

# Differences between Real and Falsified Data

Dissertation for the achievement of the doctoral degree (Dr. rer. soc.)  
of the Faculty for Social Sciences and Cultural Studies  
of Justus Liebig University Giessen

submitted by  
Uta Landrock  
Dipl. Soz.

2017

First reviewer: Prof. Dr. Jürgen H. P. Hoffmeyer-Zlotnik

Second reviewer: Prof. Dr. Peter Winker

## Content

|     |  |     |
|-----|--|-----|
| 1.  | Introduction .....   | 3   |
| 2.  | Theory .....   | 6   |
| 2.1 | Why do interviewers falsify? .....   | 6   |
| 2.2 | How do interviewers falsify? .....   | 7   |
| 3.  | Hypotheses .....   | 10  |
| 3.1 | Which effects do falsified data have on substantial findings?.....   | 10  |
| 3.2 | In what ways do interviewer effects differ in falsified data, compared to real data? .....   | 12  |
| 4.  | Database and methods .....   | 14  |
| 5.  | Validation of Theoretical Assumptions with Real and Falsified Survey Data .....  | 20  |
| 5.1 | Introduction: Falsifications in surveys .....  | 20  |
| 5.2 | Database .....   | 23  |
| 5.3 | Operationalization and Data Analysis .....   | 25  |
| 5.4 | Results .....  | 27  |
| 5.5 | Conclusion .....   | 31  |
| 6.  | Explaining Political Participation: A Comparison of Real and Falsified Survey Data .....   | 33  |
| 6.1 | Falsification in Surveys.....  | 33  |
| 6.2 | Analyses and results .....   | 36  |
| 6.3 | Conclusion and discussion .....  | 51  |
| 7.  | How Interviewer Effects Differ in Real and Falsified Survey Data: Using Multilevel Analysis to Identify Interviewer Falsifications ..... | 54  |
| 7.1 | Introduction.....  | 54  |
| 7.2 | Hypotheses .....   | 57  |
| 7.3 | Data Base and Methods.....   | 59  |
| 7.4 | Operationalization and Multilevel Regression Model.....  | 63  |
| 7.5 | Results .....  | 70  |
| 7.6 | Conclusions and Recommendations.....   | 76  |
| 8.  | Investigating Interviewer Falsifications – A Quasi-experimental Design.....  | 79  |
| 8.1 | Research Question.....   | 79  |
| 8.2 | State of Research on Interviewer Falsifications .....  | 80  |
| 8.3 | The IFiS Project - Experimental Setup and Data Base.....   | 82  |
| 8.4 | Main Findings of the IFiS Project .....  | 84  |
| 8.5 | Experimental and Quasi-experimental Research Designs.....  | 88  |
| 8.6 | Benefits of Quasi-experimental Controlled Design Implemented by IFiS .....   | 91  |
| 8.7 | Conclusion .....   | 93  |
| 9.  | Summary of results.....  | 95  |
| 9.1 | First application: Explaining healthy food consumption.....  | 95  |
| 9.2 | Second application: Explaining Political Participation .....   | 98  |
| 9.3 | Interviewer effects.....   | 101 |
| 9.4 | Methodological Considerations regarding the quasi-experimental design.....   | 105 |
| 10. | Conclusion.....  | 106 |
|     | References .....   | 113 |
|     | Appendix .....   | 120 |

## List of tables

|  |     |
|--|-----|
| Table 5-1: Regression analyses of real and falsified data – dependent variable: healthy food consumption .....   | 30  |
| Table 6-1: OLS regression analysis – determinants of political participation in real and in falsified data.....  | 49  |
| Table 7-1: Overview of variables used to analyze interviewer effects .....   | 63  |
| Table 7-2: Interviewer effects in the real data (random-intercept-only models, RIOMs).....   | 70  |
| Table 7-3: Interviewer effects in the falsified data (random-intercept-only models, RIOMs).....  | 71  |
| Table 7-4: Results of ML regression in the falsified data (intercept-as-outcome models, IOMs) .....  | 72  |
| Table 7-5: Results of ML regression in the falsified data (intercept-as-outcome models, IOMs) .....  | 74  |
| Table 8-1: Overview over independent variables that are significant either in the real or in the falsified data; dependent variable: healthy food consumption .....  | 86  |
| Table 8-2: Overview over independent variables that are significant either in the real or in the falsified data; dependent variable: political participation.....  | 86  |
| Table 8-3: Overview over the significant effects of interviewer’ characteristics on formal meta-indicators in the real and in the falsified data.....  | 87  |
| Table 8-4: Interviewer effects (i.e., intraclass correlations ICC) in the real and in the falsified data, dependent variables: income, political participation, anomy, healthy eating behavior .....                                   | 88  |
| Table 8-5: Results of ML regression in the falsified data, overview over significant independent variables on the interviewer level; dependent variables: income, political participation, anomy, eating healthy .....                 | 88  |
| Table 9-1: Overview over independent variables that are significant either in the real or in the falsified data; dependent variable: healthy food consumption .....  | 97  |
| Table 9-2: Overview over independent variables that are significant either in the real or in the falsified data; dependent variable: political participation.....  | 99  |
| Table 9-3: Overview over the tested hypotheses regarding the dependent variables: healthy eating and political participation.....  | 100 |
| Table 9-4: Interviewer effects (i.e., intraclass correlations ICC) in the real and in the falsified data, dependent variables: income, political participation, anomy, healthy eating behavior .....                                   | 102 |
| Table 9-5: Results of multilevel regression analyses in the falsified data, overview over significant independent variables on the interviewer level; dependent variables: income, political participation, anomy, eating healthy..... | 102 |

## 1. Introduction

In social science research, face to face interviews are a widely used mode of data collection. Many large and important studies, like the *German General Social Survey* (ALLBUS; GESIS, 2017), the *European Value Study* (EVS, 2015), or the *Programme for the International Assessment of Adult Competencies* (PIAAC; Rammstedt et al., 2016), employ interviewers. Interviewers play a central role: On the one hand, they generally have a positive influence on data quality (see Japtec, 2006) which is why they are often employed. They can, for example, convince respondents to participate in the study (idem), explain or probe into difficult questions, or help the respondent to understand the questionnaire correctly (Mangione et al., 1992). Research shows that, generally speaking, interviewers increase the response rate, the motivation of interviewees to participate as well as the data quality (see Japtec, 2006; Mangione et al., 1992; West and Blom, 2016). But there is also the risk that interviewers depart non-intentionally or even intentionally from their interviewer guidelines and thereby negatively affect the data quality. In the worst case, interviewers even may decide to falsify parts of interviews or entire interviews (Bredl et al., 2013). The *American Association of Public Opinion Research* (AAPOR) defines interviewer falsifications as “(...) intentional departure from the designed interviewer guidelines or instructions, unreported by the interviewer, which could result in the contamination of data” (AAPOR; 2003: 1).

The exact prevalence of falsifications is not known. One common assumption is that the proportion of falsifications in survey data does not exceed five percent (see Koch, 1995; Krejsa et al., 1999; Li et al., 2009; Schnell, 1991; Menold und Kemper, 2014). Nevertheless, under certain circumstances, this amount may be higher, as Bredl et al. (2012) have shown.

We must assume that interviewer falsifications often occur and, if they occur, often remain undetected. Therefore it is important to know which effects falsified data may have on data quality and the results of social science research: It is important, first, for estimating the potential “damage”, that is, the ways in which results from survey research could deviate from reality, under the assumption that parts of the survey should be falsified. It is important, second, because a specific pattern characteristic for falsified data could be a key for identifying data falsifications more often and more easily.

Research on interviewer falsifications identified ambiguous findings. On the one hand there are large similarities between real and falsified data regarding frequencies, means, and proportions (Reuband, 1990; Schnell, 1991; Menold and Kemper, 2014). Particularly with respect to attitudinal and behavioral questions Menold and Kemper (2014) reported only

small differences between real and falsified data. On the other hand there is no doubt that falsifications may lead to biases. These are, for example, differences regarding the magnitude of item nonresponse, extreme and middle response style, acquiescence, or primacy and recency effects (see Bredl et al., 2012; Kemper and Menold, 2014; Menold and Kemper, 2014).

