia (EST), Finland (FIN), Ireland (IRL), Italy (ITA), Latvia (LVA), Liechtenstein (LIE), Lithuania (LTU), Luxembourg (LUX), Malta (MLT), Norway (NOR), Poland (POL), Slovak Republic (SVK), Slovenia (SVN), Spain (ESP), Sweden (SWE), and Switzerland (CHE); the countries in the Asian module are Chinese Taipei (TWN), Hong Kong SAR (HKG), Indonesia (IDN), Korea (KOR), and Thailand (THA); and in the Latin American module included Chile (CHL), Colombia (COL), Dominican Republic (DOM), Guatemala (GTM), Mexico (MEX), and Paraguay (PRY). The Russian Federation (RUS) and New Zealand (NZL) are also part of the countries involved, amounting to a total of 36 countries. Greece and the Netherlands have no teacher data available.

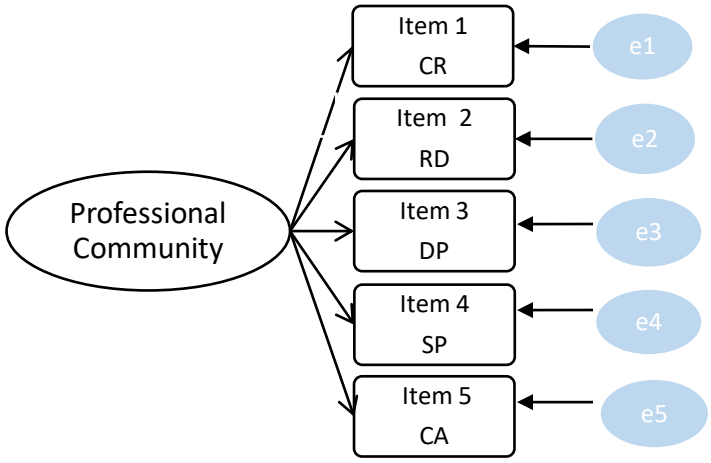

Figure 1. The model to be tested

Method and model fit indices

Multiple-Group Confirmatory Factor Analysis (MGCFA) will be used to test the measurement invariance (Billiet, 2003; Steenkamp & Baumgartner, 1998; Meuleman & Billiet, 2011) and the observed indicators are modelled as linear functions of the latent variable. The bottom-up approach will be taken (van der Veld & Saris, 2011), and full configural, metric and scalar invariance will be tested, starting from less restrictive models towards more restrictive ones. If full metric or scalar invariance is not achieved, a partial invariance will be tried as well, keeping at least two items invariant per latent construct (Byrne, Shavelson, & Muthén, 1989). Misspecifications of the model, at each level of tested invariance, will be identified and accounted for as well. As a first step, Confirmatory Factor Analysis (CFA) will be performed in each individual country to observe if the latent concept model would have a good model fit in each country. The Mplus 7.1 program was used to obtain the estimations and the model fit indices, and the JRule program (Saris, Satorra & van der Veld, 2009; van der Veld & Saris, 2011) to detect misspecifications and to consider in the model all modification indices recommended.

ICCS 2009 implemented a complex, two-stage, survey design, which implies that sampling teacher weights were taken into account to obtain unbiased estimates of the sampling error (Desa, 2014). Stratification and cluster variables were applied to the model as well to account for the complex sampling design and for correlated errors within strata and clusters.

4. Results

Invariance testing

As a first step, the individual-country CFA showed that specific decisions needed to be taken before proceeding with testing the multi-group measurement invariance. The model fit indices are satisfactory when running the CFA model across all 36 countries (RMSEA = .029, CFI = .963, SRMR = .027) and satisfactory in terms of the set level of the fit indices in other 22 countries. However, in Austria, Estonia, Ireland, New Zeeland and England, one to two model modifications were necessary in each country, allowing up to 8 error terms of two pairs of items to correlate in order to obtain a satisfactory model fit within these individual countries. The details of the modifications are presented for information in Table 1 (Table 1 can be provided on request to the author), and these modifications were allowed when running the full configural model across all countries. And finally, Liechtenstein, with only 112 participating teachers, did not show a satisfactory CFA model fit (RMSEA = .177, CFI = .799, SRMR = .061) and no specific modifications indicated by JRule, so the country was dropped from further analyses, reducing the total number of countries to 35 for all results presented next. Considering that no model modifications were recommended, we think that sample size might be the principal source of the unsatisfactory model fit for Liechtenstein.

