Working Paper: Are values in the Benelux countries comparable? Testing for equivalence with the European Social Survey 2004-5

Eldad Davidov – GESIS, Central Archive for empirical Social Research, University of Köln

Peter Schmidt – Institute for Political Science, University of Gießen

Eldad Davidov

Email address: davidov@za.uni-koeln.de

Tel (office): +49 221 4769457

Are values in the Benelux countries comparable? Testing for equivalence with the European Social Survey 2004-5

1. Introduction

Benelux is an economic union in Europe comprising three neighbouring countries, **Bel**gium, the **Net**herlands, and **Lux**embourg. It was created for the Benelux Customs Union, and was replaced by an economic union, which is still existent. Dutch and French are the commonly used languages in these countries. Out of the total of 27.2 million people living in the member states 4.7 million (17 percent) speak French and 22.5 million (83 percent) speak Dutch. Although it is not an official language of the Benelux, Belgium and Luxembourg recognise German as one of the official national languages. This union, its economic activities and the common languages reflect a shared cultural, economic, social and political background. But do these countries also share common values?

In the following chapter we are going to apply one of the currently mostly used value theory in the social sciences, the human basic values theory (Schwartz, 1992), in order to try to answer this question. More precisely, we are going to try to answer the following questions: (1) Do values as proposed by Schwartz (1992) and as measured in the European Social Survey (ESS) in the second round (2004-2005) exhibit the necessary statistical conditions of invariance, which allow comparing them across the three nations?; (2) If the answer to this question is positive, to what extent does the importance attributed to these values differ across the three countries?[1] The second round of the ESS provides us with an opportunity to conduct this test across the three countries. The cross-country comparability of the value measurements in this round has not been evaluated yet. Davidov, Schmidt and Schwartz (in press) conducted an invariance test of the values scale across 20 countries with the first round of the ESS, but it did not include Luxemburg, as Luxemburg did not collect data on values on that round[2]. We are going to start with a short presentation of the theory. Then we will provide a description of the items used in the European Social Survey to measure these values. We will proceed with statistical analyses of invariance, which is a necessary condition for comparability of constructs across cultures or groups. Finding out invariance of some values will allow us to compare only these values across the three

---

[1] Practically, in most survey analyses variable means are compared across groups such as countries without testing for the necessary conditions, which allow such a comparison. We will discuss later why such comparisons are not legitimate and may be meaningless.

[2] Mean comparison of the values collected in the European Values Study (EVS) are reported for several countries including the three Benelux countries (see for example Halman, 2001), however, without testing for the necessary statistical conditions for such a comparison.

countries. After providing results of our analyses we will finalize with some concluding remarks.

2. The theory of human basic values

Schwartz proposed in 1992 a theory which specifies value types and their content, and builds on earlier approaches (Rokeach, 1973). The theory provides a rationale for the value contents and their relationships. According to this theory, values are 'desirable, trans-situational goals, varying in importance, that serve as guiding principles in people's lives'. The theory proposes ten types of values, although it leaves their width and coverage quite open. The value types derive from three universal needs of individuals: needs as biological organisms, needs for a coordinated social interaction, and survival and welfare needs of groups (Schwartz 1992, 2005a). For example, universalism derives from the need for a coordinated social interaction. Hedonism derives from needs of a biological organism. The ten value types are: hedonism, stimulation, self-direction, universalism, benevolence, conformity, tradition, security, power and achievement. Their motivations are presented in Table 1.

The theory proposes also a structure, which reflects relationships between the values. The principal underlying this structure is that pursuing one value may be in conflict with pursuing another one. For example, pursuing conformism may be in conflict with pursuing self-direction. Pursuing universalism may be in conflict with pursuing tradition. However, pursuing tradition and conformity may be compatible. Also seeking adventures and stimulation may be compatible with hedonism. Differences in the hierarchies of values and in their relationships may distinguish between groups, societies, cultures and nations.

The structure of relations and conflicts between the values is presented in Figure 1. The theory distinguishes between four higher dimensions around a circle: self-enhancement, self-transcendence, openness to change and conservation. Self-enhancement, which reflects power and achievement values opposes self-transcendence, which reflects self-direction and stimulation values. Openness to change that represents universalism and benevolence values opposes the conservation dimension, which represents security, conformity and tradition values. This model has been tested and validated in hundreds of samples (Davidov, Schmidt and Schwartz, in press; Fontaine and Schwartz 1996; Schwartz 2005a,b; Schwartz and Boehnke 2004).

