Syllabus for course 13: “Testing survey data for measurement equivalence across countries and time”

Instructors: Prof. Dr. Jan Cieciuch
Email: jancieciuch@gmail.com

Prof. Dr. Eldad Davidov
daavidov@soziologie.uzh.ch
with Dr. Daniel Seddig
seddig@soziologie.uzh.ch

Prof. Dr. Peter Schmidt
peter.schmidt@sowi.uni-giessen.de

Date: August 24-28, 2015
Time: 14:00-18:00

About the Instructors:
Jan Cieciuch is Associate Professor of Psychology at the Cardinal Wyszynski University in Warsaw, Poland and Project Leader of the University Research Priority Program “Social Networks” at the University of Zurich. His interests are applications of structural equation modeling especially in psychology, with focus on the investigation of human values and personality traits. Recent publications appeared in the Journal of Personality and Social Psychology, Journal of Personality Assessment, Journal of Cross-Cultural Psychology, Personality and Individual Differences, Annual Review of Sociology, and Public Opinion Quarterly.

Eldad Davidov is Professor of Sociology at the University of Zurich, Switzerland and vice-president of the European Survey Research Association. His research interests are applications of structural equation modeling to survey data, especially in cross-cultural and longitudinal research. Applications include human values, national identity, and attitudes toward immigrants and other minorities. Recent publications on these topics appeared in the Annual Review of Sociology, Social Science Research, Public Opinion Quarterly, Sociological Methods and Research, International Journal of Public Opinion Research, European Sociological Review, and Political Analysis.


Short Course Description:
Determining whether people in certain countries, or at different time points score differently in measurements of interest, or whether constructs relate differently to each other across nations can indisputably assist in testing theories and advancing our sociological knowledge. However, meaningful comparisons require equivalent measurements of these constructs. This is especially true for subjective attributes such as values, attitudes, opinions, or behavior. In this course, we first discuss the meaning of the cross-group measurement equivalence, look at possible sources of nonequivalence, and suggest ways to prevent it. Next, we examine the social science methodological literature for ways to empirically test measurement equivalence and conduct such tests during the exercises with survey data using the software package Mplus. We focus on the multiple group confirmatory factor analysis (MGCFA) method. Furthermore, we consider what may be done when equivalence is not supported by the data and conclude with a discussion of recent developments such as the Bayesian estimation procedure to test for approximate invariance or multilevel structural equation modeling that can potentially be used to explain non-invariance. These methods offer exciting directions and solutions for future research in cross-group measurement equivalence.
assessment. Finally we discuss if time allows further issues, such as the newly developed alignment procedure or how multiple group structural equation modeling (rather than MGCFA) can be applied to cross-national data sets if the number of units (here countries) is too small for a multilevel analysis.

Course prerequisites:
- Knowledge of and some experience with confirmatory factor analysis and structural equation modeling. This could be acquired in the course “Introduction to Structural Equation Modeling: Confirmatory Factor Analysis with Mplus” in week 1.
- We will use the software package Mplus. A short introduction to the use of the program will be provided during the first exercise. Knowledge of Mplus could also be acquired in the course “Introduction to Structural Equation Modeling: Confirmatory Factor Analysis with Mplus” in week 1.
- Preferably (not obligatory) some knowledge and experience with measurement error in survey research. This could be acquired in the course “Understanding and Modeling Measurement Error in Social Surveys” in week 2. This course will also use the software package Mplus. Furthermore, this course will touch upon the topic of equivalence on the last day.

Target group:
Participants will find the course useful if:
- they are interested in conducting meaningful comparative research across countries, over time or across other social groups such as natives and immigrants;
- they are interested in understanding potential problems in comparative research, and in studying and applying the methodological prerequisites for meaningful comparisons across groups.

Course and learning objectives:
By the end of the course participants will:
- know how to test for measurement invariance across groups using various types of multigroup confirmatory factor analyses with the software package Mplus;
- gain practical experience in testing for measurement invariance and computing latent means for different groups (e.g. countries and time-points);
- learn how to deal with non-invariant measurements.
- If time allows, learn to run multiple group structural equation models for comparative research to test complex theories after testing for measurement invariance.

Organizational structure of the course:
First four days:
- about 120 minutes – lecture
- about 120 minutes – exercises on data prepared by the instructors
  **Note:** In some of the days we may divide the theory and exercises time differently, depending on our progress.

