**Appendix 1: Measurement Invariance, Multicollinearity, Discriminant Validity, and Common-Method Variance**

**Measurement Invariance.** We conducted measurement invariance tests for our predictors and outcome measures in all nine culture clusters. For group comparisons, parameters do not have to be exactly equal; it is sufficient if they are approximately equal (Muthén & Asparouhov, 2013). We therefore conducted an approximate measure of invariance using alignment optimization methods with the *sirt* package from R software. Results can be interpreted as follows: *R2*values vary between 0 and 1, whereby values close to 1 indicate high levels of invariance. Metric invariance is estimated through factor loadings and scalar invariance through the intercepts. As the basis for alignment optimization methods is approximate rather than exact invariance, group comparisons are still valid with up to 25% noninvariant parameters (Muthén & Asparouhov, 2014).

**Multicollinearity and Discriminant Validity.** Second, we conducted two tests to ensure that our variables were not systematically biased and were sufficiently different from each other. As a first step, we tested for multicollinearity (Daoud, 2017). The variance inflation factor (VIF) is a common estimate for this, with values up to 10 indicating no significant multicollinearity in the data (Alin, 2010). In a second step, we tested for discriminant validity between the four leadership constructs. Measures can be considered sufficiently different from one another if the average variance extracted (AVE) value of the focal measure is higher than its squared correlation with a second measure (Fornell & Larcker, 1981). We therefore compared AVE values of each leadership construct with the squared correlations of the other three leadership forms.

**Common-Method Variance.** We took additional steps to reduce potential bias in our data set. More precisely, we applied the unmeasured latent construct method technique (ULCM; Podsakoff et al., 2003; Richardson et al., 2009) to assess the presence of common-method variance. This technique implies loading all indicators assessed with the same method (e.g., self-reports) into an unmeasured latent factor. This approach allows us to distinguish the common variance shared by model indicators (e.g., items), also known as “construct-related variance,” from the common-method variance resulting from the method used to measure such indicators (e.g., self-reports) or from using responses originating from the same source (common-source variance).

Richardson et al. (2009) explain that adding a latent factor without any construct-specific indicators into a model, but instead capturing all indicators measured with the same instrument (or source), allows us to extract the common-method (or -source) variance from the model, and thus corrects the model’s correlation coefficients (and derived standardized regression coefficients). This approach allows us to empirically test the presence of common-method variance by comparing changes in the goodness-of-fit indicators of two nested models (the model without the common-method variance latent factor and with it). For common-method variance not to be a concern, the model with the common-method variance latent factor should have a better fit than the model without the common-method variance factor.

**Results**

**Measurement Invariance.** Factor loadings were *R2* = .998 and intercepts *R2* = 1.00 for transformational leadership. Each of the 63 items/culture cluster combinations (seven items and nine culture clusters) had invariant factor loadings, whereas two intercepts were noninvariant. These include item 2 for Confucian Asia and Middle East.. The degree of noninvariant factor loadings (0%) and intercepts (3.2%) was on average 1.6%, hence below the 25% threshold, confirming measurement invariance for the transformational leadership measure across the nine culture clusters.

For authentic leadership, *R2* was .994 for factor loadings and .999 for intercepts. Each of the 72 item/culture cluster combinations (8 items and 9 culture clusters) had invariant factor loadings. As regards the item intercepts, we found that 15 items were noninvariant. Those included item 1, 2 and 5 for Germanic Europe; items 5, and 6 for Confucian Asia; items 1, 2, 4, 5, and 6 for the Middle East; items 2 and 6 for Latin Europe; items 1 and 2 for Nordic Europe; and item5 for Southern Asia. The Middle East was the least invariant culture cluster whereas Anglo cultures, Eastern Europe, and Latin America were fully invariant. The degree of noninvariant factor loadings (0%) and intercepts (20.8%) was on average 10.4%, which is clearly below the 25% threshold. Therefore, we can confirm measurement invariance in all nine culture clusters for authentic leadership.