3. The measurement of values in the European Social Survey

The most widely used instrument to measure basic values is the Schwartz Value Survey (SVS). Since this instrument is intellectually demanding Schwartz developed a new

instrument, called Portrait Value Questionnaire (PVQ) (see Schwartz, Lehmann and Roccas 1999; Schwartz et al. 2001). The PVQ includes 40 verbal portraits of different people. Each portrait describes aspirations or goals of a person that point to the importance of a certain value. The European Social Survey scale to measure values derived from this PVQ scale, but was reduced to only 21 portraits, gender-matched, because of budgetary constraints. These portraits were reformulated so as to cover as wide a spectrum of each value as possible. For example, 'He thinks it is important that every person in the world be treated equally. He believes everyone should have equal opportunities in life' describes a person, for whom universalism values are important. 'It is important to him to be humble and modest. He tries not to draw attention to himself' describes a person, for whom tradition is important. Respondents have to indicate to what extent they resemble or do not resemble this person on a 1-6 scale. Table 2 summarizes the scale format, the question formulations, the item names, and their means and standard deviations in each country. Two portraits measure each value, and three measure universalism because of its broad content.

Before data collection, translation into different languages followed the demanding procedure set out in Harkness, Van de Vijver and Mohler (2003, ch. 3). Strict probability samples of the non-instituionalized population of 15 years and older completed the survey and the value questions. 1778 completed the questionnaire in Belgium, 1635 in Luxemburg and 1881 in the Netherlands. The data were taken from website http://ess.nsd.uib.no. Before proceeding with data analysis and before comparing the value priorities in the three countries we have to guarantee that values are invariant across countries. In the next section we are going to provide an explanation of the invariance test procedure, and then we will conduct it.

4. Testing invariance

In the last years there have been an increasing number of studies, which have dealt theoretically and empirically with the issue of statistical testing of invariance (Billiet 2003; Cheung and Rensvold 2000, 2002; Harkness et al. 2003; Horn and McArdle 1992; Hui and Triandis 1985; Steenkamp and Baumgartner 1998; Vandenberg 2002; Vandenberg and Lance 2000; Welkenhuysen-Gybels & Billiet 2002). The meaning of invariance is whether under different conditions of observing a phenomena (different countries, cultures, societal groups or even different time points), operationalizations yield the same measures of the same attribute. If one does not test for invariance, comparisons of a construct and its associations are problematic. Finding no difference would not necessarily mean that there are no 'real'

differences. On the other hand, finding differences may be a result of systematic response biases or different understanding of the question, rather than substantial differences.

Several techniques have been developed to test for invariance (for an overview see De Beuckelaer 2005). Probably the most powerful one is the multiple-group confirmatory factor analysis (MGCFA) approach (Jöreskog 1971**)**, and we will utilize it in the following analyses.

There are several levels of invariance, and we are going to focus on the lowest three levels, which are necessary for mean comparison (Steenkamp and Baumgartner 1998). The lowest level is configural invariance. It requires that items measuring certain constructs would be the same in the different groups. In other words, the same items should measure the same constructs in each group. Configural invariance is supported, if the model fits well to the data, there are no large modification indices[3] suggested by the program, all factor loadings are substantial and significant, and the correlations between the constructs are less than one.

However, configural invariance does not guarantee that people understand the questions in the same way in the various countries. The next level of invariance is measurement (or metric) invariance, and is a necessary condition for equivalence of meaning of the constructs. It requires that the factor loadings are the same across groups or countries. It is tested by restricting the factor loadings to be the same across nations. Measurement invariance is supported if the model fit to the data is satisfactory and does not significantly decrease, and if there are no large and substantial modification indices suggested by the program in the MGCFA.

A higher level of invariance, scalar invariance, is needed to justify comparison of construct means. Scalar invariance requires information not only on relationships (covariances) between items but also on their means, and implies that mean differences between constructs are a result of substantive differences. It is tested by restricting the intercepts of the corresponding items in each country to be equal. Scalar invariance is supported if the model fit to the data is satisfactory and does not significantly decrease, and if there are no large and substantial modification indices suggested by the program in the MGCFA. In sum, three levels of invariance are needed for mean comparison: configural, measurement and scalar. Only if the three of them are supported by the data, can we assume that constructs are equivalent and carry out mean comparison.

---

[3] A modification index suggests misspecifications in the model constraints. Large modification indices imply that releasing these constraints will produce a substantial change in parameters and an improvement in the model fit.

Testing invariance may start from the highest level, scalar invariance. This strategy is often called a 'top-down' strategy. Alternatively, it can begin with the lowest level, configural invariance. This is often called the 'bottom-up' strategy. We chose the latter in order to test whether even lower levels of invariance are not supported by the data. In the next section we begin by testing for configural invariance and proceed with testing higher levels of invariance.

5. Data analysis

Before running a MGCFA we started with single-country confirmatory factor analyses. Byrne (2001, pp. 175-176) notes the importance of analysing each country separately before analysing them all together. We used the program Amos 6.0 (Arbuckle 2005), and as input covariance matrices produced from the raw data of each country.

Analyses revealed that all factor loading were significant and substantial. However, 2 pairs of constructs correlated very highly in each country: power with achievement and conformity with tradition. This produced the problem of non-positive definite matrix of the constructs[4]. In order to solve it we had to unify each pair. As a result, we could identify eight values in each country: Hedonism, stimulation, self-direction, universalism, benevolence, security and the unified constructs conformity-tradition and power-achievement. Table 3 reports the number of distinct values found and lists the correlation between the pairs of constructs in the baseline model prior to being unified. The two pairs are adjacent in the value circle (Figure 1). From this point of view they do not falsify the theory, but only suggest that probably the number of value items in the ESS is too small in order to account for all the ten constructs.