Last day:
- presentations of participants’ work on their own data

  **Free study time and what we expect from participants:**
- Participants are encouraged to discuss the topics of the course every day with each other.
- Participants are expected to repeat the exercises conducted in class on their own.
- Participants should read the course literature.
- Performing analyses on own data individually or in groups.
- We will offer daily 60 minutes of individual consultation for interested participants after the lectures and exercises on Tuesday and Friday and from 12:00 to 13:00 on other days. Consultation could be offered for example for individual projects, for topics discussed during the theoretical session or for the exercises.

Software and hardware requirements:
SPSS and Mplus Version 7.3. Course participants will not need to bring a laptop computer for this course. This course will take place in a computer lab.
Long Course Description:
The increasing availability of large international surveys has opened a wide avenue of possibilities for sociologists interested in cross-national and longitudinal comparisons. Prominent examples of such survey projects include the European Social Survey (ESS), International Social Survey Program (ISSP), European Values Study (EVS), World Values Study (WVS), European Household Panel Study, Program for the International Assessment of Student Achievement, Global Entrepreneurship Monitor, and regional barometers (e.g., the Afrobarometer, Asian Barometer, Eurobarometer, and Latinobarometro). In all these studies, various theoretical constructs such as value priorities, attitudinal dispositions, or behavioral patterns are measured mostly by multiple indicators with the explicit goal of making comparisons across different countries, regions, and time points. Such comparisons are extremely useful for advancing social-scientific theory testing. However, the endeavor to obtain and analyze such data entails serious challenges. Human beings—the source of information in surveys—may fail to provide correct answers, may say one thing but do another, may respond in a socially desirable manner, and are easily influenced by contextual clues. In addition, respondents in cross-national surveys have been socialized in diverse economic and cultural backgrounds, speak different languages, and might understand certain ideas or constructs in varying ways or in a culturally specific manner.

We will begin the course with a discussion about the importance of measurement invariance in comparative research and define the concept of measurement invariance. Then we will present various possible sources of non-invariance. We will particularly refer to construct bias, method bias and item bias. Next we will describe various common analytical strategies to empirically test measurement equivalence focusing on the generalized latent variable framework by using multiple group confirmatory factor analysis. We will also show how to deal with ordered-categorical data when testing for measurement invariance, as such data are often used in surveys. In addition, we will pay special attention to the assessment of model fit and model modification in measurement invariance testing. We will expand our discussion by referring to alternative approaches suggested in the literature such as exploratory structural equation modeling, partial invariance, and the Bayesian procedure to test for approximate rather than exact measurement invariance. We will touch upon the alignment procedure if time allows. These topics will be accompanied by presentations of examples and applications from the literature. Finally, we will summarize possible strategies how to deal with situations in which measurement equivalence is not supported by the data. One of the strategies we will focus on is multilevel structural equation modeling. This strategy is particularly powerful, because it enables researchers to explain why measurement invariance is not given. We will use this approach to show how non-invariance may be explained. We will discuss several other recent and promising developments in the measurement invariance literature. Finally we will present multiple-group structural equation modeling for testing theories with comparative data, if the number of groups is smaller than 15 and multilevel models may not have enough power. We will exercise major topics we cover in the theoretical sessions in the PC lab using survey data from the ESS dealing with the Schwartz’ “human values” values which include several countries in a step by step process.

We warmly recommend participants to bring their own data and apply the methods we discuss and study on their data during the course. On the last day of the course participants will have the opportunity to present their findings and discuss with the class problems they encountered.

Day-to-day schedule and literature used in the course:

<table>
<thead>
<tr>
<th>Day</th>
<th>Topic(s)</th>
<th>Compulsory reading:</th>
</tr>
</thead>
</table>
| 1   | Introduction; the importance of equivalence; how equivalence can be tested; examples | • Davidov E., Meuleman B., Cieciuch J, Schmidt P., Billiet J. 2014. Measurement equivalence in cross-national research. An. R. Soc. 40: 55-75.  

Suggested reading:

2 Assessing models which test for equivalence: the chi square difference test; global fit measures; detection of misspecification; examples

Compulsory reading:
• Cieciuch, J., Davidov, E. 2014. Establishing measurement invariance across online and offline samples. A tutorial with the software packages Amos and Mplus. Studia Psychologica, 14(2).

Suggested reading:

3 Accounting for ordinality; alternative approaches: exploratory structural equation modelling; what to do when there is no invariance: explaining noninvariance using multi-level structural equation modeling; examples

Compulsory reading:

Suggested reading:
The Bayesian approach to test for approximate measurement invariance; examples

Compulsory reading:

Suggested reading:

Participants’ presentations; general summary; open questions; filling in the online evaluation questionnaire followed by a debriefing/feedback round; if time allows: When non-invariance does not matter; invariance of means and structural effects; examples

Suggested reading:

Preparatory reading:

Additional recommended literature: