

Final Report For

DENISON CONSULTING

**MEASUREMENT EQUIVALENCE OF THE
DENISON ORGANIZATIONAL CULTURE SURVEY
(DOCS) ACROSS LANGUAGE ADAPTATIONS**

AUGUST 2007

Team Members:

Nathan T. Carter, M.A.

Dalia L. Diab, M.S.

Bing C. Lin, B.A.

Shuang-Yueh Pui, B.S.

Michael J. Zickar, Ph.D.



**Institute for Psychological Research and Application
DEPARTMENT OF PSYCHOLOGY
BOWLING GREEN STATE UNIVERSITY**

Executive Summary

This report contains the results of statistical analyses used to determine the comparability of the items and scales of the Denison Organizational Culture Survey (DOCS) across language translations. More specifically, we investigated the extent to which different language adaptations could be compared to the United States (US) English version of the DOCS at both the item level as well as at the 5-item scale level.

The results of the analyses suggested that the majority of scales, or in the terms of the DOCS, trait scores, are comparable to the US English version. Overall, 88% of the items and all but three trait scale adaptations were equivalent psychometrically. At the scale level the two exceptions were **the French version of the *Empowerment* scale** and **the German version of the *Customer Focus* scale**. These scales *should not* be compared to the US English version of the corresponding scales.

Results also indicated that the same dimensions of culture are found across languages *with the exception of the Turkish version of the DOCS*. The Turkish version of the DOCS appears to rely on the same dimensions as the US English version, but is not as reliable of a measure as the US English version, thus the client is advised to take caution in comparing this language translation to the US English version.

All language translations having a sample size of less than 300 were analyzed statistically, but these analyses were based on the classical test theory approach, and hence do not warrant claims concerning the comparability of the scales. The results of these analyses are included in this report for descriptive purposes only.

Overall, the results were positive, in that 88% of items and all but three trait scale adaptations showed evidence of equivalence. Thus, nearly all scale-level comparisons to the US English version of the DOCS are psychometrically justifiable. Moreover, the empirical findings make intuitive sense, in that the only language found to be wholly comparable (i.e. at both the

item and trait levels) to the US English version was the highly similar United Kingdom English version, and that the languages with the highest rates of non-comparable items were Eastern/Asian language translations. In addition, similar languages, (i.e. Portuguese and Portuguese-Brazil; Chinese Traditional and Chinese Simplified) were found to share many of the same non-comparable items.

Background and Purpose of Study

Denison Consulting contracted the Institute for Psychological Research and Application (IPRA) at Bowling Green State University (BGSU) to assess the measurement equivalence of the DOCS across languages. The IPRA team evaluated the measurement equivalence of the DOCS using various statistical analyses such as Item Response Theory (IRT), Confirmatory Factor Analysis (CFA), and Classical Test Theory (CTT). The focus of the analyses was at the item and trait levels, as requested by Denison Consulting. The IPRA project team used the US English measure as a reference point and compared it to measures adapted to other languages. Results of this project will aid Denison Consulting in determining whether the DOCS can be used to measure organizational cultures across languages. This report details the methods and results of each of the statistical analyses used to determine the measurement equivalence of the DOCS.

In this report, we first provide an overview of the methods that we used, focusing on explaining the basics of IRT and describing the process of evaluating differential item functioning (DIF). Next we discuss the results of the IRT/DIF analyses. Then we describe the CFA procedures used to test for differences in factor structure as well as discuss the results. Then we discuss the CTT analyses. Finally, we present recommendations based on the whole body of results.

Overview of Item Response Theory and the Graded Response Model

We used IRT analyses to model how respondents in each culture answered items for the DOCS. IRT is the foundation for the DIF analyses that we used to determine whether items function the same way across language translations.

IRT is preferred when examining individual items for DIF because it allows for differences in *item discrimination* as well as individual *option endorsement* rates. Item discrimination refers to the degree that a particular item is useful in differentiating cultures that are high on a particular dimension versus items that are low on that dimension. To see this visually, please refer to Figures 1 and 2. Figure 1 presents the fifth item of the Dutch translation of the *Strategic Direction and Intent* scale, a highly discriminating item. Comparing this to Figure 2 which presents the second item of the same scale, one can see that the dominance of response probability at particular levels of the trait for each response option is more pronounced in item 5. Thus, item 5 is likely to be more sensitive in differentiating between cultures at different levels of the trait. Items may also differ in terms of option endorsement across language translations. Some options may be endorsed more frequently in a particular culture, whereas other options may be endorsed less frequently.

The particular model that we used is the Graded Response Model (GRM; Samejima, 1969), which is the preferred model to use with data that have ordered response options (i.e., put in the options) such as the DOCS. We used the computer program MULTILOG 7.03 (Thissen, 2003) to estimate the item parameters for each item using the GRM. Next, we used the Microsoft EXCEL-based program MODFIT (Stark, 2001) to determine whether the GRM fit the data. These goodness-of-fit analyses are necessary for interpreting item parameters. The data were found to have reasonable fit using the c^2/df ratio criteria established by Drasgow, Levine, Tsien, Williams, and Mead (1995), suggesting that item parameter estimates were readily interpretable. Once we established that the GRM fit the data, we were able to proceed to the DIF analyses.

Description of IRT-based Differential Item Functioning Procedure

The comparability of test scores across languages, forms, and various populations is an important consideration in their use and interpretation. In this specific situation, Denison Consulting was concerned with the comparability of test and item scores of different language adaptations to the United States English version of the DOCS. Here, the concept of differential functioning (DF) assessed at the item level (DIF) and the test level (differential *test* functioning; DTF) are outlined.

DF examines the difference between the expected scores of two examinees with the same standing on the latent trait, but who are from different groups (e.g. one examinee takes the English version of a test, and the other takes the Chinese version). If an examinee who takes one form of a test has a greater probability of endorsing an item in a particular direction (e.g. higher on the five-point scale) than that of an examinee with the same standing on the trait who takes a different form of the test, then the item response is not only dependent upon the trait being measured, but also on the form taken by the examinee. This might occur if the translation was not equivalent to the original version. Thus, one of the examinees may have a different score simply because they took a different language version of the test (Raju & Ellis, 2002).

When investigating DF, one must choose a *reference* group (e.g. the US English version), which typically consists of examinees taking the original version of the test. This reference group is then compared to various *focal* groups (e.g. all other language adaptations). To facilitate DIF analyses, the groups' item and person parameters must first be adjusted to the same metric. This is done through IRT-based linking, or equating. This procedure is an application of Stocking and Lord's (1983) test characteristic curve (TCC) method. In the TCC method, the item parameter estimates of common items (all translated items are used as anchor items in language adaptations) are used to compute constants that transform the slope and intercept of the TCC such that the focal and reference groups are on the same scale. For the current study, Baker's

(1995) equating program was used, following the example of Raju and Ellis (2002). Following the conduct of IRT equating, the test may then be examined at both the level of the item and at the level of the total test score.

The procedure used to investigate DF, introduced by Raju et al. (1995), is known as the Differential Functioning of Items and Tests (DFIT) framework. There are two primary statistics examined in determining the DF of items and tests: *Non-compensatory differential item functioning* (NCDIF) and *differential test functioning* (DTF). The NCDIF statistic represents the difference between item scores holding the level of the trait constant. If the NCDIF statistic is above .096, the item scores for that item are not considered to be comparable across forms. However, if it is less than .096, item score comparisons are considered to be justified. The DTF statistic is simply the difference between the summated item scores, or test scores, and is conceptually equivalent to the NCDIF statistic. When DTF is above .48, test-level score comparisons across groups or forms is not justified. More technical information about differential item/test functioning is presented in Appendix A.

Results of Differential Item/Test Functioning Analysis

Table 1 shows the DIF analyses for all languages at the item (DIF) and trait (DTF) levels. Items and scales that show DIF or DTF are marked in the table. Overall, the results of the 168 scale-level comparisons (DTF) in this study suggest that 166 of these alternate language adaptations of the DOCS are equivalent to the US English version of the survey, with the exception of the French adaptation of the 5-item *Empowerment* scale and the German adaptation of the 5-item *Customer Focus* scale. Other than these two scales, the comparison of scores on alternate language adaptations to the US English version of the DOCS appears to be psychometrically justified *at the scale level*, as suggested by examination of DTF statistics for each of the 12 traits for each of the 14 languages analyzed under the DFIT framework.

For item-level comparisons, the large majority of items were found to be comparable to the US English version of the DOCS. Of the 840 items compared to the US English version, 12% were found to be non-comparable, another positive finding in terms of the measurement properties of the DOCS. The language translations of Chinese Simplified, Chinese Traditional, French, German, Portuguese, Japanese, Korean, and Turkish had the highest number of items not comparable to the US English version of the items (between 8 and 16 items), with the Turkish and German languages having the highest numbers (12 and 16, respectively) of non-comparable items. Additionally, the French language translation was found to have 10 items that were not comparable to the US English DOCS. These findings are not surprising in light of the scale-level findings of the IRT and CFA analyses, in which these 3 languages were found to have non-comparable scales.

Interestingly, the languages that had high proportions of non-comparable items but were found to have comparable scales were language translations of Eastern cultures (i.e. Chinese Traditional and Simplified, Japanese, and Korean) with the exception of non-Brazilian Portuguese. Moreover, it was found that all items and scales in the United Kingdom English version of the DOCS are comparable, which lends some intuitive appeal to the results of the analyses. The remaining languages (Dutch, Italian, Portuguese-Brazil, Spanish non-Latin American, and Spanish Latin American) were found to have 5 or less items that are not comparable to the US English version of the DOCS. We recommend that these individual items be scrutinized to see if there can be improvements made in the translation.

The trait scales with the highest numbers of instances in which items were found to be non-comparable were the *Empowerment* (13 instances), *Customer Focus* (18 instances), and *Organizational Learning* (18 instances) trait scales. All other traits were found to have 8 or less instances of non-comparable items for all languages. The trait of *Goals and Objectives* was far below other traits with only 2 instances of non-comparable item findings.

Description of Confirmatory Factor Analysis Procedures

Although IRT methods are useful in determining whether individual items and scales can be compared across languages, IRT methods cannot be used to evaluate whether the structure of the DOCS, as a whole, changes across translations. This is important because even if individual items and scales remain equivalent across languages, it may be possible that for some languages, the higher-order structure of the DOCS changes. For this type of analysis, it is necessary to use CFA.

Confirmatory Factor Analysis

Under Joreskog and Sorbom's (1996) linear structural relations (LISREL) model, the CFA submodel is used to model psychological measurements. This is done by the analysis of a series of linear interrelationships among observed variables, proposed latent or explanatory variables, and the errors associated with the measurement process. Using this modeling approach allows the researcher to test a priori hypotheses concerning the statistical structure of a set of test scores.

