

Article

# The Cross-Cultural Invariance of the Servant Leadership Survey: A Comparative Study across Eight Countries

Dirk Van Dierendonck <sup>1,\*</sup>, Milton Sousa <sup>2</sup>, Sigrún Gunnarsdóttir <sup>3,4</sup>, Andrea Bobbio <sup>5</sup>, Jari Hakanen <sup>6</sup>, Armin Pircher Verdorfer <sup>7</sup>, Emin Cihan Duyan <sup>8</sup> and Raquel Rodriguez-Carvajal <sup>9</sup>

<sup>1</sup> Rotterdam School of Management, Erasmus University, Burg Oudlaan 50, Rotterdam 3000 DR, The Netherlands

<sup>2</sup> NOVA School of Business and Economics, Lisboa 1099-032, Portugal; milton.sousa@novasbe.pt

<sup>3</sup> Faculty of Business, University Bifrost, Bifröst 311, Iceland; sigrun@bifrost.is

<sup>4</sup> Faculty of Business, University of Iceland, 101 Reykjavík, Iceland

<sup>5</sup> Department of Philosophy, Sociology, Education and Applied Psychology (FISPPA), University of Padua, Via Venezia n. 14, Padova PD 35131, Italy; andrea.bobbio@unipd.it

<sup>6</sup> Työterveyslaitos, PL 40 00251 Helsinki, Finland; Jari.hakanen@ttl.fi

<sup>7</sup> School of Management, Technical University of Munich, Arcisstr. 21, 80333 Munich, Germany; armin.pircher-verdorfer@tum.de

<sup>8</sup> Department of Labour Economics and Industrial Relations, Anadolu University, 26470 Tepebaşı/Eskişehir Province, Turkey; eminci@gmail.com

<sup>9</sup> Department of Biological and Health Psychology, Autonomous University of Madrid, 28049 Madrid, Spain; raquel.rodriguez@uam.es

\* Correspondence: [dvandierendonck@rsm.nl](mailto:dvandierendonck@rsm.nl)

Academic Editor: Kathleen A. Patterson

Received: 12 January 2017; Accepted: 10 April 2017; Published: 17 April 2017

**Abstract:** This paper tests and confirms the cross-cultural equivalence of the Servant Leadership Survey (SLS) in eight countries and languages: The Netherlands, Portugal, Germany, Iceland, Italy, Spain, Turkey and Finland. A composite sample consisting of 5201 respondents from eight countries that all filled out the SLS was used. A three-step approach was adopted to test configural invariance, measurement equivalence, and structural equivalence. For the full 30-item version of the SLS, configural invariance and partial measurement equivalence were confirmed. Implications of these results for the use of the SLS within cross-cultural studies are discussed.

**Keywords:** servant leadership; cross-cultural; measurement equivalence

## 1. Introduction

We live in a cross-cultural, diverse world where globalization has become a fact of organizational life. Different countries' cultures are being brought closer together than ever before in history [1]. This process is especially noticeable in Europe, where the European Union has brought about increased economic interdependency and cultural exchange among member states [2]. This asks for greater cultural sensitivity on the part of our leaders, and also calls into question whether our leadership models, which have mostly been developed within an Anglo-Saxon context, hold across culturally diverse countries such as the members of the EU. With their mostly knowledge-based economies, European economies largely depend on bringing out the best in people. Paying attention to employees' needs is a crucial factor in staying competitive and achieving long-term success [3]. The leadership theory that is most oriented towards the needs of employees is servant leadership [4], and the focus

of servant leadership on service may also help to provide a bridge between culturally different worldviews [1].

For leadership scholars, it is important to develop an accurate understanding of the conceptual similarity of an instrument across different countries and languages. A major concern in this regard is the cross-cultural equivalence of an instrument's measurement properties (see [5,6]). In fact, only with a measure that captures the same construct across different cultural contexts we can compare results across countries and truly gain valid insights into effective cross-national leadership. This is especially true within Europe, where each country has developed its own habits, norms and values [2] and where several different cultural clusters can be distinguished [7]. Within this paper, we are therefore particularly interested in the cross-cultural validity of a servant leadership instrument. The main aim of this paper is to study the measurement invariance of the Servant Leadership Survey (SLS) [8] in different countries. Interestingly, although leadership research is a global phenomenon, explicit invariance tests are scarce (with the exception of the GLOBE leadership study; Global Leadership and Organizational Behavior Effectiveness; House et al. [9]). Additional insight into which dimensions hold cross-culturally will contribute to the servant leadership literature specifically and to the leadership literature more generally.

### *1.1. The Construct of Servant Leadership and the Servant Leadership Survey (SLS)*

Since Greenleaf [10] introduced the notion of servant leadership into the vocabulary of leadership research and practice, several conceptualizations of servant leadership have emerged, most over the past 15 years. One of the central features of servant leadership that has been established in its recent history is that servant leadership "places the good of those led over the self-interest of the leader, emphasizing leader behaviours that focus on follower development, and de-emphasizing glorification of the leader" [11] (p. 397).

