STATISTICS IN TRANSITION-new series, March 2012
Vol. 13, No. 1, pp. 143—158

CUSTOMERS RESEARCH AND EQUIVALENCE MEASUREMENT IN FACTOR ANALYSIS

Piotr Tarka[1]

ABSTRACT

Factors Analysis is often tied to specific properties of population and its culture characteristics. If measurement is applied from population to another, then extracted factors may hard to be equally compared on the reflective basic level, unless all conditions of invariance measurement are met. Hence, implementation of customers research and any inter-cultural studies require a multi-cultural model describing statistical differences in both cultures with invariance as underlying assumption. In the article we implement a model for analysis of customers' personal values pertaining to hedonic consumption aspects in two culturally opposite populations. We conducted survey in two countries and the following cities: Poland (Poznan) and The Netherlands (Rotterdam and Tilburg) with randomly prepared samples with youth representatives on both sides. This model permitted us for testing invariance measurement under cross-group constraints and thus examining structural equivalence of latent variables - values.

1. Introduction

One of the main problems in most of socio-economic research is the measurement of equivalence pertaining to samples drawn from the different populations. Equivalence as a word relates to one of the categories of *quality assessment* in studies when scores obtained from e.g., two populations are set for comparison. Equivalence of a measurement is related to the assessment of the extent to which measurements are made in the tested groups using the same units and measures, distribution scores relating to the same characteristics of respondents according to various conditions and context of made observation (e.g., based on socio-economic factors, or frame of time). The measurement is therefore characterized by invariance level. In the absence of invariance in

[1] Poznan University of Economics, Department of Marketing Research, al. Niepodległości 10, Poland. E-mail: piotr.tarka@ue.poznan.pl.

measurement, any differences between individuals and populations cannot be reasonably interpreted as comparisons in any multi-group studies. This is particularly important in studies when we use units of measurement that are relative and conventional, associated with the respondents adopted own independent system reference (Sagan, 2005; Tarka, 2010).

2. Underlying assumptions of invariance measurement

Table 1 shows the main types of invariance research occurring at all stages of the research process. The issue of measurement invariance is crucial for studies that are aimed to investigate group differences. Cross-cultural methodologists have emphasized that group comparisons assume invariance of the elements of the measurement structure (i.e., factor loadings and measurement errors) and of response biases (Billiet, 2002; Little et al., 2006). And group comparisons within a single culture also require measurement invariance to insure that potential differences (e.g., in the means or regression coefficients) can be interpreted reliably (Vandenberg and Lance, 2000).

Sub-groups within populations are often heterogeneous with regard to the parameter values of a model. Nonetheless, most within-society research continues implicitly to assume homogeneity of the population (Muthén, 1989). This is especially happening in the field research pertaining to convenience samples of social, educational, or occupational sub-groups. These groups often differ from one another or from the overall population with regard to measurement or structural parameters. In the worst case, researchers measure different constructs in the groups. Hence, within-society studies should assess possible lack of measurement invariance, when possible, to uncover potential population heterogeneity.

Researchers usually assume invariance of the structure of their measures as they compare them across the groups. The validity of this assumption is critical for any conclusions about group related differences (Vandenberg and Lance, 2000). If it is not true, one cannot even claim that the construct is the same in the different groups (Little et al., 2006). Thus, legitimate comparison of means or structural relations across groups requires invariance of the measurement structures underlying the indicators (Ployhardt and Oswald, 2004; Thompson and Green, 2006).

Table 1. General categories for invariance measurements in different populations

Categories of intercultural equivalence test	**Types of equivalence in the category**	**Description**
Invariance of the research problem	Conceptual Functional	The identity of the constructs examined The similarity function of concepts and actions, predictive validity
Translation invariance	Lexical Idiomatic Grammar Pragmatic	The importance of vocabulary terms The importance of mobile and customary terms The adequacy of grammatical structures The importance of colloquial words in everyday life and action
Measurement invariance	Global Structural Metric Scalar Measurement errors	The similarity of the covariance matrix The adequacy of measurement models Comparability of measurement units Similarity measurement scale Homogeneity of the impact of specific factors
Sample invariance	Sampling units Representativeness	Comparability of sampling units Compliance operational units, the dimensions of socio-demographic stratification
Data collection invariance	Communication with the respondent Context Style and attitude response	The similarity of behaviour patterns, the definition of private and public spheres Commonality of questions of cultural context, the areas of social taboos and permissions Consistency and similarity of responses to the posed questions and themes non-response

Source: Sagan, 2005.

