

Cross-Cultural Validation of the Work Values Scale EVAT Using Multi-Group Confirmatory Factor Analysis and Confirmatory Multidimensional Scaling

Luis M. Arciniega¹, Luis González², Vítor Soares³, Stefania Ciulli⁴, and Marco Giannini⁴

¹Instituto Tecnológico Autónomo de México (Mexico)

²Universidad de Salamanca (Spain)

³Instituto Piaget (Portugal)

⁴Università degli Studi di Firenze (Italy)

The Work Values Scale EVAT (based on its initials in Spanish: *Escala de Valores hacia el Trabajo*) was created in 2000 to measure values in the work context. The instrument operationalizes the four higher-order-values of the Schwartz Theory (1992) through sixteen items focused on work scenarios. The questionnaire has been used among large samples of Mexican and Spanish individuals reporting adequate psychometric properties. The instrument has recently been translated into Portuguese and Italian, and subsequently used in a large-scale study with nurses in Portugal and in a sample of various occupations in Italy. The purpose of this research was to demonstrate the cross-cultural validity of the Work Values Scale EVAT in Spanish, Portuguese, and Italian. Our results suggest that the original Spanish version of the EVAT scale and the new Portuguese and Italian versions are equivalent.

Keywords: work values, EVAT Scale, cross-cultural validation, multidimensional scaling.

La Escala de Valores hacia el Trabajo (EVAT) fue desarrollada en el 2000 para medir valores en el contexto del trabajo. El instrumento operacionaliza los cuatro valores de orden superior de la Teoría de Schwartz (1992) a través de dieciséis ítems focalizados en escenarios laborales. El cuestionario ha sido empleado en grandes muestras de participantes mexicanos y españoles, mostrando adecuadas propiedades psicométricas. El instrumento ha sido traducido recientemente al portugués e italiano, siendo utilizado en un estudio con una gran muestra de enfermeras en Portugal, y en una muestra conformada con sujetos de diferentes ocupaciones en Italia. El propósito de este estudio consistió en demostrar la validez transcultural de la Escala de Valores hacia el Trabajo, en español, portugués e italiano. Los resultados revelan que tanto la versión original en castellano, como las nuevas versiones en portugués e italiano son equivalentes..

Palabras clave: valores hacia el trabajo, EVAT, validación transcultural, escalamiento multidimensional.

The completion of this article was supported in part, by the Asociación Mexicana de Cultura A.C.

This article was concluded during a sabbatical stay of the first author at the Department of Psychology, at the University of Western Ontario, Canada.

Correspondence concerning this article should be addressed to Luis M. Arciniega. ITAM School of Business. Río Hondo 1. México 01080 D.F., (Mexico). Phone: +52-5556284000. Fax: +52-5556284049. E-mail: larciniega@itam.mx

Although from a semantic standpoint, work and general values are highly related constructs, they have been analyzed from distinct perspectives in the field of research. (Sagie, Elizur & Koslowsky, 1996; Schwartz, 1999). In other words, work values have been studied and measured in isolation from main stream research in general values (Rokeach, 1973; Schwartz, 1992).

The study of general values has developed a well-specified definition of the construct (e.g. Rokeach, 1973; Schwartz, 1992), and even more important, a dynamic structure that allows for classification (Schwartz, 1992).

During the last two decades a great deal of attention has been paid to the structure of both constructs. In the field of general values, the dynamic structure of values proposed by Schwartz (1992) is the most widely known theory, whereas, in the specific area of work values, the structure proposed by Elizur (1984) has shed light for many researchers in the systematic study of the construct.

Schwartz's Theory of Values (1992) establishes that the essence of a value is the motivational goal it expresses. Based on this idea, the author has derived ten types of values that form a dynamic structure (see figure 1, left), where types sharing a similar motivational goal appear closer between them (for a full description of the ten motivational types see Schwartz, 1992). On the other hand, types representing incompatible motivational goals occupy opposing sides of the continuum. As seen in figure 1 (left), the 10 types form four high-order values. It is possible to distinguish two large bipolar dimensions. Each dimension presents opposite high-order values on each of its poles.