Model-data fit in CFA is considered to be evidence of the appropriateness of the measurement model specified by the researcher in describing the structure of some set of test scores. In the case of the DOCS, the measurement model that was specified consisted of 12 separate dimensions of organizational culture. A model with these specifications was tested separately for each language form of the DOCS. These analyses should be interpreted simply as the degree of evidence that the suggested Denison model is appropriate for that language form of the DOCS.

To determine the appropriateness of the 12-factor measurement model, three primary statistics provided by the LISREL software (Joreskog & Sorbom, 2003) were examined: 1) the χ^2 statistic, which should be low and non-significant; 2) the non-normed fit index (NNFI; Tucker

& Lewis, 1973); and 3) the comparative fit index (CFI; Bentler, 1990). Both the NNFI and the CFI should be greater than .90.

Using CFA to Establish Measurement Equivalence

In using CFA to test hypotheses about the structure of test scores, one is only making assertions about one group of test-takers, or one form of the test. However, the concern of this investigation is to compare the structure of test scores across languages. To make these comparisons, the fit of each translated form must be compared to the reference form, or group, which in this case is the US English version of the DOCS. After choosing a focal group to which the reference group is compared, the a priori model is tested across both the reference and focal groups to establish a sort of baseline model, in which almost all parameters are estimated freely, and are not dependent upon the values of other parameters in the model (Joreskog & Sorbom, 1996).

After estimating this baseline model, parameters in the model are sequentially constrained to reflect our hypothetical, desired result and these constrained models are compared in sequence (Joreskog & Sorbom, 1996). Measurement equivalence implies *metric invariance*, which suggests that the strength of the interrelations between items and their respective factors, or factor loadings, are equal across language adaptations. If this criterion is met, it means that each item measures culture in the same way across languages. From there, we compare the *metric invariant* model to the more restricted *structural invariant* model, in which factor loadings as well as the factor structures are hypothesized to be equal across groups. This allows us to make statements concerning whether the same dimensions of organizational culture are observed across languages. Lastly, the most highly constrained model, the *equal error variances* model specifies that factor loadings, error variances, and intercepts are equal across groups. This allows us to make assertions about the equality of measurement reliability and error across languages.

To examine the differences in these successively constrained measurement models, a simple difference in the χ^2 statistics used to evaluate the fit of the respective models being compared is computed. If this difference is not significantly different, then this is evidence that the proposed measurement model is appropriate for both forms. NNFI and CFI were examined and used to determine the degree to which the specified measurement model differs across language adaptations. Please see Appendix B for a more detailed and technical treatment of this approach.

These analyses will help us better understand how the structure of the DOCS changes across languages. The CFA approach to measurement equivalence presented here complements the DFIT analyses which focus on individual scales, whereas this analysis focuses on how the scales relate to each other.

Results of Confirmatory Factor Analysis

Overall, the results from the CFA suggest that the alternate language adaptations of the DOCS are equivalent to the US English version of the survey for all language translations except the Turkish adaptation. These results suggest that the higher-order structure of the DOCS remains constant across all but the Turkish translation. The one exception, the Turkish adaptation, found that although the Turkish translation had the same factor loadings and factor structure, reliability of measurement was not equivalent for the two forms. Other than the Turkish adaptation, the comparison of scale scores on alternate language adaptations to the US English version of the DOCS appears to be psychometrically justified at the metric and structural levels, as suggested by examination of fit statistics for each of the constrained models in the 14 languages analyzed.

Table 2 presents results from the CFA analyses for the 14 languages by type of invariance (i.e., Metric, Structural, and Equal Errors). Fit indices that suggested non-equivalence were marked in this table. Most of the chi-squared fit indices indicated that the models for all

languages were not equivalent. Because the χ^2 fit index is highly influenced by sample size, we would expect these results. This means that the evaluation of equivalence using the χ^2 fit indices should be considered with the other fit indices (NNFI and CFI). Using this strategy, the only language found to be non-equivalent is the Turkish adaptation of the US English version of the DOCS. Further, non-equivalence was found only for the Equal Errors Model, which implies that the error terms for the observed variables in the model for the Turkish translation are not equivalent to those in the US English version, and thus the two forms are not equally reliable measures of organizational culture.

Results of Classical Test Theory Analysis

A CTT approach was used to analyze the data for languages that did not have adequate sample sizes for using an IRT or a CFA approach (i.e., $N < 300$). Fourteen languages fit that criterion. For each language, descriptive statistics were calculated for all items and traits. Complete results are presented in Tables 3a through 3f, with maximum sample sizes shown within parentheses. Moreover, the US English version results are shown in all tables for comparison purposes (with maximum sample size also shown within parentheses). These CTT results are presented to help Denison get a sense of the presence of differences across languages for these small samples; unfortunately, with the small samples, it is impossible to conduct the more stringent IRT-based and CFA-based analyses. For means and standard deviations at both the item and trait levels, please refer to Tables 3a, 3b, and 3c.

Corrected item-total correlations were calculated for each item, and means of corrected item-total correlations were calculated for each trait. Item-total correlations provide information related to item-level discrimination, similar to the IRT-based index of discrimination. The means of corrected item-total correlations ranged from .14 to .69 for the 14 languages, and from .45 to .67 for the US English version. The lowest mean (.14) was the mean of *Agreement* for the

Vietnamese translation, where item 24 had a negative correlation (-.13). The second lowest mean was the mean of *Creating Change* for the Indonesian translation, where item 34 had a negative correlation (-.14). There were no negative corrected item-total correlations in the US English version, and the lowest mean of .45 was the mean of *Core Values*. The highest mean of .69 was that of *Strategic Direction and Intent* for both the Flemish and French Canadian adaptations of the US English version of the DOCS. Similarly, the highest mean of .67 in the US English version was also that of *Strategic Direction and Intent*. For more detailed results, please refer to Tables 3d and 3e.

Internal consistency reliabilities (coefficient alphas) were calculated for each of the twelve traits, and they ranged from .28 to .87 for the 14 languages, and from .68 to .85 for the US English version. Generally speaking, alphas greater than .70 are considered adequate. Consistent with the previous findings, *Agreement* for the Vietnamese translation and *Creating Change* for the Indonesian translation had the lowest alphas of .28 and .32, respectively. The only two alphas lower than .70 for US English were .68 and .69 for *Core Values* and *Capability Development*, respectively. *Strategic Direction and Intent* had the highest alpha of .87 for both the Flemish and French Canadian versions of the DOCS. Similarly, the trait with the highest alpha of .85 for US English was also for the trait of *Strategic Direction and Intent*. Alphas for all traits are shown in Table 3f.

A note of caution should be made here. For several of the statistics reported, sample sizes were rather small. Even when using a CTT framework, it is preferred that sample sizes be relatively large. Sample sizes ranged from 50 to 57 for Croatian, 75 to 80 for Czech, 65 to 90 for Danish, 119 to 130 for Finnish, 74 to 90 for Flemish, 172 to 202 for French Canadian, 79 to 89 for Hungarian, 30 to 33 for Indonesian, 102 to 110 for Norwegian, 190 to 204 for Polish, 111 to 127 for Russian, 160 to 186 for Swedish, 163 to 195 for Thai, and 38 to 43 for Vietnamese. For US English, sample size ranged from 17475 to 19696. So, for example, the two languages with

the lowest means of corrected item-total correlations and the lowest alphas had the lowest sample sizes (Indonesian and Vietnamese). Therefore, the client is advised to interpret these results with caution, after taking into consideration sample sizes used.

Conclusions and Recommendations

This report presented the results of statistical analyses used to determine the extent to which comparisons of the items and scales of the DOCS are comparable across languages. At the scale level, the DTF statistics analyzed under the DFIT framework suggested that most of the traits were comparable to the US English version with the exception of the French version of the *Empowerment* scale and the German version of the *Customer Focus* scale. However, at the item level, the DIF statistics showed that several items cannot be compared. In addition, the CFA results indicated that the same dimensions of culture are found across languages with the exception of the Turkish version of the DOCS.

Based on the results of the DFIT analyses, the client is advised to review the translated versions of the **French *Empowerment*** and the **German *Customer Focus*** scales. More specifically, the first 4 items of the French *Empowerment* scale (items 1-4) and all 5 items of the German *Customer Focus* scale (items 36-40) showed DIF. Moreover, DIF statistics showed that there were several translations of items that were not comparable to the US English version items. Therefore, the wording used in these items needs to be revised and modified, as deemed necessary. It is also recommended that the country's culture be taken into consideration in an attempt to detect the source of the non-equivalence in these items. Special attention needs to be given to the language translations of Turkish and German, as they had the highest numbers of items not comparable to the US English version, 12 and 16 items, respectively. Based on the results of the CFA analyses, the client is advised to revise the Turkish translation of the DOCS. This finding complements the DFIT analyses, which showed that this translated version of the DOCS also had the second highest number of items displaying DIF. Therefore, by combining

these two sets of results, it is recommended that the Turkish version be completely revised and modified by retranslating it, after taking into account the Turkish societal culture and the wording used.

Finally, although the CTT analyses do not warrant any definitive conclusions concerning the comparability of the scales because results are mainly descriptive in nature, a few recommendations can be made. More specifically, regarding means of corrected item-total correlations, the lowest mean of .14 was the mean of *Agreement* for the Vietnamese translation, where item 24 had a negative corrected item-total correlation of -.13. The second lowest mean was the mean of *Creating Change* for the Indonesian translation, where item 34 had a negative corrected item-total correlation of -.14. Therefore, it is recommended that these two items and these two scales be revised. Reliability analyses complemented the corrected item-total correlations results by showing that *Agreement* for Vietnamese and *Creating Change* for Indonesian had the lowest alphas of .28 and .32, respectively. Hence, based on these results, it is also recommended that the Indonesian and Vietnamese versions of the *Agreement* and *Creating Change* items be scrutinized to see if the translations can be improved. Given the small sample sizes of these translations, however, it may be possible that these results are due to sampling error and are not indicative of the need for further translation assistance. Therefore, although the client is advised to review these specific item translations, it is recommended that more data be collected prior to making any item modifications. In general, it is recommended that more data be collected for all languages analyzed using the CTT framework to allow for IRT and CFA analyses.