3. Process of invariance measurement

Table 1 shows the main types of invariance research occurring at all stages of the research process. The issue of measurement evaluation based on invariance begins with series of conducted tests where one checks the hypotheses related to dispersions among the groups. These tests should be carried out in a sequence, because the bad model fit, makes another baseless testing measuring the level of cross-cultural equivalence (Meredith, 1993; Sagan 2005). As a result we obtain

configural invariance of the whole factorial structure and metric invariance of the factor loadings which are critical for the interpretation of constructs and are requisites for all other measurement. Configural invariance implies the same number of factors in each group and the same pattern of fixed and free parameters. It is a prerequisite for the other tests. It is the very basic form of invariance and it assesses whether we find the same patterns of loading between indicators and factors in both groups. The parameter restrictions only refer to the patterns of "loading" and "non-loading". Configural invariance is assumed if the same items measure the same factors in both groups. If configural invariance is not supported empirically, there are fundamental distinctions in the measurement structure, which means that the manifest variables measure different latent variable.

The **metric invariance** is more stringent in comparison to the configural invariance, as additional restrictions are adopted. Metric invariance means that, in addition to the conditions of configural invariance for all groups, the factor loadings are equivalent. If the model of metric invariance is maintainable, the manifest variables measure the latent variables equally well. If the model fit of the metric invariance model does not decrease significantly, metric invariance of all items can be assumed. Given a metric invariance, the contents of the factors are assumed to be equivalent. Likewise, the relations of the variables with other variables may be compared across the groups (Bollen, 1989).

The test of metric invariance is conducted by comparing the fit of the metric and configural invariance models to the data with chi-square statistics. Further 'modern' indications for invariance are differences in the indices such as *Comparative Fit Index* (CFI), *Root Mean Square Error of Approximation (RMSEA)*, and *Standardized Root Mean Residual* (SRMR). Metric invariance implies equal factor loadings across groups. For instance, the parameter λ_{ij} - representing factor loading must be the same in groups e.g., A and B. And this is tested by imposing equality constraints on the Λ - matrices that contain the factor loadings (i.e., $\Lambda^A = \Lambda^B = ...\Lambda^G$, where superscripts refer to groups A to G). Equal factor loadings indicate that the groups calibrate their measures in the same way. Hence, the values on the manifest scale have the same meaning across groups.

Metric invariance concerns a construct comparability that metric invariance is a stricter condition of construct comparability. According to the common factor perspective, the factor loadings indicate the strength of the causal effect of the latent variable ξ_j on its indicators and can be interpreted as validity coefficients (Bollen, 1989). Significantly different factor loadings imply a difference in the validity coefficients. This raises concerns about whether the constructs are the same across groups. Hence, configural invariance, by providing evidence that the construct is related to the same set of indicators, is a prerequisite for inferring that the construct has the *similar* meaning. However, metric invariance is necessary to infer that the construct has the *same* meaning, because it provides evidence about the equality of validity coefficients.

Additionally, **scalar invariance** refers to invariance of the item intercepts in the regression equations that link the indicators x_i^g to their latent variable ξ_j^g. Item intercepts can be interpreted as the systematic biases in responses of a group to an item. As a result, the manifest mean can be systematically higher or lower (upward or downward biased) than one would expect due to the groups' latent mean and the factor loading. Scalar invariance is present if the degree of up- or downward bias of the manifest variable is equal across groups. It is absent if one of the groups differs significantly in one or more of the item intercepts. To test for scalar invariance, one constrains the tau-vectors to be equal across groups $\tau^A = \tau^B = ...\tau^G$.

Then, we follow **invariance of factor variance / covariance,** which appears when groups have the same variances in their respective latent variables. This is tested by constraining the diagonal of the phi-matrices $\phi_{jj}^A = \phi_{jj}^B = ...\phi_{jj}^G$, to be equal. And invariance of the factor covariances refers to equality of the associations among the latent variables across groups. It is tested by constraining the sub-diagonal elements of the phi-matrices $\phi_{jk}^A = \phi_{jk}^B = ...\phi_{jk}^G$, to be equal. Covariances among constructs have implications for the constructs' meaning or validity (Cronbach and Meehl, 1955). Hence, unequal covariances raise concerns about the equality of construct meanings (Cole and Maxwell, 1985).

As far as the analyses of **invariance of the latent means** are concerned, they are conducted in order to test for differences between groups (or points of time) in their latent means. In contrast, traditional approaches to the analysis of mean differences use composite *manifest* scores and employ *t* tests, ANOVA, or MANOVA. The validity of testing group differences in manifest scores depends on whether the assumptions that underlie such comparisons are correct, specifically, that both the factor loadings and the item intercepts are equal (i.e., metric and scalar invariance) (Tarka, 2011). The relationship between a latent and observed mean or an expected observed value can be written as follows:

$$E\left(x_i^g\right) = \tau_i^g + \lambda_i^g \kappa_i^g. \tag{1}$$

where:

$E\left(x_i^g\right)$ - expected value of the *i*-th manifest indicator in group *g*,

κ_i^g - is the mean of factor f in group *g* to be considered in the tests related with latent means comparison of the particular groups.

It shows that a manifest mean depends not only on its latent mean but also on the factor loading and the item intercept. Thus, a manifest mean difference can be caused either by a latent mean difference or a difference in the loadings, intercepts, or both. Therefore, a test of latent mean difference requires the equality of both the factor loadings and item intercepts. The equality of the latent means is tested by constraining the kappa matrices $\kappa^A = \kappa^B = ...\kappa^G$, to be equal across groups.

Finally, a **measurement of errors invariance** concerns the hypothesis that the measurement error in the manifest indicators $\Theta^A = \Theta^B = ...\Theta^G$, is the same across groups.

4. Factor analysis model for two populations

In factor analysis model we consider a set of *m* populations $\Pi_1, \Pi_2, ..., \Pi_m$. They may be different nations, or culturally different groups, groups of individuals selected on the basis of some known or unknown selection variable, groups receiving different treatments, etc. In fact, they may be any set of exclusive groups of individuals that are clearly defined. And it is assumed that a battery of tests has been administered to a sample of individuals from each population. The battery of tests need not be the same for each group, nor need the number of tests be the same. However, since we shall be concerned with characteristics of the tests that are invariant over populations, it is necessary that some of the tests in each battery are the same or at least content-wise equivalent. A general factor analysis model in each population will be as follows (Jöreskog, 1971):

$$x_g = \mu_g + \Lambda_g f_g + e_g, \tag{2}$$

where:

x_g is random vector with mean vector μ_g (and variance-covariance matrix of population Σ_g) in group g. As a result x_g is explained by k_g common factors f_g and unique factors e_g. Furthermore, we assume that $\varepsilon\left(f_g\right) = 0$ and $\varepsilon\left(e_g\right) = 0$ and so the same with Λ_g a factor pattern. And this implies factor analytic solution as follows (Jöreskog, 1971):

$$\Sigma_g = \Lambda_g \Phi_g \Lambda'_g + \Psi_g, \tag{3}$$

where:

Φ_g - variance – covariance matrix of f_g

Ψ_g is the diagonal variance – covariance matrix of e_g.

In contrast to Jöreskog general model, Lawley and Maxwell proposed separate models for strictly two populations with variance-covariance matrices denoted as Φ_1 and Φ_2. The coefficients of factor loadings, - if invariant under the changes of populations – will cause loading matrix Λ the same for both populations. They also assumed that Ψ diagonal matrix of e, will be the same. The model can be generalized to some extent by allowing populations to have different unique factors (residual variances) on variance-covariance matrices Ψ_1 and Ψ_2, but this option complicates subsequent estimation procedures. The population variance-covariance matrices for the given x_i are thus given as follows (Lawley and Maxwell, 1963):

$$\Sigma_1 = \Lambda\Phi_1\Lambda' + \Psi, \tag{4}$$

$$\Sigma_2 = \Lambda\Phi_2\Lambda' + \Psi. \tag{5}$$

Such being the case, certain loadings are a priori zero, and that number of and positions of these are such as to determine factors uniquely. The factors are arbitrary and for computational convenience, they can be chosen in such a way that the matrix will be (Lawley and Maxwell, 1963):

$$\Phi = \frac{\left(n_1\Phi_1 + n_2\Phi_2\right)}{\left(n_1 + n_2\right)}. \tag{6}$$

and has unit diagonal elements. As a result there are k factors, where certain specified elements of the loading matrix Λ are zero and that the population variance-covariance matrices satisfy assumptions of the Eq. (4) and (5).