Servant leadership has been shown to have a positive impact on follower well-being as well as the overall effectiveness of individuals and teams. Moreover, it has passed the test of showing incremental validity beyond what other popular leadership approaches, most notably transformational leadership, have offered (for a comprehensive overview, see [4,12]).

More recently, Van Dierendonck introduced an integrative model of servant leadership and developed the Servant Leadership Survey (SLS) [8], which represents a psychometrically sound measure covering the key aspects of servant leadership. The SLS is a 30 item eight-dimensional scale of servant leadership, including empowerment, accountability, standing back, humility, authenticity, courage, forgiveness and stewardship. In essence, this measure emphasizes that servant leaders empower and develop people, are willing to retreat into the background and let others shine, hold followers accountable for their work, are willing to let bygones be bygones, dare to take risk, are willing to show what they stand for, have an openness to learn and a willingness to admit mistakes, and work for the good of the whole.

### *1.2. Cultural Influences on the Perception of Servant Leadership*

The countries in our sample are quite diverse and well-representative of Europe and its broad range of national cultures. We were able to include data from the Netherlands, Portugal, Iceland, Italy, Finland, Germany, Turkey and Spain. These countries are all part of the European Region and each has its own culture, language and history.

With regard to the impact of national culture on servant leadership effectiveness, Van Dierendonck [4] suggested that individualism versus collectivism may be influential. Our sample includes collectivistic countries such as Portugal and Turkey on the one end of the spectrum and individualistic countries such as the Netherlands and Italy on the other end, with the remaining countries scoring in between [13]. Additionally, with respect to power distance, the sample is quite diverse: it includes countries with a high-power distance such as Turkey and Portugal and those with a low-power distance such as Finland and Germany. Additionally, it should be noted that each country has a different native language,

which allows us to compare eight different language versions of the same measure. Translating words into different languages has its challenges, given that a literal translation is not always the best method of conveying meaning. What appear to be the same words may hold very different meanings [14]; this is certainly the case for the term 'leader', where the context often determines which term may be the best translation. More specifically, for what constitutes good leadership, there are also clear differences within Europe. According to the GLOBE leadership study (Global Leadership and Organizational Behavior Effectiveness; House et al. [9]), the Netherlands and Germany belong to the Germanic cluster, which scores high on future orientation and performance orientation and low on humane orientation and institutional collectivism. Italy, Spain and Portugal belong to the Latin Europe cultural cluster, which is characterized by a comparatively low value placed on humane orientation and on institutional collectivism. Turkey belongs to the Middle East cluster, which scores high on in-group collectivism and low on uncertainty avoidance. Finland (and probably Iceland, although no data are available) belongs to the Nordic European cluster, which scores high on future orientation and gender egalitarianism and low on assertiveness orientation and in-group collectivism.

This cultural diversity within Europe in general and the countries represented in our combined sample can also be illustrated by Ronen and Shenkar's [7] recent cultural clustering of similarity and dissimilarity in work-related attitudes based on the most influential cultural models, including Hofstede's model and the GLOBE study. According to their analysis, the Netherlands, Finland, and Iceland are part of the Nordic cluster, Germany belongs to the Germanic cluster, Portugal, Spain and Italy belong to the Latin Europe cluster, and Turkey is part of the Near East cluster.

Given the overall differences in culture and ideal leadership types in the eight countries, one could wonder to what extent a single leadership theory may be helpful in furthering our understanding of leadership in Europe. The GLOBE research program provides insight into leader behaviours that are perceived by almost everyone around the world as effective and desirable. This ideal leadership type represents values and behaviours that have a clear and wide overlap with servant leadership, such as high integrity, having strong interpersonal skills, and lacking egocentric and self-focused behaviours [9,15]. According to the GLOBE study, people throughout the world perceive trustworthiness, fairness and honesty as essential elements of good leadership. Thus, it seems that although the perception of leader effectiveness is influenced by national context and culture [16], globalization and interdependence between nations have led to a growing awareness of the importance of leaders who show in their everyday behaviour a focus on growth, development and innovation among their followers [17]. As such, it might be expected that the essence of servant leadership will be viewed similarly throughout Europe.

### *1.3. Establishing Measurement Equivalence and Invariance*

Ideally, we would like to have a measure where the factor structure is fully equivalent, which means that the items are related to their respective dimensions in the same way and the interrelations between the dimensions as exemplified by their relative contributions to one underlying servant leadership dimension are identical. Several researchers in this field have stressed, however, that such full measurement invariance frequently does not hold, and thus represents a too-strict requirement for group comparisons [18,19]. Especially with an eight-dimensional measure such as the SLS, it is unlikely that eight different survey versions will be realized and experienced identically. Hence, in our study we focus on partial measurement invariance, which allows some model parameters to vary across groups while still ensuring that the overall comparison is meaningful [18,19].

Combining the recommendations of Byrne [20] on testing cross-cultural multi-group equivalence with those of Chen et al. [21] on testing the measurement invariance of second-order factor models, we distinguish three levels in testing the measurement invariance of the SLS.