Figure 1 – Example of a Highly Discriminating Item

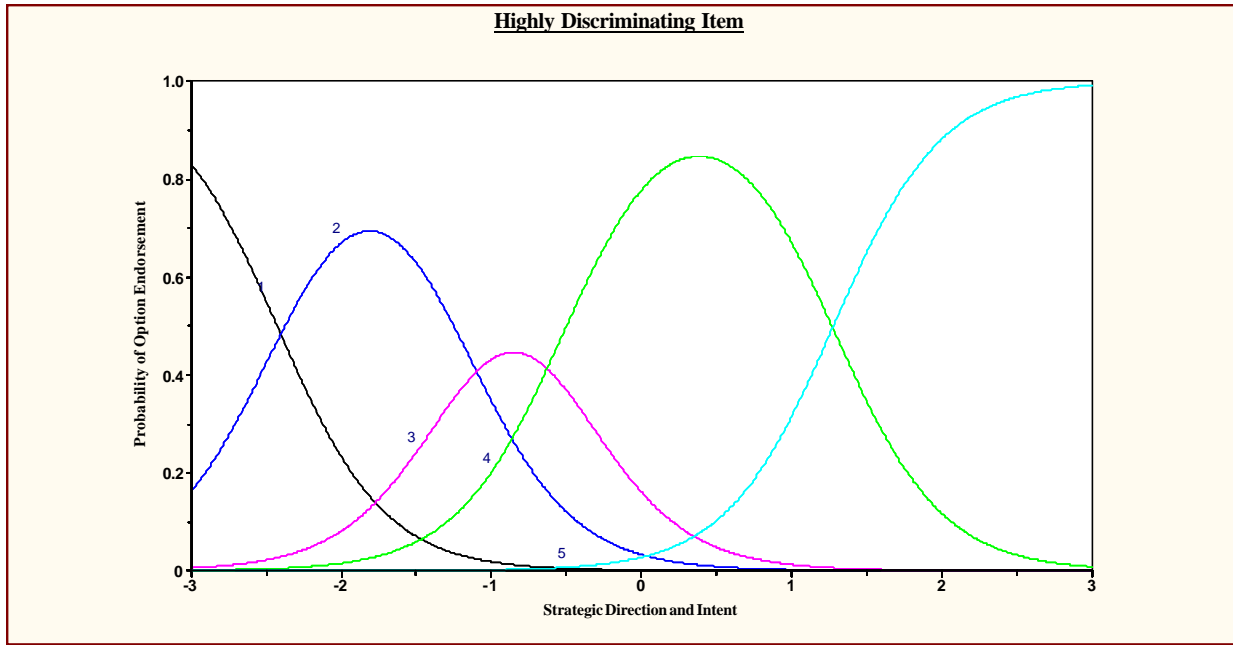


Figure 2 – Example of a Poorly Discriminating Item

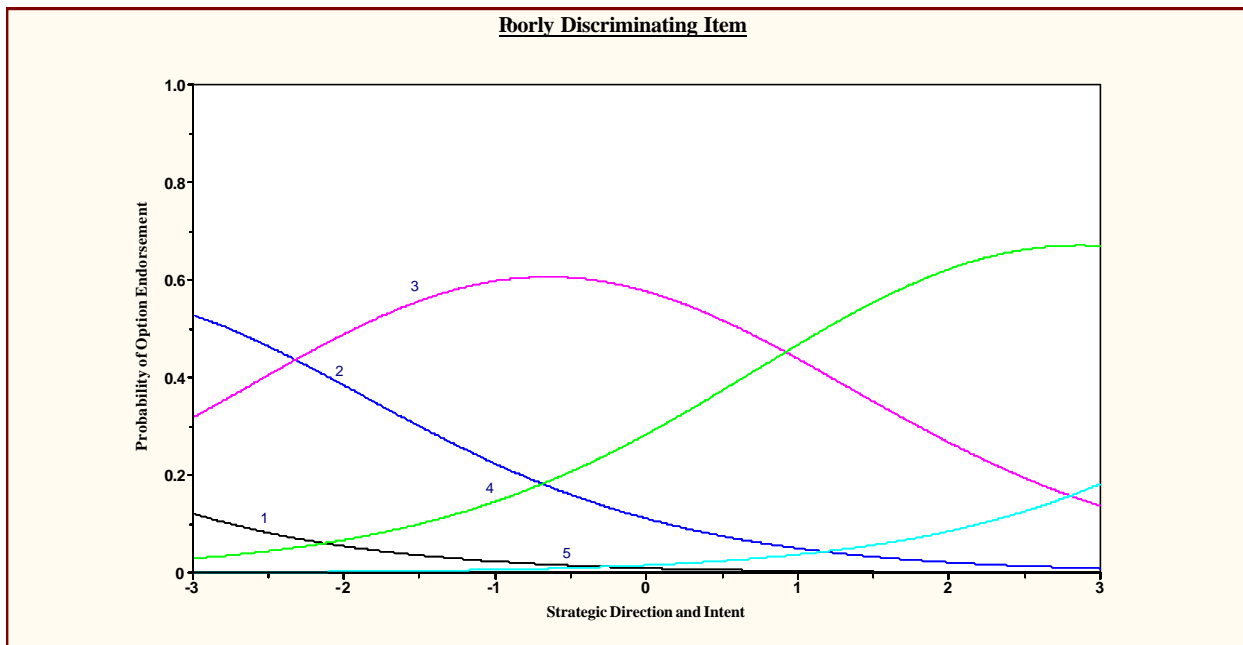


Table 1 – Results of Differential Item Functioning Analysis

	<i>Chinese Simplified</i>	<i>Chinese Traditional</i>	<i>Dutch</i>	<i>UK English</i>	<i>French</i>	<i>German</i>	<i>Italian</i>	<i>Japanese</i>	<i>Korean</i>	<i>Portuguese</i>	<i>Portuguese -Brazil</i>	<i>Spanish-Latin America</i>	<i>Spanish-Non-Latin America</i>	<i>Turkish</i>
Item 1					?		?	?				?		?
Item 2					?		?	?	?					
Item 3					?									
Item 4					?	?								
Item 5														?
Trait 1					?									
Item 6						?								
Item 7						?								
Item 8													?	?
Item 9														?
Item 10														
Trait 2														
Item 11									?					
Item 12														
Item 13														?
Item 14								?						
Item 15								?	?			?		?
Trait 3														
Item 16	?	?												
Item 17		?												
Item 18														
Item 19	?	?												
Item 20	?					?								
Trait 4														
Item 21														
Item 22								?						?
Item 23								?						
Item 24					?			?	?	?	?			
Item 25														
Trait 5														

	<i>Chinese Simplified</i>	<i>Chinese Traditional</i>	<i>Dutch</i>	<i>UK English</i>	<i>French</i>	<i>German</i>	<i>Italian</i>	<i>Japanese</i>	<i>Korean</i>	<i>Portuguese</i>	<i>Portuguese-Brazil</i>	<i>Spanish-Latin America</i>	<i>Spanish-Non-Latin America</i>	<i>Turkish</i>
Item 26					?	?		?						
Item 27						?								
Item 28														
Item 29			?			?								
Item 30								?					?	
Trait 6														
Item 31		?					?	?						
Item 32									?					
Item 33										?				
Item 34														?
Item 35														
Trait 7														
Item 36	?	?				?		?						?
Item 37	?	?				?				?				?
Item 38		?				?			?					?
Item 39					?	?								
Item 40	?					?								
Trait 8						?								
Item 41										?		?	?	
Item 42														
Item 43	?	?	?		?	?				?	?	?		
Item 44					?	?								
Item 45	?	?	?		?				?					
Trait 9														
Item 46														
Item 47									?	?	?			
Item 48														
Item 49														
Item 50									?				?	
Trait 10														

	<i>Chinese Simplified</i>	<i>Chinese Traditional</i>	<i>Dutch</i>	<i>UK English</i>	<i>French</i>	<i>German</i>	<i>Italian</i>	<i>Japanese</i>	<i>Korean</i>	<i>Portuguese</i>	<i>Portuguese -Brazil</i>	<i>Spanish-Latin America</i>	<i>Spanish-Non-Latin America</i>	<i>Turkish</i>
Item 51														
Item 52		?												
Item 53						?								
Item 54														
Item 55														
Trait 11														
Item 56							?				?			
Item 57						?				?				?
Item 58							?							
Item 59										?	?			
Item 60														
Trait 12														

Notes:

‡ Values indicate differential item functioning.

‡ Values indicate differential test functioning at the trait level.

Trait 1 = Empowerment; Trait 2 = Team Orientation; Trait 3 = Capability Development; Trait 4 = Core Values; Trait 5 = Agreement; Trait 6 = Coordination and Integration;

Trait 7 = Creating Change; Trait 8 = Customer Focus; Trait 9 = Organizational Learning; Trait 10 = Strategic Direction and Intent; Trait 11 = Goals and Objectives; Trait 12 = Vision

Table 2 – Results of Confirmatory Factor Analysis

<i>Language</i>	<i>Model</i>	<i>χ²</i>	<i>NNFI</i>	<i>CFI</i>
Chinese Simplified	Equivalent Forms Model (M0)			
	Multiple Groups Baseline Model (M1)			
	Metric Invariance Model (M2)	•		
	Structure Invariance Model (M3)	•		
	Equal Error Variances Model (M4)	•		
Chinese Traditional	M0			
	M1			
	M2	•		
	M3	•		
	M4	•		
Dutch	M0			
	M1			
	M2	•		
	M3	•		
	M4	•		
UK English	M0			
	M1			
	M2	•		
	M3	•		
	M4	•		
French	M0			
	M1			
	M2	•		
	M3	•		
	M4	•		
German	M0			
	M1			
	M2	•		
	M3	•		
	M4	•		

<i>Language</i>	<i>Model</i>	<i>???</i>	<i>NNFI</i>	<i>CFI</i>
Italian	M0			
	M1			
	M2	•		
	M3			
	M4	•		
Japanese	M0			
	M1			
	M2	•		
	M3	•		
	M4	•		
Korean	M0			
	M1			
	M2	•		
	M3	•		
	M4	•		
Portuguese	M0			
	M1			
	M2	•		
	M3	•		
	M4	•		
Portuguese-Brazil	M0			
	M1			
	M2	•		
	M3	•		
	M4	•		
Spanish-Latin America	M0			
	M1			
	M2	•		
	M3	•		
	M4	•		

<i>Language</i>	<i>Model</i>	χ^2	<i>NNFI</i>	<i>CFI</i>
Spanish-Non-Latin America	M0			
	M1			
	M2	•		
	M3	•		
	M4	•		
Turkish	M0			
	M1			
	M2	•		
	M3	•		
	M4	•	•	•

Note: • Values indicate non-equivalence.