5. From the model of two populations towards the model of two samples

For general model, we have N_g respondents in the sample from g-th population, $\bar{x}_g$ as the usual sample mean vector and S_g - sample variance-covariance matrix with $n_g = N_g - 1$ degrees of freedom. Thus, we obtain independent measurements for different groups.

We may thus assume that S_1 and S_2 are the separate variance-covariance matrices for e.g., two groups with respectively n_1 and n_2 degrees of freedom, obtained by taking a random sample from each population. Then, general log-likelihood function for S_g sample will be:

$$\log L_g = -\frac{1}{2} n_g \left\{ \log_e \left|\Sigma_g\right| + \operatorname{tr}\left(S_g \Sigma_g^{-1}\right) \right\}. \tag{7}$$

So, if the samples are independent, the log-likelihood for all the samples is:

$$\log L_g = \sum_{g=1}^{m} \log L_g. \tag{8}$$

And the log-likelihood function for two separate groups will be (without function with observations) (Lawley and Maxwell, 1963):

$$-\frac{1}{2} n_1 \left\{ \log_e \left|\Sigma_1\right| + \operatorname{tr}\left(S_1 \Sigma_1^{-1}\right) \right\} - \frac{1}{2} n_2 \left\{ \log_e \left|\Sigma_2\right| + \operatorname{tr}\left(S_2 \Sigma_2^{-1}\right) \right\}. \tag{9}$$

To estimate unknown parameters we should have maximized it with respect that non-zero elements of Λ, the elements of Ψ, and the elements of Φ_1 and Φ_2 are subject to restriction that Φ has diagonal elements. The resulting equations of estimations may be simplified and solved iteratively. The hypothesis will be tested by means of the criterion (Lawley and Maxwell, 1963):

$$n_1 \log_e \left(\frac{\left|\hat{\Sigma}_1\right|}{\left|S_1\right|} \right) + n_2 \log_e \left(\frac{\left|\hat{\Sigma}_2\right|}{\left|S_2\right|} \right). \tag{10}$$

which for large samples is distributed approximately χ^2 with $\left(p^2 - k^2 - m\right)$ degrees of freedom, where *m* is the number of non-zero loadings.

If we want to administer the same test/measurement within different populations, we must follow conditions of invariance as previously discussed. In particular we need to consider invariance of:

- Λ in factorial pattern over populations,
- $\psi_1^2 = \psi_2^2$ of with variances of regression.

Then, we identify parameters where Λ in $\Sigma_g = \Lambda \Phi_g \Lambda' + \psi_g^2$, will be replaced by $\Phi_g^* = T \Phi_g T'$, $g = 1, 2, ..., m$, and where T is an arbitrary non-singular matrix of order $k \times k$. Then, each Σ_g remains the same so that the function F is unaltered.

$$F = \frac{1}{2} \sum_{g=1}^{m} n_g \left[\log\left|\Sigma_g\right| + \mathrm{tr}\left(S_g \Sigma_g^{-1}\right) - \log\left|S_2\right| - p_g \right]. \tag{11}$$

Since the matrix T has k independent elements, this means that at least k^2 independent conditions must be imposed on $\Lambda, \Phi_2, \Phi_2, ..., \Phi_m$ to make them uniquely defined. And the most convenient way of doing this is to let all the Φ_g, be free and to fix one non-zero element and at least $k-1$ zeros in each column of Λ. In an exploratory study one can fix exactly $k-1$ zeros in almost arbitrary positions. Jöreskog (1971) claims that one may choose zero loadings where one thinks there should be "small" loadings in the factor pattern. The resulting solution may be rotated further, if desired, to facilitate better interpretation. In a confirmatory study, on the other hand, the positions of the fixed zeros, which often exceed $k-1$ in each column, are given *a priori* by a hypothesis and the resulting solution cannot be rotated without destroying the fixed zeros.