The fit indices indicate a satisfactory model fit for the full scalar invariance model (Model A, Table 2), and when applying the additional alternative model evaluation procedure (following van der Veld & Saris, 2011) to inspect possible misspecifications of the model, no such misspecifications for the full configural model were detected.

Table 2. Results of Configural invariance testing

Model		Compared Model	χ^2 (df)	RMSEA	ΔRMSEA	CFI	ΔCFI
A	Full configural invariance (CFA – with 8 error terms correlated for 5 countries)		1533.40 (167)	.070		.968	

Note: Minimum CFI acceptable fit .90; Maximum RMSEA acceptable fit .08; *ΔCFI* <=|.01|; *ΔRMSEA* <=|.01| ; N = 58259; 35 countries; Details about the freed parameters can be requested from the author.

In a subsequent step, full metric invariance was tested and the fit indices are presented in Table 3 (Model B). Looking at the absolute differences in fit indices, we see that they are within satisfactory margins. However, when applying the additional alternative model evaluation procedure (following van der Veld & Saris, 2011) to inspect possible misspecifications of the model at the level of metric invariance, one such misspecification was detected (details in Table 1 in the Appendix, MGCFA Measurement invariance testing, Metric).

Table 3. Results of Metric invariance testing

Model		Compared Model	χ^2 (df)	RMSEA	ΔRMSEA	CFI	ΔCFI
B	Full metric invariance	A	2161.77 (303)	.061	.009	.956	.012
B_1	Partial metric invariance (free loading Item_RD for Switzerland)		2128.74 (302)	.060	-	.957	-

Note: Minimum CFI acceptable fit .90; Maximum RMSEA acceptable fit .08; *ΔCFI* <=|.01| ; *ΔRMSEA* <=|.01| ; N = 58259; 35 countries; Details about the freed parameters can be requested from the author.

Giving more detail into the model modifications identified for metric invariance, Table 4 presents the estimated slope parameters (factor loadings) and we can see that only one slope parameter deviates from the common solution for the final model. The common solution is specific for these countries considered in this model. It is the slope parameter for the *Reflective Dialogue* (RD) item in Switzerland, the low factor loading indicating that *working collaboratively with one another in devising teaching activities* is hardly related to the latent concept of professional community. It seems that in Switzerland the identification of the *Reflective Dialogue* (RD) item with the latent concept of professional community is weaker than in the other countries.

Table 4. Slope parameter estimates for the final model

	Item CR*	Item RD	Item DP	Item SP	Item CA
Switzerland		0.775			
Common solution for the model					
Slope parameters	1.000	1.232	1.400	1.577	1.379

Note: *Marker item (fixed to identify the model); N = 58259; 35 countries;

As regards metric invariance, we can see that partial metric invariance holds for 35 countries (Model B_1, Table 3). Looking at the model fit indices, we see that freeing the deviant factor loading for Item_RD for Switzerland does not make a considerable difference compared to the more restrictive metric model, but freeing it is nevertheless recommended.

In addition to the aforementioned model modifications, the intercept of Item_RD for Switzerland was freed (following van der Veld & Saris, 2011) and the scalar invariance model was run. Table 5 (Model C) clearly shows that full scalar

invariance does not show a satisfactory model fit. Applying the additional alternative model evaluation procedure (following van der Veld & Saris, 2011) to inspect the misspecifications of the model, 37 additional noninvariant intercepts were found across the countries (details in Table 1 which can be provided on request to the author, MGCFA Measurement invariance testing, Scalar).