The first level is configural invariance, which means that the same items are indicators of a dimension in each group. For the SLS, this level has already been confirmed through confirmatory factor analysis for the eight-dimensional structure in the survey version for the Netherlands and the

UK [8], Finland [22], Italy [23], Germany [24], Portugal [25], Spain, Argentina and Mexico [26]. In the present research, we will test the configural equivalence of the instrument in eight European countries simultaneously, which we expect to be stable.

**Hypothesis 1.** *The factor structure of the SLS is configurally equivalent across samples.*

In addition to testing the factorial structure of the SLS, this model will also provide the baseline for the subsequent invariance models. The second level refers to measurement equivalence. This tests whether the factor loadings connecting the items with the dimensions are similar across groups. This level provides insight into the extent to which the items are interpreted in a similar way, and has also been called metric invariance. With metric invariance, the relations between the factors and external variables can be compared. Thus, it allows for the comparison of results between studies performed in different countries, given that the measurement unit underlying a factor is the same. This gives our second hypothesis.

**Hypothesis 2.** *The factor loadings connecting the items to the dimensions will be invariant across countries, indicating measurement equivalence.*

The third level is structural equivalence, which focuses on the unobserved (i.e., latent) variables. In our case, we focus on the model covering the eight servant leadership dimensions and their interrelations. The SLS covers different attributes that together represent servant leadership. Specifically, we aim to test whether the eight-dimensional structure of servant leadership that is supposed to underlie servant leadership as measured by the SLS is similar across countries. In other words, the parameters of interest are the relationships between the eight dimensions of the SLS as well as the factor loadings between the eight dimensions. In line with the different leadership preferences between cultures, which make full measurement equivalence unlikely, we focus here on partial metric invariance [18]. This requires cross-country invariance of most but not necessarily all intercorrelations in order for subsequent tests with the SLS to be meaningful.

**Hypothesis 3.** *The SLS will show partial structural equivalence across countries.*

In their developmental article, Van Dierendonck and Nuijten [8] showed that their eight-dimensional instrument covers both the 'leader' side (i.e., empowerment, accountability, courage, stewardship) and 'servant' side elements of servant leadership (i.e., standing back, humility, authenticity). In the present research, we are also examining the possibility of shortening the measure so that the remaining items and dimensions are perceived as equivalent across countries, while simultaneously representing the 'leader' and the 'servant' aspects of the original SLS.

Taken together, this study provides insight into the extent to which servant leadership as a whole can be interpreted similarly across different samples from eight different countries. By explicitly testing three levels of measurement invariance, we are able to gain insight into the extent that empirical results of the SLS can be compared.

## 2. Method

### 2.1. Participants

For this study, we used a composite sample of 5201 respondents from eight countries. We included those countries where we had data of at least 250 respondents having filled out the full survey (after listwise deletion of the missing data). It was a convenient sample with a very diverse occupational background. Part of the data was from samples in already published articles; in those cases, the appropriate reference is added. The average age of respondents was 41.3 years (SD = 10.6) and gender was known for 4226 respondents; 56.8% respondents were male and 43.2% were female.

The division across countries was: 1107 finance professionals from Portugal working in the banking sector (35% women; age 40.0, SD = 10.5) [25], 1049 persons from Iceland working in health care, primary education, university and at a service desk (67% women; age = 47.0, SD = 9.3) [27]. 811 persons from The Netherlands from a diverse range of occupations (40% women; age 40.8, SD = 10.5), 753 persons from Italy from a diverse range of sectors, mainly profit (48% women; age = 38.7, SD = 10.2) [23], 491 judges from Finland working at the general courts (51% women; age = 50.6, SD= 10.8) [22], 395 people from Germany working in a broad range of sectors, both profit and non-profit (47% women; age = 37.3, SD = 10.8), 337 persons from Turkey working in the automotive and textile industry (34% women, age = 33.3, SD = 8.8) [28], and 258 persons from Spain from diverse work settings (53% women, age = 32.6, SD = 8.5) [26].

2.2. Measure

*Servant leadership* was measured with the 30 items from the Servant Leadership Survey (SLS) developed by Van Dierendonck & Nuijten [8]. The eight dimensions are empowerment (seven items), accountability (three items), standing back (three items), humility (five items), authenticity (four items), courage (two items), forgiveness (three items) and stewardship (three items). Ratings were given on a six-point Likert scale, ranging from ‘fully disagree’ to ‘fully agree’ with no middle category. All participants filled out the survey in their native language. A translation-back translation procedure was used for all translations, as recommended by Brislin [29]. At least two persons were involved with all translations. The English version depicted by Van Dierendonck and Nuijten [8] was translated into the other languages (i.e., Portuguese, Icelandic, Italian, Finnish, German, Turkish and Spanish), with the exception for Dutch, which was the original language of the survey. Specifically, a series of forward and backward translations focusing on cross-cultural and conceptual accuracy rather than on literal equivalence was used to reach the best possible translation. In case of doubt, the first author of the survey development article was consulted on the meaning of a particular item.