In order to make observable variables comparable, according to different units of measurement in different samples, one can rescale these variables before beginning the factor analysis. As a result we assume (Jöreskog, 1971):

$$S = \left(\frac{1}{n}\right) \sum_{g=1}^{m} n_g S_g, \tag{12}$$

where: $n = \sum_{g=1}^{m} n_g$, and

$$D = \left(\mathrm{diag}\ \hat{\Phi}\right)^{-\frac{1}{2}}. \tag{13}$$

Then, the variance-covariance for the rescaled variables is:

$$S_g^* = DS_g D. \tag{14}$$

The weighted average of S_g^* is a correlation matrix. The advantage of rescaling is that, when combined with an option of rescaling the factors, factor loadings are of the same order of magnitude as usual when correlation matrices are analyzed and when factors are standardized to unit variances. This makes it easier to choose starting values for the minimization and interpretation of the results. It should be indicated further that it is not permissible to standardize the variables in each group and to analyze the correlation matrices instead of the variance-covariance matrices. This violates the likelihood function (7-8) which is based on the distribution of the observed variances and covariances. Invariance of factor patterns is expected to hold only when the standardization of both tests and factors are relaxed.

6. Example: values system analysis in Polish and Dutch youth

We drew basic ideas and developed our research on Rokeach (1973) and Schwartz (1992) definition of values, describing them as *"desirable, transsituational goals, varying in importance, that serve as guiding principles in the life of a person or other social entity"*. As a result values are driven by different motivations (Schwartz and Sagiv, 1995) (Table 2).

The theory postulates 10 different types of values and two main value dimensions. The 10 types of values are arranged in a circumplex structure around the following dimensions: *self-transcendence* versus *self-enhancement* and *openness to change* versus *conservation.* Figure 1 displays the circular structure of the types of values as well as the two dimensions behind them (Schwartz and Boehnke, 2004; Schwartz, 2005).

The dimension of *self-transcendence/self-enhancement* describes the possible conflict between the acceptance of others as equal entities and the concern for their well-being (types of values: universalism and benevolence) versus the tendency to try to achieve personal success as well as predominance over others (types of values: power and achievement). The second dimension reflects the possible conflict between independent thought and action and preference for an exciting life (types of values: self-direction and stimulation) versus the tendency to seek stability, security, and attachment to customs, traditions, and conventions (types of values: security, conformity, and tradition). Virtually different types of values correlate differently. And the value type related to *hedonism*, forms a link between *openness to change* and *self-enhancement* (Tarka, 2010).

Table 2. The 10 types of values with motivational goals and the higher-order dimensions

Value	Motivation	Dimension
Self-direction	Independent thought and action-choosing, creating, exploring	Openness to change
Stimulation	Excitement, novelty, and challenge in life	Openness to change
Hedonism	**Pleasure and sensuous gratification for oneself**	**Between self-enhancement and openness to change**
Achievement	Personal success through demonstrating competence according to social standards	Self-enhancement
Power	Social status and prestige, control and dominance over people and resources	Self-enhancement
Security	Safety, harmony, and stability of society, of relationships, and of self	Conservation
Conformity	Restraint of actions, inclinations, and impulses likely to upset or harm others and violate social expectations or norms	Conservation
Tradition	Respect, commitment, and acceptance of the customs and ideas that traditional culture or religion provide the self	Conservation
Benevolence	Preservation and enhancement of the welfare of people with whom one is in frequent personal contact	Self-transcendence
Universalism	Understanding, appreciation, tolerance, and protection for the welfare of all people and for nature	Self-transcendence

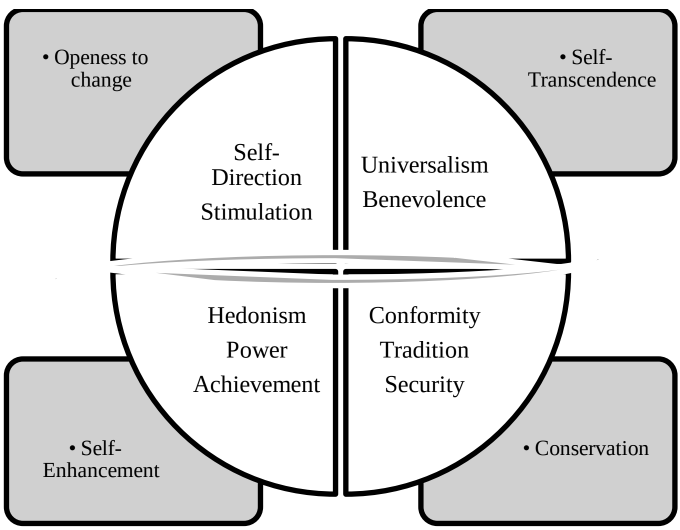
Source: Sagiv and Schwartz, 1995.