Table 5. Results of Scalar invariance testing

Model		Compared Model	χ^2 (df)	RMSEA	ΔRMSEA	CFI	ΔCFI
C	Full scalar invariance	B_1	9478.20 (437)	.111	.051	.787	.170
C_1	Partial scalar invariance (38 free intercepts for 30 countries)		3913.49 (400)	.073	-	.917	-

Note: Minimum CFI acceptable fit .90; Maximum RMSEA acceptable fit .08; *ΔCFI* <=|.01|; *ΔRMSEA* <=|.01| ; N = 58259; 35 countries; Details about the freed parameters can be requested from the author.

As regards scalar invariance, some countries have higher or lower intercepts for the different items, as seen from the deviating parameters from the common solution for the final model presented in Table 6.

Table 6. Intercept parameter estimates for the final model

Country	Sample size	Item_CR*	Item_RD	Item_DP	Item_SP	Item_CA
AUT	949					
BFL	1582		2.981			
BGR	1813					
CHL	1698				3.014	2.917
TWN	2335					
COL	1954	2.847				2.792
CYP	875	3.369		2.512		
CZE	1557	3.356				
DNK	882					
DOM	715			2.495		2.896
ENG	1408					
EST	1745			2.336		2.822
FIN	2247				2.452	
GTM	1013				2.974	
HKG	1413	2.731				
IDN	2035			2.055	2.428	
IRL	1810					2.768

ITA	2846		2.909				
KOR	2266			2.477			
LVA	1994		2.459	2.474			
LTU	2669		2.290				
LUX	272						
MLT	862	2.806					2.787
MEX	1816	2.817					2.735
NZL	1297						
NOR	482		3.288	2.501			
PRY	1096				2.955		2.820
POL	2044						
RUS	2964			2.278			
SVK	1948			2.456			
SVN	2698						2.781
ESP	1934						2.796
SWE	1864						2.873
CHE	1416		2.445				
THA	1742			2.479			
Common solution for the final model							
Intercepts		3.110	2.681	2.643	2.684		2.594

Note: *Marker item (fixed to identify the model); N = 58259; 35 countries;

A higher/lower intercept value indicates, in general, that given the common level of participation in professional community across all countries involved, whether teachers are more or less initially involved in a specific practice, in some countries more than in other countries (following Meuleman & Billiet, 2011). Usually the intercepts could reflect initial country differences, determined by other characteristics different than the professional community practices. To be more specific, for the item reflecting a *Collective Responsibility*, we see that fewer teachers in Colombia, Hong Kong, Malta and Mexico, given the common level of participation in professional community, indicate initially to support good discipline throughout the school even with students not belonging to their own class than in the other countries; while more teachers in Cyprus and the Czech Republic support good discipline throughout the school than in other countries. The deviating intercepts for *Reflective Dialogue* indicate that only in Latvia, Lithuania, and Switzerland, fewer teachers report an initial level of working collaboratively in devising teaching activity than is expected from the average level of participation in professional community. In Italy, Belgium (Flemish), and Norway, the reversed pattern is found for the *Reflective Dialogue* item. For the *Deprivatisation of Practice* item, which refers to taking on tasks and responsibilities in addition to teaching, such as tutoring or school projects, we see that fewer teachers than expected considering the general level of involvement in professional community, take on such tasks,

in all countries with deviating intercepts presented in Table 6, majority being Eastern European and Asian countries. This is the only item for which all deviating intercepts indicate less initial perceived level of involvement of teachers in *Deprivatisation of Practice*, than it is expected from the cross-cultural level of participation in professional community. For the item representing a *Shared sense of Purpose*, the pattern is more diverse, with more teachers in Chile, Guatemala, and Paraguay, reporting a higher perceived initial level of taking actively part in school development plans and improvement activities, while the reversed pattern occurs for Indonesia and Finland. For the *Collaborative Activity* item, as for the *Deprivatisation of Practice* item but in a reversed direction, the pattern is the same for all countries in Table 6, all deviant intercepts indicating this time that more teachers than expected considering their general level of involvement in professional community, report a higher initial level of collaboration in defining and drafting the school development plan or other development schemes. It seems that it is quite common for almost all Latin American countries, namely Chile, Paraguay, Dominican Republic, Colombia and Mexico to have this pattern, together with a few European countries.