3. Results

Mplus 7.11 was used for data analysis [30]. As a preliminary step, data from all samples were combined and subjected to confirmatory factor analysis (CFA) to test the eight-dimensional model, allowing all dimensions to intercorrelate. The model fit the data well; that is  $\chi^2 = 6296.351$ ,  $df = 377$ , CFI = 0.94, TLI = 0.93, RMSEA = 0.06. With one underlying second order factor, the fit was still good with  $\chi^2 = 8220.788$ ,  $df = 397$ , CFI = 0.92, TLI = 0.91, RMSEA = 0.06. This confirms the underlying eight-dimensional model for the sample as a whole. Table 1 depicts the correlations between the eight dimensions for the overall sample. There may, however, still be differences between countries, which we tested subsequently.

**Table 1.** Descriptions and intercorrelations of short and long versions of the Servant Leadership Survey dimensions (N = 5201).

SLS Dimension	M	SD	1	2	3	4	5	6	7	8	9	10	11
1. Empowerment, short	4.09	1.19	(0.92)										
2. Empowerment, long	4.11	1.16	0.99	(0.92)									
3. Humility, short	3.81	1.22	0.76	0.76	(0.91)								
4. Humility, long	3.83	1.14	0.77	0.77	0.97	(0.92)							
5. Standing back	3.58	1.13	0.64	0.65	0.69	0.70	(0.77)						
6. Stewardship	4.04	1.22	0.77	0.78	0.68	0.71	0.58	(0.81)					
7. Authenticity, short	3.64	1.07	0.62	0.62	0.68	0.69	0.59	0.57	(0.71)				
8. Authenticity, long	3.62	1.01	0.59	0.59	0.65	0.67	0.59	0.57	0.96	(0.74)			
9. Courage	3.66	1.18	0.31	0.30	0.27	0.27	0.29	0.28	0.26	0.39	(0.64)		
10. Accountability	4.58	1.12	0.59	0.62	0.43	0.43	0.33	0.60	0.37	0.33	0.12	(0.88)	
11. Forgiveness	2.61	1.23	0.34	0.35	0.35	0.35	0.29	0.27	0.17	0.13	0.12	0.21	(0.77)

Note: Internal consistency between brackets at the diagonal. All correlations,  $p < 0.001$ .

We evaluated the measurement invariance across the different countries and languages in three steps (Table 2). First, we tested the configural invariance where the underlying model is the same in all countries, and all factor loadings were set to be free across the sub-groups (i.e., countries). Our results show that the eight-dimensional configural invariance model, allowing all latent dimensions to correlate, was clearly a better fit for the data than the one-dimensional model where all items load on the same factor. Accordingly, the configural invariance model obtained good relative fit indices with 0.92 for CFI (Bentler Comparative Fit Index), 0.91 for TLI (Tucker Lewis Index), 0.06 for RMSEA (Root Mean Square Error of Approximation) and 0.05 for SRMR (Standardized Root Mean Square Residual). This confirms Hypothesis 1, that the proposed factorial structure can be confirmed in the different samples.

In the next step, we tested for multi-group invariance by analysing the measurement equivalence model, which was then compared with the configural model. In the measurement equivalence model, not only is the underlying model the same, additionally, the factor loadings linking the items to their latent factors are set to be the same across sub-groups (i.e., countries). Results show the  $\chi^2$  difference to be statistically significant ( $\Delta\chi^2 = 1643.203$ ,  $p < 0.001$ ), which is to be expected given that the chi-square is sensitive to sample size. With a sample size of 5201 persons, even very small differences are likely to become significant. More important, therefore, are the differences in comparative fit indices. These are limited: only 0.02 for CFI and TLI, similar for RMSEA and 0.03 for SRMR. This pattern indicates some differences (see [31]), but overall and in terms of absolute values, three out of four fit indices are still within established levels for CFI, RMSEA and SRMR [20]. These results indicate that there are variations between the different samples with regard to the interpretation of some of the dimensions, thus we show only partial measurement invariance and Hypothesis 2 could not be confirmed; however, the overall interpretation of servant leadership as measured by the SLS is still quite similar.

**Table 2.** Measurement invariance cross-national fit indices of the SLS.

Model Tested	$\chi^2$	<i>df</i>	CFI	TLI	RMSEA	SRMR
One-dimensional model	30,262.319	3443	0.72	0.72	0.11	0.14
Eight-dimensional model						
Configural invariance	8880.289	3016	0.92	0.91	0.06	0.05
Measurement equivalence	10,523.492	3170	0.90	0.89	0.06	0.08
Structural equivalence	13,450.214	3324	0.86	0.86	0.07	0.10
Eight-dimensional model with one underlying factor						
Configural invariance	13,493.904	3179	0.89	0.88	0.07	0.08
Measurement equivalence	15,119.764	3330	0.88	0.87	0.07	0.10
Structural equivalence	16,510.405	3380	0.86	0.86	0.08	0.14

Finally, the structural model was analysed. This is an even stricter model given that the correlations between the latent factors are also set to be the same between the countries. Here, the fit indices dropped below the 0.90 threshold for CFI and TLI and the 0.08 threshold for SRMR, which indicates that the relationships between the dimensions are different between the countries. It should be noted, however, that full structural invariance seldom occurs and that partial invariance suffices to allow for a comparison of the findings between studies.