Table 3 – Results of Classical Test Theory Analysis**Table 3a: Means and Standard Deviations of Items and Traits by Language (US, Croatian – Flemish)**

	US English (N=19696)		Croatian (N=57)		Czech (N=80)		Danish (N=90)		Finnish (N=130)		Flemish (N=90)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Item 1	4.15	0.73	4.25	0.61	4.22	0.69	4.55	0.52	4.10	0.68	3.95	0.76
Item 2	3.09	1.03	3.68	0.91	3.53	0.87	3.44	0.99	3.46	0.97	3.18	1.06
Item 3	3.09	1.05	2.95	0.99	3.49	0.92	3.39	0.95	3.31	0.90	3.33	1.03
Item 4	3.36	0.97	3.26	0.90	3.35	0.82	3.77	0.85	3.25	0.93	3.64	0.90
Item 5	3.25	1.04	3.77	0.89	3.39	0.97	3.13	1.01	3.50	0.96	3.33	1.02
Empowerment	3.39	0.97	3.58	0.86	3.60	0.85	3.65	0.87	3.53	0.89	3.49	0.96
Item 6	3.56	0.99	3.16	0.95	3.41	0.91	3.39	1.10	3.38	1.00	3.09	1.01
Item 7	3.53	0.96	3.73	0.82	4.01	0.69	3.98	0.69	3.79	0.82	3.69	0.88
Item 8	3.38	1.05	3.43	0.87	3.56	0.87	3.63	0.87	3.62	0.93	3.13	1.03
Item 9	3.62	0.97	3.68	0.92	3.90	0.71	3.76	0.83	4.10	0.75	3.97	0.95
Item 10	3.35	0.98	3.55	0.87	3.99	0.61	3.69	0.83	3.75	0.82	3.51	1.00
Team Orientation	3.49	0.99	3.51	0.89	3.77	0.76	3.69	0.86	3.73	0.86	3.48	0.98
Item 11	3.15	1.05	3.46	0.87	3.75	0.95	4.05	0.73	3.63	0.77	3.43	0.89
Item 12	3.39	0.91	3.37	1.08	3.98	0.75	3.67	0.89	2.68	0.95	3.40	0.91
Item 13	3.38	1.05	3.25	1.02	3.86	0.92	3.74	0.90	3.80	0.87	3.26	1.06
Item 14	3.59	0.99	3.68	1.02	3.84	0.89	4.02	0.80	3.84	0.92	3.69	0.97
Item 15	3.45	1.03	3.67	0.89	3.71	0.98	3.57	0.80	3.76	0.91	3.49	0.88
Capability Development	3.39	1.01	3.48	0.98	3.83	0.90	3.81	0.82	3.54	0.88	3.45	0.94
Item 16	3.14	1.06	3.44	0.85	3.39	0.93	3.53	0.84	3.50	0.87	3.19	0.77
Item 17	3.41	0.96	3.82	0.73	3.58	0.79	3.49	0.93	3.76	0.81	3.51	0.94
Item 18	3.87	0.93	3.91	0.66	4.22	0.67	3.86	0.86	3.87	0.78	3.85	0.95
Item 19	4.19	0.85	4.23	0.76	4.16	0.71	3.65	0.87	3.89	0.86	4.15	0.71
Item 20	4.35	0.72	4.25	0.61	4.43	0.59	4.35	0.59	4.43	0.65	4.23	0.68
Core Values	3.79	0.90	3.93	0.72	3.96	0.74	3.78	0.82	3.89	0.79	3.79	0.81
Item 21	3.46	0.89	3.60	0.93	3.70	0.78	3.35	0.76	3.82	0.70	3.73	0.75
Item 22	3.76	0.92	3.60	0.74	3.97	0.83	3.86	0.80	3.88	0.92	3.70	0.92
Item 23	2.91	0.93	3.04	0.74	3.04	0.85	3.22	0.73	3.09	0.82	3.09	0.83
Item 24	3.08	0.96	3.35	0.91	3.45	0.91	3.42	0.77	3.70	0.77	3.11	0.86
Item 25	3.39	0.99	3.75	0.82	3.14	0.88	3.57	0.90	3.67	0.93	3.72	0.92
Agreement	3.32	0.94	3.47	0.83	3.46	0.85	3.48	0.79	3.63	0.83	3.47	0.86
Item 26	3.49	1.00	3.92	0.67	3.87	0.70	3.79	0.83	3.74	0.75	3.40	0.98
Item 27	2.95	1.05	3.10	0.93	3.69	0.84	3.29	0.93	3.53	0.86	3.15	0.94
Item 28	2.66	0.97	2.76	0.92	2.89	0.86	2.96	0.84	2.70	0.93	2.69	0.92
Item 29	2.82	1.04	3.54	0.99	3.29	0.97	3.46	0.94	3.45	1.03	2.86	0.95
Item 30	3.19	0.95	3.64	0.75	3.48	0.83	3.37	0.84	3.73	0.79	3.21	0.91
Coordination and Integration	3.02	1.00	3.39	0.85	3.45	0.84	3.37	0.88	3.43	0.87	3.06	0.94
Item 31	2.25	0.92	2.30	0.75	2.84	1.04	2.29	0.97	2.43	0.84	2.42	0.91
Item 32	2.81	1.07	3.40	0.97	3.30	1.05	3.53	0.86	3.23	1.05	2.90	1.06
Item 33	3.18	1.01	3.25	0.87	3.65	0.81	3.53	0.84	3.40	0.82	3.17	0.95

Item 34	2.56	0.97	2.92	0.81	3.12	0.83	3.40	0.95	2.84	0.95	2.98	0.85
Item 35	3.17	0.87	3.32	0.83	3.45	0.84	3.36	0.80	2.99	0.89	3.19	0.83
Creating Change	2.79	0.97	3.04	0.84	3.27	0.91	3.22	0.88	2.98	0.91	2.93	0.92
Item 36	3.06	0.97	3.04	0.87	3.00	0.89	2.66	0.94	3.10	0.92	2.76	0.87
Item 37	3.13	0.97	3.06	0.94	3.28	0.95	2.80	0.98	3.22	0.91	2.97	0.79
Item 38	2.88	1.01	3.26	0.89	3.39	0.91	3.54	0.98	3.12	0.98	3.24	0.95
Item 39	3.13	1.05	3.26	0.89	3.38	0.95	3.18	1.01	3.60	0.93	3.16	0.86
Item 40	3.56	1.00	3.91	0.62	3.89	0.78	4.01	0.93	3.65	1.00	3.82	0.82
Customer Focus	3.15	1.00	3.30	0.84	3.39	0.90	3.24	0.97	3.34	0.95	3.19	0.86
Item 41	3.18	1.05	3.75	0.83	3.84	0.71	3.37	0.87	3.55	0.85	3.67	0.74
Item 42	2.81	1.05	2.88	0.96	3.19	0.89	3.49	0.85	3.65	1.06	3.26	0.91
Item 43	2.91	1.02	3.26	0.86	2.91	0.92	3.52	0.93	3.18	0.91	2.59	0.75
Item 44	3.75	0.96	4.05	0.85	4.15	0.56	3.96	0.85	3.99	0.74	3.91	0.94
Item 45	2.83	1.02	3.65	0.88	3.32	1.01	3.41	0.88	3.14	0.95	3.40	0.82
Organizational Learning	3.10	1.02	3.52	0.88	3.48	0.82	3.55	0.87	3.50	0.90	3.37	0.83
Item 46	3.63	0.96	3.91	0.71	4.12	0.61	3.89	0.83	3.94	0.81	3.72	0.90
Item 47	3.05	1.01	3.91	0.71	3.33	0.91	3.17	0.78	3.13	0.89	3.09	0.74
Item 48	3.65	0.93	3.75	0.62	4.05	0.71	3.86	0.70	4.13	0.70	3.67	0.83
Item 49	3.46	1.01	3.66	0.90	3.92	0.78	3.63	0.86	3.83	0.97	3.46	0.97
Item 50	3.28	1.11	3.30	1.15	3.73	1.00	3.68	0.87	3.62	1.14	3.41	1.04
Strategic Direction and Intent	3.42	1.00	3.71	0.82	3.83	0.80	3.65	0.81	3.73	0.90	3.47	0.90
Item 51	3.32	0.89	3.31	0.81	3.72	0.70	3.75	0.82	3.80	0.74	3.21	0.76
Item 52	3.38	0.92	3.29	0.80	3.68	0.82	3.70	0.71	3.54	0.93	3.28	0.89
Item 53	3.79	0.77	3.48	0.80	3.97	0.64	3.97	0.82	4.00	0.64	3.64	0.83
Item 54	3.69	0.83	3.92	0.65	4.13	0.52	3.91	0.72	4.18	0.58	3.66	0.68
Item 55	3.40	0.95	3.52	0.83	3.96	0.59	3.85	0.69	3.85	0.76	3.61	0.80
Goals and Objectives	3.52	0.88	3.50	0.78	3.89	0.65	3.84	0.75	3.87	0.73	3.48	0.79
Item 56	3.16	1.01	3.19	0.82	3.53	0.87	3.30	0.82	3.30	0.94	3.17	0.91
Item 57	3.55	0.93	3.58	0.98	3.71	0.76	3.63	0.85	3.72	0.79	3.62	0.91
Item 58	2.61	0.98	3.25	0.88	3.14	0.89	2.80	0.88	3.02	1.00	2.85	0.92
Item 59	3.01	1.01	3.19	0.84	3.29	0.91	3.42	0.69	3.51	0.77	3.10	0.91
Item 60	3.23	0.87	3.50	0.67	3.68	0.62	3.39	0.79	3.55	0.75	3.22	0.81
Vision	3.11	0.96	3.34	0.84	3.47	0.81	3.31	0.81	3.42	0.85	3.19	0.89

Note: Maximum sample sizes are shown within parentheses.

Table 3b: Means and Standard Deviations of Items and Traits by Language (US, French Canadian – Polish)