Previous research has also ascertain that in analyses on statistical correlations and causal influences, falsified data may lead to biased results regarding the explained variances and effect sizes, even when the share of falsifications is low (Reuband, 1990; Schnell, 1991; Schraepfer and Wagner, 2003). Schraepfer and Wagner (2003) analyzed data of the German Socio Economic Panel (SOEP) that include interviews which were evidentially falsified by interviewers. They estimated regression analyses on log gross income and reported that the inclusion of falsified data reduces the explained variances and affects the effect sizes. The inclusion of falsified data leads to an overestimation of the effects of age and gender and to an underestimation of the effects of duration of training and working hours per week (Schraepfer and Wagner, 2003).

Schnell (1991) and Reuband (1990) analyzed data fabricated in the lab and reported that falsifications lead to higher explained variances. Schnell (1991) as well as Reuband (1990) used subsamples from existing real datasets and asked their study participants to take on the role of falsifying interviewers. Their task was to invent answers to attitudinal and behavioral questions that were also used in the survey questionnaire. For that purpose the “falsifiers” were informed about basic sociodemographic characteristics that were known from the real datasets, such as gender and age. Afterwards real and falsified data were compared (Reuband, 1990; Schnell, 1991). Reuband (1990) conducted regression analyses and reported that falsifiers overestimate the differences of the provided sociodemographic variables. For the example of attitudes towards abortion, he estimated a regression analysis. In the falsified data, the influence of age and education was higher than in the real data (Reuband, 1990). Schnell (1991) analyzed the effects of subjective self-definition of social stratum, top-bottom-scale, and net income on subjective political competence. In the falsified data, all three explaining variables had significant effects. By contrast, in the real data only two of the three explaining variables were statistically significant (Schnell, 1991).

Given the partly contradictory findings, a better understanding of the differences between real and falsified data must be found to assess the quality of survey data. Particularly, the question

arises under which conditions statistical correlations are overestimated and under which conditions they are underestimated. Also, little is known so far on how the differences between real and falsified data look like when theory-driven assumptions are tested in more elaborate and complex models, such as it is realistic for actual social science research.

This research gap is targeted by the present research: The question my PhD thesis aims to answer is how real and falsified survey data differ with respect to substantial social science research, that means, with respect to results of theory-driven analyses. In other words: Which influence do falsifications have on findings of substantial social science research?

In the work presented here, theory-driven models of statistical correlations and influences are tested. These models are elaborate and designed in a way as they would be for researching adequately a common content-related research question. The central methodological question behind these models is how falsified data affect the consistency of the models, the direction of effects as well as the effect sizes of the determinants. Furthermore, the thesis presents analyses, testing the influence of interviewers' sociodemographic characteristics, attitudes and personality traits on real and on fabricated survey data. These broaden the understanding of the mechanisms leading to the differences between real and falsified data: In addition to the evidence on *how* the two kinds of data deviate they gather evidence on *why* they deviate. Thereby the analyses contribute to the theoretical knowledge on interviewer falsifications and provide potential indications for identifying falsified data.

In the next sections, the theoretical perspectives are described on why and how interviewers may decide to falsify. Afterwards, the hypotheses are derived and the database and methods are introduced. Subsequently, the results from the own empirical research are presented. The empirical work consists of three analyses, each presented in one paper that contributes to my PhD thesis. In the first two analyses, the effects of falsifications on the results of substantial theory-driven analyses are shown: in the first case analyses on healthy eating behavior, in the second case analyses on political participation. The third analysis examines differences between real and falsified survey data with respect to interviewer effects. The fourth paper that contributes to the PhD thesis discusses methodological considerations of the implemented quasi-experimental design. Afterwards, a short summary of the results will be given. In the final section conclusions are drawn with respect to the here raised research questions.

## 2. Theory

### 2.1 Why do interviewers falsify?

Interviewers may decide to intentionally depart from the guidelines and instructions. This decision is assumed to usually be the result of a cost-benefit consideration, based on which the interviewers try to achieve their goals with as few resources as possible. The goals may be, for example, conducting the interviews, achieving a high response rate, or maximizing the remuneration. The resources that need to be invested are time and travel expenditures, but also, for example, cognitive effort. The different tasks of an interviewer may, in that respect, be costly: It may be difficult, for example, to identify and find access to target persons, to motivate target persons to participate, or the average duration of an interview may be time-consuming. Insofar the interviewers balance between the expected benefit and the necessary effort (Sodeur, 2007): The endeavor to realize interviews depends on this cost-benefit calculation and the question to what extent conducting a real interview and to what extent falsifying an interview (or parts of it) promises the most favorable relation between goal achievement and the investment of resources.

Under certain circumstances – when the benefit of following the interviewers’ instructions is considered as low or when there are more promising alternatives to conducting the interviews – interviewers may decide to violate the rules to reduce effort (Sodeur, 2007). Interviewers may decide, for example, to interview another person than the target person, to skip questions in the questionnaire or to manipulate answers to filter questions. In the most extreme case, interviewers may decide to falsify parts of interviews or even entire interviews. What might make the interviewers’ decision easy is that they are usually not interested in the scientific outcomes of a study or in a particularly high data quality (Bredl et al., 2013). Insofar, a rational-choice perspective is likely to model the interviewer behavior adequately.

These theoretical considerations have implications on how to reduce the risk of interviewer falsifications. The design of a study and of the questionnaire as well as the organization of the fieldwork define the opportunities, restrictions, costs and benefits under which the interviewers make decisions to falsify or not to falsify (see, for example, Winker et al., 2015). That means that it is in the hands of the researchers to provide preconditions for the interviewers that make a decision not to falsify easy and likely. An easy access to the field, a study in which target persons are motivated to participate, a well-designed, short and easy questionnaire, or a payment of interviewers by the invested working hours instead of by

completed interviews are examples of how the risk of interviewer falsifications can be reduced. The first and maybe most important insight from research on interviewer falsifications therefore is that the proportion of falsified interviews is presumably, not only but also a result from decisions researchers can make. It may not be possible to completely avoid interviewer falsifications; but it is possible to keep their prevalence low.

## **2.2 How do interviewers falsify?**

The falsifying process itself, that is, the question in what way interviewers decide to falsify survey data, also underlies a cost-benefit calculation: On the one hand falsifiers have to falsify plausible data to prevent detection. That means they have to make effort to minimize the risk of detection (Kemper and Menold, 2014). On the other hand the fabrication of survey data should not be too effortful, since otherwise falsifying will not fulfill its purpose of saving resources. Falsifying will fulfill its purpose as long as the effort to falsify a certain interview is lower than the effort to conduct the interview (Menold et al., 2013). Falsifying interviewers work as carefully as necessary in order to prevent detection. Nevertheless, within the limits of this goal, they try to reduce time and effort as much as possible to save resources (Menold and Kemper, 2014). The optimal balance of these two goals determines their falsification strategy. If the risk of being detected seems too high to falsify an entire interview, this balance may also be achieved by conducting a real interview, skipping certain time-consuming parts of it and falsifying these parts only.

The strategy which optimizes the cost-benefit balance and which falsifiers apply is *satisficing*. The concept of satisficing is usually used to describe cognitive processes of respondents when answering survey questions. Respondents do not try to give a precise answer, but they reduce their cognitive effort and try to give an acceptable answer that satisfies themselves (see Krosnick and Alwin, 1987). (They may, for example, just give a good estimate of their net household income because thinking of the exact number would take them much time and effort.)

Optimizing is the process of searching for the optimal answer and means that the answering process goes through four stages: (1) understanding the question, (2) retrieving the relevant information from memory, (3) evaluating the completeness and relevance of the retrieved information, and finally (4) giving the answer (Tourangeau et al., 2000; Schwarz, 2007; see also Bogner and Landrock, 2016). By contrast, in the case of satisficing the



respondents reduce this cognitive effort: One possibility is that the answering process may be shortened. This is, for example, the case when respondents only give rough estimations of their income instead of retrieving a more precise or even the correct answer. In more pronounced cases of satisficing, respondents may skip one of the four stages completely, for example, when they do not even try to understand the question. Respondents who apply satisficing strategies typically report the first answer that comes to their minds and that promises to be plausible or acceptable (Krosnick et al., 1996).