Method of data collection

Initially the 19-item question battery was applied in the study to measure value priorities among Polish and Dutch youth representatives in the academic environment. The interviewee was confronted with a five-point Likert scale (where: 1 = totally disagree; 5 = totally agree). This type of scale is a parallel in which each item represents an alternative and equivalent tool for measuring a latent trait. Evaluation of reliability and measurement invariance in scales of this type is made in context of classical theory of the test.

Data collection was based on paper and pencil interviews. In the course of empirical research, printed questionnaires had been handed out to a number of individuals (respondents) at Universities in Poland (in Poznan Universities e.g., Poznan University of Economics, Adam Mickiewicz University of Poznan and Poznan University of Technology) and Universities in the Netherlands in Rotterdam and Tilburg for final evaluation of the prepared sets of items. Sampling frame was derived and prepared according to internal universities database including complete list of participants attending three introductory classes of undergraduate level. Respondents were selected on the rules of simple random sampling. Only a small percentage (less than 5%) of those contacted refused to participate in a study. Thus a collected sample was n = 285. The data was collected between May and June 2009.

Figure 1. Dimensions of value systems



Source: own construction

<u>Data analysis and empirical results</u>

At first the factorial structure was tested and then measurement invariance of the instruments that were operationalized according to the value theory. Sometimes one may also apply the other model for comparison of measurements at different points in time. Such being the case, growth curve models of latent dimension are used (e.g. latent growth curve models). But the latter option was beyond objective of this article.

The assumptions for the assessment of invariance were as follows:

- scale consisted of multiple items,

- items were of reflective form (they reflected latent dimension otherwise latent variable),
- the measurement was performed in two groups at one time.

All analyses were conducted using the computer program LISREL, where a maximum likelihood was applied as the estimation method.

Single group analysis

At first we measured directly the higher-order dimensions of the values by their corresponding items. The two higher dimensions self-transcendence/self-enhancement and openness to change/conservation constituted four factors. The remaining 19 items were attributed to these four factors.

The models required several modifications. At first, items that did not achieve adequate factor loadings were eliminated. The criterion we set for an item to load on a factor was 0.49 and higher. Some loadings were too low for the conservation, self-transcendence, and self-enhancement factors. As the invariance test should be performed on the same measurement model, we eliminated the same items in both samples. Consequently, the final model that we tested for invariance included 15 items.

Multi-group analysis

Next we turned to multiple-group comparison. This model enabled us to test to what extent the value measurements were invariant across the samples. To test it we used the model that included 4 constructs, 15 items, 1 cross-loading and 2 errors. We compared two groups (e.g. Polish and Dutch nationalities). The empirical covariance matrix of the items for each group served as the input. Variance-covariance matrix allowed for comparing outcomes in terms of intergroup value of the original units of measurement. If the variance-covariance matrices are not significantly different in both groups, one can perform further analysis of individual aspects of measurement invariance.

The evaluation of the **structural invariance** of latent variables (also called configural invariance) was conducted in both groups. This reflected a test of the hypothesis of equality variance-covariance matrix based on the degree of goodness of fit of structural independent models made on the basis of data from individual cultures. Good models and their fit to data proved the existence of a configuration invariance and enabled us further comparison between the constructs. The degree of fit was tested using the statistics such as χ^2 index, CFI Bentler, PCLOSE probability of close fit, coefficient RMSEA. Values close to .95 for CFI and below.06 for RMSEA suggest a good fit (Hu and Bentler, 1993).