	US English (N=19696)		French Canadian (N=202)		Hungarian (N=89)		Indonesian (N=33)		Norwegian (N=110)		Polish (N=204)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Item 1	4.15	0.73	3.87	0.92	3.86	0.84	3.93	0.78	3.98	0.77	4.25	0.68
Item 2	3.09	1.03	3.19	1.07	3.64	0.93	3.90	0.76	3.07	0.87	3.33	1.00
Item 3	3.09	1.05	2.90	1.02	3.34	1.16	4.10	0.88	3.01	0.90	3.13	1.13
Item 4	3.36	0.97	3.62	0.91	3.52	1.09	4.03	0.85	3.53	0.79	3.10	1.03
Item 5	3.25	1.04	3.10	1.00	3.69	1.04	3.77	0.57	3.15	0.93	3.39	1.00
Empowerment	3.39	0.97	3.33	0.98	3.61	1.01	3.95	0.77	3.35	0.85	3.44	0.97
Item 6	3.56	0.99	3.38	1.05	3.27	1.03	3.75	0.67	3.25	0.97	3.23	0.96
Item 7	3.53	0.96	3.27	1.13	4.03	0.80	3.75	0.92	3.58	0.89	3.75	0.89
Item 8	3.38	1.05	3.00	1.11	3.76	1.00	3.94	0.84	3.21	0.99	3.05	1.15
Item 9	3.62	0.97	3.51	1.06	3.85	0.87	3.97	0.90	3.62	0.98	3.75	0.88
Item 10	3.35	0.98	3.27	0.96	3.82	0.94	3.88	0.91	3.40	0.87	3.66	0.85
Team Orientation	3.49	0.99	3.28	1.06	3.75	0.93	3.86	0.85	3.41	0.94	3.49	0.95
Item 11	3.15	1.05	3.09	1.06	3.72	0.92	3.67	1.02	3.07	1.02	3.80	0.83
Item 12	3.39	0.91	2.91	1.02	3.67	0.91	3.88	0.82	3.56	0.80	3.51	1.00
Item 13	3.38	1.05	3.34	1.03	4.13	0.86	3.94	0.75	3.54	0.92	2.75	1.09
Item 14	3.59	0.99	3.54	1.05	3.76	0.99	4.42	0.61	3.77	0.81	3.54	1.08
Item 15	3.45	1.03	3.44	0.96	3.84	1.02	2.33	0.99	3.62	0.85	3.41	1.02
Capability Development	3.39	1.01	3.26	1.02	3.82	0.94	3.65	0.84	3.51	0.88	3.40	1.01
Item 16	3.14	1.06	2.96	1.03	3.91	0.92	3.45	0.87	3.27	0.87	3.39	0.99
Item 17	3.41	0.96	3.66	0.78	4.12	0.85	3.42	0.90	3.11	0.92	4.00	0.68
Item 18	3.87	0.93	3.49	0.98	4.16	0.85	3.91	0.95	3.70	0.81	4.19	0.70
Item 19	4.19	0.85	3.55	1.08	4.49	0.61	4.24	0.83	3.54	0.75	4.42	0.69
Item 20	4.35	0.72	4.06	0.81	4.40	0.62	4.58	0.50	4.17	0.63	4.39	0.65
Core Values	3.79	0.90	3.54	0.94	4.21	0.77	3.92	0.81	3.56	0.79	4.08	0.74
Item 21	3.46	0.89	3.46	0.91	3.73	0.86	3.97	0.73	3.61	0.77	3.60	0.80
Item 22	3.76	0.92	3.72	0.99	4.16	0.67	3.91	1.04	3.92	0.69	3.78	0.75
Item 23	2.91	0.93	3.09	0.93	3.25	0.92	3.48	0.87	3.01	0.83	3.21	0.93
Item 24	3.08	0.96	3.01	0.99	3.22	1.08	2.67	1.05	3.45	0.70	3.43	0.89
Item 25	3.39	0.99	3.25	1.04	3.96	0.82	3.94	0.79	3.24	0.81	3.82	0.78
Agreement	3.32	0.94	3.31	0.97	3.66	0.87	3.59	0.90	3.45	0.76	3.57	0.83
Item 26	3.49	1.00	3.10	0.98	4.01	0.78	3.97	0.81	3.39	0.95	3.89	0.74
Item 27	2.95	1.05	2.66	0.88	3.25	0.98	2.94	1.00	3.00	0.90	3.09	0.96
Item 28	2.66	0.97	2.49	0.95	3.10	0.93	3.18	0.77	2.78	0.89	2.78	0.81
Item 29	2.82	1.04	2.64	1.02	3.78	0.84	3.27	1.10	3.11	0.96	3.19	1.03
Item 30	3.19	0.95	2.93	0.93	3.49	0.87	3.67	0.82	3.31	0.77	3.66	0.69
Coordination and Integration	3.02	1.00	2.76	0.95	3.53	0.88	3.41	0.90	3.12	0.89	3.32	0.85
Item 31	2.25	0.92	2.29	0.98	2.56	1.05	2.85	1.06	2.11	0.81	2.23	0.89
Item 32	2.81	1.07	2.76	1.13	3.43	1.03	3.85	0.76	2.77	0.91	3.20	1.03
Item 33	3.18	1.01	3.13	1.05	3.66	0.89	4.06	0.70	2.97	0.95	3.37	0.95
Item 34	2.56	0.97	2.99	0.96	3.08	0.97	3.06	1.14	3.02	0.93	3.06	0.96
Item 35	3.17	0.87	3.31	0.87	3.18	0.90	3.58	0.61	3.06	0.78	3.27	0.87

Creating Change	2.79	0.97	2.90	1.00	3.18	0.97	3.48	0.86	2.79	0.88	3.03	0.94
Item 36	3.06	0.97	3.01	1.03	3.11	0.87	3.70	0.81	2.79	0.80	2.64	0.90
Item 37	3.13	0.97	3.14	1.04	3.05	0.86	3.24	0.97	3.01	0.86	2.54	0.93
Item 38	2.88	1.01	2.74	1.08	3.94	0.74	3.45	1.00	2.97	0.92	3.11	0.96
Item 39	3.13	1.05	3.07	1.08	3.77	0.86	3.18	1.04	3.28	0.95	3.16	0.96
Item 40	3.56	1.00	3.35	0.98	3.99	0.81	4.03	0.81	3.95	0.86	3.96	0.77
Customer Focus	3.15	1.00	3.06	1.04	3.57	0.83	3.52	0.93	3.20	0.88	3.08	0.90
Item 41	3.18	1.05	3.45	0.98	3.68	0.77	4.33	0.74	3.51	0.86	3.78	0.74
Item 42	2.81	1.05	2.94	1.02	3.16	0.85	3.76	0.79	3.29	1.02	2.77	1.00
Item 43	2.91	1.02	2.74	1.03	3.43	1.01	2.94	1.09	2.69	0.73	2.97	0.88
Item 44	3.75	0.96	4.02	0.84	4.21	0.78	4.48	0.62	3.83	0.81	3.93	0.95
Item 45	2.83	1.02	2.89	1.10	3.78	0.93	3.91	0.72	3.02	0.83	3.06	1.03
Organizational Learning	3.10	1.02	3.21	0.99	3.65	0.87	3.88	0.79	3.27	0.85	3.30	0.92
Item 46	3.63	0.96	3.49	1.10	4.15	0.70	4.30	0.73	3.83	0.82	4.01	0.65
Item 47	3.05	1.01	3.06	0.94	3.16	0.90	3.27	0.88	2.93	0.81	3.27	0.88
Item 48	3.65	0.93	3.37	0.94	4.06	0.69	3.91	0.88	3.95	0.65	4.10	0.66
Item 49	3.46	1.01	3.23	0.99	3.91	0.80	4.18	0.81	3.44	0.91	3.78	0.84
Item 50	3.28	1.11	3.03	1.09	4.04	0.85	3.64	0.86	3.06	1.08	3.64	0.97
Strategic Direction and Intent	3.42	1.00	3.24	1.01	3.86	0.79	3.86	0.83	3.44	0.85	3.76	0.80
Item 51	3.32	0.89	3.16	0.80	3.62	0.86	3.74	0.73	3.34	0.75	3.65	0.74
Item 52	3.38	0.92	3.15	1.01	3.84	0.83	3.42	0.81	3.27	0.92	3.62	0.87
Item 53	3.79	0.77	3.54	0.96	4.14	0.66	3.97	0.55	3.60	0.85	3.94	0.63
Item 54	3.69	0.83	3.54	0.85	4.18	0.74	4.03	0.48	3.87	0.73	3.92	0.71
Item 55	3.40	0.95	3.36	0.91	3.86	0.85	3.77	0.96	3.39	0.86	3.69	0.84
Goals and Objectives	3.52	0.88	3.35	0.91	3.93	0.79	3.79	0.70	3.49	0.82	3.76	0.76
Item 56	3.16	1.01	3.22	0.99	2.97	1.01	3.85	0.83	3.02	0.93	3.55	0.86
Item 57	3.55	0.93	3.25	1.02	3.85	0.86	4.03	0.77	3.21	0.92	3.57	0.74
Item 58	2.61	0.98	2.34	0.99	2.51	1.15	3.03	0.95	2.85	0.79	2.81	0.96
Item 59	3.01	1.01	2.73	1.03	3.33	0.98	3.88	0.70	2.96	0.86	3.30	0.83
Item 60	3.23	0.87	3.08	0.89	3.46	0.80	3.61	0.66	3.14	0.72	3.49	0.79
Vision	3.11	0.96	2.92	0.98	3.22	0.96	3.68	0.78	3.04	0.84	3.34	0.84

Note: Maximum sample sizes are shown within parentheses.

Table 3c: Means and Standard Deviations of Items and Traits by Language (US, Russian – Vietnamese)

	US English (N=19696)		Russian (N=127)		Swedish (N=186)		Thai (N=195)		Vietnamese (N=43)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Item 1	4.15	0.73	4.24	0.74	4.35	0.58	4.18	0.69	4.14	0.77
Item 2	3.09	1.03	3.81	0.80	3.28	0.88	3.90	0.84	3.79	0.67
Item 3	3.09	1.05	3.71	0.96	3.35	0.88	3.58	0.98	3.67	0.75
Item 4	3.36	0.97	4.06	0.85	3.61	0.82	3.75	0.74	3.72	0.73
Item 5	3.25	1.04	4.06	0.96	3.49	0.87	3.82	0.82	3.72	0.59
Empowerment	3.39	0.97	3.98	0.86	3.62	0.81	3.85	0.81	3.81	0.70
Item 6	3.56	0.99	3.48	0.95	3.29	0.93	3.52	0.86	3.83	0.76
Item 7	3.53	0.96	3.96	0.88	3.52	0.91	3.72	0.83	3.81	0.80
Item 8	3.38	1.05	3.66	0.95	3.25	1.01	3.66	0.90	3.95	0.70
Item 9	3.62	0.97	3.86	0.87	3.21	0.94	4.04	0.80	3.98	0.60
Item 10	3.35	0.98	3.94	0.81	3.44	0.80	3.65	0.81	3.81	0.63
Team Orientation	3.49	0.99	3.78	0.89	3.34	0.92	3.72	0.84	3.88	0.70
Item 11	3.15	1.05	3.92	0.81	3.24	0.96	3.40	1.00	3.40	0.94
Item 12	3.39	0.91	4.18	0.88	3.68	0.79	3.87	0.72	3.55	0.80
Item 13	3.38	1.05	4.03	0.94	3.70	0.91	3.86	0.85	3.52	0.80
Item 14	3.59	0.99	4.28	0.83	4.05	0.86	4.38	0.72	4.00	0.91
Item 15	3.45	1.03	3.12	1.14	3.67	0.90	2.44	1.07	2.83	0.99
Capability Development	3.39	1.01	3.91	0.92	3.67	0.88	3.59	0.87	3.46	0.89
Item 16	3.14	1.06	3.77	0.83	3.19	0.88	3.38	0.89	3.05	0.90
Item 17	3.41	0.96	4.14	0.72	3.05	0.85	3.86	0.91	3.70	0.61
Item 18	3.87	0.93	4.44	0.61	3.77	0.83	4.13	0.82	3.83	0.59
Item 19	4.19	0.85	4.43	0.68	3.14	0.84	3.73	0.83	3.78	0.80
Item 20	4.35	0.72	4.41	0.65	4.05	0.74	4.12	0.72	4.13	0.79
Core Values	3.79	0.90	4.24	0.70	3.44	0.83	3.85	0.83	3.70	0.74
Item 21	3.46	0.89	3.65	0.84	3.31	0.92	3.55	0.98	3.75	0.78
Item 22	3.76	0.92	4.18	0.72	4.20	0.61	3.63	0.85	3.73	0.75
Item 23	2.91	0.93	3.32	0.78	3.19	0.78	3.06	0.90	3.35	0.74
Item 24	3.08	0.96	3.44	0.89	3.54	0.84	2.53	0.83	3.03	0.73
Item 25	3.39	0.99	4.24	0.68	3.50	0.82	3.40	0.86	3.90	0.67
Agreement	3.32	0.94	3.76	0.78	3.55	0.79	3.23	0.88	3.55	0.73
Item 26	3.49	1.00	4.10	0.76	3.49	0.88	3.69	0.83	3.71	0.93
Item 27	2.95	1.05	3.53	0.83	3.25	0.98	3.13	0.94	3.41	0.77
Item 28	2.66	0.97	3.25	0.99	2.69	0.88	2.94	0.90	3.34	0.66
Item 29	2.82	1.04	3.14	0.94	3.38	0.95	3.05	0.89	3.15	0.82
Item 30	3.19	0.95	3.80	0.85	3.12	0.74	3.53	0.76	3.51	0.68
Coordination and Integration	3.02	1.00	3.56	0.87	3.19	0.89	3.27	0.86	3.42	0.77
Item 31	2.25	0.92	3.23	1.04	2.26	0.82	2.55	1.14	2.63	1.00
Item 32	2.81	1.07	3.50	1.12	3.34	0.94	2.56	1.06	2.50	0.99
Item 33	3.18	1.01	3.87	0.81	3.31	0.86	3.46	0.91	3.10	0.87
Item 34	2.56	0.97	3.09	0.92	3.20	0.93	2.49	0.89	3.13	0.61
Item 35	3.17	0.87	3.53	0.77	2.88	0.79	3.41	0.77	3.30	0.61
Creating Change	2.79	0.97	3.44	0.93	3.00	0.87	2.90	0.95	2.93	0.82