In the case of falsifying, interviewers may apply satisficing strategies to reduce cognitive effort in a very similar way: They will report the first answer that comes to their minds and that promises to be acceptable for their supervisor or the project leader. One reason why falsifiers apply this strategy is that they have only little information about the respondents, namely the information from the responses to the survey questionnaire. An optimizing falsifying strategy would require very difficult and time-consuming investigations about the respondents and very likely be much more effortful than conducting the interview honestly. Hence, falsifiers – like respondents – will apply satisficing strategies, particularly, when the risk of an imprecise answer being detected is low and when the cognitive costs for ascertaining the optimal answer are high (see Krosnick and Alwin, 1987). Consequently, falsifiers will limit their effort to find an acceptable answer. Falsifying interviewers consider answers as acceptable when they appear to be plausible (Krosnick et al., 1996; Krosnick and Alwin, 1987; Menold et al., 2013). Then falsifiers assume that also researchers do not doubt the reported answers, and their falsification remains undetected (Menold et al., 2013).

Although there are similarities between real respondents and falsifying interviewers with respect to satisficing, there is also an important difference: Falsifiers have to prevent detection to avoid sanctions. Consequently, falsifying interviewers show weaker satisficing than survey respondents (Menold et al., 2013). Falsifiers make, for example, more effort to understand the question correctly and think a bit longer, whether the answer that is coming in their mind could be realistic. By contrast, respondents do not bear the risk of detection, because even in doubtful cases their answer will be accepted as correct. And even if not, there are little sanctions they have to fear. Therefore, the respondents have stronger tendencies to reduce effort than falsifying interviewers (Menold et al., 2013; Kemper and Menold, 2014; Menold and Kemper, 2014). This may lead to respondents giving less consistent answers in comparison with falsifying interviewers (Menold and Kemper, 2014; Reuband, 1990). This also implies that, if falsified data appears to be more consistent than real data, as it was the

case in the research of Schnell (1991) and Reuband (1990), this may not only provide insights on how data is falsified. It may also point towards another problem with respect to data quality, independent from falsifications: Data based on correctly conducted interviews may be affected by imprecise or wrong answers given by real interviewees. For example, real respondents tend to show a stronger extreme response style than falsifiers, as Kemper and Menold (2014) report. When comparing real and falsified survey data and assessing the effects of interviewer falsifications on data quality, it has to be kept in mind that the data quality not only depends on the interviewers but also on the respondents.

Previous research has shown that falsifiers are able to invent plausible answers even when they have only basic information about respondents (Reuband, 1990; Menold et al., 2013). Reuband (1990) reported that most of the falsifying interviewers apply stereotypes to fabricate interviews. To invent plausible answers with limited effort falsifiers rely on *implicit everyday knowledge and general stereotypes* on how people would behave (Reuband, 1990; Schnell, 1991). Thus, falsifiers are able to reproduce the means and proportions surprisingly exactly (Schraepfer and Wagner, 2003; Schnell, 1991; Reuband, 1990). The application of stereotypes and everyday knowledge allows the falsifiers to complete the questionnaire like the respondent would do, even with only little sociodemographic information about the respondent (Reuband, 1990; Menold et al., 2013). Furthermore, the falsifiers seem to be able to reproduce the correlations between these sociodemographic characteristics and attitudes or behaviors, simply because they are aware of them or because they share stereotypes that correspond to existing correlations. Falsifiers even overestimate such sociodemographic differences. That may lead, as mentioned above, to higher explained variances in the falsified data compared to real data (Reuband, 1990). Although – or because – falsifying interviewers usually have very little information about the respondent they develop more consistent models of behavior, by consequently applying their stereotypical knowledge.

Additionally, falsifying interviewers answer the survey questions more carefully than real respondents and take greater care in avoiding contradictions between the answers; this also leads to more consistent models in falsified than in real data (Kemper and Menold, 2014; Menold and Kemper, 2014). By contrast, real respondents often report inconsistent attitudes and behaviors (Reuband, 1990). Therefore, the implicit models of behavior that the falsifiers apply may lead to a higher consistency of the falsified interviews (Schnell, 1991). The results

of regression analyses give evidence: Falsified data tend to show higher explained variances than real data (Reuband, 1990; Schnell, 1991).

Despite of incomplete information, implicit or explicit knowledge allows the falsifiers to produce consistent data (Reuband, 1990). According to Reuband (1990) this knowledge is generated and shared via communications and interactions: In everyday communication with friends, relatives, or peers people learn how society works. Thus, a basic knowledge is generated which allows people to draw conclusions about attitudes and behaviors of persons on the base of their sociodemographic characteristics (Reuband, 1990).

### **3. Hypotheses**

#### **3.1 Which effects do falsified data have on substantial findings?**

The following empirical analyses assess differences between real and falsified data. The first two empirical analyses aim to answer the question which effects falsified data have on substantial findings of social science analysis. Social science research usually aims to test theoretical assumptions and hypotheses that were deduced from complex theories. Therefore the question arises whether falsifying interviewers – who, as mentioned above, are able to invent consistent patterns of answers to survey questions – also may produce data that meet the predictions of established social science theories, tested in elaborate complex multivariate statistical models for explaining dependent variables. In other words: The question is whether interviewers apply actual social science theories. Are the effects of falsification on results from data analyses different if substantive, content-related, theory-driven multivariate analyses are compared, instead of less complex analyses?

The underlying assumption is that falsifiers do *not* understand the interrelations described by complex social science theories. That would imply that the distinction between more stereotypical interrelations, as they are experienced in everyday life and have found their way into everyday knowledge, on the one hand and theoretically predicted interrelations on the other hand may be decisive for whether or not falsified data shows stronger correlations than real data. If this should be the case, this distinction would considerably improve our understanding of how falsifications affect data. In order to test this assumption not only theoretically predicted determinants but also independent variables that refer to stereotypical relations are analyzed in the here presented empirical work.

Thus, in the first two empirical analyses the results of theory-driven multivariate regressions should be analyzed, always comparing real and falsified data based on identical models. Realistic content-related and theory-driven social science research is conducted to assess the influences of interviewer falsifications. Two examples for such content-related research are investigated, rather than merely one, in order to generate a minimum of reliability of the findings. Two cases are, of course, the minimum number for a test series and far away from ensuring that the findings can be generalized. However, they provide a first chance for a comparison and for checking how robust findings are. If findings are similar in at least two examples of content-related, theory-driven social science research then the assumption seems appropriate that these findings can be generalized. The two examples investigated in the following are the determinants of eating healthy on the one hand and those of political participation on the other hand. These two examples correspond to actual branches of social science research. They can draw on established social science theories that are supported by a number of empirical studies. And they can be operationalized by variables in the data at hand. Based on these two examples of content-related, theory-driven social science, the following four hypotheses should be tested:

H1.1: For the theoretically predicted determinants there are more significant effects in the real than in the falsified data.

Falsifying interviewers probably do not know the underlying theoretical base of a study. Thus, they should in the majority of cases not be able to reconstruct the theoretically predicted correlations. Consequently, in the real data more and stronger statistically significant effects of the theory-driven explaining variables should occur than in the falsified data.

H1.2: For real sociodemographic information known to falsifying interviewers, there are more significant effects in the falsified than in the real data.

Reuband (1990) reports that falsifiers overestimate sociodemographic differences between respondents. Furthermore, falsifiers have only little information about the respondents and their living situations. So, if they receive any basic sociodemographic information, such as age or gender, this information provides their only chance to improve their guessing of which answers the respondent would have given. Therefore, falsifying interviewers overestimate the relevance of these sociodemographic characteristics. Accordingly, the effects of the

sociodemographic variables that were provided to the falsifiers should be stronger in the falsified than in the real data.

H1.3: For further correlates (either based on common stereotypes or found in research), there are more significant effects in the falsified than in the real data.

As would have been done in actual social science research, not only theory-driven explaining variables are included in the analyses but also independent variables that either seem to correspond with stereotypes and everyday knowledge plausibly or that have been proven as relevant in specific contexts in previous studies. Falsifiers apply stereotypes and implicit everyday knowledge to construct relations and therefore overestimate the influence of these further correlates. The real respondents, by contrast, apply stronger satisficing and offer less consistent models of behavior. Thus, statistically significant effects of further correlates can be expected to occur more often in falsified than in real data.