The model in each country showed a good fit to the data, but also suggested a few modification indices. In Belgium, the program required another path from the unified construct conformity-tradition to the item measuring importance to get respect from others. In Luxemburg and the Netherlands the program required an additional path from universalism to the item measuring importance to do different things. Modifying the three models according to the proposed changes produced an acceptable fit to the data. The modifications and the global fit measures in each of the single-country CFA's are presented in Table 4.

Next we tested for configural invariance. We conducted a MGCFA for the three countries simultaneously starting with the basic model with ten values. Once again, we had the problem of non-positive definite matrix of the constructs. This time, three pairs of

---

[4] In non-technical terms this problem means that the two constructs are linearly dependent; the two constructs are associated with each other so strongly that they cannot be separated.

constructs correlated too highly: power with achievement (.90), conformity with tradition (.94), and universalism with benevolence (.83). In order to solve the problem they were unified. This resulted in a model comprising of seven values: hedonism, stimulation, self-direction, security and the unified constructs power-achievement, universalism-benevolence and conformity-tradition. The three pairs represent values which are adjacent in the value circle presented in Figure 1. From this point of view, unifying them does not falsify the postulated value model. However, they suggest once again that there are probably not enough value items in the ESS in order to measure ten values.

This model with seven values showed an acceptable fit to the data, but suggested some large modification indices. We continued to improve it by introducing new paths according to the proposed modifications. First, we added a path from the unified construct conformity-tradition to the item measuring importance to get respect from others. Second, we introduced an additional item from the unified construct universalism-benevolence to the item measuring importance to do different things. The model with seven values and the two additional paths is presented in Figure 2. As a result, this model indicated a good fit to the data (Pclose=1.0, RMSEA=.03, chi-square/degrees of freedom=2.8, AIC=1899, BCC=1913). In other words, the model and its specifications are invariant across the three countries. The unstandardized and standardized factor loadings are reported in the Appendix.

Configural invariance does not guarantee that people understand the questions similarly. Next we tested for measurement (metric) invariance. Finding measurement invariance will guarantee that the meaning of the constructs was equivalent across countries. We tested it by constraining the factor loadings of each construct to be the same across nations. Table 5 presents the fit measures of the invariant models. According to the global fit measures we cannot reject the measurement invariant model (Pclose=1.0, RMSEA=.03, chi-square/degrees of freedom=2.8, AIC=1933, BCC=1944). The increase in chi-square (by 98 per 32 more degrees of freedom) is significant, but not large. Even the AIC and the BCC criteria, often used for model comparison, do not increase substantially. Therefore we conclude that items load equally on their corresponding constructs in the three countries. People in the three nations seem to understand the questions in a similar way.

In the next step we tested for scalar invariance, a necessary condition for comparing latent means across nations. To test it we constrained the intercepts of corresponding indicators to be equal. The fit indices suggested that one should reject this model (Pclose=1.0, RMSEA=.03, chi-square/degrees of freedom=3.9, AIC=2625, BCC=2634). Although Pclose

and RMSEA are acceptable, the increase in chi square is significant and substantial (by 811 per 42 more degrees of freedom). As Table 5 shows, also the AIC and BCC increased substantially compared with the measurement invariance model. Therefore we conclude that the full set of seven values may not be compared across countries. However, it could still be the case, that the scale of a few values meets the requirements of scalar invariance. After looking at the modification indices we realised, that misspecification in the equality constraints on the intercepts were highest for the values hedonism, security and the unified values power-achievement and conformity-tradition. Releasing these constraints resulted in an acceptable model fit (Pclose=1.0, RMSEA=.03, chi-square/degrees of freedom=2.9, AIC=1987, BCC=1998). The increase in chi square compared with the measurement invariant model was significant (by 78 per 12 more degrees of freedom) but not as substantial anymore. Also the increase in the AIC and BCC criteria was not as high. Consequently, we concluded that whereas the scale of hedonism, security and the unified values power-achievement and conformity-tradition does not meet the requirements of scalar invariance, the scale of the values self-direction, stimulation and the unified construct universalism-benevolence does. This allows mean comparison of those values across the three countries.

Sörbom (1978) has shown that in order to be able to compare means across groups, mean levels have to be constrained to zero in one group in order to get the model identified. This is the mostly used approach, and we apply it here. Therefore we constrained mean values to be zero in one reference country (Belgium). Mean differences between the countries will be reflected in the mean values in Luxemburg and the Netherlands. Results are presented in Table 6. It turns out that there are no differences in the mean levels of the values between Belgium and Luxemburg. There is also no significant difference in the mean level of the self-direction value across the three countries. However, stimulation is slightly (.09) but significantly (p<0.05) higher in the Netherland. Furthermore, universalism-benevolence is significantly *less* pronounced in the Netherlands (-.12). This finding suggests that the three countries share at least to some extent common values with respect to the values that may be compared.