Given that partial invariance may be the best that can be reached with a multi-dimensional measure such as the SLS (in our case: measurement equivalence), we checked the relative contribution of the chi-square of the different samples for the structural equivalence model. If the relative chi-square is higher, this indicates that the lesser fit is due to some kind of misfit within the sample from that country as compared to the overall full-sample fit. The relative chi-square was highest for Iceland (2251.287), followed by Portugal (1438.097), Finland (1046.610), the Netherlands (967.304), Germany (859.880), Italy (852.997), Turkey (777.433) and Spain (687.681). We therefore checked for invariance by testing the model without the data from Iceland. The relative fit of the configural invariance model was similar (CFI = 0.92, TLI = 0.91, RMSEA = 0.05, SRMR = 0.05). It did, indeed, result in

a better fit for the measurement equivalence model, with CFI = 0.91, TLI = 0.90, RMSEA = 0.05, SRMR = 0.07. The fit of the structural invariance model was similar as with the full dataset with CFI = 0.87, TLI = 0.86, RMSEA = 0.06, SRMR = 0.08, thereby confirming partial measurement invariance, as stated in Hypothesis 3.

#### *Development of the SLS Short Version*

Next, we examined the possibility of improving cross-cultural stability by shortening the SLS. The previous analysis confirmed configural invariance and partial measurement invariance. Structural invariance, however, remained below recommended values. We aimed for a measure that would fulfil all three invariance standards and where the subscales would be internally consistent with a Cronbach's alpha of at least 0.70. Our basic premise was to choose those dimensions and items that best reflected the underlying servant leadership concept [8]. Thus, we aimed for cross-cultural stability by removing those dimensions and items that were interpreted differently between countries. We started by removing those dimensions that least reflected servant leadership with the greatest variance across countries. Next, within the remaining dimensions, items with a relatively low item-total correlation were removed to retain those items that best reflected the total underlying variance of a subscale. In this way, we aimed to generate a shorter measure that would reflect the core dimensions of servant leadership with a stable and invariant underlying factorial model across countries and with reliable subscales.

To gain better insight into the different interpretations of servant leadership across different countries, we used the configural invariance model with eight dimensions and one underlying second-order factor. The outcome provides the standardized factor loadings for the eight dimensions of servant leadership for all countries. This gave the following pattern: Empowerment = 0.94 (range between 0.81 and 0.96); humility = 0.89 (range between 0.83 and 0.94); standing back = 0.84 (range between 0.82 and 0.97); stewardship = 0.94 (range between 0.44 and 0.99); authenticity = 0.83 (range between 0.68 and 0.98) courage = 0.34 (range between -0.33 and 0.75); forgiveness = 0.40 (range between 0.21 and 0.63); accountability 0.64 (range between 0.23 and 0.84). From these average values and ranges, we may conclude that empowerment, humility, standing back, stewardship and authenticity are clearly incorporated within servant leadership across countries. Courage is a problematic dimension, as the overall factor loading is the lowest of all eight dimensions; the negative value is due to the results of the Icelandic sample where the two items have a low intercorrelation ( $r = 0.11$ ). We therefore decided to remove both courage items. Forgiveness was the other dimension with a low overall factor loading. In some countries, this value dropped below 0.30. The items belonging to this dimension were therefore also dropped. The third dimension we decided to remove was accountability, given that its average factor loading was lower than that of the other five and that its range was broader, with a lower limit of 0.23 in one of the countries, indicating that accountability was not perceived as a central part of servant leadership overall.

We aimed for a shorter scale that combined cross-cultural invariance with good internal consistency for the subscales. For the internal consistency, at least three items in each subscale are preferred. Two of the remaining five dimensions consisted of three items (i.e., standing back and stewardship) and were therefore retained in their original forms. For empowerment, one item's factorial loading was lower in several countries whereas the remaining six items loaded equally well; thus, this item was removed. Given that empowerment has two underlying aspects; encouragement of followers' growth and providing autonomy and freedom to do one's job, the remaining six items were kept to represent both elements within this subscale. The correlation between the seven- and six-item versions is 0.99. A similar process was used to decide which three of the five humility items best reflected the essence of this subscale. Here, the correlation between the five-item and three-item scale was 0.97. For authenticity, one item was removed to enhance the overall fit of this dimension with the other dimensions, resulting in a three-item scale with a correlation of 0.96 with the four-item scale. Together, the end result is a five-dimensional survey with 18 items (see Table 3).

**Table 3.** Confirmatory standardized factor loadings of the SLS, short version (text of the items from Van Dierendonck & Nuijten [8] (p. 256)).