H1.4: The falsifiers construct more consistent cases than real life; therefore, the explained variance is higher in the falsified than in the real data.

Falsified survey data follow more stereotypical and therefore less complex assumptions than empirical reality. Also, real respondents show stronger satisficing than falsifying interviewers. Consequently, the regression models of explaining behavior should be more consistent in falsified data than in real data – the explained variances should be higher in the falsified than in the real data.

### **3.2 In what ways do interviewer effects differ in falsified data, compared to real data?**

The third empirical analysis focusses on possible differences regarding interviewer effects in real and falsified survey data which may be used to identify interviewer falsifications. Obviously an interviewer has a particularly strong impact on the reported answers in an interview if she or he gives these answers her- or himself. This may mean that interviewer effects in falsified data should be stronger than in real data. Furthermore the ways in which this direct influence affects the data and the correlations within the data may be very different from the ways interviewers affect actual answers in real interviews, as it is usually described by the concept of “interviewer effects”. For interviewer effects in this stricter sense it has been shown, for example, that they are larger among experienced interviewers than among inexperienced interviewers (Olson and Bilgen, 2011). This may or may not be true for the

influence of falsifying interviewers on the data. Accordingly the third empirical analysis compares interviewer effects in real and in falsified data. The following two general hypotheses were tested:

H2.1: Interviewer effects occur both in real and in falsified data.

Interviewer effects may occur in real fieldwork settings when characteristics and behaviors of interviewers influence the responses of the respondent (see Groves and Magilavy, 1986). I assume that interviewer falsifications may be considered as extreme form of interviewer effects, because the falsifying interviewer influences the responses directly. Thus, in real as well as in falsified data interviewer effects can be expected to occur.

H2.2: The interviewer effects in falsified data are larger than in real data.

If falsifying responses to survey questions is considered as extreme form of interviewer effects, interviewer effects in falsified data should be clearly larger than in real survey data.

Furthermore, more specific hypotheses were tested with respect to certain interviewer characteristics that are prone to lead to interviewer effects. They are each tested for real and for falsified data separately:

H2.3a: The core sociodemographic characteristics of the interviewers affect the reported responses.

As reported in literature (see, for example, West and Blom, 2016), interviewers' characteristics like gender, age, and education and additionally income of the interviewers are expected to cause interviewer effects.

H2.3b: The magnitude of interviewer effects depends on the interviewer's experience.

Following Olson and Bilgen (2011) I assume that experienced interviewers show stronger interviewer effects than interviewers without experience.

H2.3c: Associations exist between the behaviors and attitudes of interviewers and the reported behaviors and attitudes of the respondents they interview.

Schanz (1981) found associations between the respondents' answer and the interviewers' answer to the same survey question, that is, in real interviews the respondent may be affected

by interviewer-specific social expectations. For falsified data it is even more plausible to assume such an association since a falsifying interviewer might take his own attitudes and behavioral routines as inspiration to invent plausible answers. This hypothesis therefore aims to test whether the interviewers' own response to a survey question affects the reported response of the respondent to the same question.

H2.3d: The occurrence and magnitude of interviewer effects depends on the personality traits of the interviewer.

I expect that personality traits of interviewers, such as extraversion, self-confidence, conscientiousness, and self-efficacy, may impact the occurrence and magnitude of interviewer effects (West and Blom, 2016; Winker et al., 2015).

H2.3e: The magnitude of interviewer effects depends on the interviewer payment scheme used (payment per completed interview vs. payment per hour).

As reported by Winker et al. (2015) payment per completed interview may, in contrast to payment per hour, cause or foster interviewer effects. Consequently, interviewer effects should be stronger when interviewers are paid per completed interview and not per hour.

#### **4. Database and methods**

For the analyses data of the research project "IFiS – Identification of Falsifications in Surveys" are used. This project was funded by the German Research Foundation (DFG) and conducted by Prof. Dr. Peter Winker from the University of Giessen and Dr. Natalja Menold from GESIS Mannheim. The IFiS project aims to research strategies and methods to identify interviewer falsifications. A quasi-experimental design was applied because datasets with proven falsified interviews rarely exist (Winker et al., 2015).

The database consists of three datasets that were collected in summer 2011. For that purpose, 78 interviewers were recruited among students at the campus of the University of Giessen. These 78 interviewers conducted 710 real face-to-face interviews. The respondents were students at the University of Giessen as well. The interviewers recruited the respondents without any quota restriction. All interviews were audio recorded and checked afterwards to assure that all interviews were conducted correctly and to eliminate the possibility that interviews or parts of interviews were falsified. Every interviewer conducted 9 interviews on

average. Half of the interviewers were paid per completed interview (8 Euros), the other half were paid per hour (12 Euros). The payment scheme was randomly assigned. The average interview duration was 30 minutes. The questionnaire consisted of 62 factual, attitudinal, and behavioral questions about political, economic, and social themes. The items were mainly adopted from ALLBUS, the German General Social Survey (Koch et al., 1999).

In a second step, the same interviewers falsified survey data in the lab. Similar to the approach of Reuband (1990) the falsifying interviewers received a short description of real respondents that had been interviewed before (but not by the same interviewer). The description informed about core sociodemographic characteristics. The information was of that kind that a falsifying interviewer in a real fieldwork setting could have obtained easily with a short interview with the respondent: sex, age, studied subject, number of semesters enrolled, marital status, place of residence, living situation, and country of origin. The interviewers were instructed to fill in the questionnaire like the described person probably would have done.

The exact instruction was:

Please read carefully the description of the person whose interview you are to falsify. Please complete the attached questionnaire as if you had really conducted a personal interview with the respondent. During falsification, please place the description of the respondent next to the questionnaire, so that you are always aware of the characteristics of that person.

The person whose interview you are to falsify...

- is female,
- is 20 years old,
- studies teaching,
- is enrolled in her second semester at a university.
- She is unmarried, in a steady relationship,
- lives in Huettenberg, a rural village in Hesse,
- with her parents or relatives.
- Country of birth: Germany.



This procedure allowed capturing 710 falsified interviews; to each real interview a corresponding falsified interview was collected. Again, half of the falsifying interviewers were paid per completed falsified interview (3 Euros), the other half per hour (9 Euros). As additional incentive to enhance the motivation of the interviewers, a lottery game was implemented: Three of the interviewers whose falsifications remained undetected won 100 Euros each (see Kemper and Menold, 2014). The interviews were fabricated after conducting the real interviews. That is consistent with a real fieldwork setting where a falsifying interviewer would typically conduct a few interviews before deciding to falsify survey data.

At last, the interviewers filled in the survey questionnaire for themselves. They additionally answered questions about their falsifying strategies. This approach allows collecting much information about the interviewers that can be used, for example, to analyze interviewer effects.

The strength of this quasi-experimental design lies in the fact that data is obtained that are *confirmed* real or *confirmed* falsified. Furthermore, there was control for the number of falsified interviews. The design ensured not only equal sample sizes (of 710 interviews) in each of the two groups, but even an identical sociodemographic composition, so that real and falsified data are directly comparable to each other: Since the sociodemographic profiles of the real study participants (as described above) were used also for the second part of the study in which interviewers invented the answers for study participants, the composition of the two groups according to these profiles is identical. A particularly relevant strength of this quasi-experimental design with instructed falsifiers is that there was full control over and transparency of the falsifying process: It is known precisely who the falsifiers are and which information they had available for executing their task.

An obvious limitation of the applied quasi-experimental approach and the data used is that the interviewers as well as the respondents were all students. Thus, the interviewers are familiar with the living situation of the respondents. That may lead to smaller differences between real and falsified survey data because the student interviewers are probably more able to imagine how a student respondent would answer the survey question. Furthermore, for the same reason, there is only little variation regarding age and education of respondents and interviewers (Winker et al., 2015).