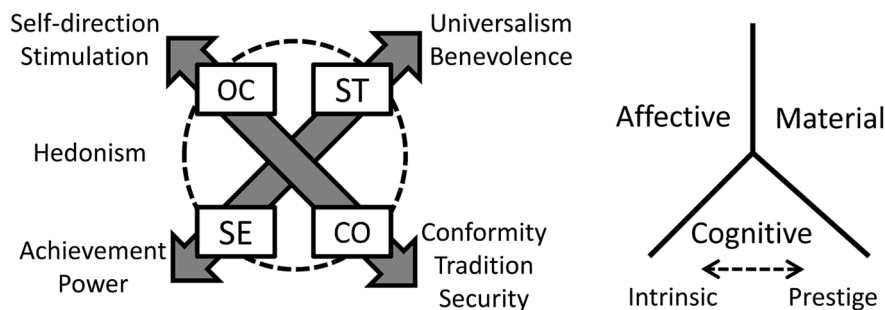
The four high-order values are labeled: *self-enhancement*, *self-transcendence*, *openness to change* and *conservation*. The first two are part of a bipolar dimension that refers to opposite motivational objectives: one, to enhance personal interests even at the expenses of others and the other, to transcend selfish concerns and promote the welfare of others. The other bipolar dimension clusters two different objectives:

one refers to the extent to which they motivate individuals to follow their own and unique intellectual and emotional interests, and the other, centered on preserving the *status quo* and the stability in relations with persons and institutions (Schwartz, 1992). Figure 1 shows which value types are contained by each high-order value.

Considering the modality of the outcomes, Elizur (1984) proposed a structure for the construct of work values based on three dimensions (see figure 1, right). He labeled these dimensions as instrumental (or material), cognitive and affective. The proposed structure has been validated in different countries (Elizur, 1984; Elizur, Borg, Hunt and Beck, 1991) and seems to be robust.

Ros and collaborators (Ros, Schwartz and Surkiss, 1999) suggested and proved that the four higher-order values of the Schwartz theory and the three dimensions of the modality facet of work values established by Elizur (1984) can converge, if the cognitive dimension is divided into two sub-dimensions: prestige and intrinsic (see Figure 1 right). Once the cognitive dimension is divided, each dimension of the modality facet encounters its parallel in one of the four higher-order values. That is, the higher-order value *openness to change* parallels the intrinsic dimension, *conservation* the instrumental or material dimension, *self-enhancement* the prestige, and finally, the high-order value *self-transcendence* parallels the affective dimension. Based on this convergent theoretical construct structure, Arciniega and González (2000) developed an instrument for measuring work values.

The Work Values Scale EVAT (based on its initials in Spanish: *Escala de Valores hacia el Trabajo*) operationalizes the four higher-order-values of the Schwartz Theory (1992) through sixteen items focused on work scenarios (Arciniega & González, 2000). The questionnaire has been used among large samples of Mexicans and Spaniards to assess the impact of work values on job attitudes, such as job satisfaction (Arciniega & González, 2005), organizational



OC=Openness to change, CO= Conservation, SE= Self-enhancement, ST= Self-transcendence

Figure 1. The four high-order values of the Schwartz's theory (1992) and the three dimensions of the modality facet of Elizur's model (1984).

commitment (Arciniega & González, 2006) and career anchors (González & Arciniega, 2005), reporting adequate psychometric properties. Recently the instrument has been translated into Portuguese and Italian, and used in a large-scale study with nurses in Portugal (Soares, 2008), and in a sample of various occupations in Italy (Ciulli, 2005).

The purpose of this research was to demonstrate the cross-cultural validity of the Work Values Scale EVAT in Spanish, Portuguese, and Italian, using two techniques of measurement equivalence: multi-group confirmatory factor analysis (MGCFA) and confirmatory multidimensional scaling (CMDS).

Method

Participants.

Participants in the present study were drawn from three different countries. Each sample was composed of 220 employees. As a prerequisite to testing measurement invariance across samples it is necessary to ensure sample comparability. A common and suitable procedure for this is to use samples that are matched on the basis of a predetermined set of characteristics (Steenkamp & Baumgartner, 1998). Therefore, samples in the present study consisted of working adults with at least 13 years of formal education, and with equivalent age composition (as described further). The years of formal education criterion was obtained from the Italian sample, and fixed on the Mexican and Portuguese samples.

