

Summary

This methodological study focuses on the international comparability of two of the most frequent robust attitude scales: namely social capital and political disaffection scales as they are measured in the European Social Survey first data set. The first goal of this study is to evaluate the levels of cross-country and cross-time attitude scales comparability. The key analytic part of the study is based on structural modeling which is the most popular technique used in this context today. It can provide a statistical test for the measurement invariance of factors over groups or rounds of survey. Furthermore, it enables researches to decide which type of statistics may be compared, given different levels of invariance. The second, rather theoretical goal is to introduce the concept of equivalence in the context of cross-national surveys to the Czech sociology. This issue hasn't been analysed and published in a systematic way in the Czech social sciences yet, and it is still overshadowed by the social survey quality issues and by the concept of measurement error in cross-national surveys.

The text is divided into several parts which focus successively on the theoretical background, methodological issues and empirical findings. The first chapter introduces the development of the cross-national surveys in the 20th century. The second chapter summarizes the concept of equivalence in the cross-national surveys, introduces basic terms, suggests brief categorization of the basic types of equivalence and outlines the portfolio of statistical techniques which are typically used for testing the level of scales comparability. The third chapter summarizes theoretical background of social capital and political disaffection concepts and shows the ways of their measurement as attitude scales in the European Social Survey. In this context, there are all systematic measurement errors which can jeopardize scales comparability pointed out. These theoretical and methodological parts of the study are used for switching to the second analytic part where the level of attitude scales comparability is actually analysed.

Therefore the fourth chapter describes the basic techniques, which verify the consistency and reliability of social capital and political disaffection attitude scales, and brings an overview of first outcomes of scales comparability level across 22 countries. The fifth chapter introduces structural modeling which is a more technically advanced method for the evaluation of scales

equivalence. The way and the analysis procedure are described there in more detail. In the sixth chapter the basic model of attitude scales is illustrated, including the description of all relations between the selected measurement indicators and latent variables. Then, it is the subject of further testing in comparison analyses. The seventh, the eighth and the ninth chapters represent separate multi-group analyses of selected attitude scales comparability complemented by rich empiric material on the basis of which detailed evaluation of the level of scales and item equivalence are carried out. The tenth chapter deals with the explanation of the character of time dependent multi-group analysis based on interpersonal general trust and institutional trust attitude scales, and it intentionally complements the outcomes of previous analyses. The last chapter summarizes the most important outcomes proceeding from all comparability analyses and universal recommendations for substantive comparative work options suitable for dealing with these attitude scales are suggested.

The results of the data analysis are the following. The attitude scales were verified as configural and at least partially metric equivalent in all participating countries. It is obvious that the social capital and political disaffection concepts are understood and interpreted in the same way in all countries. The scalar equivalence of these scales hasn't been verified in these European countries. Therefore it is probable that extent of the measurement scales of items and individual answer categories of these scales aren't perceived by respondents identically. At the same time, respondents use these measurement scales in different way. Therefore, even though respondents have the same opinion about trust in people or institutions, their score at measurement scale doesn't have to be the same. Because plenty of culturally and linguistically various countries were chosen to this analyses, obviously there is certain kind of the measurement error which caused the deviation in the intercept parameter of the item. High sensitivity of the scalar equivalence to the combination of slight cultural, linguistic and historic differences between countries and to the gentle changes of public opinion in time can cause this problem.

The implications of the given conclusions into practise is then of such a nature that in all the participating countries in ESS (2002) there has been no problem in comparing the relations between these two attitude scales, or relations of these attitude scales toward other metric equivalent attitude scales, or, perhaps the relations between the scales attitudes and other metric equivalent variable out of the data analysis set. In this process, we can use correlation, or regression analyses, or structural modeling. This type of meritory analyses is applied at cross-national level more frequently. The complications

arises when we want to compare latent means or item means across countries because outcomes show that it is not possible due to absence of the level of scalar equivalence. It would be an invalid comparison which is not allowed. Only in case of social capital factors the comparison of averages is possible among ten European countries. However, a detailed multi-group analysis has revealed that latent means can be mutually compared in small homogenous groups of countries. Mostly, these countries are homogenous from the point of view of geographical location, culture, common history and even linguistic origins. Apparently, it is because the subtle cultural and linguistic deviations make the comparison of variable means impossible.

Not surprisingly, the multi-group analyses of scalar equivalence show that items which measure the latent concept of social capital are much more problematic as regards cross-country comparability than items which measure latent concept of political disaffection or institutional trust. These results correspond to results of other empirical research projects and methodology studies. Structural modeling has revealed, as the most problematic the following three items in the social capital attitude scale (*HELPFUL*, *VOLUNORG*, and *LEGAL*) and two items in the political disaffection attitude scale (*OPINION* and *CARE*). These items probably deviate due to methodological systematic error but these outcomes should be confirmed in the subsequent cognitive testing or by a more appropriate statistical technique such as hierarchical structural modeling or item response theory.

General conclusion of this study is that it is not possible to rely exclusively on high-quality work of the international research team which plans and prepares the cross-national survey. The fact that any attitude scales are thoroughly formulated, pretested and analysed because of their validity and reliability and that every item is expertly translated into various languages and cognitive tested doesn't mean that these scales are fully comparable across all countries in the data set. It is always necessary to verify attitude scales comparability over and over in every new round of data collection.