What also can be regarded as a limitation of the database used in the following is the artificial situation of the falsifying process: The interviewers have not decided to falsify based on a calculation of costs and benefits, as theoretically described in the previous section (2.1). The design of the study, the length of the questionnaire, the difficulty of field access or the interviewers themselves are not related to the fact that the here assessed interviewer falsifications have occurred. Merely the *instruction* to falsify is the reason for the occurrence of falsifications. In that sense the falsifying interviewers acted very responsibly and in complete accordance with the instructions which they had been given; and that means: they acted differently from how real falsifiers – who in fact depart from instructions – would act in a real fieldwork setting.

The question must be raised whether or not the so generated interviewer falsifications are comparable to real falsifications and whether the findings based on instructed falsifications can be generalized for all interviewer falsifications. This question is discussed in the fourth and last paper. It raises and discusses reasons why the results from quasi-experimental designs, like the one used in the following, might or might not be generalizable. It also introduces means taken to minimize the discussed limitations, such as a lottery game among those falsifying interviewers that are not detected, as a motivation for generating falsifications that cannot be identified easily. In the end, when interpreting the findings from the following three empirical papers, it must be clear that instructed falsifications are not the same as real falsifications in real fieldwork settings. But there are convincing reasons to assume that they are comparable enough to trust the results and consider them as generalizable. The underlying theories used for implementing the research design, for deducting hypotheses as well as for reflecting findings increase the chance that findings are generalized appropriately. (“Appropriately” means that there are good reasons to assume that the findings would hold for real falsifications in real fieldwork settings.) Furthermore, there is no proof that actual falsifiers in a real fieldwork setting would falsify interviews in a different way than the participants in the study at hand. Finally, it also must be considered that there are probably no better alternatives of data collection for research on interviewer falsifications than instructed falsifications.

Based on this data, the following chapters aim to answer the research question, which influence do falsifications have on findings of substantial social science research; or more

concrete, which differences exist between real and falsified data with respect to the results of substantive, theory-driven multivariate analyses.

In **chapter 5 (1<sup>st</sup> publication)** possible differences between real and falsified survey data were investigated on the example of explaining healthy eating behavior, applying the theory of planned behavior (TPB). The theoretically predicted determinants are intention and perceived behavioral control as well as attitudes and subjective norms towards healthy food consumption. The sociodemographic variables prescribed to the falsifiers (like gender, age, living situation, or income) serve as control variables. Thus, one is able to answer the question on which information the falsifiers rely while fabricating interviews. Leisure activities (like TV-consumption or doing sports) and characteristics like BMI and the preference of healthy desserts were included to ascertain whether falsifiers apply stereotypes and construct stereotypical correlations. The focus lies on comparisons of effect sizes and explained variances between real and falsified data.

**Chapter 6 (2<sup>nd</sup> publication)** aims to research whether the findings of chapter 5 (1<sup>st</sup> publication) can be confirmed and hence might be considered as robust. The second application of researching differences between real and falsified data with respect to theory-driven multivariate analyses uses the example of explaining political participation. In contrast to the first case of healthy food consumption, where the application of the TPB and the operationalization of dependent and explaining variables were developed particularly for this study, the case of explaining political participation is grounded on concepts and variables that are well approved in many studies and since decades of social science research. A further difference between the two applications consists of the assumption that explaining political participation is more complex than explaining healthy eating behavior; in the latter case some correlations seem to be more intuitively guessable by non-social scientists. An instrumental approach is used to explain political participation. The determinants of political participation are dissatisfaction with the political, economic or social situation as well as the political efficacy, that is, the perceived influence on political decisions. (The existence of norms of political participation and social incentives for political engagement could not be included in the analyses because the corresponding items are missing in the questionnaire.) Additionally – as in the case of explaining healthy eating – the control variables provided to falsifiers (age and gender) and further correlates found in research (self-placement on the left-right-dimension, self-reported social class, TV-consumption and attractiveness of respondent,

reported by the interviewer) were included in the analyses. Also in this second application the effect sizes and explained variances between real and falsified data were compared.

In these first two empirical chapters multivariate OLS regressions were estimated to analyze the effects of falsifications on the results of substantial theory-driven research. Real and falsified survey data were analyzed separately; afterwards the results were compared.

**Chapter 7 (3<sup>rd</sup> publication)** investigates further sources of differences between real and falsified survey data, namely the influence of the interviewers and their characteristics and behaviors on the reported responses. In real as well as in falsified data one can expect that interviewer effects may occur. The selected dependent variables are income as an example for a sensitive and open-ended question, political anomy as an attitudinal question as well as healthy eating and political participation as two examples for behavioral questions; these variables are known to be prone to interviewer effects. The independent, content-related variables on the respondents' level serve as control variables. The explaining variables on the interviewers' level are known to potentially cause or lead to interviewer effects. These are particularly the interviewer's gender and experience, but also certain personality traits of the interviewer as well as his or her own attitudes and behaviors. Additionally the effects of the applied payment scheme (per hour or per completed interview) on the reported responses shall be investigated.

To analyze differences regarding interviewer effects in the third empirical paper, multilevel regression analyses were conducted separately for real and for falsified data; afterwards the results were compared.

**Chapter 8 (4<sup>th</sup> publication)** focusses on the question whether the applied methodological approach is suitable and appropriate for research on interviewer falsifications. Potentials and limitations of quasi-experimental research designs were discussed to allow statements on the generalizability of the reported findings.

## 5. Validation of Theoretical Assumptions with Real and Falsified Survey Data<sup>1</sup>

**Abstract:** Falsification of survey data in face-to-face surveys has been intensively discussed in the literature. The results about the impact of falsifications on survey data are equivocal. While some authors report a strong impact, others find only little differences between real and falsified data. We argue that the impact of falsifications cannot be neglected, particularly when theory-driven analyses are conducted and not ad hoc analyses. The latter reproduce stereotypes used by both, researchers and falsifiers. To test this assumption we compare the results of multivariate regression analyses with real and falsified data by using a) theory-driven predictors and b) ad hoc predictors. As an example of theory-driven analyses we used the theory of planned behavior (TPB) for predicting self-reported healthy eating behavior. As ad hoc predictors we included sociodemographic information about the respondents known to the falsifiers as well as variables, which are indicated by everyday theories. The results show that theory-driven relationships were more strongly pronounced in the real data. In contrast, stereotypical and non-theory-driven relationships were more strongly pronounced in the falsified data. The results provide insights in the area of social cognition when predicting the behavior of others.

### 5.1 Introduction: Falsifications in surveys

Face-to-face interviews are a widely used mode of data collection. The assistance of interviewers can enhance data quality, since they can help ensure that survey questions are correctly understood (Mangione et al., 1992). On the other hand, there is the risk that interviewers intentionally depart from the instructions and falsify parts of an interview or even the entire interview (Bredl et al., 2013). Interviewers want to save effort and time and therefore they sometimes decide to falsify (Sodeur, 2007). In addition, it is sometimes a difficult task to obtain participation and an interviewer may try to solve this problem by falsifying data (Turner et al., 2002). Different authors indicate that the proportion of falsified interviews in surveys usually does not exceed five percent (e.g., Koch, 1995; Krejsa et al. 1999; Li et al., 2009). However, under certain circumstances the percentage of falsifications

---

<sup>1</sup> Already published: Landrock, Uta and Menold, Natalja (2016): Validation of Theoretical Assumptions with Real and Falsified Survey Data. *Statistical Journal of the IAOS* 32(3): 305-12. The final publication is available at IOS Press through <http://dx.doi.org/10.3233/SJI-161020>.

may be higher: This may be, for example, the case if a survey does not apply extended field control procedures (Bredl et al., 2012) or in the case of duplication of valid cases (Kuriakose and Robbins, 2016; Koczela et al., 2015). In any case, the spread of falsifications makes it important to determine how falsified data may impact the results of analyses of survey data.

With respect to the impact of falsifications, large similarities between real and falsified data have been reported regarding proportions and means (Reuband, 1990; Schnell, 1991). Menold and Kemper (2014) identified only small differences in means and proportions between real and falsified data, especially in attitudinal and behavioral questions. However, falsifications may have a strong impact on the results of multivariate analyses. Schraepler and Wagner (2003) provided an example with data of the German Socio Economic Panel, where the inclusion of falsified data strongly reduced the estimated effects in a multivariate regression analysis. They calculated a linear regression on the log gross income. In the sample, which included fabricated data, the results were biased as compared to those in the sample with real data. The effects of age and gender were overestimated and the effects of duration of training and working hours per week were underestimated in the data when including falsifications (Schraepler and Wagner, 2003). Finally, Schraepler and Wagner (2003) found higher adjusted R-squared in the real sample than in the sample that included real and fabricated data.

