“Understanding and Modelling Measurement Error in Social Surveys”

Instructors: Jaak Billiet, Melanie Revilla
Email: Jaak.Billiet@soc.kuleuven.be, melanie.revilla@hotmail.fr

Date: August 18-22, 2014
Time: 9:00-13:00

About the instructors:
Prof. Billiet Jaak, PhD, is emeritus professor in social methodology at KU Leuven since 2007. He was in 2001-2011 a member of the central co-ordination team of the European Social Survey.
Methodological research: validity assessment, interviewer and response effects, and the modelling of measurement error in social surveys. Substantive research: longitudinal and comparative research in the domains of ethnocentrism, political attitudes and religious orientations. He played a central role in the implementation of the fourth wave of the European Value Study in 2008. Jaak Billiet has published widely with research fellows from various countries in peer-reviewed journals and specialized books on survey methodology and cross-cultural analysis.

Melanie Revilla, PhD, is a researcher at the Research and Expertise Centre for Survey Methodology (RECSM) and associate professor at Universitat Pompeu Fabra (UPF, Barcelona, Spain). She received her PhD from Universitat Pompeu Fabra (2012) in the areas of statistics and survey methodology, under the supervision of professors Willem Saris (UPF) and Peter Lynn (Essex University). Her dissertation dealt with the effects of different modes of data collection on the quality of survey questions. Besides the impact of the mode of data collection, she is interested in all aspects of survey methodology.

Short course description:
This course provides a comprehensive introduction to measurement error (ME) in survey research. This is a crucial topic because there are always errors in the measurement and if we do not take this into account, substantive results may be completely wrong without realizing it. First, we will discuss some basic concepts in the field of ME and discuss the epistemological background of the distinction between conceptual validity and measurement validity. Examples and ways to reduce measurement error will be discussed, such as basic rules for question wording. Then, attention shifts from reduction of error to measuring error. In a first step the findings of “split ballots” will be critically reviewed, and this approach will be critically evaluated. This is an occasion to introduce the cognitive insights in question wording and order effect that may provide deeper understanding of what is going on in the question-answer process. In next step we will discuss the modeling of measurement error using models for assessing reliability and validity such as the MTMM model. The course will also have a detailed look at response styles in survey questions, as it is an important source of measurement error. We will use acquiescence as a generic example and model it using Structural Equation Modelling (SEM). Finally, specific problems of measurement in the context of cross-cultural research are shortly reviewed. This is an opportunity to inform the participants about a related course on approaches to equivalence testing in cross-cultural studies which is a key requirement if one wants to compare data across countries or groups. The practical exercises will guide participants to apply the concepts and procedures learned in the theoretical sessions. The software Mplus and data from the ESS will be used.
Course prerequisites:

- students must be motivated to apply what they will learn in their own research and should be interested in the quality of survey data.
- basic knowledge of statistical analysis and structural equation modelling. The examples in the theoretical part are based on LISREL® and participants should be able to interpret the parameters but the concepts used are general applicable in empirical analysis of ME. The practical exercises assume a minimum practical knowledge of Mplus. For those not familiar with it, these requirements can be fulfilled by attending course "Introduction to the Structural Equation Modelling framework" in the previous week.

Target group:
Participants will find the course useful if they:

- are PhD students who will use survey data in their work;
- employees in survey or marketing institutes who are engaged in the analysis of survey data, or in preparation of survey questionnaires;
- are exceptionally specialized research master students interested in measurement error if this topic is not covered in their program.

Course and learning objectives:
By the end of the course participants will:

- have acquired a thorough understanding of the different kinds of measurement error related to the instrument (questionnaire) in social surveys;
- be aware of ways to detect, measure and control for errors;
- be familiar with the most recent developments and literature in the field of measurement error, especially in the context of large scale cross-cultural surveys;
- be familiar with the best practices in the field;
- be able to apply the strategies in practice.

Organizational structure of the course:
The course structure includes 4 hours of in-class instruction in the morning, of which 3 hours will be a lecture (9:00 - 12:00) and 1 hour will comprise practical exercises (12:00 – 13:00). In addition, the participants are expected to take part in 2 hour exercises session in presence of a course instructor (14:00-16:00), which will take place in the PC-pool the afternoon. In the afternoon, participants can also make appointments with the course instructors for individual consultations. The presence in the PC-pool during this time is considered to be an important part of the learning for this course, as we will learn how to work with different software than the one used for the lectures. The instructor will propose different exercises and help the participants when needed.

Software and hardware requirements:
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:
This course provides a comprehensive introduction to measurement error in survey research in the context of the Total Survey Error concept. To start with, the basic concepts and sources of measurement error as well as the concepts of validity and reliability will be introduced in the context of measurement quality, measurement validity, and conceptual (construct) validity. We will discuss examples of and ways to reduce measurement error, such as basic rules for question wording as well as strategies for single indicators. Then the course will introduce participants to the measurement of measurement error. In this context, the use and findings of ‘split ballots’ will be discussed and critically reviewed. Measurement error will be related to the concept of ‘attitude strength’, classical as well as recent cognitive models of the question-answer process (e.g. Zaller’s RAS model) and the concept of response instability. The next step concerns the modelling of measurement error using models for assessing reliability and validity such as the test-retest model, simplex model, RMM model, congeneric model, and the MTMM model as well as their application. We will study their advantages and
disadvantages as well as the true-score MTMM model. The course also covers meta analysis of MTMM models and how to adjust their correlations.