After freeing 38 intercept-related equality constraints which caused substantial misfit, we see a satisfactory model fit for the partial scalar invariance (Model C_1). A main point in Table 6 is that each country has at least two invariant intercepts per construct, indicating that the latent mean difference is identifiable. To summarize, in Table 6 we can see that ten countries have a deviant intercept for the *Deprivatisation of Practice* (DP) proxy item and 11 countries have a deviant intercept for the *Collaborative Activity* (CA) proxy item. Items proxy for *Reflective Dialogue* (RD) and *Collective Responsibility* (CR) have around 6 countries each with a deviant intercept. Choosing *Collective Responsibility* item (CR) (the fixed loading) and *Reflective Dialogue* item (RD) (the free loading and intercept for Switzerland) as the two invariant items across all groups (Byrne, Shavelson, & Muthén, 1989), partial scalar invariance holds for only 23 countries, as it can be seen in Table 6.

Respecting the results of the measurement invariance testing, we are able to compare the average scores of the professional community concept only in 23 countries[3] and not for all 36 countries initially considered.

Conclusion and discussion

The aim of the present article was to establish whether the measurement of the professional community concept, as present in this data and in accordance with the theory on the concept, can be meaningfully discussed in 36 countries and whether its metric and scalar equivalence allows for meaningful comparisons of these practices across the mentioned groups.

As regards measurement equivalence, we have established that the latent concept of professional community can be meaningfully discussed in all 35 countries, excluding Liechtenstein based on the country-specific CFA results.

[3] The countries are AUT BGR CHL TWN DNK DOM ENG EST FIN GTM IDN IRL KOR LUX NZL PRY POL RUS SVK SVN ESP SWE and THA (see Footnote 2 for the abbreviations).

Moreover, in 34 countries all the items representing the specific dimensions are strongly related with the latent concept of professional community, being now confident that these dimensions are appropriate in measuring the concept as it is defined in this study. The only exception was Switzerland, where the item *working collaboratively with one another in devising teaching activities* that represents the *Reflective Dialogue* dimension seems to have a weaker relation with the latent concept, and we therefore allowed this item to vary for Switzerland when estimating the partial metric equivalence model. This exception could be explained by the fact that teachers in Switzerland receive reflective supporting measures for lesson development to structure teaching separately from team development to promote cooperation, communication, and collegiality (Eurydice, 2013). Achieving partial metric equivalence for all 35 countries, will allow further research to validly compare relationships of the latent concept of professional community, measured through these five items in this data, with other relevant related organisational characteristics.

