9th GESIS Summer School in Survey Methodology
Cologne, August 2020

Syllabus for Course 4:
“Introduction to Survey Design”

Instructors:  Prof. Dr. Bella Struminskaya  Dr. Ulrich Krieger
E-mail:       b.struminskaya@uu.nl  ulrich.krieger@uni-mannheim.de
Homepage:     https://bellastrum.com  https://www.sowi.uni-mannheim.de/blom/team/krieger/

Date: 10-14 August 2020
Time: 09:00-13:00 + 16:00-17:30 + time for individual consultations
Time zone: CEST, course starts on Monday at 09:00
Venue: Online via Zoom

About the Lecturers:
Dr. Bella Struminskaya is an assistant professor in methods and statistics at Utrecht University. Her research focuses on the design and implementation of online, mixed mode, and smartphone surveys. She has published on various aspects of data quality, nonresponse and measurement error, including panel conditioning and device effects. Her current research interests focus on passive data collection, in particular smartphone sensor measurement and recruitment and maintenance of online panels.

Dr. Ulrich Krieger serves as the head of operations at the German Internet Panel and is a survey researcher at the University of Mannheim. In his research he examines survey nonresponse and nonresponse bias predictive analytics and survey operations. He and his team are currently examining ways to prevent respondents to drop out of panel studies.

Selected Publications:

Short Course Description:
This course gives an overview of the design and implementation of surveys from the initial planning phase to the data preparation as a final step. Topics include survey mode assessment and selection, sampling frames and designs, nonresponse, questionnaire design, cognitive pretesting, assessing measurement errors and data editing. The course is taught from a Total Survey Error perspective weighing data quality at each step of the process against associated costs.
The course is taught through formal online lectures in which the theoretical foundation in the literature is discussed, less formal presentations and discussions of survey design in existing survey research, as well as personal tutorial meetings that give participants the opportunity to discuss exercises and their own survey designs. Each day we will discuss a specific topic that each focuses on one or more aspects of survey design within the Total Survey Error framework. First, the choice of the survey mode is discussed, and how different ways to sample respondents follow from that choice. On the second day, we focus on the issue of survey nonresponse – how to prevent, analyze, and correct for it. On the third and fourth day, the actual survey content is discussed – how to write survey questions, make sure that they measure what they are intended to measure, test them, and finally, how to assess whether survey data are of good quality. On the final day, we focus on data coding and maximizing quality. We conclude with an overview perspective of all survey errors and their interaction with survey costs. The course will be applicable to surveys of individuals, households, and organizations in different survey modes: mail, face-to-face, web, and paper-and-pencil surveys.

*Please note that this is an introductory course (see also course prerequisites and target group).*

**Keywords:**
Introduction, survey methods, survey design

**Course Prerequisites:**
- No previous experience in survey research is needed; however, some practical experience in conducting surveys and analyzing data will be beneficial.
- A basic understanding of statistics is assumed, at the level of basic inferential statistics (t-tests).
- All students need to send a brief summary of their experience with surveys (about 0.5 page) and the questions they have about how to design surveys before the start of the course to the instructors, at the latest on August 7, 2020.

**Target Group:**
Participants will find the course useful if they:
- are thinking about conducting a quantitative survey themselves;
- use survey data and wish to understand its potential errors;
- are Master or PhD students preparing their own survey;
- are researchers who collaborate within a survey research project.

*The course is tailored to those relatively new to the area of survey methodology* possibly planning to later follow more advanced and specialized courses at the GESIS Summer School. The course does not provide an introduction into data analysis of survey data. Rather, it is focused on the design of surveys.

**Course and Learning Objectives:**
By the end of the course participants will:
- have a good grasp of the complexities of interacting survey errors;
- be able to design a survey project themselves taking the possibility of survey errors into account;
- be prepared for more specialized courses at the GESIS Summer School.

**Organizational Structure of the Course:**
The course days contain about three hours of classroom instruction (morning sessions) and up to three hours of individual tutoring, discussions and group exercises (afternoon sessions). There is no obligatory afternoon session planned for the last afternoon (August 14).