In other studies falsifications consistently led to higher explained variances in regression analyses (Reuband, 1990; Schnell, 1991). Reuband (1990) as well as Schnell (1991) used artificially produced falsifications. They drew a subsample from an existing real data set and asked their study participants to take on the role of falsifiers. The task of the falsifiers was then to produce data in response to survey questions, using some of the demographic information of respondents derived from the existing data set. Then, falsified and real data were compared. In an example of the attitude towards abortion, Reuband (1990) conducted a regression analysis with age and education as independent sociodemographic variables and came to the conclusion that in the falsified data the effects of these determinants and the explained variance were higher than in the real data (Reuband, 1990). Since Reuband (1990) provided sociodemographic variables as information falsifications are supposed to be based on, falsifiers seemed to rely on this information and to produce data in which information given to falsifiers resulted in significant relationships with other variables.

Falsifiers may also be able to strongly predict some relationships. Schnell (1991) used a multiple regression analysis to examine the effects of subjective self-definition of social

stratum, top-bottom-scale, and net income on subjective political competence. In the real data, two of the predictors were significant, while in the falsified data, all three predictors were significant. The explained variance was higher in the falsified than in the real data (Schnell, 1991). It seems that falsifiers produced more consistent results when providing answers to related questions.

As shown above the results of previous studies point out that falsifiers seem to be able to produce means and distributions, which are comparable with real data. However, previous research does not provide a clear picture with respect to the question of which relationships can be expected to be strongly pronounced in falsified data and which not. In this article we address this question and test the following assumptions about the results of multivariate analyses one can expect to differ between real and falsified data.

Firstly, we expect falsifiers to not be able to predict relationships in the data which are driven from complex theoretical models. To test these assumptions, we look for an established and reasonable empirically tested theory. In the present article, we apply the theory of planned behavior (TPB, Ajzen, 1991) to identify differences between real and falsified data. The TPB was used in our study to predict healthy food consumption.

Previous studies did not use a priori theoretically driven models to predict the relationships but rather used ad hoc models developed on the basis of available data. Such ad hoc models can also be plausible for laymen so that interviewers may be able to provide data which are consistent with everyday theories about social life and society, even with a higher consistency than in the real data. Therefore, in addition to theory-driven predictions, we also compared effects and correlations that were not predicted by a social science theory but seemed plausible according to common stereotypes. In doing so, we aim to replicate the results found by Schnell (1991). Regarding the term “stereotypes”, we follow the Hilton and von Hippel (1996) definition, stated as “the standard viewpoint that stereotypes are beliefs about the characteristics, attributes, and behaviors of members of certain groups” (Hilton and von Hippel, 1996: 240).

Third, we would like to show that falsifiers strongly use sociodemographic information about respondents when providing falsifications which is plausible to assume when considering the results obtained by Schraepfer and Wagner (2003) and by Reuband (1990), described above. For our analyses, we assume to obtain stronger relationships between the sociodemographic variables in falsified than in real data.

When addressing the question, which information is used by falsifiers while producing data we research cognitive falsification strategies. Considering these strategies contributes to the area of social cognition, which encompasses “(...) the mental processes involved in perceiving, attending to, remembering, thinking about, and making sense of the people in our social world” (Moskowitz, 2005: 3). With respect to social cognition the results can help to understand which kind of information falsifiers predominantly use for predicting the respondents’ responses, e.g. on opinions and self-reported behavior.

Furthermore, the research on falsifications in surveys and the impact of falsifications on data analyses is important for the image of survey data and its reliability and usability in society.

## **5.2 Database**

Our database consisted of two datasets collected in 2011 in an experimental study. The study and the procedure were described by, e.g., Menold et al. (2013) and Kemper and Menold (2014)<sup>2</sup>. For the first dataset, 39 interviewers conducted 365 real face-to-face interviews. The interviewers and the respondents were students at the University of Giessen, and they were recruited on the campus.

Among the interviewers, 69% were female. The interviewers’ mean age was 25.46 years (SD = 2.45). They studied social sciences or psychology (56.4%), language and art (12.9%) and other disciplines such as biology, chemistry, medicine and economic sciences (30.7%). The interviewers were, on average, in their fifth semester (SD = 2.97). Among the respondents, 60.8% were female, and the average age was 26.06 years (SD = 2.45). The respondents studied social sciences or psychology (30.4%), language and art (23.5%), natural sciences (13.1%) and other disciplines such as engineering, medicine or economic sciences (33%). The respondents were also, on average, students in their fifth semester (SD = 3.1). All interviews were audio-recorded to assure that they were actually conducted and not falsified. The questionnaire contained 62 questions on attitudes and behavior and sociodemographic

---

<sup>2</sup> We use a subsample from this study. We used only data collected by GESIS, the Leibniz Institute for the Social Science (n = 730), in which predominantly social science students participated. In the other part of this study, collected by the ZEU (Center for international Development and Environmental Research) participants are more heterogeneous. The results for the entire data set are comparable to those presented in this article. However, some relationships are not as strong as in the subsample we used here, which we explain by the fact that the students in the two samples differed with respect to their study discipline. The results for all data are available on request by the authors.



information, mainly taken from the German General Social Survey (ALLBUS<sup>3</sup>) and discussing political, economic and social topics. A payment either per interview (8 Euros) or per hour (12 Euros) was provided to the interviewers (e.g., Winker et al., 2015).

For the real respondents, an incentive of two Euros was provided. On average, each interviewer conducted 9.1 (SD = 1) interviews. An interview took approximately 30 minutes.

For the second dataset, the same interviewers fabricated survey data in the lab so that a data set of 365 falsified interviews corresponding to each of the N = 365 real interviews was obtained. These falsifications were done after conducting the real interviews. As a consequence the falsifiers have a better idea of the typical responses interviewees would give. This is comparable to a real field setting, where the interviewers would typically conduct a few true interviews before they falsify interviews partially or completely.

For the falsifications, either 9 Euros per hour or 3 Euros per falsified interview were paid. The falsifying interviewers received basic sociodemographic information about the real survey participants, for whom they were supposed to invent data and who were interviewed by a colleague (not by themselves) in the first part of the study. The information the interviewers was given was of such kind a falsifying interviewer could easily have researched himself with a short contact, e.g. by phone: sex, age, studied subject, number of semesters enrolled, marital status, place of residence, living situation (with whom the respondent lives together in a household), and country of origin (e.g., Menold et al., 2013). The interviewers were instructed to falsify the data as if it were collected in a real survey setting. Interviewers who took the role of falsifiers were further instructed to imagine a person according to these basic characteristics and to answer the questionnaire as this person would in a face-to-face interview. The instruction was:

Please read the description of the person, whose interview you are supposed to falsify, carefully. Please fill in the attached questionnaire as if you had conducted a personal interview with the respondent in reality.

---

<sup>3</sup> German General Social Survey: <http://www.gesis.org/en/allbus/allbus-home/>

### 5.3 Operationalization and Data Analysis

We compared the outcomes of multivariate analyses for falsified and real data using the example of healthy food consumption. Our dependent behavioral variable measured healthy food consumption in days per week on which the respondent reports healthy eating. The question in the questionnaire was: “On how many days per week do you eat healthy?”

Our explaining variables followed the theory of planned behavior (Ajzen, 1991; Ajzen and Krebs, 1994). The TPB has often been used in social science research with respect to a variety of topics (Ajzen, 1991) and claims to predict a certain behavior with the help of four determinants. The TPB assumes a two-step causal influence. Behavior is directly influenced by the actor’s intention to act, mediated by their perceived behavioral control. One step before, the actor’s intention is influenced by their attitudes, by their subjective norms and by their perceived behavioral control. These determinants thereby influence behavior indirectly. Ajzen (1991) refers to attitude as to the actor’s expectation that a behavior will have certain consequences, combined with their evaluation of these consequences as positive or negative. Subjective norm is the perception that other people expect a certain behavior, combined with the motivation to fulfill these people’s expectations. The perceived behavioral control is the actor’s belief that he or she is actually capable of showing the respective behavior.

In the survey, the intention was operationalized by two items:

In the future I will eat healthy at least 4 days per week.

In the following weeks I will eat healthy at least 4 days per week.

The attitude regarding healthy food consumption was measured with three items that asked the participant to provide an evaluation of healthy food consumption as good, useful or advantageous:

It would be (bad/good) if I ate healthy at least 4 days per week.

It would be (useless/useful) if I ate healthy at least 4 days per week.

It would be (advantageous/disadvantageous) if I ate healthy at least 4 days per week.  
(reversed item)

Two statements on how friends and how important others evaluate healthy food consumption measured the subjective norms:

My friends would approve if I ate healthy at least 4 days per week.

People who are important to me would like it if I ate healthy at least 4 days per week.

The perceived behavioral control was also measured with two items. The first was about the thinking that there is the possibility to eat healthy and the second was of being convinced that to eat healthy is completely in one's own hands:

It is possible for me to eat healthy at least 4 days per week.

It is completely in my own hands to eat healthy at least 4 days per week.

All above-mentioned items were measured with 7-point rating scales. This operationalization was needed, because the TPB is a frame theory which describes the relationship between any attitudes and corresponding behaviors. Therefore, researchers have to provide operationalization for a concrete kind of behavior by themselves. For the operationalization of the items described above we followed strongly the examples given by Ajzen and Krebs (1994).

We separately used both datasets, real and falsified, to test the postulated relationships in terms of regression analyses. The differences between the real and falsified data were inspected with respect to the explained variances as well as the strength and direction of the single effects. We expected that the overall explained variance is higher in the real data, when only the TPB variables were included (cp. model 1). We also expected stronger theoretically predicted relationships in the real than in the falsified data. These predictions correspond to our expectation that falsifiers are not able to predict rather complex theoretically driven assumptions.

In a next step we expected to see the impact of the information about the respondents that was provided to the falsifiers (cp. model 2). This information pertained to gender, age, the living situation, and the relationship status. Since we expected that falsifiers rely on this information when providing falsified data, the corresponding variables are to be expected to be more strongly related to healthy food consumption in the falsified data than in the real data. In model 2 these variables were added to the model 1 which included variables regarding TPB. We expected to obtain significant relationships of personal variables as well as an

increase of explained variances in model 2 in the case of falsified data but not in the real data case.

To see how everyday knowledge and stereotypes influence the results of prediction of healthy food consumption, we additionally used other sociodemographic variables, such as income, vocational education and training, and seeing university as center of activities in model 3. Here, we expected that such information would have a greater influence on the falsifiers' fabrication of responses compared to the effect the characteristics have in the real data. In addition, we expect a higher predictive power of those variables in the falsified than in the real data. We also added variables on TV consumption (measured in minutes per day), the preference of healthy desserts, participating in sports, and the interviewee's BMI (body mass index, measured by self-reported height and weight of respondents) and obtained the final model 3. The favorite dessert was derived from a list of 12 desserts. We differentiated between healthy desserts (fruit curd, fruit salad or yogurt) and unhealthy desserts (mousse au chocolate, tiramisu, chocolate pudding or pancakes). We assumed that falsifiers used the answers they had invented for these questions about TV consumption, BMI, doing sports and preferring healthy desserts when trying to invent plausible answers to the questions regarding healthy eating behavior by using stereotypes. This information should be irrelevant for the prediction of healthy eating behavior when applying TPB. However, according to stereotypes, preferring unhealthy desserts, watching a lot of television, being obese and not participating in sports would typically be assumed to correlate with unhealthy food consumption. Therefore, we expect these variables to be more relevant in the falsified than in the real data. Therefore, the use of these variables in the final model 3 was expected to lead to an increase of explained variances in the false data but not in the real data.

To compare the outcomes of analyses, we calculated identical regression models for the false and real data. SPSS 22 software was used for the analysis.

## 5.4 Results

Table 5-1 shows the results of regression analyses in the real and the falsified data. First we look at the real data; in model 1, all of the theory-driven variables showed significant effects on the reported eating behavior<sup>4</sup>. The correlations of healthy food consumption with the subjective norm, perceived behavioral control, and the intention were positive, as expected. A

---

<sup>4</sup> We also checked the values of the VIFs and tolerances, the values are not problematic concerning multicollinearity.

result not congruent with the assumptions of TPB was that attitude had a negative relationship with healthy eating behavior, meaning that the more respondents believed that healthy food consumption was good and positive, the fewer days of healthy eating they reported. This result can be explained by the fact that students who thought their eating behavior is not healthy evaluated healthy food consumption even more valuable than students who reported to eat healthy. The continuous effects of the theory-driven variables could be proven in all three models in the real data. The explained variance of these variables amounted to 50% (adjusted  $R^2 = 0.5$ , Table 5-1) in model 1 and was very reasonable.

The variables additionally included in model 2, which utilized information about real respondents provided to falsifiers to be used while producing falsifications, were all not significantly related to the dependent variable in the real data. Adding these variables to the variables of TPB did not change the explained variance of  $R^2 = 0.5$ . These variables also did not show significant relationships with the dependent variable in model 3.

The final model 3 is the most interesting because it includes also explaining variables, which are not related to the theory of TPB, but rather indicate stereotypical relationships with eating behavior. In model 3, we added additional sociodemographic variables, which were not given to falsifiers but could also be used by them while predicting healthy eating behavior: disposable income, consideration of the university as the center of interests and activities and vocational education. In the real data, a significant relationship was found between the university as the center of activities and the dependent variable. However, neither significant relationships between the variables on TV consumption, the preference of healthy desserts, participating in sports, and the interviewee's BMI and healthy eating behavior as a dependent variable nor a change of explained variance were observed.

Next, we look at the falsified data. Model 1, which included all theory-driven variables regarding the TPB, explained 40 per cent of variance, which was lower than that observed in the real data. Attitude did not have a significant effect on the healthy eating behavior. However, other independent variables – subjective norm, perceived behavioral control and intention – were significantly and positively associated with the dependent variable. The falsifiers supposed that people with a strong subjective norm, high perceived behavioral control, and high intention reported eating healthy more often than people with lower values for the subjective norm, perceived behavioral control, and intention. However, in model 3, in which variables describing stereotypical assumptions were additionally included, subjective norm lost significance and was no longer relevant for predicting healthy eating behavior.

In model 2, gender was a significant predictor of healthy eating behavior. In the falsified data, females were reported to eat healthy more often than males. This relationship is not significant in the real data. The amount of explained variance is slightly higher in model 2 than in model 1.

In the final model 3, the other sociodemographic variables were included. These variables did not have significant effects on the dependent variable, a result that is largely similar to the result found for the real data. However, within the real data, the university as the center of activities had a significant effect, while that was not the case in the falsified data. In the case of falsified data, all the additional included variables on TV consumption, the preference of healthy desserts, participating in sports, and the interviewee's BMI have significant effects, as expected, unlike the results obtained from the real data. In the falsified data, there were negative relationships between healthy eating behavior on the one hand and TV-consumption and BMI on the other hand. In contrast, preferring healthy desserts and doing sports were significantly and positively related to the dependent variable. When compared with models 1 and 2, the relationships of several other independent variables with the dependent variable changed. As mentioned above, the effect of the subjective norm was no longer significant in model 3, so that only two of the four variables testing TPB were significantly related with the dependent behavior variable. The effect of gender lost its significance, while the living situation gained significance. Living with parents and/or being single yielded a positive relationship with healthy eating behavior. Adding variables related to stereotypical relationships led to an increase of explained variance up to 48% in model 3.