6. Summary and conclusions

The Benelux countries share economic, political, social and cultural background. Our goal was to test whether they also share common values. We tried to answer this question in a two-step procedure:

-Firstly, by testing whether values as proposed by Schwartz (1992) and as measured in the European Social Survey (ESS) in the second round (2004-2005) exhibit the necessary statistical conditions of invariance, which are required in order to compare them across the three nations. In other words, we tested to what extent they are equivalent;

-Secondly, by comparing values, which meet the required statistical conditions (metric and scalar invariance).

We started by single-country CFA's in Belgium, Luxemburg and the Netherlands. In these analyses we could identify eight rather than the ten theoretically postulated values. Two pairs of values, power and achievement and conformity and tradition had to be unified, because their correlation was too high. It seems that the ESS does not include enough items (21 rather than the full set of 40 PVQ items) in order to identify all ten values. In addition, following the modification indices, we introduced one additional path in each model from one construct to another item, which was originally not intended to measure it. As a result, the three models displayed a good fit to the data, and we could proceed with the MGCFA.

The MGCFA displayed configural invariance of the values across the three countries. However, we had to unify three pairs of values due to their high correlation: power and achievement, universalism and benevolence, and conformity and tradition. Furthermore, we introduced due to high modification indices two additional paths in the model from two unified constructs to items, which were not originally intended to measure them. As a result, the model displayed an acceptable fit to the data. Also the fit of the measurement (metric) invariance model was acceptable. In spite of the different languages in the three countries (French, Dutch and to some extent German), people in Belgium, Luxemburg and the Netherlands seem to understand the value items similarly. This is *not* a sufficient condition to allow value means comparison across these countries. However, it allows comparing associations of these values with other variables across these nations. In order to test whether means may also be compared we continued with the scalar invariance test. We found out that only three constructs met this test: self-direction, stimulation and the unified construct universalism-benevolence. Therefore, only these values were compared.

The comparison did not yield substantially significant differences between the countries. Whereas no significant differences at all were found between Belgium and Luxemburg, people in the Netherlands assigned a little more importance to stimulation and a little less importance to universalism and benevolence compared with Belgium and Luxemburg. It may be the case, that immigration-related social, economic and political problems that the

Netherlands has encountered in recent years have caused these lower universalism and benevolence scores, which emphasize openness to other nations and cultures.

We found out that the values hedonism, security and the unified values power-achievement and conformity-tradition do not meet the scalar invariance test and may not be compared using MGCFA among the three countries. The question, which arises, is whether there are other methods available, which could allow such a comparison. Indeed, Little, Slegers and Card (2006) proposed two other methods for mean comparison. The first is called the marker variable method and the second the effects-coding method. Especially the latter method is interesting in our case, since it does not require constraining the intercepts across groups to be equal, which was a problem in this study. However, these methods have not been tried out with real data yet, and this might be an interesting topic for future research.

7. References

Arbuckle, J. L. (2005). Amos 6.0 User's Guide. Chicago, IL: SPSS.

Billiet, J. (2003). Cross-Cultural Equivalence with Structural Equation Modeling. Pp. 247-264, in Cross-Cultural Survey Methods, edited by J. A. Harkness, F. J. R. Van de Vijver and P. Ph. Mohler. New York, NY: John Wiley.

Byrne, B. M. (2001). Structural Equation Modeling with AMOS. Basic Concepts, application and programming. London: Lawrence Erlbaum Associates.

Cheung, G. W. & Rensvold, R. B. (2000). Assessing Extreme and Acquiescence Response Sets in Cross-Cultural Research Using Structural Equations Modeling. Journal of Cross-Cultural Psychology, 31, 187-212.

Cheung, G. W. & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. Structural Equation Modeling, 9, 233-255.

Davidov, E., Schmidt, P. & Schwartz, S. (in press). Bringing values back in: The adequacy of the European Social Survey to measure values in 20 countries. Public Opinion Quarterly.

De Beuckelaer, A. (2005). Measurement Invariance Issues in International Management Research. PhD dissertation. Limburg: Limburgs Universitair Centrum.

Fontaine, J. & Schwartz, S. H. (1996). Universality and Bias in the Structure of Psychological Questionnaire Data. Paper presented at the XIII Congress of the International Association of Cross-Cultural Psychology, Montreal, Canada.

Halman, L. (2001). The European Values Study: A Third Wave. Source book of the 1999/2000 European Values Study Surveys. Tilburg: EVS,WORC.

Harkness, J. A., Van de Vijver, F. J. R. & Mohler, P. Ph. (Eds.) (2003). Cross-Cultural Survey Methods. New York, NY: John Wiley.

Horn, J. L. & McArdle, J. J. (1992). A Practical and Theoretical Guide to Measurement Invariance in Aging Research. Experimental Aging Research, 18, 117-144.

Hui, C. H. & and Triandis, H. C. (1985). Measurement in Cross-Cultural Psychology: A Review and Comparison of Strategies. Journal of Cross-Cultural Psychology, 16, 131-152.

Jöreskog, K. G. (1971). Simultaneous Factor Analysis in Several Populations. Psychometrika, 36, 409-426.

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 Modeling, 13, 59-72.

Rokeach, M. (1973). The Nature of Human Values. New York: Free press.

Schwartz, S. H. (1992). Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. Advances in Experimental Social Psychology, 25, 1-65.

Schwartz, S. H. (2005a). Basic Human Values: Their Content and Structure across Countries. Pp. 21-55, in Valores e Comportamento nas Organizações [Values and Behavior in Organizations], edited by A. Tamayo and J. B. Porto. Petrópolis, Brazil: Vozes.

Schwartz, S. H. (2005b). Robustness and Fruitfulness of a Theory of Universals in Individual Human Values. Pp. 56-95, in Valores e comportamento nas organizações [Values and behavior in organizations], edited by A. Tamayo and J. B. Porto. Petrópolis, Brazil: Vozes.

Schwartz, S. H. & Boehnke, K. (2004). Evaluating the Structure of Human Values with Confirmatory Factor Analysis. Journal of Research in Personality, 38, 230-255.

Schwartz, S. H., Lehmann, A., & Roccas, S. (1999). Multimethod probes of basic human values. In J. Adamopoulos & Y. Kashima (Eds.), Social Psychology and Culture Context: Essays in Honor of Harry C. Triandis (pp. 107-123). Newbury Park, CA: Sage.

Schwartz, S. H., Melech, G., Lehmann, 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.

Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. British Journal of Mathematical and Statistical Psychology, 27, 229-239.

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. (2002). Towards a Further Understanding of and Improvement in Measurement Invariance Methods and Procedures. Organizational Research Methods, 5, 139-158.

Vandenberg, R. J. & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices and Recommendations for Organizational Research. Organizational Research Methods, 3, 4-69.

Welkenhuysen-Gybels, J. & Billiet, J. (2002). A Comparison of Techniques for Detecting Cross-Cultural Inequivalence at the Item Level. Quality & Quantity, 36, 197-218.

Table 1. *Definitions of the Motivational Types of Values in Terms of their Core Goal*

POWER: Social status and prestige, control or dominance over people and resources

ACHIEVEMENT: Personal success through demonstrating competence according to social

    standards

HEDONISM: Pleasure and sensuous gratification for oneself

STIMULATION: Excitement, novelty, and challenge in life

SELF-DIRECTION: Independent thought and action-choosing, creating, exploring

UNIVERSALISM: Understanding, appreciation, tolerance and protection for the welfare of

    all people and for nature

BENEVOLENCE: Preservation and enhancement of the welfare of people with whom one is

    in frequent personal contact

TRADITION: Respect, commitment and acceptance of the customs and ideas that

    traditional culture or religion provide the self

CONFORMITY: Restraint of actions, inclinations, and impulses likely to upset or harm

    others and violate social expectations or norms

SECURITY: Safety, harmony and stability of society, of relationships, and of self

* From Davidov, Schmidt and Schwartz (in press).

Table 2: The ESS Human Values Scale: Format, item wording, item names, means and standard deviations (N(BE)=1,778, N(LUX)=1,635, N(NL)=1,881, total N=5,294)

Here we briefly describe some people. Please read each description and think about how much each person is or is not like you. Tick the box to the right that shows how much the person in the description is like you.

HOW MUCH LIKE YOU IS THIS PERSON?

	Very much like me	Like me	Some-what like me	A little like me	Not like me	Not like me at all
1. Thinking up….	6	5	4	3	2	1

Value	Item # (according to its order in the ESS questionnaire), item name (in parentheses) and wording (male version)	Mean (Std. deviation)		
		BE	LUX	NL
Self-Direction (SD)	1. Thinking up new ideas and being creative is important to him. He likes to do things in his own original way (ipcrtiv).	2.6 (1.2)	2.5 (1.2)	2.5 (1.2)
	11. It is important to him to make his own decisions about what he does. He likes to be free to plan and not depend on others (impfree).	2.1 (1.0)	2.3 (1.2)	2.1 (1.0)
Universalism (UN)	3. He thinks it is important that every person in the world be treated equally. He believes everyone should have equal opportunities in life (ipeqopt).	2.0 (0.9)	1.9 (0.9)	2.1 (0.9)
	8. It is important to him to listen to people who are different from him. Even when he disagrees with them, he still wants to understand them (ipudrst).	2.3 (1.0)	2.2 (1.1)	2.4 (1.0)
	19. He strongly believes that people should care for nature. Looking after the environment is important to him (impenv).	2.1 (0.9)	2.1 (1.1)	2.2 (0.9)
Benevolence (BE)	12. It's very important to him to help the people around him. He wants to care for their well-being (iphlppl).	2.2 (0.9)	2.2 (1.1)	2.3 (0.9)
	18. It is important to him to be loyal to his friends. He wants to devote himself to people close to him (iplylfr).	1.8 (0.8)	1.9 (0.9)	2.1 (0.9)
Tradition (TR)	9. It is important to him to be humble and modest. He tries not to draw attention to himself (ipmodst).	2.5 (1.2)	2.7 (1.3)	3.3 (1.3)
	20. Tradition is important to him. He tries to follow the customs handed down by his religion or his family (imptrad).	2.8 (1.4)	2.8 (1.4)	2.8 (1.3)

Table 2. (contd.)

Value	Item # (according to its order in the ESS questionnaire), item name (in parentheses) and wording (male version)	Mean (Std. deviation)		
		BE	LUX	NL
Conformity (CO)	7. He believes that people should do what they're told. He thinks people should follow rules at all times, even when no-one is watching (ipfrule).	3.3 (1.4)	3.0 (1.4)	2.8 (1.2)
	16. It is important to him always to behave properly. He wants to avoid doing anything people would say is wrong (ipbhprp).	2.4 (1.2)	2.5 (1.3)	2.7 (1.1)
Security (SEC)	5. It is important to him to live in secure surroundings. He avoids anything that might endanger his safety (impsafe).	2.4 (1.2)	2.3 (1.2)	2.7 (2.0)
	14. It is important to him that the government insures his safety against all threats. He wants the state to be strong so it can defend its citizens (ipstrgv).	2.5 (1.1)	2.3 (1.2)	2.6 (1.1)
Power (PO)	2. It is important to him to be rich. He wants to have a lot of money and expensive things (imprich).	4.3 (1.2)	4.3 (1.2)	4.4 (1.9)
	17. It is important to him to get respect from others. He wants people to do what he says (iprspot).	3.1 (1.4)	3.5 (1.5)	3.5 (1.2)
Achievement (AC)	4. It's important to him to show his abilities. He wants people to admire what he does (ipshabt).	3.1 (1.4)	3.4 (1.4)	3.4 (1.3)
	13. Being very successful is important to him. He hopes people will recognize his achievements (ipsuces).	3.5 (1.3)	3.3 (1.4)	3.4 (1.2)
Hedonism (HE)	10. Having a good time is important to him. He likes to "spoil" himself (ipgdtim).	2.5 (1.2)	2.8 (1.3)	3.1 (1.2)
	21. He seeks every chance he can to have fun. It is important to him to do things that give him pleasure (impfun).	2.7 (1.2)	3.1 (1.5)	2.6 (1.1)
Stimulation (ST)	6. He likes surprises and is always looking for new things to do. He thinks it is important to do lots of different things in life (impdiff).	2.9 (1.3)	2.7 (1.4)	2.9 (1.3)
	15. He looks for adventures and likes to take risks. He wants to have an exciting life (ipadvnt).	4.0 (1.4)	4.0 (1.5)	3.9 (1.3)

Table 3. Number of values found in each country and values unified to solve the problem of non-positive definite matrix of the constructs in single-country analyses (The correlation between the constructs that had to be unified in the baseline model appears in parentheses)

Country	Number of values	Values unified
1. Belgium	8	COTR (0.94), POAC (0.91)
2. Luxemburg	8	COTR (0.93), POAC (0.89)
3. The Netherlands	8	COTR (1.02), POAC (0.97)

Table 4. Global fit measures in the single country CFA's after model modifications

Country	Modifications	Chi square/df	RMSEA	Pclose
1. Belgium	COTR→iprspot	3.1	0.05	.62
2. Luxemburg	UN→impdiff	2.8	0.05	.51
3. The Netherlands	UN→impdiff	2.9	0.05	.91

df = the number of degrees of freedom; RMSEA = root mean square error of approximation; PCLOSE = probability of close fit. For details see for example Arbuckle (2005).

Table 5. Global fit measures for the configural, measurement and scalar invariance models in the multiple-group CFA's

Model	Chi square	df	RMSEA	Pclose	AIC	BCC
1. Configural invariance	1411	512	.03	1.00	1899	1913
2. Measurement invariance	1509	544	.03	1.00	1933	1944
3. Scalar invariance	2273	580	.03	1.00	2625	2634
4. Partial scalar invariance (for UNBE, SD and ST only)	1587	556	.03	1.00	1987	1998