Item 36	3.06	0.97	3.04	0.84	3.01	0.81	2.99	1.11	2.66	0.73
Item 37	3.13	0.97	3.45	0.87	3.13	0.82	3.16	1.01	2.73	0.81
Item 38	2.88	1.01	3.39	0.94	3.22	1.00	3.15	1.01	3.80	0.81
Item 39	3.13	1.05	3.59	0.87	3.75	0.91	2.89	1.05	2.80	0.98
Item 40	3.56	1.00	3.69	0.85	3.96	0.85	3.83	0.74	3.80	0.81
Customer Focus	3.15	1.00	3.43	0.87	3.41	0.88	3.20	0.99	3.16	0.83
Item 41	3.18	1.05	3.80	0.79	3.40	0.90	4.15	0.71	4.05	0.60
Item 42	2.81	1.05	3.18	0.88	3.14	0.88	3.34	0.87	3.23	0.87
Item 43	2.91	1.02	3.31	0.97	3.19	0.88	2.51	0.95	2.59	0.85
Item 44	3.75	0.96	4.27	0.75	3.90	0.72	4.15	0.67	4.05	0.69
Item 45	2.83	1.02	3.68	0.96	3.15	0.88	3.46	0.82	3.56	0.82
Organizational Learning	3.10	1.02	3.65	0.87	3.36	0.85	3.52	0.81	3.50	0.77
Item 46	3.63	0.96	4.27	0.73	3.94	0.83	4.13	0.67	3.92	0.82
Item 47	3.05	1.01	3.47	1.06	3.17	0.65	3.44	0.99	3.13	0.66
Item 48	3.65	0.93	4.12	0.72	4.04	0.59	4.04	0.67	3.76	0.71
Item 49	3.46	1.01	4.09	0.78	3.84	0.84	3.91	0.81	3.79	0.87
Item 50	3.28	1.11	4.08	0.83	3.54	1.03	3.27	1.09	3.18	0.80
Strategic Direction and Intent	3.42	1.00	4.01	0.82	3.71	0.79	3.76	0.85	3.56	0.77
Item 51	3.32	0.89	3.92	0.59	3.20	0.91	3.55	0.75	3.60	0.58
Item 52	3.38	0.92	3.85	0.84	3.52	0.89	3.20	0.99	3.12	0.88
Item 53	3.79	0.77	4.05	0.77	3.63	0.82	3.79	0.75	3.81	0.55
Item 54	3.69	0.83	4.17	0.67	3.70	0.75	3.92	0.63	3.77	0.68
Item 55	3.40	0.95	3.90	0.81	3.64	0.79	3.79	0.74	3.70	0.64
Goals and Objectives	3.52	0.88	3.98	0.74	3.54	0.83	3.65	0.77	3.60	0.67
Item 56	3.16	1.01	4.04	0.70	3.31	0.92	3.56	0.85	3.44	0.73
Item 57	3.55	0.93	4.16	0.70	3.39	0.99	3.65	0.84	3.74	0.66
Item 58	2.61	0.98	3.60	0.88	2.72	0.92	2.50	0.91	2.21	0.74
Item 59	3.01	1.01	3.61	0.88	3.44	0.79	3.10	0.93	3.53	0.74
Item 60	3.23	0.87	3.61	0.81	3.55	0.74	3.38	0.84	3.21	0.99
Vision	3.11	0.96	3.80	0.79	3.28	0.87	3.24	0.88	3.23	0.77

Note: Maximum sample sizes are shown within parentheses.

Table 3d: Corrected Item-Total Correlations for Each Trait by Language (US, Croatian – Hungarian)

	US English (N = 19696)	Croatian (N=57)	Czech (N=80)	Danish (N=90)	Finnish (N=130)	Flemish (N=90)	French Canadian (N=202)	Hungarian (N=89)
Item 1	0.36	0.40	0.36	0.31	0.28	0.53	0.47	0.48
Item 2	0.56	0.49	0.50	0.50	0.59	0.71	0.50	0.57
Item 3	0.56	0.58	0.50	0.61	0.49	0.58	0.49	0.56
Item 4	0.53	0.41	0.48	0.32	0.58	0.52	0.56	0.57
Item 5	0.51	0.63	0.40	0.52	0.55	0.57	0.57	0.62
Empowerment	0.51	0.50	0.45	0.45	0.50	0.58	0.52	0.56
Item 6	0.54	0.55	0.53	0.47	0.36	0.43	0.62	0.42
Item 7	0.68	0.66	0.59	0.63	0.55	0.70	0.71	0.62
Item 8	0.70	0.69	0.62	0.49	0.50	0.56	0.72	0.71
Item 9	0.62	0.58	0.66	0.56	0.45	0.51	0.51	0.60
Item 10	0.54	0.60	0.49	0.41	0.42	0.43	0.54	0.63
Team Orientation	0.62	0.62	0.58	0.51	0.46	0.53	0.62	0.60
Item 11	0.39	0.46	0.23	0.28	0.39	0.48	0.47	0.49
Item 12	0.57	0.60	0.59	0.73	0.34	0.65	0.63	0.47
Item 13	0.57	0.62	0.69	0.72	0.61	0.65	0.65	0.51
Item 14	0.55	0.59	0.49	0.59	0.48	0.72	0.64	0.51
Item 15	0.19	0.07	0.18	0.37	0.12	0.28	0.32	-0.04
Capability Development	0.46	0.47	0.44	0.54	0.39	0.56	0.54	0.39
Item 16	0.44	0.52	0.34	0.37	0.45	0.49	0.35	0.63
Item 17	0.33	0.46	0.25	0.45	0.38	0.52	0.31	0.67
Item 18	0.57	0.54	0.46	0.56	0.36	0.70	0.51	0.64
Item 19	0.37	0.52	0.39	0.24	0.18	0.29	0.42	0.41
Item 20	0.52	0.49	0.44	0.44	0.36	0.55	0.35	0.68
Core Values	0.45	0.51	0.38	0.41	0.34	0.51	0.39	0.61
Item 21	0.52	0.48	0.54	0.46	0.33	0.33	0.57	0.56
Item 22	0.29	0.38	0.40	0.45	0.55	0.26	0.38	0.52
Item 23	0.59	0.63	0.47	0.71	0.45	0.36	0.49	0.64
Item 24	0.45	0.48	0.44	0.59	0.50	0.21	0.43	0.53
Item 25	0.48	0.25	0.36	0.48	0.50	0.30	0.57	0.59
Agreement	0.47	0.44	0.44	0.54	0.47	0.29	0.49	0.57
Item 26	0.39	0.35	0.56	0.29	0.46	0.32	0.44	0.50
Item 27	0.64	0.61	0.56	0.67	0.50	0.61	0.53	0.61
Item 28	0.63	0.67	0.50	0.64	0.50	0.62	0.64	0.60
Item 29	0.50	0.54	0.01	0.62	0.63	0.04	0.45	0.37
Item 30	0.61	0.54	0.57	0.56	0.46	0.61	0.62	0.64
Coordination and Integration	0.56	0.54	0.44	0.56	0.51	0.44	0.54	0.54
Item 31	0.55	0.36	0.62	0.49	0.37	0.52	0.56	0.58
Item 32	0.59	0.44	0.72	0.51	0.52	0.55	0.63	0.55
Item 33	0.63	0.59	0.63	0.43	0.56	0.63	0.58	0.63
Item 34	0.43	0.48	0.46	0.50	0.43	0.42	0.29	0.43
Item 35	0.50	0.45	0.54	0.44	0.46	0.43	0.50	0.59

Creating Change	0.54	0.46	0.59	0.47	0.47	0.51	0.51	0.56
Item 36	0.62	0.69	0.65	0.64	0.60	0.54	0.73	0.61
Item 37	0.66	0.64	0.78	0.63	0.66	0.63	0.78	0.36
Item 38	0.54	0.30	0.76	0.41	0.47	0.44	0.59	0.36
Item 39	0.57	0.58	0.59	0.79	0.63	0.55	0.54	0.28
Item 40	0.26	-0.03	0.49	0.10	0.34	0.45	0.28	0.26
Customer Focus	0.53	0.44	0.65	0.51	0.54	0.52	0.58	0.37
Item 41	0.52	0.46	0.49	0.27	0.47	0.42	0.55	0.35
Item 42	0.51	0.47	0.55	0.40	0.59	0.18	0.46	0.46
Item 43	0.37	0.28	0.57	0.46	0.59	0.24	0.44	0.41
Item 44	0.49	0.55	0.30	0.37	0.39	0.38	0.34	0.36
Item 45	0.55	0.54	0.52	0.36	0.57	0.39	0.59	0.53
Organizational Learning	0.49	0.46	0.49	0.37	0.52	0.32	0.48	0.42
Item 46	0.69	0.66	0.57	0.57	0.55	0.70	0.72	0.67
Item 47	0.48	0.08	0.40	0.19	0.38	0.52	0.53	0.34
Item 48	0.71	0.73	0.71	0.59	0.59	0.70	0.73	0.73
Item 49	0.79	0.84	0.75	0.67	0.73	0.84	0.80	0.78
Item 50	0.67	0.61	0.68	0.40	0.58	0.70	0.68	0.61
Strategic Direction and Intent	0.67	0.58	0.62	0.49	0.56	0.69	0.69	0.62
Item 51	0.60	0.69	0.42	0.61	0.51	0.67	0.50	0.49
Item 52	0.53	0.49	0.64	0.46	0.47	0.41	0.47	0.45
Item 53	0.51	0.71	0.65	0.64	0.61	0.64	0.59	0.47
Item 54	0.55	0.34	0.49	0.47	0.34	0.54	0.56	0.58
Item 55	0.59	0.70	0.46	0.51	0.46	0.57	0.57	0.62
Goals and Objectives	0.56	0.58	0.53	0.54	0.48	0.57	0.54	0.52
Item 56	0.62	0.61	0.75	0.53	0.49	0.73	0.67	0.62
Item 57	0.63	0.80	0.66	0.62	0.60	0.57	0.71	0.67
Item 58	0.32	0.51	0.26	0.35	0.46	0.55	0.18	0.21
Item 59	0.58	0.49	0.70	0.35	0.49	0.71	0.63	0.60
Item 60	0.57	0.67	0.33	0.46	0.53	0.72	0.66	0.42
Vision	0.54	0.62	0.54	0.46	0.51	0.66	0.57	0.50

Notes:

Maximum sample sizes are shown within parentheses.

Means of corrected item-total correlations are shown in bold in the rows corresponding to traits.

Table 3e: Corrected Item-Total Correlations for Each Trait by Language (US, Indonesian – Vietnamese)

	US English (N = 19696)	Indonesian (N=33)	Norwegian (N=110)	Polish (N=204)	Russian (N=127)	Swedish (N=186)	Thai (N=195)	Vietnamese (N=43)
Item 1	0.36	0.63	0.32	0.42	0.36	0.31	0.44	0.49
Item 2	0.56	0.53	0.35	0.51	0.55	0.49	0.14	0.62
Item 3	0.56	0.48	0.52	0.44	0.56	0.43	0.44	0.45
Item 4	0.53	0.50	0.33	0.48	0.50	0.42	0.41	0.50
Item 5	0.51	0.70	0.43	0.37	0.51	0.46	0.46	0.37
Empowerment	0.51	0.57	0.39	0.44	0.50	0.42	0.38	0.49
Item 6	0.54	0.51	0.33	0.53	0.50	0.49	0.55	0.34
Item 7	0.68	0.62	0.62	0.54	0.66	0.69	0.53	0.65
Item 8	0.70	0.64	0.68	0.58	0.60	0.70	0.59	0.63
Item 9	0.62	0.66	0.68	0.58	0.64	0.59	0.49	0.37
Item 10	0.54	0.63	0.49	0.44	0.47	0.51	0.52	0.65
Team Orientation	0.62	0.61	0.56	0.53	0.57	0.60	0.54	0.53
Item 11	0.39	0.37	0.07	0.30	0.40	0.27	0.27	0.44
Item 12	0.57	0.59	0.57	0.36	0.70	0.64	0.52	0.47
Item 13	0.57	0.59	0.52	0.47	0.62	0.64	0.48	0.48
Item 14	0.55	0.50	0.41	0.36	0.61	0.50	0.11	0.36
Item 15	0.19	0.01	0.17	0.05	0.22	0.26	-0.24	-0.02
Capability Development	0.46	0.41	0.35	0.31	0.51	0.46	0.23	0.35
Item 16	0.44	0.42	0.49	0.34	0.51	0.50	0.20	0.28
Item 17	0.33	0.18	0.52	0.34	0.59	0.25	0.38	0.37
Item 18	0.57	0.57	0.48	0.56	0.62	0.52	0.54	0.38
Item 19	0.37	0.00	0.26	0.42	0.28	0.16	0.27	0.18
Item 20	0.52	0.21	0.37	0.61	0.41	0.42	0.45	0.48
Core Values	0.45	0.28	0.42	0.45	0.48	0.37	0.37	0.34
Item 21	0.52	0.35	0.47	0.57	0.50	0.38	0.46	0.22
Item 22	0.29	0.60	0.27	0.29	0.41	0.21	0.39	0.25
Item 23	0.59	0.49	0.44	0.59	0.56	0.59	0.46	0.23
Item 24	0.45	0.34	0.37	0.55	0.35	0.46	0.25	-0.13
Item 25	0.48	0.31	0.37	0.33	0.41	0.24	0.25	0.12
Agreement	0.47	0.42	0.38	0.47	0.45	0.38	0.36	0.14
Item 26	0.39	0.28	0.38	0.30	0.49	0.24	0.36	0.40
Item 27	0.64	0.33	0.53	0.38	0.42	0.52	0.52	0.36
Item 28	0.63	0.25	0.45	0.31	0.44	0.46	0.54	0.59
Item 29	0.50	0.32	0.53	0.35	0.28	0.53	0.38	-0.20
Item 30	0.61	0.37	0.62	0.25	0.58	0.36	0.49	0.50
Coordination and Integration	0.56	0.31	0.50	0.31	0.44	0.42	0.46	0.33
Item 31	0.55	0.06	0.46	0.36	0.56	0.51	0.42	0.58
Item 32	0.59	0.36	0.47	0.56	0.67	0.50	0.52	0.57
Item 33	0.63	0.46	0.48	0.55	0.52	0.49	0.43	0.63
Item 34	0.43	-0.14	0.40	0.42	0.35	0.48	0.14	0.09
Item 35	0.50	0.37	0.34	0.46	0.49	0.33	0.44	0.35
Creating Change	0.54	0.22	0.43	0.47	0.52	0.46	0.39	0.44

Item 36	0.62	0.43	0.59	0.62	0.58	0.62	0.58	0.60
Item 37	0.66	0.24	0.56	0.55	0.60	0.63	0.47	0.31
Item 38	0.54	0.53	0.55	0.41	0.33	0.57	0.31	0.13
Item 39	0.57	-0.12	0.60	0.52	0.43	0.58	0.30	0.20
Item 40	0.26	0.28	0.24	0.12	0.05	0.34	0.12	0.06
Customer Focus	0.53	0.27	0.51	0.44	0.40	0.55	0.36	0.26
Item 41	0.52	0.62	0.48	0.41	0.45	0.39	0.37	0.22
Item 42	0.51	0.59	0.49	0.39	0.35	0.28	0.28	0.46
Item 43	0.37	0.42	0.28	0.44	0.41	0.22	0.03	0.23
Item 44	0.49	0.43	0.36	0.39	0.47	0.43	0.21	0.31
Item 45	0.55	0.43	0.40	0.52	0.53	0.43	0.31	0.63
Organizational Learning	0.49	0.50	0.40	0.43	0.44	0.35	0.24	0.37
Item 46	0.69	0.56	0.63	0.43	0.68	0.56	0.70	0.55
Item 47	0.48	0.38	0.32	0.21	0.46	0.37	0.45	0.39
Item 48	0.71	0.76	0.49	0.46	0.78	0.53	0.67	0.66
Item 49	0.79	0.74	0.80	0.66	0.76	0.73	0.72	0.69
Item 50	0.67	0.45	0.69	0.50	0.43	0.48	0.50	0.38
Strategic Direction and Intent	0.67	0.58	0.59	0.46	0.62	0.53	0.61	0.53
Item 51	0.60	0.59	0.57	0.48	0.45	0.23	0.54	0.48
Item 52	0.53	0.12	0.35	0.31	0.50	0.36	0.50	0.44
Item 53	0.51	0.53	0.69	0.46	0.70	0.40	0.55	0.56
Item 54	0.55	0.42	0.48	0.46	0.65	0.41	0.44	0.64
Item 55	0.59	0.66	0.41	0.49	0.68	0.36	0.47	0.62
Goals and Objectives	0.56	0.46	0.50	0.44	0.60	0.35	0.50	0.55
Item 56	0.62	0.31	0.65	0.50	0.49	0.52	0.40	0.59
Item 57	0.63	0.51	0.77	0.51	0.47	0.68	0.42	0.37
Item 58	0.32	0.37	0.36	0.18	0.28	0.42	-0.01	-0.28
Item 59	0.58	0.65	0.45	0.53	0.54	0.54	0.48	0.57
Item 60	0.57	0.15	0.57	0.44	0.36	0.47	0.41	0.24
Vision	0.54	0.40	0.56	0.43	0.43	0.53	0.34	0.30

Notes:

Maximum sample sizes are shown within parentheses.

Means of corrected item-total correlations are shown in bold in the rows corresponding to traits.

Table 3f: Alphas for Each of the 12 Traits by Language

	US English (N=19696)	Croatian (N=57)	Czech (N=80)	Danish (N=90)	Finnish (N=130)	Flemish (N=90)	French Canadian (N=202)	Hungarian (N=89)	Indonesian (N=33)	Norwegian (N=110)	Polish (N=204)	Russian (N=127)	Swedish (N=186)	Thai (N=195)	Vietnamese (N=43)
Empowerment	0.74	0.73	0.69	0.69	0.74	0.80	0.75	0.78	0.78	0.64	0.68	0.73	0.67	0.61	0.73
Team Orientation	0.82	0.82	0.79	0.74	0.70	0.75	0.82	0.81	0.82	0.78	0.76	0.79	0.81	0.77	0.75
Capability Development	0.69	0.70	0.66	0.76	0.62	0.78	0.77	0.61	0.61	0.57	0.54	0.73	0.69	0.35	0.57
Core Values	0.68	0.74	0.61	0.65	0.58	0.74	0.63	0.81	0.48	0.67	0.68	0.72	0.61	0.61	0.57
Agreement	0.71	0.68	0.68	0.76	0.71	0.52	0.72	0.78	0.66	0.63	0.71	0.69	0.62	0.61	0.28
Coordination and Integration	0.78	0.77	0.66	0.78	0.74	0.67	0.76	0.77	0.54	0.73	0.55	0.68	0.67	0.70	0.52
Creating Change	0.77	0.70	0.80	0.72	0.71	0.75	0.75	0.78	0.32	0.67	0.71	0.75	0.71	0.63	0.69
Customer Focus	0.76	0.68	0.85	0.74	0.77	0.75	0.80	0.61	0.45	0.74	0.69	0.63	0.77	0.60	0.46
Organizational Learning	0.73	0.70	0.72	0.62	0.75	0.56	0.72	0.67	0.72	0.65	0.67	0.68	0.59	0.43	0.61
Strategic Direction and Intent	0.85	0.77	0.81	0.72	0.78	0.87	0.87	0.82	0.79	0.80	0.68	0.81	0.75	0.80	0.76
Goals and Objectives	0.78	0.80	0.76	0.77	0.71	0.78	0.77	0.75	0.68	0.73	0.68	0.81	0.59	0.73	0.76
Vision	0.77	0.82	0.76	0.70	0.74	0.85	0.78	0.72	0.63	0.78	0.66	0.67	0.76	0.57	0.47

Note: Maximum sample sizes are shown within parentheses.

