

Editorial

Bart Meuleman, Eldad Davidov & Daniel Seddig

Over the last decades, the increasing availability of comparative survey data has opened up a wide avenue of research opportunities for social scientists. International survey projects -such as the European Social Survey (ESS), the European (EVS) and World Values Studies (WVS), or the European Household Panel Study (EHPS)- measure a wide range of attitudes and behaviors with the explicit purpose of making comparisons across countries, regions or time points (Lynn, Japenc, & Lyberg, 2005). The potential relevance of such comparisons is paramount. Besides identifying differences between contexts and cultures, comparative data is helpful in testing theories about social change and contextual influences on individual characteristics. The insight that comparison is a crucial methodological tool is not new, but is as old as social science itself. After all, Durkheim (1964, p. 139) already argued that “*comparative sociology is not a particular branch of sociology: it is sociology itself*”.

The advantages of the comparative design come at a methodological price, however. Collecting and analyzing cross-national survey data brings along additional methodological challenges (Berry et al., 1992; Harkness et al., 2003; Harkness et al., 2010; van de Vijver & Leung 1997). Among a great many pressing methodological issues, comparative research hinges crucially on the assumption that measurements are comparable or equivalent (Horn & McArdle, 1992; Johnson 1998; Davidov et al., 2014). Respondents in international surveys were socialized in different cultural backgrounds, speak different languages and have cultural-specific understandings of certain ideas and concepts. Therefore, it is not guaranteed that survey measurements travel successfully across national and cultural borders (Jowell et al., 2007). Equally important is to guarantee that measurements travel successfully across groups within countries (Davidov & Siegers, 2010; Sarrasin, Green, Berchtold & Davidov, 2012), across modes of data collection (Ciecuch & Davidov, 2016), or across time (Widaman, Ferrer, & Conger, 2010). Therefore, the validity of comparisons of survey measurements across groups and time is of great concern (Jowell, 1998).

Fortunately, in recent years comparative researchers have increasingly acknowledged the importance of the comparability of measurements. A variety of methodologies have been proposed to assess to what extent survey measurements are cross-culturally equivalent (Davidov, Schmidt, Billiet & Meuleman, 2018). This special issue has the ambition to contribute to the contemporary debates on the

comparability of survey measures. By providing new tools, novel insights and original applications in the field of measurement equivalence, this collection of papers advances our current knowledge on measurement equivalence.

A first set of three papers shows how measurement equivalence of multiple-item scales can be tested using a multiple-group factor analytic approach. *Wiebke Breustedt* argues that the generalizability of theories on political trust requires that this concept should be measured in a comparable way. Analyzing data from various rounds of the WVS by means of multiple group confirmatory factor analysis (MGCFA), Breustedt shows that this assumption should indeed not be taken for granted: Only in 19 out of 32 investigated democracies, configural invariance could be established. This important finding calls for a further development of cross-culturally robust instruments to gauge citizens' trust in public institutions. *Maksim Rudnev* and colleagues extend the popular MGCFA equivalence test to higher-order factor models. This paper explains in detail which model constraints are necessary to operationalize various levels of equivalence in second-order factor models. In addition, an empirical illustration evaluating the equivalence of Seeman's second-order concept of alienation across eight countries is provided. The study by *Vera Lomazzi* addresses an important weakness in the MGCFA strategy, namely that the requirements for equivalence are very strict, especially when a comparison involves a large number of groups. Lomazzi proposes to use the recently introduced alignment optimization procedure as an alternative for the common MGCFA model. Analyzing the gender role attitudes scale in the WVS across 59 countries, the results indicate that the alignment procedure is less strict and suggests that valid comparisons are possible across a wider range of countries than when the classical MGCFA model is used.