As regards scalar invariance, the large number of noninvariant intercepts indicate that teachers in some countries report a higher or a lower perceived initial level of involvement in specific professional community practices than expected, given the common level of participation in each specific practice across all 35 countries. To mention that these deviant intercepts are specific to this group of countries, the common solution might differ if a different group of countries is tested. Item 3 in Figure 1 proves noninvariant in its intercept in most countries, item referring to teachers taking on tasks and responsibilities in addition to teaching, implying feedback through tutoring or school projects, as a proxy for the *Deprivatisation of Practice* (DP) dimension. We see that fewer teachers than expected report taking on such tasks, considering the general level of involvement in professional communities, usually in term of intercepts determined by other characteristics than the professional community practices. The majority of countries that present a lower initial participation level in *Deprivatisation of Practice* practices are European countries (Cyprus, Slovakia, Latvia, Estonia, Russian Federation, Norway) and a few are Asian countries (Thailand, Korea, Indonesia). This finding is in accordance with the outcomes of many other studies performed within individual countries (De Neve, Devos, & Tuytens, 2015; Lomos, Hofman, & Bosker, 2011; Lomos, 2012) or across many countries (Isac, da Costa, Araújo, Soto Calvo, & Albergaria-Almeida, 2015; OECD, 2014; Vieluf, Kaplan, Klieme, & Bayer, 2012), *Deprivatisation of Practice* being the professional community dimension less practiced by teachers. This clear pattern might be influenced by the specific regulations in different countries, where such tasks could be compulsory/not compulsory, implying/not implying a reduction in teaching time, or are remunerated/not remunerated (Eurydice, 2013). In some countries, mainly in Western European countries, being tutored by an experienced teacher is compulsory for beginner teachers, which implies that experienced teachers are involved in such deprivatisation activities, even if not in collaboration with all teachers. It might also be determined by the specific school culture, where most teachers understand their activities as individual and do not feel comfortable in sharing or deprivatising their practice and knowledge through school projects, tutoring, or other such activities.

In addition, the item referring to *cooperating in defining and drafting the school development plan*, as a proxy for the *Collaborative Activity* dimension, proved also noninvariant in many countries, indicating that more teachers than expected report higher their general initial level of involvement compared to the general level of all countries considered. It seems that this is quite common for almost all Latin American countries, such as Chile, Paraguay, Dominican Republic, Colombia, and Mexico, together with a few European countries. Again, this result is comparable with results from other studies, where collaboration appears to be the most common professional community practice, next to reflection (Isac, da Costa, Araújo, Soto Calvo, & Albergaria-Almeida, 2015; OECD, 2014; Vieluf, Kaplan, Klieme, & Bayer, 2012). Even if the Latin American countries do not have a common framework for operating schools, there seem to be specific measures in place in each country which support schools in creating their own syllabus and school development plan, with teachers having significant autonomy to create such a common school plan in cooperation with each other. Such programs are the "Escuelas de calidad" in Mexico, 2001, where the participating schools had to train their teachers in accordance with the school project (Secretaria de Education Publica, 2015); or, for example, in Colombia, law 0115 from 1994 (General Law of Education, 2015) which defined the institutional school projects, establishing that schools will have to advance their own development plan, based on the nationally defined performance indicators for each level.

To summarise the results of the scalar invariance test, in most of the deviating situations, we tend to see a pattern, mainly related to the continent the country belongs to. For example, we see for the items referring to *Reflective Dialogue* and *Deprivatisation of Practice* a clear European presence, and for the items referring to a *Shared sense of Purpose* and *Collaborative Activity*, a clear presence of the Latin American countries. In addition, for the *Reflective Dialogue* dimension we see only European countries. In this case, partial scalar invariance was established for 23 countries out of 35, indicating that the latent average participation in professional community can be meaningfully compared across 23 countries, taking into consideration the model modification indices considered.

References

Billiet, J. (2003). Cross-cultural equivalence with structural equation modelling. In: J. Harkness, F. Van de Vijver, & P. Mohler, (Eds.), *Cross-cultural survey methods* (pp. 247-265). New York: Wiley.

Billiet, J., & Welkenhuysen-Gybels, J. (2004). Assessing cross-national construct equivalence in the ESS: the case of six immigration items. In C. van Dijkum, C. Blasius, & C. Durand (Eds.), *Recent development and applications on research methodology*. Amsterdam: Barbara Pudrich Pub. Proceedings of the six international conference on social science methodology.

Bryk., A. S., Camburn, E., & Louis, K. S. (1999). Professional community in Chicago Elementary schools: facilitator factors and organizational consequences. *Educational Administration Quarterly, 35*, 751-781.

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.

De Neve, D., Devos, G., & Tuytens, M. (2015). The importance of job resources and self-efficacy for beginning teachers' professional learning in differentiated instruction. *Teaching and Teacher Education, 47*, 30-41.