df = the number of degrees of freedom; RMSEA = root mean square error of approximation; PCLOSE = probability of close fit; AIC = the Akaike information criterion; BCC = the Browne-Cudeck criterion. For details see for example Arbuckle (2005).
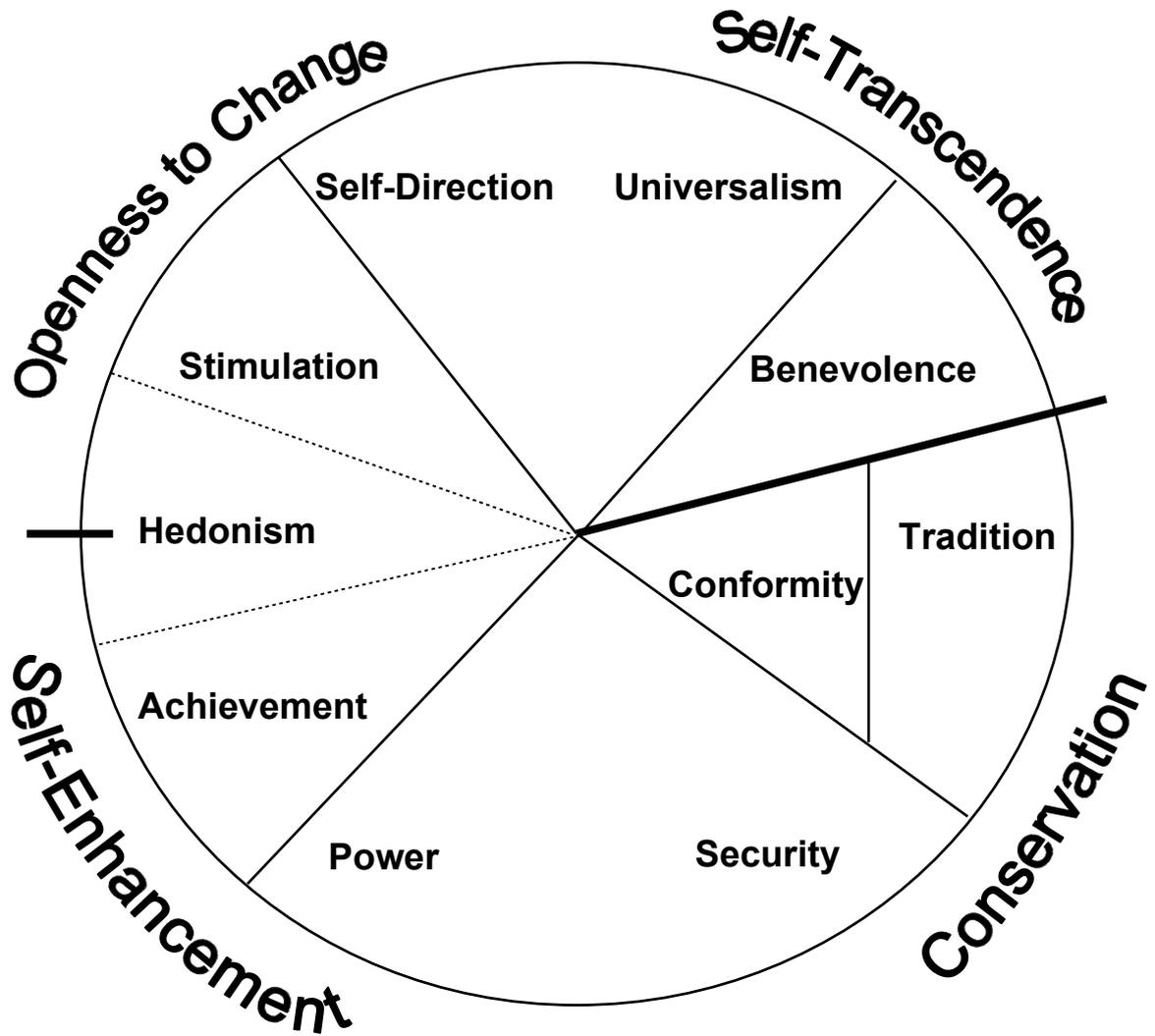
Table 6. Mean differences of the values across three countries

Country\Value	Self-direction	Stimulation	Universalism-Benevolence
1. Belgium	-	-	-
2. Luxemburg	-.01	.08	.00
3. The Netherlands	.06	.09*	-.12*