Two papers investigate how particularities of languages and writing might affect cross-cultural comparability. *Dagmar Krebs* and *Yaacov G. Bachner* tackle the intriguing question how the direction of writing – left to right vs. right to left – interacts with the way in which respondents use response scales. After all, respondents can pick up information from response scales (incremental or decremental) and factor this in their response behavior. To test this expectation, the authors analyze data from a split-ballot design among German and Israeli students. The results indicate that clear response-order effects are present, but that they are very similar in left-to-right (German) and right-to-left (Hebrew) reading directions. *Diana Zavala-Rojas* studies if the language in which a survey was conducted has a noticeable impact on measurements of various political attitudes among bilingual citizens. Concretely, bilingual respondents' institutional trust and satisfaction with politics and economy were measured twice in a different language. Within-subject equivalence tests show measurements are largely equivalent across the language of survey administration, even if the correlation between two language-versions of a latent variable is not identical to 1. Summarizing, the message of these two papers is opti-

mistic: If the necessary precautions are taken, characteristics of languages are not insurmountable for comparative researchers.

Finally, the paper by *Silke L. Schneider* draws our attention to the important message that equivalence not only matters for subjective concepts measured by multiple items. Also objective social-structural characteristics, such as educational attainment, need to be measured in a comparable way. Schneider assesses the comparability of the education variable included in PIAAC (Programme for the International Assessment of Adult Competencies). Equivalence is evaluated from the perspective of construct validity, that is, by looking at the relationship with respondents' general skills. The study shows that especially decisions to collapse the detailed education variable into a smaller number of categories challenge comparability, and identifies several pitfalls in the educational attainment variables currently used in comparative research (such as the lack of differentiation between general and vocational training).

References

- Berry, J.W., Poortinga, Y.H., Segall, M.H., & Dasen, P.R. (Eds.) (1992). *Cross-Cultural Psychology. Research and Applications*. Cambridge: University Press.
- Cieciuch, J. & Davidov, E. (2016). Establishing measurement invariance across online and offline samples. A tutorial with the software packages Amos and Mplus. *Studia Psychologica* 15 (2), 83-99.
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement Equivalence in Cross-National Research. *Annual Review of Sociology*, 40, 55-75.
- Davidov, E., Schmidt, P., Billiet, J. & Meuleman, B. (2018). *Cross-cultural analysis: Methods and applications*. New York: Routledge.
- Davidov, E., & Siegers, P. (2010). Comparing basic human values in east and west Germany. In T. Beckers, K. Birkelbach, J. Hagenah & U. Rosar (Eds.), *Komparative empirische Sozialforschung [Comparative empirical Social Research]* (pp 43-63). Wiesbaden: VS Verlag.
- Durkheim, E. (1964). *The Rules of the Sociological Method*. New York: The Free Press.
- Harkness, J.A., Braun, M., Edwards, B., Johnson, T.P. & Lyberg, L. et al. (2010). *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. Hoboken (NJ): John Wiley & Sons.
- Harkness, J.A., van de Vijver, F.J.R. & Mohler, P.P. (Eds.) (2003). *Cross-Cultural Survey Methods*. New York: John Wiley
- Horn, J.L. & McArdle, J.J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Exp. Aging Res.* 18, 117-44.
- Johnson, T. (1998). Approaches to equivalence in cross-cultural and cross-national survey research. In J.A. Harkness (Ed.), *Zuma-Nachrichten Spezial Volume 3. Cross-Cultural Survey Equivalence*, (pp. 1-40). Mannheim: Zuma.
- Jowell, R., Roberts, C., Fitzgerald, R. & Eva, G. (2007). *Measuring Attitudes Cross-Nationally. Lessons from the European Social Survey*. London: Sage.

- Jowell, R. (1998). How comparative is comparative research? *Am. Behav. Sci.* 42(2), 168-77.
- Lynn, P., Japac, L. & Lyberg, L. (2005). What's so special about cross-national surveys? In J.A. Harkness (Ed.) *Zuma-Nachrichten Spezial* Volume 12. *Conducting Cross-National and Cross-Cultural Surveys*, (pp. 7-20). Mannheim: Zuma.
- Sarrasin, O., Green, E.G.T., Berchtold, A. & Davidov, E. (2012). Measurement equivalence across subnational groups: an analysis of the conception of nationhood in Switzerland. *International Journal of Public Opinion Research* 25(4): 522-534.
- Widaman, K.F., Ferrer, E. & Conger R.D. (2010). Factorial invariance within longitudinal structural equation models: measuring the same construct across time. *Child Development Perspectives* 4(1): 10-18.