**Appendix 1**

**Table A1: Measurement Items Used in the Study**

*Unit Level Transformational Leadership (TFL):* (García-Morales *et al*., 2008; p. 211) **Factor Loading**

1. The unit leader has a clear common view of its final aims and is able to transmit them and achieve the commitment

of the rest of the unit’s employees 0.85

1. The unit leader succeeds in motivating and guiding unit employees on the job 0.87
2. The unit leader is always on the lookout for new opportunities for the unit 0.86
3. The unit leader always acts as the unit’s leading force 0.89

Eigenvalue = 5.36; variance explained = 80.23%

*Unit Knowledge-Sharing Climate (KSC):* (Husted and Michailova, 2002)

1. Time spent on knowledge sharing in this unit is generally time well spent 0.85
2. Employees in the unit do not consider knowledge sharing risky because people are generally open to new ideas 0.83
3. The unit provides employees with support for knowledge sharing 0.88
4. Employees in the unit have a sense of security when sharing knowledge with colleagues 0.86

Eigenvalue = 4.52; variance explained = 41.22%

*Internal to unit Knowledge-Sharing (IKS):* (Lee, 2001; Lu *et al*., 2006)

1. Meet with my colleagues in this unit and exchange ideas with them regularly 0.82
2. Access my colleagues in this unit and exchange new ideas and developments with them 0.81
3. Interact with my colleagues in this unit to discuss suggestions and ideas 0.87
4. Make sure to be available for sharing experiences with my colleagues in this unit 0.84

Eigenvalue = 6.23; variance explained = 78.55%

*Unit Innovation Performance* (OECD/Eurostat 2005 and López-Cabrales, *et al*., 2009 T1 T2

1. Market introduction of technologically new product (good or service) developed by this unit (totally or in part); 0.80 0.81
2. Market introduction of technologically improved product (good or service) developed by this unit (totally or in part); 0.84 0.85
3. *Extensions of existing product (good or service) lines (that do not only entail changes to aesthetic aspects); 0.35 0.36*
4. *Changes introduced to existing product (good or service), entailing significant improvement; 0.31 0.32*
5. Development of new lines/ranges of product (good or service); 0.86 0.87
6. Frequency of replacement of old product (good or service) by others with important changes; 0.88 0.89
7. Proportion of technologically new or improved product (good or service) in the turnover of the company; 0.83 0.83
8. Product (good or service) innovation performed by the unit. 0.86 0.88

Eigenvalue T1 = 5.91; variance explained = 73.29%; Eigenvalue T2 = 5.93; variance explained = 73.44%

Note: Items 3 and 4 were deleted due to low factor loadings.

**Appendix 2**

This technical appendix reports on tests conducted for: (a) Sample Representativeness; (b) Construct Validity and Model Fit; (c) Data Duality Checks; and (d) Data Aggregation.

**(a) Sample Representativeness**

We conducted several tests to check the representativeness of our sample. We assessed response bias, utilising the GRS database to assess differences in net sales, employment levels and firm age between responding and non-responding firms for both Time 1 and Time 2. The Kolmogorov-Smirnov test (Kleinbaum, Kupper & Muller, 1988) showed no significant differences in any of these variables in either time period (at the *p*<0.10 level). We also computed a two-stage Heckman Test to investigate response bias. No significant biases were found. Given that we used self-report measures of unit innovation performance, we were able to cross-validate the unit leaders’ responses on innovation with just over half of the units that had an R&D specialist. We found no statistically significant differences within these units between the unit leaders and R&D specialists’ responses on the innovation items (at the *p*<0.10 level).

**(b) Construct Validity and Model Fit**

We used confirmatory factor analysis (CFA) to establish construct validity (Table A2.1). Model fit was evaluated using a number of criteria: The *X2* statistic and its degrees of freedom, the robust root mean square error of approximation (RMSEA) (90% CI), the robust comparative fit index (CFI) and the robust Tucker-Lewis index (TLI). Values of CFI ≥.90, TLI ≥ .90 and RMSEA ≤ .10 indicate an acceptable fit of the research model with the data, while values of CFI ≥.95, TLI ≥ .95 and RMSEA ≤ .06 indicate a good fit (Hu & Bentler, 1999).