The course will also have a detailed look at response styles in survey questions, as it is an important source of measurement error: What are response styles, and what are the different kinds of response styles and what are the ways of detecting and measuring response styles? We will use acquiescence as a generic example and model it using Structural Equation Modelling (SEM). What are the advantages of modelling acquiescence in a cross-cultural context? Do we really measure an acquiescence method effect? We will also examine whether this kind of modelling is useful for other response styles.

Finally, the course will shortly discuss the importance of equivalent measurement in cross-cultural research, and offer a flavour of it by showing some examples of consequences that non-equivalent measures may have for conclusions in cross-nation research. This introduction to equivalence is conceived as a stepping stone for course 13 Testing Survey Data for Measurement Equivalence across Countries and Time.

The practical exercises are focused on practical measurement problems in large datasets and will guide participants to apply the concepts and procedures they have learned in the theoretical morning sessions. The software Mplus and data from the ESS will be used. The course will be applicable to surveys of individuals.

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

<table>
<thead>
<tr>
<th>Day</th>
<th>Topic(s)</th>
</tr>
</thead>
</table>
| 1   | **Introduction to the course and basic concepts:**  
- The concepts of data quality, Total Survey Error, and measurement error  
- Sources of measurement error  
- The idea of a latent variable measured by multiple indicators;  
- The distinction between measurement validity and conceptual  
- (theoretical) validity  
- Necessary conditions for measurement validity and what with conceptual validity?  
- Approaches to validity and reliability assessment (overview)  
- How to reduce measurement error?  
- Discussion of several strategies used for single indicators  
- Basic rules for question wording  

**Introduction to first set of exercises:**  
- Discuss different existing questions that have been used or proposed to be used in the ESS. Propose improvements.  
- Operationalize some complex concepts.  
- Formulate questions for single items related to own research.  

**Individual work and supervision**  

Compulsory reading: Billiet,  


Suggested reading:  
### Approaches to reliability and validity: the measurers approach (I) split ballots:
- Reduction of ME (part 2). Basic rules applicable in a cross-cultural context (example: ESS).
- Classical approach to measurement of ME: the split ballot approach.
- What are split ballots? An overview of findings from split ballots.
- Pro's and con's of split ballots in kinds of question wording effects detected.
- A typology: question constraint, unintended information in the question text and response scales, leading question and loaded questions, context effects in questions.
- Kinds of context effects (assimilation and contrast). How to recognize them.
- Rules for specific kinds of question wording effects
- The concept of “attitude strength” and its methodological relevance
- How to explain question wording effect: cognitive models of question-answering (classical and new approaches)
- Zaller and the well known problem of response instability (implications…)

### Introduction to second set of exercises:
- Introduction to the SQP program (version 2)
- Use it to look at the impact of different aspects usually studied by split-ballot and to improve the questions made on Monday
- Use Mplus to estimate simple measurement models: 1 factor with 1, 2, 3 indicators (to see identification issue), 2 correlated factors with several indicators

### Individual work and supervision

**Compulsory reading:**

**Suggested reading:**

### Approaches to reliability and validity: the measurers approach (II) Modeling measurement error:
- Finalization of findings from split ballots
- Understanding of question wording effects found (cognitive explanations included).
- Discussion on what to do?
- The MTMM model: introduction. Estimation and interpretation of parameters from MTMM models.
- The true score MTMM model.
- How to use the MTMM parameters for adjusting observed correlations
- Meta analysis of MTMM models and how to adjust correlations (SQP)
- The application of MTMM experiments in ESS (examples)

### Introduction to third set of exercises:
- Use Mplus to estimate several MTMM models (“classical”, true score, multiple group)
- Use the quality estimates from SQP to compute the true correlations from a set of observed correlations.

### Individual work and supervision

**Compulsory reading:**

**Suggested reading:**

### 4 The MTMM approach (part 2)
- Applications of MTMM models
- Deviations of classical 3x3 model: combining MTMM and split ballot approach
- Advantages and disadvantages of MTMM models
- Assumptions behind the SQP
- The application of MTMM in ESS: findings and conclusions.

**Modelling response styles (part 1):**
- What are response styles?
- Kinds of response styles
- Generic example: acquiescence and approaches to acquiescence
- Short introduction into SEM
- Modeling acquiescence with SEM, necessary conditions, possible models, how to interpret the parameters?
- Advantages of modeling acquiescence in a cross-cultural context?

### Introduction to fourth set of exercises:
- Use Mplus to estimate some of the models presented in the lecture
- Go through the questionnaire given and found a set of questions that can be used to be analyzed in a model including an acquiescence factor
- Estimate the model using Mplus.

**Individual work and supervision**

**Compulsory reading:**

**Suggested reading:**

### 5 Measuring response styles (part 2):
- Do we really measure acquiescence method effects?
- Is this kind of modeling useful for other response styles? Examples
- Other ways of modeling response style (recent approaches and finding)

**The assessment of equivalent measurement in cross-cultural research (introduction):**
- Examples of measurement error related to translation of items in cross-cultural context…
- Kinds of equivalence and procedures to assess equivalence
- What in case of single indicators?
- Equivalence of social-background variables
- The research in Gesis on the equivalence of background variables (Hoffmeyer-Zlotnik et al.)
- The assessment of equivalence for multiple indicator latent variables: the concept of DIF.
An example of implications of non-equivalence (measurement of religion)

**Introduction to fifth set of exercises**
- Do an example of multiple-group analysis in Mplus
- Testing using JRule for Mplus

**Individual work and supervision**

**Compulsory reading:**

**Suggested reading:**

**Preparatory reading:**

**Additional recommended literature:**