For LMX quality, *R2* was .997 for factor loadings and .999 for intercepts. All 63 items/culture cluster combinations were invariant for the factor loadings, whereas nine intercepts were noninvariant. These include items 1, 3 and 5 for Eastern Europe, as well as item 3 for Confucian Asia, items 3 and 5 for Latin Europe, items 1 and 5 for Nordic Europe, and lastly, item 2for Southern Asia. Eastern Europe was the least invariant cluster, whereas Germanic Europe, the Middle East, Anglo cultures, and Latin America were completely invariant. The average noninvariant factor and intercept loadings were 7.2%, leading us to conclude that there was measurement invariance for LMX quality in the nine culture clusters.

Identity leadership showed high factor loadings, *R2* = .998, and intercepts, *R2* = .999. Only two of 135 items/culture cluster combinations (15 items and 9 culture clusters) for factor loadings were noninvariant. These were item 2 for Confucian Asia and item 8 for Southern Asia. In regard to the item intercepts, we found that 19 items were noninvariant. Those included items 8 and 13for Germanic Europe, as well as items 7, 8, 13, and 14 for Eastern Europe, items 2 and 13 for Confucian Asia, and items 9, 13, and 14 for the Middle East. Furthermore, for Latin Europe, items 6, 8, and 14 were noninvariant, and for Southern Asia, items 6 and 8. In Anglo cultures, Latin America, and Nordic Europe, all items were invariant. Germanic Europe was the least invariant cluster. The degree of noninvariant factor loadings (1.5%) and intercepts (14.1%) was on average 7.8%, which is clearly below the 25% threshold. On this basis, we could confirm measurement invariance in all nine culture clusters for identity leadership.

Exploring measurement invariance for FIB, the factor loadings were *R2* = .994 and intercepts *R2* = .999. As we measured innovative behavior with nine items, we tested for 81 items/culture cluster combinations. All of the factor loadings were invariant, but 16 intercept loadings were noninvariant. The noninvariant intercept loadings included items 5 and 6 for Eastern Europe, items 1, 2, 3, and 8 for Confucian Asia, items 5 and 8 for Anglo cultures, as well as items 1, 2, and 9 for Latin Europe, item 4 for Nordic Europe, and items 3, 4, 5 and 9 for Southern Asia. These results indicate that Confucian Asia and Southern Asia were the least invariant culture clusters, whereas Germanic Europe, Middle East, and Latin America were the most invariant. Average noninvariant loadings were 9.9%, leading us to conclude that we can assume measurement invariance for FIB in the nine culture clusters.

In sum, we found measurement invariance for all five constructs across the nine culture clusters. This, then, served as a basis for proceeding with further analyses as planned.

**Multicollinearity and Discriminant Validity.** Results for tests of multicollinearity and discriminant validity also generally indicated no significant problems. More precisely, all but two VIF scores were below 10. Furthermore, the majority of squaredcorrelations/AVE comparisons were above .1, and hence did not indicate an issue with discriminant validity. We therefore considered it appropriate to proceed with further analysis. Detailed results can be found in Appendix 1, Table 1 (multicollinearity) and Appendix 1, Table 2 (discriminant validity).