Next, we assessed **metric invariance,** e.g. the factor loadings of all items that were constrained to be identical across two groups. This assessment was based on a comparison of relative fitting between two structural models. In the first model, corresponding factor loadings were set as equal in all groups (factor loading λ_1 in the first group was equal to the value of the factor loading λ_1 in the second group of respondents). In the second model, factor loadings in both groups were the free parameters. If the fit of the model with defined (fixed) factor loadings was not significantly worse, as compared to model with free – released loadings, then items would measure the latent variables (factors) in a comparable way in both analyzed groups. However, if the degree of data fit in a model, pertaining to fixed factor loadings was significantly worse, then comparison of factor loadings between groups could be only made on partial invariance measurement between the groups.

Table 3. Fit measures for the model assessing configural, metric, and scalar invariance

	Configural invariance	**Metric invariance**	**Scalar invariance**
Chi-square	179,56	192,10	205,60
CFI – comparative fit index	0,935	0,940	0,948
RMSEA – root mean square error of approximation	0,046	0,049	0,040
PCLOSE – probability of close fit	0,495	0,515	0,540
SRMR – standardized root mean square of residuals	0,081	0,082	0,086

Source: own calculation in LISREL.

Next, we turned to the test of ***scalar invariance.*** It allowed us to compare the mean values for the latent variables, especially to detect: 1). inter-group differences in the responses (according to particular statements which determined latent dimensions) and also 2). effects of respondents attitudes and differences in their style of giving responses to these statements.

The global fit measures of the configural invariance model, which are displayed in Table 3, suggest that the model should not be rejected. The results indicate that the metric invariance model is supported by the data. A chi-square difference test between the configural and the metric invariance model revealed that there was no significant difference in the model fit. Also, the fit indices of CFI, RMSEA, and SRMR are further indications for invariance. In case of scalar invariance, we may observe that the constrained intercepts of the items are equal across the samples. As the results we cannot reject the scalar invariance model.

As configural, metric, and scalar invariance has been confirmed, the comparison of latent mean values between the two samples was easy to conduct.

And because one intended to compare the ***latent means* in both groups,** therefore we added a vector of manifest means as input. With regard to the parameter matrices, the τ_x -vector and the κ -vector were added. The results are presented in Table 4.

Table 4. Latent mean differences of four constructs (reference group: Polish sample)

	Means for Polish group	**Means for Dutch group**	**Effect sizes (r)**
Openness to change	3,95	3,24	,28∗
Self-enhancement	4,87	4,10	,31∗
Self-transcendence	2,67	2,65	,00
Conservation	3,16	2,53	,12*

Note: Effect sizes of r in the latent means at $*\ p < 0{,}5$*; Source: own calculation in LISREL.*

Results show significant mean differences for the constructs *openness to change* (estimate = 0,28), *self-enhancement* (estimate = 0,31), and *conservation* (estimate = 0,12). For the construct *self-transcendence* we found no significant mean difference between groups (estimate = 0,0). As a result differences have been found for the latent means of both samples for the constructs ***openness to change, self-enhancement***, the hypothesis that the latent means for value questions were identical in both groups must be rejected. Individuals in the Dutch sample displayed lower levels of ***openness to change*** and ***self enhancement*** (which were also in their own part of hedonic senseless and excessive consumption of the market goods) as compared to Polish sample.

From the above results and application of model it is quite interesting to infer that Polish youth as compared to Dutch youth (being derived from agglomerations such as: Poznan, Rotterdam and Tilburg) exposes more interest towards types of values such as Hedonism in general. Apparently young Poles look now for more pleasure and enjoyable life (also pertaining to products and services consumption) than their foreign colleagues from already developed countries. Events from the past and hard rules of socialism and limitation in access for years to free market goods left their strong impact on young people’s life and behaviour. Being kept too long away from open market sources, citizens of eastern block of Europe, e.g. Poland, seem to recoup their delays and catch up with latest trends arising on the market. In contrast, Dutch youth, being too long exposed to wide markets, virtually grew accustomed to its products and services. As a result this situation affected their life style, lowering also their interests in Hedonism that is senseless and excessive consumption of the market goods. And these facts simply reveal a new perspective for companies business activities that is either to point on new directions associated with entry on new ascending markets.

7. Conclusions

Discussed in article a model of factor analysis model was strongly based on the examination of measurement invariance and specifically, factor invariance. Researcher when using such a type of model avails of the opportunity to detect invariance for tested items and simultaneously generate reliable and valid constructs. If these assumptions are not satisfactory then making further inferences becomes pointless. In a consequence the model requires certain parameters (e.g., factor loadings) to be constrained in the process of identification, which means they need to be invariant across groups, and act as referent variables. If this invariance assumption for some reason would be violated, then location of the parameters that actually differ across groups would become difficult. In case of the conducted analysis and implemented model, it simply turned out to be a satisfactory solution regarding the researched problem and final calculated scores.