**Table A2. 1. Data Sources, Time Periods, Confirmatory Factor Analysis (CFA), Reliability and Agreement of Scales**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable and Respondents** | **Time Period(s)** | **χ2** | **Robust**  **Root Mean Squared Error of Approximation (RMSEA)** | **Robust**  **Comparative Fit Index (CFI)** | **Robust**  **Tucker Lewis Index (TLI)** | **Internal Consistency**  **Cronbach’s Alpha (α)** | **Average Variance Extracted** | **Maximum Squared Correlation** | **Rwg** | **ICC(1)** | **ICC(2)** |
| Unit-Level Transformational Leadership (TFL)  (employees) | T1 | χ2 = 791.44, *df* = 326, *p*<0.001 | 0.043 | 0.982 | 0.976 | 0.873 | 0.623 | 0.376 | 0.78 | 0.24 | 0.81 |
| Unit Knowledge-Sharing Climate (KSC) (employees) | T1 | χ2 = 744.06, *df* = 314, *p*<.001 | 0.042 | 0.993 | 0.985 | 0.923 | 0.856 | 0.412 | 0.88 | 0.29 | 0.86 |
| Internal to unit Knowledge Sharing (IKS) (employees) | T1 | χ2 = 830.42, *df* = 319, *p*<0.001 | 0.034 | 0.980 | 0.972 | 0.893 | 0.782 | 0.356 | 0.80 | 0.30 | 0.83 |
| Unit Innovation Performance (IPerf) (unit leader) | T1  T2 | χ2 = 124.62, *df* = 36, *p*<.001  χ2 = 120.83, *df* = 36, *p*<.001 | 0.032  0.031 | 0.982  0.991 | 0.976  0.983 | 0.823  0.857 | 0.698  0.792 | 0.349A  0.362 | n/a  n/a | n/a  n/a | n/a  n/a |

To test for model fit we first estimated the hypothesised model with four factors: TFL, KSC, IKC and IPerf (Time 2). This model (Model A, Table A2.2) showed a good fit to the data (*X2* = 1265.19, *df* = 330; *p* < 0.001; robust RMSEA = 0.0.018 (0.000, 0.041) 6; robust CFI = 0.953; and robust TLI = 0.926). Next, we estimated three alterative models (Models B to D containing between three and one factors respectively) and compared those with Model A. In conclusion, the CFA results provide support for the construct validity of our study measures.

**Table A2. 2. Model Fit Indices and Model Comparisons for Estimated CFA Models**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **χ2** | **Robust**  **Root Mean Squared Error of Approximation (RMSEA)** | **Robust**  **Comparative Fit Index (CFI)** | **Robust**  **Tucker Lewis Index (TLI)** | **Model Comparison** |
| Four Factor Model A | χ2 = 1265.19, *df* = 330, *p*<0.001 | 0.018  (0.000, 0.045) | 0.953 | 0.926 |  |
| Three Factor Model B | χ2 = 1767.12, *df* = 334, *p*<0.001 | 0.155  (0.066, 0.091) | 0.794 | 0.661 | Vs 4-Factor Model |
| Two Factor Model C | χ2 = 2867.12, *df* = 338, *p*<.001 | 0.178  (0.123, 0.239) | 0.530 | 0.495 | Vs 4-Factor Model |
| One Factor Model D | χ2 = 3766.58, *df* = 342, *p*<.001 | 0.195  (0.221, 0.332) | 0.388 | 0.292 | Vs 4-Factor Model |

**(c) Data Quality Checks**

We tested for discriminant validity using the Heterotrait-monotrait (HTMT) ratio criterion using the formula suggested by Henseler, Ringle &Sarstedt (2015).