Empowerment	Standardized Factor Loadings
1. My manager gives me the information I need to do my work well.	0.79
2. My manager encourages me to use my talents.	0.89
3. My manager helps me to further develop myself.	0.84
4. My manager encourages his/her staff to come up with new ideas.	0.86
5. My manager gives me the authority to take decisions which make my work easier to me.	0.73
6. My manager offers me abundant opportunities to learn new skills.	0.75
Humility	
7. My manager learns from criticism.	0.87
8. My manager learns from different views and opinions of others.	0.88
9. If people express criticism, my manager tries to learn from it.	0.90
Standing Back	
10. My manager keeps himself/herself at the background and gives credits to others.	0.74
11. My manager is not chasing recognition for the things he/she does for others.	0.72
12. My manager appears to enjoy his/her colleagues' success more than his/her own.	0.71
Stewardship	
13. My manager emphasizes the importance of paying attention to the good of the whole.	0.81
14. My manager has a long-term vision.	0.79
15. My manager emphasizes the societal responsibility of our work.	0.70
Authenticity	
16. My manager is open about his/her limitations and weaknesses.	0.65
17. My manager is often touched by the things he/she sees happening around him/her.	0.59
18. My manager shows his/her true feelings to his/her staff.	0.75

Again, we conducted an overall test of this five-dimensional model by combining all samples from different countries into one overall sample. This yielded an excellent overall fit of  $\chi^2 = 1749.325$ ,  $df = 125$ , CFI = 0.97, TLI = 0.97, RMSEA = 0.05, SRMR = 0.02, which was better than the fit associated with the one-dimensional model with all 18 items loading on one factor ( $\chi^2 = 8049.607$ ,  $df = 152$ , CFI = 0.88, TLI = 0.87, RMSEA = 0.10, SRMR = 0.05). Table 4 shows the measurement invariance fit indices, which were excellent for the configural invariance model. The overall fit was still good for the five-dimensional model taking country into account as a subgroup and holding the factorial loadings across countries constant (measurement equivalence). Even the structural equivalence model showed a good fit to the data. A comparison of the relative indices of the short version (Table 4) with those of the full version (Table 2) shows that the overall invariance fit increased considerably.

**Table 4.** Measurement invariance cross-national fit indices of the five-dimensional version of the SLS.

Model Tested	$\chi^2$	$df$	CFI	TLI	RMSEA	SRMR
Configural invariance	3371.873	1000	0.95	0.94	0.06	0.04
Measurement equivalence	4234.748	1091	0.93	0.92	0.07	0.08
Structural equivalence	5812.286	1182	0.90	0.90	0.08	0.09

Note: All  $p < 0.0001$ .

Not surprisingly, the overall five-dimensional model, including a second order SLS-factor using the full sample and not taking country into account, confirmed the excellent fit of the SLS short version ( $\chi^2 = 2227.886$ ,  $df = 130$ , CFI = 0.97, TLI = 0.96, RMSEA = 0.06, SRMR = 0.03), whereby all dimensions contributed strongly to the underlying servant leadership factor: 0.91 for empowerment, 0.91 for humility, 0.86 for standing back, 0.91 for stewardship and 0.89 for authenticity. The factor loadings of the items on the subscales can be found in Table 3. The internal consistencies of these subscales ranged between 0.71 and 0.92 (see Table 1). For the scale as a whole, the internal consistency is 0.95 with item-total correlations ranging between 0.58 and 0.79. We can therefore safely conclude that this 18-item version is factorially valid, internally consistent and holds cross-nationally.



#### 4. Discussion

This paper set out to study the cross-cultural equivalence of the Servant Leadership Survey. With a composite sample of 5201 individuals, we were able to establish the factorial validity, configural invariance and measurement equivalence of the measure across eight countries and languages. Taken as a whole, the present paper contributes to the servant leadership literature because it shows that it is possible to measure servant leadership in a similar way, internationally and cross-culturally, with samples from a broad range of sectors and professions. Additionally, we also found indications that the relations underlying the original dimensions as formulated by Van Dierendonck and Nuijten [8] may differ between countries. To tackle this issue, we introduced a shorter version of the SLS that retains the essential elements of the longer version.

Cross-cultural research into servant leadership has been growing over the last few years, as have the number of measures available [4]. Servant leadership theory originated within the United States, as have most of the measurement instruments. Although research on servant leadership throughout the world is growing (for an overview see, [1]), our study is the first to explicitly explore and confirm that servant leadership can be measured across languages, nations and cultures in Europe. This paper builds on the encouraging results of the SLS in the original developmental paper [8]. Including the British sample in that article and the eight samples in the present paper, it has now been shown that the SLS can be reliably measured in nine languages.