REFERENCES

BILLIET, J. (2002), *Cross-cultural equivalence with structural equation modeling*, [in:] Mohler, P.P. (Ed.) *Cross-cultural survey methods*, New Jersey: John Wiley & Sons Inc., pp. 247 - 264.

BOLLEN, K.A., (1989), *Structural equations with latent variables*, New York: Wiley.

COLE, D.A., MAXWELL, S.E. (1985), *Multitrait-multimethod comparisons across populations: a confirmatory factor analytic approach.* Multivariate Behav. Res. 20, pp. 389 – 417.

CRONBACH, L.J., MEEHL, P.E. (1955), *Construct validity in psychological tests,* Psychol. Bull. 52, pp. 281 – 302.

HU, L.-T., BENTLER, P.M., (1999), *Cut-off criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives.* Structural Equation Model. 6, pp. 1–55.

JÖRESKOG, K.G. (1971), *Simultaneous factor analysis in several populations,* Psychometrika, 36, pp. 409 – 426.

LAWLEY, D.N., MAXWELL, A.E. (1963), *Factor analysis as a statistical method.* London: Butterworths.

LITTLE, T.D., SLEGERS, D.W., CARD, N.A. (2006), *A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models,* Structural Equation Model, 13(1), pp. 59 – 72.

MEREDITH, W. (1993), *Measurement invariance, factor analysis and factorial invariance*, Psychometrika, Vol. 58, pp. 525 - 543.

MUTHÉN, B., (1989) *Latent variable modeling in heterogeneous populations,* Psychometrika, 54(4), pp. 557 – 585.

PLOYHARDT, R.E., OSWALD, F.L. (2004), *Applications of mean and covariance structure analysis: integrating correlational and experimental approaches,* Organ. Res. Methods, 7(1), pp. 27 – 65.

ROKEACH, M., (1973), *The Nature of Human Values,* New York: Wiley.

SAGAN, A. (2005), *Analysis of equivalence off measurement scales in inter-cultural studies,* Scientific Papers, Vol. 659, pp. 59 – 73.

SCHWARTZ, S.H. (1992), *Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries,* [in:] Zanna, M. (Ed.), *Advances in experimental social psychology,* vol. 25, Orlando: Academic Press, pp. 1 – 65.

SCHWARTZ, S., SAGIV, L. (1995), *Identifying culture-specifics in the content and structure of values,* J. Cross-Cult. Psychol. 26(1), pp. 92 – 116.

SCHWARTZ, S.H., BOEHNKE, K. (2004), *Evaluating the structure of human values with confirmatory factor analysis,* J. Res. Pers. 38(3), pp. 230 – 255.

SCHWARTZ, S.H. (2005), *Basic human values: their content and structure across countries,* [in:] Tamayo, A., Porto, J.B. (Eds.), *Values and Behavior in Organizations,*. Vozes, Petrópolis, pp. 21 – 55.

TARKA, P. (2010), *Measurement scales for customers hedonic values – comparison of reliability techniques* – Scientific papers “Econometrics” UE Wrocław, 29, pp. 23 – 38.

TARKA, P. (2010), *Latent variable models - issues on measurement and finding exact constructs in customers values* – Polish Statistical Review (Przegląd Statystyczny), 4, pp. 142 – 167.

TARKA, P. (2011), *Statistical analysis of youth’s value systems in Poland and Netherlands – an approach to LOV and RVS scale* [in:] Józef Pociecha (Ed.) *„Methods of data analysis”*, Academic papers UE Kraków, 2011, pp. 86 – 94.

THOMPSON, M.S., GREEN, S.B. (2006), *Evaluating between-group differences in latent means.* [in:] Hancock, G.R., Mueller, R.O. (Eds.) *Structural Equation Modeling: A Second Course,* Greenwich: Information Age, pp. 119 - 169.

VANDENBERG, R.J., LANCE, C.E., (2000), *A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research,* “Organ. Res. Methods” 3(1), pp. 4 – 69.