**Table A2. 3. HTMT Results**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **TFL** | **KSC** | **IKS** | **IPerf**  **T1** | **IPerf**  **T2** |
| Unit-Level Transformational Leadership (TFL)  (employees) | - |  |  |  |  |
| Unit Knowledge-Sharing Climate (KSC) (employees) | 0.77 | - |  |  |  |
| Internal to unit Knowledge Sharing (IKS) (employees) | 0.81 | 0.79 | - |  |  |
| Unit Innovation Performance (IPerf) (unit leader) | 0.70 | 0.72 | 0.73 | - | - |

From the HTMT results, the values in Table A2.3 indicate there are no discriminant validity problems according to the HTMT = 0.85 criterion. In a subsequent step, we utilised CFA to establish convergent validity. Convergent validity is accepted when factor loadings are higher than 0.4 and t coefficients are significant. We follow Jackson, Gillaspy & Purc-Stephenson, (2009) on reporting practices in CFA and report multiple fit indices for each scale (Table A2.2).

We assessed the presence and potential influence of common method variance. We used Harman’s one factor text, CFA and post-hoc analysis to test for the presence of common method variance. We entered the scales into an exploratory factor analysis using principle component analysis with variance rotation and principle axis analysis with varimax rotation to determine the number of factors that are necessary to account for a substantial proportion of the variance. We also loaded all three scales on one factor to investigate the fit of the CFA model. These various analyses all revealed the presence of three distinct factors with eigenvalues greater than 1.0, rather than one single factor. The three factors together accounted for 55.4 per cent of the total variance, while the first (largest factor) did not account for the majority of the variance (28.2 per cent). Thus, no general factor is apparent. The results of the CFA revealed that the single factor model did not fit the data well (*X2* = 3.617, *df* = 315; *p* < 0.001; RMSEA = 0.279; CFI = 0.582, TLI = 0.439). These analyses do not preclude the possibility of common method variance; however, they indicate that common method variance is not a major concern and is unlikely to confound the interpretation of the results.

We tested for autocorrelation and computed a Wald test based on the null hypothesis of the independence of residual terms. The *p* value for this test was 0.365. We, therefore, fail to reject the hypothesis that the residuals are independent, indicating that autocorrelation is not an issue in our estimations. In addition, we conducted a Breusch-Pagan test for heteroscedasticity in the control model and the results reveal that this was an issue for these estimates [*X2* (10) = 11.35, *p* <0.001]. Therefore, we corrected for this using robust standard errors (reported for the estimations).

Finally, we investigated longitudinal invariance. We used the same unit innovation performance measure at Time 1 and Time 2. The invariance test requires that the data are analysed by fitting the two waves of data simultaneously with two separate models. We used the four-step method proposed by Van de Schoot, Lugtig & Hox (2012). We found that the factor loadings were not significant across the models estimated. We concluded that the innovation measure used in the estimations is time invariant.

**(d) Data Aggregation**

As we were interested in investigating unit-level TFL, unit knowledge-sharing climate and internal knowledge sharing, we established within-group agreement and between-group variability to decide whether individual level scores could be aggregated to the unit level. In line with previous research (Ehrhart & Naumann, 2004) we estimated Rwg scales for each unit, and ICC (1) and ICC (2) values (see Table 1, last 3 columns). First, Rwg scores provide an indication of the level of agreement within each unit and the ideal cut-off is .70 (Kline, 2005). The average Rwg scores in our sample were 0.78 (range 0.66 – 1.00) for TFL, 0.88 (range 0.72-1.00) for unit KSC; and 0.80 (range 0.72 – 1.00) for IKS. One unit fell below the 0.70 threshold for unit KSC and we thus excluded it from the analysis (Chen, Mathieu, & Bliese, 2005). This reduced the final number of units used in the analysis to 124.

We computed interclass correlation coefficients using ICC (1). These results indicated that 24 per cent of the variance in TFL, 29 per cent of the variance in unit KSC and 30 per cent for unit IKS could be explained by unit membership. Kline (2005) explain that when using ICC (1), if the F test for between groups from the ANOVA is significant, the aggregation of participants within each group is considered justified. The relevant F scores were all statistically significant. Finally, the ICC (2) values for TFL of (0.81), unit KSC (0.86), IKS (0.83), all exceed the recommended 0.70 cut-off.