The partial invariance of the measurement invariance model confirms the findings of factorial validity of the separate samples combined into our composite sample, as shown in recent articles [22–24,26]. Both the factorial structure and the individual item loadings are similar across most countries and languages. It took only the removal of one country (i.e., Iceland) to reach generally acceptable values. This encouraging result allows for future testing and interpretation of the relationships between servant leadership and other concepts across countries using the SLS. The lack of fit of the Icelandic sample may have been caused by the sample's specific language and culture, but it is just as likely due to the specific characteristics of the sample, given that our composite sample consisted of convenient subsamples with different occupational backgrounds. Specific for Iceland is the relative high percentage of woman and an on-average older age. Future research will have to determine to what extent the lack of fit for the Icelandic sample is due to cross-cultural factors or due to specific sample characteristics in terms of profession, sector, age and gender.

Regretfully, for the eight-dimensional model, the fit indices of the structural equivalence model were below recommended values. This may be due to several reasons. One is the broad spectrum measured by the dimensions of the SLS. Although aspects such as empowerment and humility are conceptually defined as part of servant leadership just like forgiveness and courage, it may be clear that their focus is different. For example, empowerment can be seen as a more general approach to encouraging followers to become the best they can be, whereas forgiveness is more conditional in that it is only shown after followers make mistakes. It should also be acknowledged that even for a valid and reliable measure, structural differences between countries are usually real and interpretable given measurement invariance of the constructs, which is the case for the SLS. Another possible explanation is the specific content of the operationalization of the dimensions. With only three or two items, respectively, aspects such as accountability, forgiveness and courage cannot be measured in their full complexity. To acknowledge these aspects more comprehensively, it may be necessary to formulate additional items. Another option is to use already existing measures when one is interested in studying those particular dimensions, instead of the SLS. For example, for leadership forgiveness, researchers may want to review Verdoold and Van Dierendonck [32].

The results of the shorter, five-dimensional measure are very encouraging. The fit indices confirmed configural invariance, measurement equivalence and structural equivalence. This version allows us to test servant leadership in different countries with the same measure, knowing that the underlying concept is experienced in a similar way. It also allows for a comparison of mean levels. The latter is very helpful because it means that we can now study the extent to which servant leadership

is experienced in different countries. One of the limitations of the present study is that we had to rely on convenient samples for our composite sample. The samples in this paper came from different occupations and sectors, so no mean values are given. As such, we cannot make any claim with regard to generalizability of the average values in the different countries, and we strongly encourage further research to overcome this limitation. It would be very insightful for a future study to measure servant leadership within the same sector(s) in different countries, similar to the setup used in the original study by Hofstede [33]. Another limitation is that we did not have similar outcome variables available across samples to test the similarity in predictive validity. Additionally, in some countries the sample size was relatively small compared with other countries. Nevertheless, by combining the different datasets, our analyses have the unique feature of being able to test the factorial validity across different countries, which has a strong additional value above separately testing the factorial validity in one single country.

In conclusion, this article builds on the earlier paper by Van Dierendonck and Nuijten [8] by also confirming the configural and measurement cross-cultural equivalence of the SLS. It is the first servant leadership measure where such an extensive insight is available, thus rendering this measure a serious contender for researchers throughout the world interested in using a valid and reliable measure to operationalize servant leadership.

**Author Contributions:** The study was designed by the first author, who analyzed the data and wrote the first draft of the paper. The other authors were responsible for the translation of the items into their language and the data collection in their country. The order of these authors is based on country sample size. All commented on the text and helped improve it. All authors read and approved of the final manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Irving, J.A. Cross-cultural perspectives on servant leadership. In *Servant Leadership: Developments in Theory and Research*; van Dierendonck, D., Patterson, K., Eds.; Palgrave MacMillan: London, UK, 2010; pp. 118–129.
2. Starren, A.; Luijten, K.; Drupsteen, L.; Vilkevicius, G.; Eeckelaert, L. *Diverse Cultures at Work: Ensuring Safety and Health through Leadership and Participation*; European Agency for Safety and Health at Work: Brussels, Belgium, 2013.
3. O’Leary, B.S.; Lindholm, M.L.; Whitford, R.A.; Freeman, S.E. Selecting the best and the brightest: Leveraging human capital. *Hum. Resour. Manag.* **2002**, *41*, 325–340. [[CrossRef](#)]
4. Van Dierendonck, D. Servant leadership: A review and synthesis. *J. Manag.* **2011**, *37*, 1228–1261. [[CrossRef](#)]
5. Chen, F.F. What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *J. Personal. Soc. Psychol.* **2008**, *95*, 1005–1018. [[CrossRef](#)] [[PubMed](#)]
6. Van de Vijver, F.; Leung, K. Methodological issues in psychological research on culture. *J. Cross-Cult. Psychol.* **2000**, *31*, 33–51. [[CrossRef](#)]
7. Ronen, S.; Shenkar, O. Mapping world cultures: Cluster formation sources and implications. *J. Int. Bus. Stud.* **2013**, *44*, 867–897. [[CrossRef](#)]
8. Van Dierendonck, D.; Nuijten, I. The Servant Leadership Survey: Development and validation of a multidimensional measure. *J. Bus. Psychol.* **2011**, *26*, 249–267. [[CrossRef](#)] [[PubMed](#)]
9. House, R.J.; Hanges, P.J.; Javidan, M.; Dorfman, P.W.; Gupta, V. *Culture, Leadership and Organizations: The GLOBE Study of 62 Societies*; Sage Publications: Thousand Oaks, CA, USA, 2004.
10. Greenleaf, R.K. *Servant Leadership: A Journey into the Nature of Legitimate Power and Greatness*; Paulist Press: New York, NY, USA, 1977.
11. Hale, J.R.; Fields, D.L. Exploring servant leadership across cultures: A study of followers in Ghana and the USA. *Leadership* **2007**, *3*, 397–417. [[CrossRef](#)]
12. Liden, R.C.; Panaccio, A.; Hu, J.; Meuser, J.D. Servant leadership: Antecedents, consequences, and contextual moderators. In *The Oxford Handbook of Leadership and Organizations*; Day, D.V., Ed.; Oxford University Press: Oxford, UK, 2014; pp. 357–379.
13. Hofstede, G.; Hofstede, G.J.; Minkow, M. *Cultures and Organizations: Software of the Mind*, 3rd ed.; McGraw-Hill: New York, NY, USA, 2010.