* P<0.05

Figure 1: Structural relations among the ten values and the dimensions.
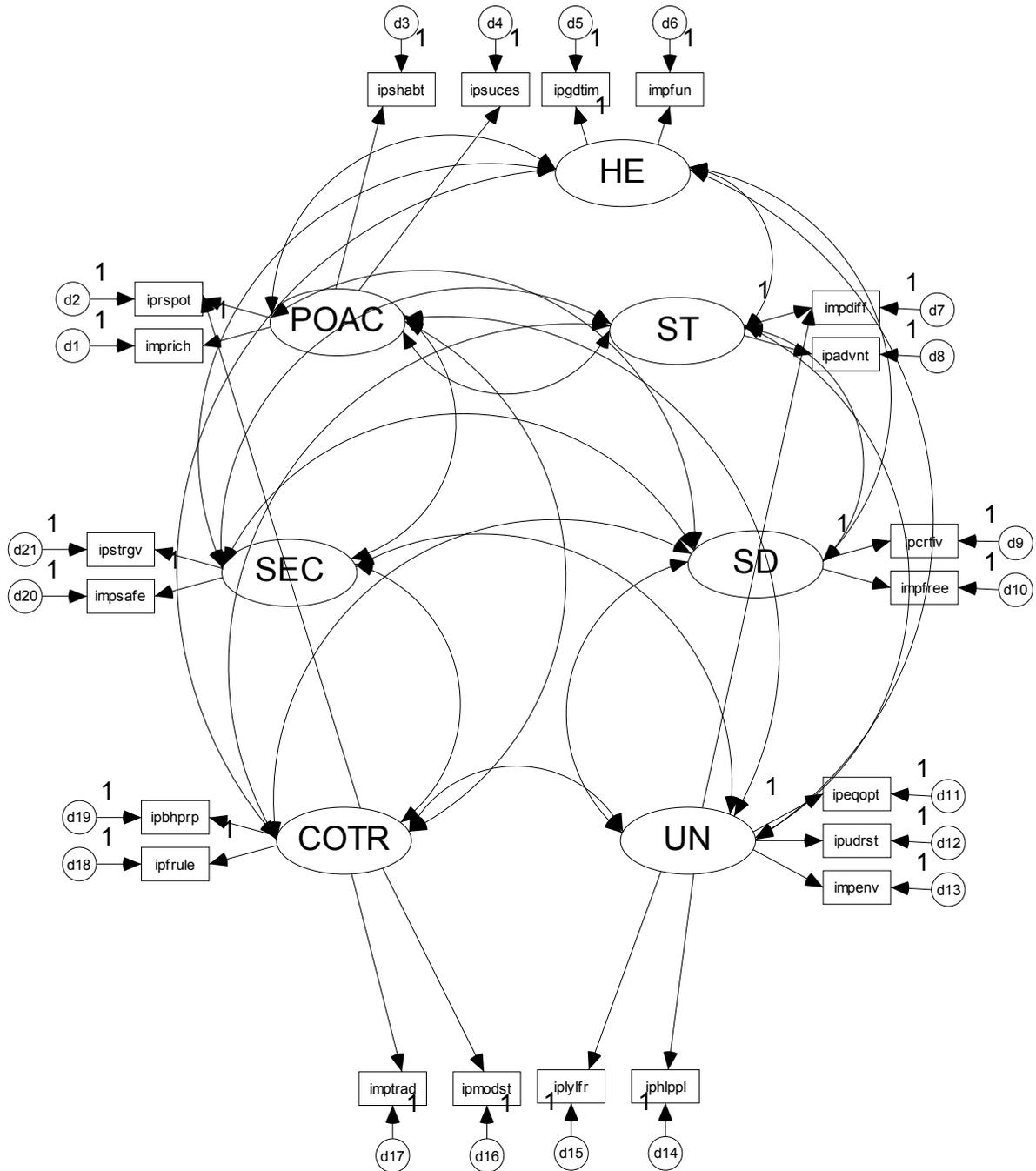Adopted from Davidov, Schmidt and Schwartz (in press).

Figure 2: Confirmatory factor analysis-the configurally invariant model across Belgium, Luxemburg and the Netherlands (For item and value names see table 2)

Appendix:

A. Unstandardized regression coefficients for Belgium, Luxemburg and the Netherlands in the

partially scalar invariant model (all coefficients are significant, p<0.01)

	SD	UNBE	COTR	SEC	POAC	HE	ST
ipcrtiv	1.00						
impfree	.75						
ipeqopt		1.00					
ipudrst		1.10					
impenv		1.20					
iphlppl		1.38					
iplylfr		1.12					
ipmodst			.76				
imptrad			.97				
Ipfrule			1.00				
Ipbhprp			1.10				
Impsafe				1.00			
Ipstrgv				.98			
Imprich					1.00		
Iprspot			**.42**		.97		
Ipshabt					1.49		
Ipsuces					1.65		
Ipgdtim						1.00	
Impfun						1.05	
Impdiff		**.86**					1.00
Ipadvnt							1.31

The path coefficients added to the original model in the modification process are in bold.

Empty cells represent no direct relation between the values and the indicators.

For item and value names see Table 2.

B. Standardized factor loadings and cross loadings for each country (based on the partially

scalar invariance model)

Item	Value Factor	BE	LU	NL
ipgdtim	HE	,629	,561	,637
impfun	HE	,654	,549	,686
impdiff	ST	,605	,584	,615
ipadvnt	ST	,700	,684	,769
imprich	POAC	,480	,483	,516
iprspot	POAC	,436	,400	,479
impsafe	SEC	,620	,601	,623
ipstrgv	SEC	,627	,637	,664
ipfrule	COTR	,536	,521	,616
ipbhprp	COTR	,678	,651	,728
ipeqopt	UNBE	,489	,493	,478
ipudrst	UNBE	,486	,479	,517
ipcrtiv	SD	,591	,583	,606
impfree	SD	,516	,448	,531
impenv	UNBE	,553	,519	,571
ipmodst	COTR	,474	,437	,438
imptrad	COTR	,508	,508	,571
ipshabt	POAC	,658	,653	,703
ipsuces	POAC	,738	,718	,805
iprspot	**COTR**	**,232**	**,213**	**,255**
impdiff	**UNBE**	**,300**	**,290**	**,305**
iplylfr	UNBE	,617	,598	,578
iphlppl	UNBE	,661	,618	,656

The path coefficients added to the original model in the modification process are in bold.

For item and value names see Table 2.