**Appendix 1, Tab1e 1**

*VIF Scores of Multicollinearity Tests*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Country Cluster** | **TFL** | **AL** | **LMX** | **IL** | **SID** | **PIL** |
| All | 7.97 | 5.56 | 4.02 | 4.65 | 1.84 | 3.57 |
| Anglo cultures | 10.93 | 7.06 | 4.96 | 6.82 | 1.95 | 4.17 |
| Middle East | 7.73 | 6.37 | 4.12 | 3.3 | 2.70 | 4.05 |
| Confucian Asia | 9.71 | 9.29 | 3.35 | 4.31 | 4.21 | 5.89 |
| Eastern Europe | 8.73 | 5.31 | 3.84 | 5.15 | 1.70 | 3.86 |
| Germanic Europe | 5.85 | 3.78 | 3.91 | 4.05 | 1.42 | 3.25 |
| Latin America | 10.52 | 6.06 | 3.77 | 6.76 | 1.68 | 4.23 |
| Latin Europe | 8.64 | 4.73 | 4.72 | 5.34 | 1.54 | 3.47 |
| Nordic Europe | 7.49 | 4.44 | 2.95 | 5.77 | 1.54 | 3.15 |
| Southern Asia | 6.31 | 4.55 | 3.90 | 4.05 | 1.81 | 2.72 |

*Note*. LMX = LMX quality; TFL = Transformational Leadership; AL = Authentic Leadership; IL = Identity Leadership; PIL = Personal Identification with the Leader; SID = Social Identification with the Team and Organization.

**Appendix 1, Table 2**

Average Variance Extracted (AVE)/ Shared Variance (SV) Ratios (Fornell - Larcker Criterion) for Discriminant Validity

| **AVE/SV** |  | **All** | **Anglo Cultures** | **Middle East** | **Confucian Asia** | **Eastern Europe** | **Germanic Europe** | **Latin America** | **Latin Europe** | **Nordic Europe** | **Southern Asia** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| IL Prototypicality | TL | 1.06 | 1.03 | 1.27 | 1.01 | 1.02 | 1.13 | 1.00 | 1.02 | 1.00 | .98 |
| AL | 1.14 | 1.10 | 1.34 | 1.04 | 1.13 | 1.23 | 1.11 | 1.13 | 1.08 | 1.09 |
| LMX | 1.27 | 1.18 | 1.53 | 1.17 | 1.28 | 1.26 | 1.31 | 1.21 | 1.38 | 1.13 |
| IL Advancement | TL | 1.08 | .99 | 1.27 | 1.22 | 1.00 | 1.12 | .95 | 1.07 | 1.02 | .91 |
| AL | 1.16 | 1.05 | 1.33 | 1.26 | 1.11 | 1.22 | 1.05 | 1.19 | 1.10 | 1.00 |
| LMX | 1.29 | 1.13 | 1.52 | 1.43 | 1.25 | 1.25 | 1.24 | 1.27 | 1.41 | 1.04 |
| IL Entrepreneurship | TL | 1.07 | 1.00 | 1.25 | 1.13 | 1.08 | 1.13 | .95 | 1.01 | .99 | .97 |
| AL | 1.15 | 1.06 | 1.31 | 1.17 | 1.19 | 1.24 | 1.04 | 1.12 | 1.07 | 1.08 |
| LMX | 1.28 | 1.15 | 1.49 | 1.32 | 1.35 | 1.27 | 1.23 | 1.20 | 1.37 | 1.11 |
| IL Impresarionship | TL | 1.24 | 1.13 | 1.37 | 1.21 | 1.28 | 1.41 | 1.13 | 1.31 | 1.09 | 1.14 |
| AL | 1.33 | 1.20 | 1.44 | 1.24 | 1.42 | 1.54 | 1.25 | 1.45 | 1.18 | 1.26 |
| LMX | 1.48 | 1.29 | 1.64 | 1.41 | 1.60 | 1.58 | 1.47 | 1.55 | 1.51 | 1.31 |
| TL | IL Prototypicality | 1.07 | 1.02 | 1.30 | 1.15 | 1.00 | 1.04 | .97 | 1.04 | .98 | .98 |
| IL Advancement | 1.06 | .97 | 1.32 | 1.16 | .99 | .99 | .93 | 1.01 | .99 | .91 |
| IL Entrepreneurship | 1.00 | .95 | 1.19 | 1.06 | 1.02 | .97 | .90 | .95 | .92 | .88 |
| IL Impresarionship | 1.17 | 1.14 | 1.28 | 1.12 | 1.20 | 1.36 | 1.05 | 1.22 | 1.06 | 1.01 |
| AL | .88 | .89 | .93 | .83 | .89 | .86 | .86 | .88 | .87 | .79 |
| LMX | 1.01 | .99 | 1.07 | 1.15 | 1.03 | .89 | .99 | .93 | 1.14 | .92 |
| AL | IL Prototypicality | .96 | .93 | 1.22 | 1.16 | .87 | .90 | .93 | .89 | .82 | .93 |
| IL Advancement | .95 | .89 | 1.23 | 1.18 | .86 | .86 | .89 | .86 | .82 | .87 |
| IL Entrepreneurship | .90 | .87 | 1.12 | 1.07 | .89 | .83 | .86 | .81 | .76 | .83 |
| IL Impresarionship | 1.06 | 1.04 | 1.20 | 1.13 | 1.05 | 1.17 | 1.01 | 1.04 | .88 | .96 |
| TL | .74 | .76 | .83 | .82 | .70 | .68 | .75 | .68 | .67 | .68 |
| LMX | .87 | .90 | .96 | 1.11 | .82 | .77 | .88 | .75 | .89 | .78 |
| LMX | IL Prototypicality | 1.10 | 1.03 | 1.33 | 1.17 | .98 | 1.06 | 1.08 | 1.09 | 1.12 | 1.01 |
| IL Advancement | 1.09 | .99 | 1.35 | 1.18 | .97 | 1.01 | 1.03 | 1.06 | 1.13 | .94 |
| IL Entrepreneurship | 1.03 | .97 | 1.22 | 1.07 | 1.00 | .99 | 1.00 | 1.00 | 1.05 | .90 |
| IL Impresarionship | 1.21 | 1.15 | 1.31 | 1.14 | 1.18 | 1.39 | 1.17 | 1.28 | 1.21 | 1.04 |
| TL | .87 | .87 | .92 | 1.00 | .81 | .81 | .85 | .82 | .94 | .83 |
| AL | .90 | .92 | .92 | .99 | .82 | .89 | .86 | .86 | .95 | .82 |