14. Van de Vijver, F.; Hambleton, R.K. Translating tests: Some practical guidelines. *Eur. Psychol.* **1996**, *1*, 89–99. [[CrossRef](#)]
15. Mittal, R.; Dorfman, P.W. Servant Leadership across Cultures. *J. World Bus.* **2012**, *47*, 555–570. [[CrossRef](#)]
16. Hamilton, F.; Bean, C. The importance of context, belief and values in leadership development. *Bus. Ethics Eur. Rev.* **2005**, *14*, 336–347. [[CrossRef](#)]
17. Northouse, P.G. *Leadership: Theory and Practice*, 6th ed.; Sage Publications, Inc.: Thousand Oaks, CA, USA, 2013.
18. Byrne, B.M.; Shavelson, R.J.; Muthén, B. Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychol. Bull.* **1989**, *105*, 456–466. [[CrossRef](#)]
19. Steenkamp, J.B.; Baumgartner, H. Assessing Measurement Invariance in Cross-National Consumer Research. *J. Consum. Res.* **1998**, *25*, 78–107. [[CrossRef](#)]
20. Byrne, B.M. Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema* **2008**, *20*, 872–882. [[PubMed](#)]
21. Chen, F.F.; Sousa, K.H.; West, S.G. Testing measurement invariance of second-order factor models. *Struct. Equ. Model.* **2005**, *12*, 471–492. [[CrossRef](#)]
22. Hakanen, J.; van Dierendonck, D. Servant Leadership and life satisfaction: the mediating role of justice, job control, and burnout. *Int. J. Servant Leadersh.* **2011**, *7*, 253–261.
23. Bobbio, A.; van Dierendonck, D.; Manganelli, A.M. Servant leadership in Italy and its relation to organizational variables. *Leadership* **2012**, *8*, 229–243. [[CrossRef](#)]
24. Pircher Verdorfer, A.; Peus, C. The Measurement of Servant Leadership: Validation of a German Version of the Servant Leadership Survey (SLS). *Z. Arbeits Organisationspsychol.* **2014**, *58*, 1–16. [[CrossRef](#)]
25. Sousa, M.; van Dierendonck, D. Servant Leadership and Engagement in a Merge Process. *J. Organ. Chang. Manag.* **2014**, *27*, 877–899. [[CrossRef](#)]
26. Rodríguez-Carvajal, R.; de Rivas, S.; Herrero, M.; Moreno-Jiménez, B.; van Dierendonck, D. Leading People Positively: Cross-Cultural Validation of the Servant Leadership Survey (SLS). *Span. J. Psychol.* **2014**, *17*, e63. [[CrossRef](#)] [[PubMed](#)]
27. Gunnarsdóttir, S.; Jónsdóttir, B.G. Servant leadership and research in Iceland. *Icel. Rev. Polit. Adm.* **2013**, *9*, 415–438. [[CrossRef](#)]
28. Duyan, E.C.; van Dierendonck, D. Hizmetkar Liderligi Anlemal: Teorisen ampirk arastirmaya dogru. *Sosyol. Konf.* **2014**, *49*, 1–32.
29. Brislin, R.W. Back-translation for cross-cultural research. *J. Cross Cult. Psychol.* **1987**, *1*, 185–216. [[CrossRef](#)]
30. Muthén, L.K.; Muthén, B. *Mplus 7.11 User's Guide*; Muthén & Muthén: Los Angeles, CA, USA, 2013.
31. Chen, F.F. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model.* **2007**, *14*, 464–504. [[CrossRef](#)]
32. Verdoodt, A.; van Dierendonck, D. Development of a leadership forgiveness measure. *Int. J. Servant Leadersh.* **2010**, *6*, 285–292.
33. Hofstede, G. *Culture's Consequences: International Differences in Work-Related Values*; Sage Publications: Beverly Hills, CA, USA, 1980.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).