Desa, D. (2014). Evaluating measurement invariance of TALIS 2013 complex scales: comparison between continuous and categorical multiple-group confirmatory factor analyses. *OECD Educational Working Papers, 13*, 1-39. Retrieved from OECD Publishing http://www.oecd-ilibrary.org/education/evaluating-measurement-invariance-of-talis-2013-complex-scales_5jz2kbbvlb7k-en.

Eurydice (2013). Switzerland. Teachers and education staff. Retrieved from European Commission, Eurydice, \ Retrieved from https://webgate.ec.europa.eu/fpfis/mwikis/eurydice/index.php/Switzerland:Teachers_and_Education_Staff

General Law of Education. (2015). *Ley 115 de Febrero 8 de 1994 (Law 115 of February 8th 1994)*. Retrieved from MINEDUCATION, http://www.mineducacion.gov.co/1621/articles-85906_archivo_pdf.pdf

Isac, M.M., da Costa, P.D., Araújo, L., Soto Cavo, E., & Albergaria-Almeida, P. (2015). *Teaching practices in primary and secondary schools in Europe: Insights from large-scale assessments in education*. European Commission, JRC Science and Policy Report.

Kruse, S. D., Louis, K. S., & Bryk, A. S. (1995). An emerging framework for analysing school-based professional community. In K. S. Louis, & S. D. Kruse (Eds.), *Professionalism and community: Perspectives on reforming urban schools* (pp. 23-44). Thousand Oaks, CA: Corwin Press.

Lomos, C., Hofman, R.H., & Bosker, R.J. (2011). The relationship between departments as professional communities and student achievement in secondary schools. Teaching and Teacher Education, 27, 722-731.

Lomos, C. (2012). Professional community and student achievement. Doctoral dissertation. The Netherlands: Ipskamp Drukkers B.V.

Louis, K. S., Marks, H. M., & Kruse, S. (1996). Teachers' professional community in restructuring schools. *American Educational Research Journal, 33*(4), 757-798.

Meuleman, B., & Billiet, J. (2011). Religious Involvement: its relation to values and social attitudes. In Eldad Davidov, Peter Schmidt, & Jaak Billiet (Eds.), *Cross-cultural analysis. Methods and applications*. (pp. 173 - 206). New York: Routdledge, Taylor & Francis Group.

Meuleman, B., & Billiet, J. (2012). Measuring attitudes toward immigration in Europe: the cross-cultural validity of the ESS immigration scales. *Research and Methods, 21*(1), 5-29.

OECD. (2014). *TALIS 2013 results: An international perspective on teaching and learning*. TALIS: OECD Publishing.

SEP (Secretaria de Education Publica) (2015). *Escuelas de Calidad (Quality schools)*. Retrieved from Secretaria

de Education Publica (Board of Public Education), http://basica.sep.gob.mx/pec/start.php?act=Poblacion

Steenkamp, J.B., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national research, *Journal of Consumer Research, 25*, 78-90.

Saris, W.E., Satorra A., & van der Veld, W.M. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling: A Multidisciplinary Journal, 16*(4), 561 – 582.

Toole, J. C., & Louis, K. S. (2002). The role of professional learning communities in international education. In K. Leithwood & P. Hallinger (Eds), *Second international handbook of educational leadership and administration*. (pp. 247-279). Dordrecht, NL: Kluwer Academic Publishers.

van der Veld, W.M., & Saris, W.E. (2011). Causes of generalized social trust. In Eldad Davidov, Peter Schmidt, & Jaak Billiet (Eds.), *Cross-cultural analysis. Methods and applications*. (pp. 207 – 248). New York: Routdledge, Taylor & Francis Group.

Vieluf, S., Kaplan, D., Klieme, E., & Bayer, S. (2012). Teaching practices and pedagogical innovation: Evidence from TALIS, OECD Publishing. Retrieved from OECD: http://dx.doi.org/10.1787/9789264123540-en