Italy sample.- This sample was specially collected by the fourth author to conduct a study to validate the Italian version of the EVAT scale (Ciulli, 2005). Sampling was based on the *snowball* method. Volunteers were solicited by the fourth author to participate in the study and were encouraged to recruit their acquaintances to participate as well. The 220 individuals were Italian citizens living in the Tuscany region, 38.6 % were male. Age composition of the 220 individuals is shown on Table 1. Based on these percentages, we tried to stratify the Mexican and Portuguese samples. Participants voluntarily and anonymously completed the Italian version of the EVAT scale.

Mexico sample.- A 220 subjects sample was extracted randomly from a large database of 3,201 employees from 30 different companies belonging to the same holding, from 11 different cities in central and northern Mexico (Arciniega & González, 2005). This process was constrained by the matching criteria described above. Sixty-one percent (61%) of the participants were male. Participants voluntarily and anonymously completed the original Spanish language version of sixteen items of the EVAT scale.

Portugal sample.- The 220 individuals of this sample were selected randomly from a database of a large scale study with nurses in Portugal (N= 1,140: Soares, 2008).

Table 1
Age composition by sample.

	< 25 years	26-35 years	36-45 years	> 45 years
Mexico	30%	47%	12%	11%
Portugal	29%	50%	14%	7%
Italy	36%	42%	11%	11%

Twenty percent (20%) of the participants were male. Participants voluntarily and anonymously completed the Portuguese language version of the EVAT scale.

Since we were unable to obtain perfect matching on each of the four age ranges in the stratification process described earlier (see Table 1), a χ^2 test was computed between the percentages of age ranges, suggesting that samples were equivalent with respect to this demographic variable ($\chi^2_{(6)} = 3.09, p = 0.80$).

Measures

The *Escala de Valores hacia el Trabajo* (EVAT: Arciniega & González, 2000), is a sixteen item self-report measure that operationalizes the four higher-order values of the Schwartz's theory (Schwartz, 1992), contextualized to the work environment. The sixteen items of the instrument are based on the *Portraits Values Questionnaire* (Schwartz, Melech, Lehman, Burgess, Harris, & Owens, 2001). The questionnaire uses short verbal portraits that describe the goals and wishes of sixteen employees, implicitly expressing their work values (e.g. *He always strives to make sure that all employees receive the same treatment and opportunities*). Respondents are asked to rate themselves in terms of each of the sixteen portraits, and use a 7-point scale (1= *not like me at all*, 7= *very much like me*) to score their comparisons.

To create the Portuguese and Italian versions of the EVAT, back translation processes were followed based on the method proposed by Brislin (1986). Both the Portuguese and Italian versions of the instrument are available from the authors.

Analysis

Firstly, we ran Confirmatory Factor Analyses (CFA's) in order to assess the EVAT construct's structure on each of the three samples independently. We next turned into testing the measurement equivalence across the data from the three different countries.

Although a number of approaches have been used to evaluate measurement equivalence across samples (Vandenberg and Lance, 2000), there is general agreement that the multisample covariance structure analysis (Joreskog, 1971) provides the most powerful and versatile technique for testing cross-group measurement equivalence.

One of the most robust methods for assessing measurement equivalence across samples using the MGCFA technique is the one proposed by Vandenberg and Lance (2000). The method suggests two complementary sets of tests of measurement equivalence. The first is named measurement invariance, and the second, structural invariance. Measurement invariance tests are concerned on assessing the equivalence of the relationships between observed variables and latent variables across samples. If these tests are validated, it should imply that the instrument and its underlying latent variables conceptually mean the same thing across samples. Meanwhile, the structural invariance set of tests evaluate the relationships among the latent variables (*e.g.* unique variances, covariances, means). If these constraints are demonstrated, it would indicate, that the instrument not only conceptually means the same thing across cultures, it would also support the idea, that individuals from the three samples respond on the same manner, that is, same degree of reliability, variance, range of ratings, and mean level of ratings, as if they were all part of a large sample.

Since the goal of the present study is focused on demonstrating that the construct structure of the EVAT scale is conceptually equivalent across the samples from the three different countries, we concentrated on evaluating the tests related to measurement invariance: *configural* and *metric* invariance.

Firstly, a model with an equivalent factor structure (*i.e.*, items relate to the same factors) across groups was evaluated in order to test for *configural* invariance. Secondly, a nested sequence of Model 1 was assessed, adding the constraint that factor pattern coefficients for like items are equal across groups, providing a test of *metric* invariance.

If the conditions are demonstrated, it could be said that across samples there is the same number of underlying dimensions, the items load on the same factors, and the items estimated factor loadings have equivalent patterns

and equivalent magnitudes. In other words, the EVAT scale and its underlying constructs, conceptually mean the same thing across individuals from the three different countries.

Secondly, we computed three independent multidimensional scaling configurations, one for each of the samples. We then selected the configuration that best represented the theoretical construct structure. Afterward, we used the coordinates of this representation as the base for computing the other two configurations. This process consisted of fixing constraints to a MDS configuration that had been labeled by Borg and Groenen (2005) as a confirmatory MDS. The essence of the method is comparing the goodness of fit indices of a free configuration (*i.e.*, without constraints) versus a MDS configuration where the coordinates in one or two axis are fixed (*i.e.*, base configuration), and evaluating the degree of the reduction in the goodness of fit (*e.g.* S-stress).

Results

Concerning the independent CFA's for each of the three samples, we found adequate goodness of fit indices between each sample and the proposed four oblique factors model. The CFI for the Mexican sample was of 0.962 (RMSEA=0.053), that for the Portuguese was of 0.950 (RMSEA=.053), and a CFI of 0.942 (RMSEA=0.048) was obtained with the data from Italy.

For the multi-group confirmatory factor analyses, the fit indices for each of the two models are presented in Table 2. As expected, Model 1 provides the best overall level of fit relative to the others models. It appears that the 4 oblique factor model, with 4 items per factor, provides an adequate representation of the data in each of the three samples.

Model 2 indicates relatively little decrement in fit relative to Model 1. Specifically, though the change in χ^2 from Model 1 to Model 2 is statistically significant, the additional

Table 2
Results for the sequence of measurement invariance tests for the EVAT scale.

Model	χ^2	Df	$\Delta\chi^2$	Δdf	P	RMSEA	CFI	ΔCFI
Mexico - Italy - Portugal								
1 Configural invariance	384.86	239				0.053	0.955	
2 Metric invariance	466.46	261	81.6	22	0.00	0.060	0.936	-0.02
Mexico - Italy								
1b Configural invariance	253.7	158				0.053	0.957	
2b Metric invariance	274.3	169	20.7	11	0.04	0.055	0.950	-0.01
Mexico - Portugal								
1c Configural invariance	265.49	160				0.055	0.958	
2c Metric invariance	318.72	171	53.2	11	0.00	0.063	0.941	-0.02
Italy - Portugal								
1d Configural invariance	250.34	160				0.051	0.947	
2d Metric invariance	308.82	171	58.5	11	0.00	0.061	0.919	-0.03

fit indices are nearly identical to those for Model 1. In addition, the change in fit from Model 1 to Model 2 is within the critical values suggested by Cheung and Rensvold (2002) for the CFI. Models 2 b-d represent tests of *configural* and *metric* invariance across all possible two groups comparisons in the present study (*i.e.*, Mexico vs. Italy, Mexico vs. Portugal, and Italy vs. Portugal). For each of these models all of the fit measures indicate that the null hypothesis of invariance should not be rejected. Thus, it appears that the EVAT measure is conceptually equivalent across all three groups, and between each possible pair of groups.

We computed independent MDS configurations for each of the three samples (see figure 2). Then we selected the configuration obtained from the Italian sample as the base configuration, considering it to be the best representation of the theoretical structure. Finally, we computed confirmatory MDS's for the Mexican and Portuguese samples fixing the coordinates obtained from the Italian sample using the Proxscal algorithm included in the version 15.0 of the SPSS package. Table 3 shows the raw goodness of fit indices for the three samples, and the constrained indices for the Mexican and Portuguese samples. The light decrement in the goodness of fit indices allows us to say that the three configurations are equivalent.

A snapshot of the raw MDS configurations (Figure 2), can supply clues as to why one or more items, in one or more factors, are not clustering properly with the rest of the items operationalizing a specific construct or constructs. For instance, the four items operationalizing *self-enhancement* (SE) are: 3, 7, 11 and 15. In figure 2, most of the points representing these items appear in the top-left region. The triangles represent the configuration from the data from Mexico, the circles from Portugal and the squares from Italy. It is evident that the points associated with item 7 from the Mexican, Italian and Portuguese configurations appear to be close to one another. A similar pattern occurs with items 3, and 11, and in a certain way, with item 15. However, an item that was not supposed to appear in this region, item 14, should have appeared closer to the rest of the items operationalizing the higher-order value *conservation* (CO: bottom left on Figure 2). Apparently, participants from Italy and Mexico, clearly associated the content of this item to the meaning of the *higher-order* value *conservation*, but the subjects from the Portuguese sample also associated the content of the item with the core concept of *self-enhancement*,

Table 3
Goodness of fit of the free and constrained configurations.

Configuration with the data from	S-stress of the free configuration	S-stress fixing the coordinates to the base configuration
Italy (base configuration)	0.067	-
Mexico	0.074	0.204
Portugal	0.047	0.108

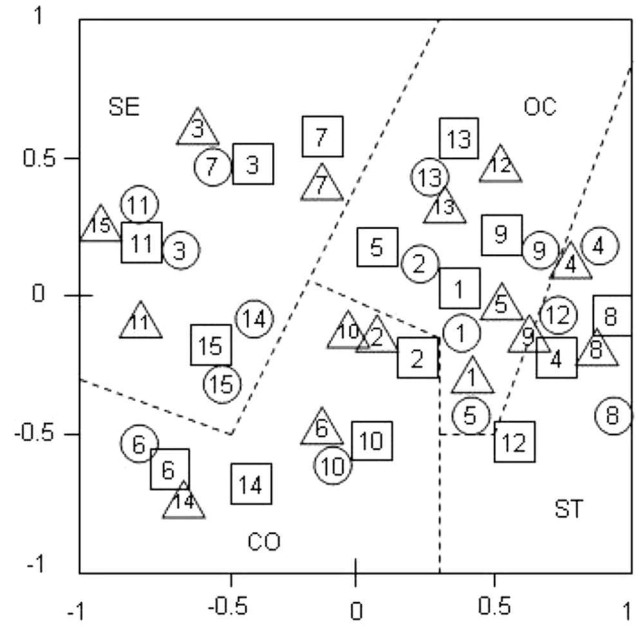


Figure 2. -MDS configurations of the three samples.
Notes: SE= self-enhancement, OC= openness to change, CO= conservation, ST= self-transcendence. Triangles represent points associated to the items of the configuration from the Mexican sample, circles from the Portuguese, and squares from the Italian.

explained by the marginal position of the point in the border line of the two regions. Even when the S-stress, and the *configural* and *metric* invariance tests, across the three samples and between Portugal vs. Mexico, and Portugal vs. Italy, do not suggest a problem with the equivalence of meaning of the 16 items in general, this particular type of analysis could suggest specific actions, in a particular item, in one of the three versions of the instrument. Therefore, we propose that CMDS is a helpful complementary tool of MGCFA when assessing measurement equivalence across groups.

Discussion

The measurement equivalence of constructs across cultures is a critical and often underestimated assumption in cross-cultural applied research. Surprisingly, many multinational corporations tend to show comparative studies about values, attitudes and perceptions of their employees between the countries where they have branches and/or operations, without being certain they are really measuring the same constructs across cultures.

This study provides researchers and practitioners of the field of work and organizational Psychology, and of Psychology in general, a specific instrument for measuring values in the work context that can be used for conducting reliable comparisons between Spanish, Italian and Portuguese speakers. The construct's structure of the instrument is based on the most robust theory of values in the Social Psychology field.

Our results suggest that the original Spanish version of the EVAT scale and the recently developed Portuguese and Italian versions were equivalent across the three cultures we sampled. The results also suggest that the four oblique factor structure of the EVAT scale proposed by Arciniega and González (2000) is stable across Portuguese and Italian populations. It seems that the EVAT scale is an effective instrument for operationalizing the four higher-order values of the Schwartz Theory in the work context, and in different languages.

While our results indicate measurement equivalence across three Latin cultures, it is yet to be determined whether such equivalence exists across additional countries from different continents and religions (e.g. China, Germany, India, Indonesia).

The article also provides an example of a particular procedure for assessing measuring equivalence across cultures, using two complementary multivariate statistical techniques.

Multi-group confirmatory factor analysis is a widely used technique for assessing measurement equivalence across groups. As can be seen in table 2, changes in the goodness of fit are evaluated probabilistically, giving the process objectivity. Researchers can clearly determine if they are measuring with *configural* or *metric* invariance, and decide if they can compare measurements across groups or not. If *configural* invariance is validated, but *metric* invariance is rejected, it is hard to know where the lack of fit is. CMDS can be a helpful tool on discovering those items that are being interpreted in a different way, as was described in detail in the results section based in one specific example.

Our results suggest that CMDS is a serviceable technique for assessing measurement equivalence, but requires improvements to provide precise fit indices. Future research should concentrate on running parallel simulations of MGCFA and CMDS, in order to determine possible cut-off values for the S-stress index. In other words, to establish a range for the S-stress that could be equivalent to *configural* and *metric* invariance.

Referentes

- Arciniega, L.M., & González, L. (2000). Desarrollo y validación de la escala de valores hacia el trabajo EVAT 30. *Revista de Psicología Social, 15*, 281-296.
- Arciniega, L.M., & González, L. (2005). Other-oriented values and job satisfaction. *Problems and Perspectives in Management, 4*, 128-132.
- Arciniega, L.M., & González, L. (2006). What is the influence of work values relative to other variables in the development of organizational commitment? *Revista de Psicología Social, 21*, 35-50.
- Borg, I., & Groenen, P.J.F. (2005). *Modern multidimensional scaling: Theory and applications*. New York, NY.: Springer.
- Brislin, R. W. (1986). The wording and translation of research instruments. In W. J. Lonner & J. W. Berry (Eds.), *Field methods in cross-cultural research* (pp. 137-164). Newbury Park, CA.: Sage.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233-255.
- Ciulli, S. (2005). *Uno strumento per la misura dei valori professionali: proprietà psicometriche della scala de valores hacia el trabajo EVAT*. Unpublished manuscript. Università degli studi di Firenze. Florence, Italy.
- Elizur, D. (1984). Facets of work values: A structural analysis of work outcomes. *Journal of Applied Psychology, 69*, 379-389.
- Elizur, D., Borg, I., Hunt, R. & Beck, I.M. (1991). The structure of work values: A cross cultural comparison. *Journal of Organizational Behavior, 12*, 21-38.
- González, L., & Arciniega, L.M. (2005). Anclajes de carrera profesional y valores en el trabajo. In J. Romay, J. Salgado, M. Romani y D. Robla (Comps.). *Psicología de las Organizaciones, del Trabajo y de los Recursos Humanos y de la Salud, 4*. 69-77. Madrid: Biblioteca Nueva.
- Joreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika, 36*, 409-426.
- Rokeach, M. (1973). *The nature of human values*. New York: Free Press.
- Ros, M., Schwartz, S.H., Surkiss, S. (1999). Basic individual values, work values, and the meaning of work. *Applied Psychology: An International Review, 48*, 49-71.
- Sagie, A., Elizur, D. & Koslowsky, M. (1996). Work values: a theoretical overview and a model of their effects. *Journal of Organizational Behavior, 17*, 503-514.
- Schwartz, S. H. (1999). A theory of cultural values and some implications for work. *Applied Psychology: An International Review, 48*, 23-47.
- Schwartz, S.H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna. (Ed.). *Advances in Experimental Social Psychology, 25*, Academic Press, New York, pp. 1-65.
- Schwartz, S.H., Melech, G., Lehman, A., Burgess, S., Harris, M. & Owens, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross-Cultural Psychology, 32*, 519-542.
- Soares, V.M. (2008). *Valores e Burnout em enfermagem*. Unpublished Doctoral Dissertation. Universidad de Salamanca, Salamanca, Spain.
- Steenkamp, J. B. E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*, 78-90.
- Vandenberg, R.J., & Lance, C.E. (2000). A review and synthesis of measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*, 4-69.

Received April 22, 2008

Revision received October 20, 2008

Accepted November 24, 2008