Appendix A: Technical Appendix for the DFIT Framework

The basic concept behind Raju et al.'s (1995) DFIT framework is that for a person, s , with some level of a trait, or \mathbf{q} , there is an expected true score for an item, i , for that person when they are viewed as either a member of the reference group, t_{isR} , as well as when they are viewed as a member of the focal group, t_{isF} . Two true scores for each person are based on the two sets of parameter estimates derived from the reference and focal samples. These true scores are compared by computing a difference statistic:

$$d_{is} = t_{isF} - t_{isR}. \quad (1)$$

When $d_{is} = 0$ at all values of \mathbf{q} the expected true scores for the item are considered to be identical, and so are the item parameters. Further, a conceptually equivalent statistic is used to determine DF at the test level:

$$D_s = (T_{sF} - T_{sR}) = d_{is} + \dots + d_{ns}, \quad (2)$$

where T_{sF} and T_{sR} , are the expected true scores for the test. Mathematically, T_{sF} and T_{sR} are simply the sum of their respective expected true scores at the item-level (i.e., t_{isR} and t_{isF}).

These equations lay the groundwork for the following two statistical indices which are used in the determination of DF items and tests. The first, termed non-compensatory DIF, or *NCDIF*, is used to determine DF at the item level, and is expressed as:

$$NCDIF = \mathbf{e}(d^2) = \mathbf{m}_{d^2} = \mathbf{s}_d^2 + \mathbf{m}_d^2. \quad (3)$$

Raju et al. (1995) have suggested using .096 as the cutoff for NCDIF of polytomous items with 5 response options. Values greater than this value suggest DF of items. The statistic used for examining DF at the test level (*DTF*) is expressed as:

$$DTF = \mathbf{e}(D^2) = \mathbf{m}_{D^2} = \mathbf{s}_D^2 + \mathbf{m}_D^2, \quad (4)$$

where \mathbf{e} signifies an expected value. Both NCDIF and DTF are simply difference measures between the reference and focal groups' expected true scores. In fact, taking the square root of

these statistics gives an approximation of the absolute true score difference between groups at the item and test level, respectively (Flowers, Oshima, & Raju, 1999).

It is important to note that DTF is compensatory in nature, meaning that on a two-item test, if one item shows DF such that the focal group is favored, and the other shows DF such that the reference group is favored, DTF will not suggest DF at the test level, because one item has compensated for the others' bias. This characteristic of DTF is important in that, as Raju et al. (1995) note, decisions are typically made at the level of the test, and not at the level of the item. However, as its name implies, NCDIF, is non-compensatory, and thus one item cannot compensate for DF found in another item on that test. Thus, a test with several DF items may not necessarily show significant amounts of *DTF*.

In summation, $NCDIF > .096$ signifies a DF item. The product of the number of items (n) and the NCDIF cutoff (.096) is suggested by Raju et al. (1995) as the cutoff for DTF. Thus, in the case of the Denison Culture Survey, for each of the twelve traits, the cutoff for DTF is: $n(NCDIF \text{ cutoff}) = 5(.096) = .48$, where n is the number of items in the test. All DFIT analyses in this report were conducted using Raju's (1999) DFITP5 program.

Appendix B: Technical Appendix for the LISREL Framework

Confirmatory factor analysis (CFA) is a measurement model which is a submodel of Jorakeskog & Sorbom's (1996) LISREL framework for structural equation modeling. The CFA model is written:

$$\mathbf{x} = \mathbf{L}_x \boldsymbol{\xi} + \mathbf{d}, \quad (5)$$

where \mathbf{x} is a matrix containing observed item responses, $\boldsymbol{\xi}$ is a matrix containing the specified factor pattern, or latent variables, \mathbf{L}_x is a matrix of factor loadings, or partial regression coefficients, estimated through the regression of x_i on the elements of $\boldsymbol{\xi}$, and \mathbf{d} is an error or residual matrix. By using Equation 5, and assuming that all x_i are measured in deviations from their mean, we have the covariance matrix of \mathbf{x} ,

$$\mathbf{S} = \mathbf{L}_x \mathbf{F} \mathbf{L}'_x + \mathbf{Q}_d \quad (6)$$

where \mathbf{F} is the variance-covariance matrix of the latent factors, $\boldsymbol{\xi}$, and, \mathbf{Q}_d is a diagonal matrix containing measurement error variances. The right-hand side of Equation 6 (for the covariance matrix, \mathbf{S}), contains all parameter matrices of interest in examining measurement equivalence through CFA.

This examination of measurement equivalence is applied through testing the equality of three separate parameters by sequentially constraining the measurement model estimated freely across groups. We are concerned with the equality of: 1) factor loadings, \mathbf{L}_x , to examine the feasibility of comparing mean scores across groups or measures; 2) factor structures, \mathbf{F} , to examine if identical factor structures are found across forms; and 3) error terms, \mathbf{Q}_d which would suggest that measurement reliability is equal across groups.

The test for equality of these three components of the CFA model was conducted using a sequential nested models approach, starting with the most complex measurement model, one in which the parameters noted above are estimated freely. This model is then compared to a model

where the elements of L_x are constrained to be equal across groups. The model constraining L_x is then compared to a model with the additional constraint that the elements of F are constrained to be equal. Finally, the model constraining both L_x and F is then compared to a model which imposes the additional constraint that the Q_d be equal for each item across groups. Differences in the chi-square statistic for each model, Δc^2 , is then used in reference to its associated change in degrees of freedom (Δdf) to determine the statistical significance of Δc^2 . However, as noted in the *Results of Confirmatory Factor Analysis* section of this report, this statistic is known to be highly sensitive to sample size, in the sense that significant c^2 values are commonly an artifact of large samples (Joreskog & Sorbom, 2003). Thus, differences in other fit indices such as the NNFI and CFI were the primary determinants in examining measurement equivalence. Both of these indexes are based on c^2 and its associated df . Differences in NNFI and CFI greater than .01 are considered to be indicative of differences between the fit of the specified measurement model for different forms of the survey.

References

- Baker, F.B.(1995). *EQUATE 2.1: A computer program for equating two metrics in item response theory* [Computer program]. Madison: University of Wisconsin, Laboratory of Experimental Design.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238-246.
- Drasgow, F., Levine, M.V., Tsien, S., Williams, B.A., & Mead, A.D. (1995). Fitting polytomous item response theory models to multiple-choice tests. *Applied Psychological Measurement*, *19*, 143-165.
- Flowers, C.P., Oshima, T.C., & Raju, N.S. (1999). A description and demonstration of the polytomous DFIT framework. *Applied Psychological Measurement*, *23*, 309-326.
- Joreskog, K.G., & Sorbom, D. (1996). *LISREL 8: User's Reference Guide*. Chicago, Scientific Software International.
- Joreskog, K.G., & Sorbom, D. (2003). *LISREL (Version 8.54)* [Computer software]. Lincolnwood, IL: Scientific Software International.
- Raju, N.S. (1999). *DFITP5: A fortran program for calculating polytomous DIF/DTF* [Computer program]. Chicago, IL: Illinois Institute of Technology.
- Raju, N.S., van der Linden, W.J., & Fleer, P.F. (1995). IRT-based internal measures of differential functioning of items and tests. *Applied Psychological Measurement*, *19*, 353-368.
- Raju, N.S., & Ellis, B.B. (2002). Differential functioning of items and tests. In F. Drasgow & N. Schmitt (Eds.), *Measuring and analyzing behavior in organizations: Advances in measurement and data analysis* (pp. 156-188). San Francisco, CA: Josey-Bass.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores.

Psychometrika Monograph Supplement, 34, 100-114.

Stark, S. (2001). *MODFIT: A computer program for model-data fit*. Urbana-Champaign:

University of Illinois, Urbana-Champaign.

Stocking, M.L., & Lord, F.M. (1983). Developing a common metric in item response

theory. *Applied Psychological Measurement*, 7, 201-210.

Thissen, D. (2003). *MULTILOG 7.0: Multiple, categorical item analysis and test scoring using*

item response theory [Computer Program]. Chicago, IL: Scientific Software, Inc.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum-likelihood factor

analysis, *Psychometrika*, 35, 417-437.

Principal Team Members

Michael J. Zickar, Ph.D. (Faculty Sponsor)

Michael received his Ph.D. (1997) in Industrial-Organizational (I-O) Psychology from the University of Illinois at Urbana-Champaign. He is an associate professor of psychology and chair of the doctoral program in I-O Psychology at BGSU. His research interests include application of psychometric analysis to organizational issues and development of psychological tests. His research has been published in leading scientific journals such as *Journal of Applied Psychology*, *Personnel Psychology*, and *Applied Psychological Measurement*.

Nathan T. Carter, M.A.

Nathan earned his bachelor's degree in psychology and master's degree in I-O Psychology at Western Kentucky University in 2004 and 2006. He is currently a doctoral student in I-O Psychology at BGSU. His research interests focus on the application of psychometric techniques, particularly item response theory and factor analysis, in organizational and educational settings. He also has interests in the role of personality in the workplace and the history of applied psychology.

Dalia L. Diab, M.S.

Dalia holds a bachelor's degree in psychology from the American University of Beirut and a master's degree in I-O Psychology from Indiana University-Purdue University Indianapolis. She is currently a third-year doctoral student in I-O Psychology at BGSU. Her research interests include judgment and decision making in the workplace, organizational attraction, recruitment, psychometrics, and cross-cultural issues at work.

Bing C. Lin, B.A.

Bing received his bachelor's degree in psychology from Purdue University, West Lafayette, Indiana, in 2006. He is currently a doctoral student in I-O Psychology at BGSU. His research interests include item response theory and general occupational health psychology issues, such as work stress, unwinding experiences, and counterproductive work behavior.

Shuang-Yueh Pui, B.S.

Shuang received her bachelor's degree in psychology from University of Idaho, Moscow, Idaho in 2004. She was a doctoral student in I-O Psychology at Wayne State University for two years and transferred to BGSU in 2006. Her research interests involve judgment and decision making in the workplace, organizational staffing, and cross-cultural psychology.