**Software and Hardware Requirements:**
None. This course does not include the use of statistical software.
Long Course Description:

Why take a course on survey design?
Surveys are everywhere. Within the social sciences, the majority of empirical studies rely on surveys to collect data on demographics, attitudes, and behavior. Setting up a survey may seem to be a relatively simple process. Everyone can ask questions! In practice, however, conducting a survey often turns out to be hugely complicated, for various reasons. First of all, the number of choices for the basic design of your survey can seem overwhelming. You have to think about the choice of survey mode, obtaining a good sample, limiting nonresponse, asking good questions and analyzing data, all within time and cost constraints.
To make things even more complicated, each individual design choice affects other aspects of the survey design. For example, choosing to do a survey online is generally cheap and quick, but it will be difficult or impossible to obtain a representative sample, and some questions are less suitable to be asked online. Moreover, the correct choice of a survey design depends on your study population and your research question.
This course introduces you to survey design. We will discuss the various stages that you encounter in doing a survey and will evaluate the trade-offs between different design choices you may face. We do this from the perspective of Total Survey Error. The overall goal is to limit the overall error of your survey in order to enable you to give the best answer possible to your research question.

Focus of the course
The course aims to give an overview of the survey design and survey processes from a Total Survey Error perspective. It prepares students to take more specialized courses in one of the later weeks of the GESIS summer school (see learning objectives).

What will not be covered?
In this course, we will not cover the analysis of survey data. That is, how to analyze your survey data once you have collected it. For this, you will need to take (or have taken) a general statistics course for social scientists.
We will also not cover how to work with software for implementing Internet surveys or surveys in other modes.
Finally, the course will not cover qualitative interviews. Our course focuses on doing a survey with structured, closed-ended questions. If you would like to learn more about qualitative research and the combination of qualitative and quantitative research, we refer to the course on mixed-methods research. Finally, those individuals with a specific interest in mixing survey modes (Internet, Face-to-Face, phone and mail) are referred to the course on mixed-mode surveys. We do cover mixed-mode surveys, but only shortly. We do discuss how to do survey research in different modes, with about equal amount of attention given to face-to-face and web surveys, and less attention to paper-and-pencil and telephone surveys.

How will the course work?
During the morning session from 9 am to about 1 pm, the two instructors will give interactive lectures introducing the topics of the day. During these sessions there will be time for Q&A sessions and interactive elements.
There will be ample room for discussion, and we encourage students to contribute their own experiences and questions. There will be room for discussing specific topics you may wish to know more about. The lectures include some practical group exercises and demonstrations.
During the afternoon session from 4 pm to about 5:30 pm Monday to Thursday exercises provided by instructors and that are linked to the materials discussed in the lectures will be discussed. Students are expected to work on those problems in groups. On day 1, you will design your own sample and choose a survey mode. On day 2, you will design a fieldwork strategy aimed to minimize nonresponse. On days 3 and 4, you will work on designing and evaluating your own questionnaire. We encourage students to work on their own projects during the afternoon sessions but will provide example questionnaires and datasets for those students that do not have their own survey project (yet) to work on. On Friday no exercise is provided. There will be time for students to present their survey-methods-related work and receive feedback from the instructors and peers. Also, the lecturers will be available for one-on-one consultations on students’ ongoing projects.
### Day-to-day Schedule and Literature:

<table>
<thead>
<tr>
<th>Day</th>
<th>Topic(s)</th>
</tr>
</thead>
</table>
| 1   | **Survey processes in various interviewer-assisted and self-completion modes: Sampling strategies and coverage**  
The first day sets the scene of the course. First, we discuss the strengths and weaknesses of surveys as a research design in comparison to other research methods. We introduce the Total Survey Error (TSE) framework and discuss how the survey mode affects the potential for different survey errors. Dimensions of survey mode – computer- vs. paper-based, interviewer-assisted vs. self-completion and aural vs. visual – are contrasted. During the second part of this day, we dive into the availability of sampling frames and their coverage of the population. We shortly contrast probability and non-probability samples.  
On this first day, participants receive the opportunity to develop a survey design (mode and sampling strategy) for their own research question and to be guided therein. |
| Suggested reading: |
| 2   | **Nonresponse processes, prevention and correction**  
On day two we cover the various types of nonresponse in survey data and how optimizing the data collection processes may minimize them. Specifically, we look at the effects of incentives on nonresponse rates and nonresponse bias. Surveys always contain some degree of nonresponse. The course will show participants how they can correct for unit nonresponse by means of weighting. If time permits, we will talk about the role of interviewers and fieldwork procedures in general to monitor fieldwork. The effect of these treatments on analyzies is demonstrated.  
On this second day, participants design a fieldwork strategy for their survey that aims to minimize coverage and nonresponse errors. |
| Suggested reading: |
| - Groves, R.M. et al. (2009), chapters 6, 10.5 & 10.6 | |
| 3   | **Questionnaire design and data accuracy**  
The third day looks into the survey response process. We assess why operationalizing our research questions into survey questions can be intricate, which survey factors may affect responses and how different respondents differentially understand our questions. Finally, a bit of time is spent discussing effective lay-out of questionnaires, especially in light of designing web and designing or adapting surveys for mobile devices and using new possibilities of collecting smartphone sensor data.  
On day three, participants develop an own short questionnaire taking into account the design principles covered during class. |
<p>| Suggested reading: |
| - Groves et al. (2009), chapter 7 | |</p>
<table>
<thead>
<tr>
<th>Page</th>
<th>Questionnaire testing, mode, interviewers, and cross-national comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Day four continues with questionnaire development. Today, however, we will look into methods of pre-testing survey questions, including the qualitative technique of cognitive interviewing. We will focus on the effects of survey modes on measurement error. If time permits, we will discuss eye tracking as a method to test questionnaires and detect errors and problems in questionnaire design. Finally, if time permits, we discuss how to do surveys in different cultures, with the goal of comparing countries within the framework of Total Survey Error. Course participants will practice the pre-testing techniques learned in class with the questionnaires they developed on day three.</td>
</tr>
<tr>
<td></td>
<td><strong>Suggested reading:</strong></td>
</tr>
<tr>
<td></td>
<td>- Groves et al. (2009), chapter 8</td>
</tr>
<tr>
<td>5</td>
<td>Data preparation: assessing measurement quality, survey quality vs. costs</td>
</tr>
<tr>
<td></td>
<td>On the final day, we look at how to develop and assess measurement errors after we have collected data. This includes topics such as data preparation, coding, and editing. To conclude the course, we look back at the Total Survey Error framework and evaluate how different costs associated with survey design decisions might affect data quality. We discuss the methods that can be used to assess the different components of the Total Survey Error framework and discuss how we may trade-off survey quality against survey costs. Optionally, we can discuss questions of data protection and research ethics in survey research.</td>
</tr>
<tr>
<td></td>
<td>On day 5, students will have an opportunity to present their current or future survey projects in the online class meeting, and to receive feedback from all participants and the instructors on how to improve the survey design. For this, students are encouraged to send details of their future project to the instructors before 11 August. Upon the discussion with the instructors, participants are encouraged to prepare at least two powerpoint slides: (1) the description of the design/problem statement, (2) questions to the participants they would like to discuss.</td>
</tr>
<tr>
<td></td>
<td><strong>Suggested reading:</strong></td>
</tr>
<tr>
<td></td>
<td>- Groves et al. (2009), chapters 2 (again) &amp; 10</td>
</tr>
</tbody>
</table>

**Preparatory Reading:**
No preparatory reading is necessary. We assume participants to have basic knowledge of statistics (descriptive, graphs, and inferential statistics at the level of at least the t-test). During the course, students will find the suggested reading in the course manual helpful in explaining the topics covered in greater detail.

**Additional Recommended Literature:**
Based on your specific research questions and research projects, we will be able to recommend additional literature during the course.