*Note*. IL = Identity Leadership; LMX = Perceived LMX quality; AL = Authentic Leadership; TL = Transformational Leadership. In order to satisfy the criterion, AVE/SV should be > 1.00.

**Appendix 2**

*Rescaled Relative Weights by Culture Cluster (Relative Importance Analysis)*

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Anglo**  **cultures** | **Confucian Asia** | **Eastern Europe** | **Germanic Europe** | **Latin America** | **Latin Europe** | **Middle East** | **Nordic Europe** | **Southern Asia** |
| TFL | 1.01 | 0.95 | 1.88 | 3.06 | 1.88 | 4.10 | 2.99 | 1.45 | 2.34 |
| AL | 13.38 | 2.71 | 7.49 | 10.34 | 9.09 | 14.45 | 1.12 | 2.41 | 3.70 |
| LMX | 2.06 | 33.50 | 12.98 | 5.49 | 14.25 | 11.74 | 7.34 | 8.46 | 11.07 |
| IL | .75 | 3.11 | 4.64 | 2.65 | 1.32 | 4.17 | 7.02 | .86 | 10.15 |
| SID | 60.32 | 51.52 | 56.55 | 48.98 | 60.46 | 56.03 | 63.77 | 70.99 | 56.60 |
| PIL | 10.57 | 5.25 | 8.05 | 17.07 | 6.73 | 4.97 | 9.37 | 8.38 | 9.26 |
|  |  |  |  |  |  |  |  |  |  |
| CMV | 11.92 | 2.96 | 8.41 | 12.41 | 6.27 | 4.54 | 8.38 | 7.46 | 6.88 |

*Note*. IL = Identity Leadership; TFL = Transformational Leadership; AL = Authentic Leadership; SID = Social Identification; PIL = Personal Identification with the Leader; CMV = Common-Method Variance Factor.