**Table 5-1: Regression analyses of real and falsified data – dependent variable: healthy food consumption**

| Included variables  | Real data               |                         |                         | Falsified data          |                         |                         |
|---|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
|   | Model 1<br>Beta (sign.) | Model 2<br>Beta (sign.) | Model 3<br>Beta (sign.) | Model 1<br>Beta (sign.) | Model 2<br>Beta (sign.) | Model 3<br>Beta (sign.) |
| <i>Theory-driven:</i>   |                         |                         |                         |                         |                         |                         |
| Attitude  | -.11 (.010)             | -.11 (.014)             | -.11 (.010)             | -.03 (.505)             | -.07 (.175)             | -.08 (.081)             |
| Subjective norm   | .13 (.003)              | .12 (.009)              | .11 (.016)              | .12 (.013)              | .10 (.047)              | .06 (.188)              |
| Perceived behavioral control  | .18 (.000)              | .19 (.000)              | .18 (.000)              | .26 (.000)              | .26 (.000)              | .24 (.000)              |
| Intention   | .58 (.000)              | .57 (.000)              | .55 (.000)              | .45 (.000)              | .42 (.000)              | .33 (.000)              |
| <i>Control variables prescribed to falsifiers:</i>  |                         |                         |                         |                         |                         |                         |
| Gender = female (reference: male)   |                         | .08 (.067)              | .06 (.220)              |                         | .18 (.000)              | .09 (.093)              |
| Age   |                         | .03 (.544)              | .02 (.629)              |                         | -.01 (.832)             | .02 (.649)              |
| Living situation: Living with parents (reference: all others <sup>1</sup> )                   |                         | -.01 (.919)             | .01 (.864)              |                         | .07 (.180)              | .10 (.040)              |
| Living in a shared apartment (reference: all others <sup>2</sup> )                            |                         | -.03 (.548)             | -.02 (.698)             |                         | .03 (.502)              | .06 (.242)              |
| Being single (reference: having a partner)  |                         | -.00 (.951)             | -.00 (.987)             |                         | .08 (.071)              | .09 (.047)              |
| <i>Further sociodemographic variables:</i>  |                         |                         |                         |                         |                         |                         |
| Disposable income   |                         |                         | .05 (.294)              |                         |                         | .03 (.458)              |
| University as the center of interests and activities<br>(reference: university not as center) |                         |                         | .08 (.049)              |                         |                         | .08 (.060)              |
| Vocational education and training (reference: no)   |                         |                         | -.01 (.906)             |                         |                         | .04 (.364)              |
| <i>Leisure activities:</i>  |                         |                         |                         |                         |                         |                         |
| TV consumption in min/day   |                         |                         | -.05 (.241)             |                         |                         | -.11 (.015)             |
| Doing sports (reference: no)  |                         |                         | -.02 (.570)             |                         |                         | .12 (.009)              |
| Favorite dessert is healthy (reference: unhealthy)  |                         |                         | .06 (.142)              |                         |                         | .13 (.004)              |
| BMI   |                         |                         | -.06 (.164)             |                         |                         | -.14 (.010)             |
| Adj. R <sup>2</sup> (sign.)   | .50 (.000)              | .50 (.000)              | .50 (.000)              | .40 (.000)              | .42 (.000)              | .48 (.000)              |

<sup>1</sup>) all others = in a shared apartment, as subtenant in a family, in a students' dormitory, in a rented apartment, in a freehold apartment, in their own house

<sup>2</sup>) all others = living with parents, as subtenant in a family, in a students' dormitory, in a rented apartment, in a freehold apartment, in their own house

## 5.5 Conclusion

The aim of the present study was to investigate how falsified survey data influence the results of theory-driven data analyses and which correlations in a multivariate analysis differ between real and falsified data. The influence of theoretically relevant variables was compared to the influence of variables that seem relevant according to stereotypes as well as to other variables, which are not relevant from the theoretical point of view. As a framework for the theory-driven assumptions, we used TPB (Ajzen, 1991), which is an established theory in the social sciences that has often been utilized to predict different types of behavior (Billari et al., 2009; Prapavessis et al., 2015). As a behavior of interest, we used healthy food consumption, which was assumed to be predicted by the attitude towards this behavior, subjective norm, intention and perceived behavior control. We found support for the assumed relationships in the real data (model 1), even though the direction of the relationship between the attitude and behavior differed from the assumed relationship. This difference can be explained by the kind of behavior we used, so that even a positive attitude towards this behavior (eating healthy) may be associated with less frequent healthy food consumption. Next, we used self-reports of respondents about the behavior, which can be a source of bias. However, attitude was a relevant predictor of behavior in the real data, while in the falsified data a significant relationship between the attitude and behavior was not observed. In addition, the subjective norm was not a relevant predictor of behavior in the falsified data after controlling for the effect of possible stereotyping of assumed relationships. The model, which was supposed to test a priori assumptions, therefore found support in the real data but not in the falsified data. In addition, the results of higher explained variances of regressions models in the falsified data, as reported by Reuband (1990) and Schnell (1991), could not be observed in our data with respect to the model used to test the TPB assumptions; a contradictory result was obtained instead. We postulated that higher explained variances can be expected with respect to ad hoc developed models, which may also consist of relationships, which could be rather stereotypical assumptions. We did not expect such a result when theory-driven models were used, and our results supported this expectation.

In contrast to the theory-driven assumptions, which could not be consistently supported in the falsified data, the results show that theoretically non relevant variables played a role in predicting the behavior under investigation in the falsified data but not in the real data (model 2 and 3). In the real data only one of twelve variables was significantly related to the dependent variable, while in the falsified data, variables of living situation as



well as other variables that were included to test possible stereotypical assumptions were significant. This applies in particular to stereotypical relationships of other behavioral variables related to healthy food consumption (model 3). Thus, the falsifiers seem to apply stereotypes to invent consistent cases – people who watch less TV, prefer healthy desserts, have a lower BMI and play more sports are people who eat healthier than “couch potatoes” with higher values of TV consumption and BMI, who do fewer sports and prefer unhealthy desserts.

Our results help to clarify and to put into perspective the results obtained by previous research (Reuband, 1990; Schnell 1991; Schraepfer and Wagner, 2003) with respect to the differences between real and falsified data when using multivariate analyses. Our results show that one can expect that theory-driven, rather complex relationships are difficult for falsifiers to construct. In such analyses, one can expect to obtain inconsistent results and lower explained variances in falsified data. Moreover, when considering stereotypical and theoretically non relevant variables in explanations, they will gain on significance and explain the variance of a dependent variable better with falsified than with real data.

Our results also provide hints on which information falsifiers use to predict behavior of others. Falsifiers firstly use information available to them about respondents, which can be elucidated through brief contact. Additionally, they use stereotypical assumptions and laymen knowledge. However, falsifiers are less able to predict relationships, as postulated by a scientific theory, which is a new finding presented in this article.

It should be mentioned in this context that our falsifiers were students who were mainly studying social sciences. This fact could lead to the result that theory-driven assumptions could be supported to some extent in the falsified data. However, although our falsifiers are rather familiar with social science theories such as TPB, they were not able to reproduce the strength of the correlations of the real data set.

In addition, we have one dataset of 100 per cent real and one dataset of 100 per cent falsified interviews, which is seldom the case in reality, where, fortunately, the proportion of falsifications is much lower. However, simulation analyses showed that differences such as those presented in this article are also applicable for the detection of falsifiers when the proportion of falsifications is low (Storfinger and Winker, 2013). Therefore, our results can be used to better understand the impact of falsified data on the results obtained with survey data and to improve methods for the ex post detection of falsifications presented by other researchers (e.g., Bredl et al., 2013; Menold et al., 2013).

## 6. Explaining Political Participation: A Comparison of Real and Falsified Survey Data<sup>5</sup>

**Abstract:** This paper examines differences between real survey data and data falsified by interviewers. Previous studies show that there are only small differences between real and falsified data which implies that falsifying interviewers are able to (re-)produce realistic frequency distributions. The question this paper aims to answer is whether they are also able to produce multivariate results in accordance with the assumptions of established social science approaches. As an example for a realistic theory-driven data analysis, real and falsified data are compared in terms of the identified determinants of political participation. I use an experimental data set with data partly collected in real interviews and partly by interviewers being instructed to falsify; that is, to fill in the questionnaire based on little information about the respondent. The questionnaire measures twelve political activities, based on which I calculate an index for political participation. There are differences in the models between the real and the falsified data: The explained variances are higher in the regression models of the falsified data. There are some variables significant in both data sets and some that are significant only in the real or in the falsified data. These differences can be explained by our theoretical assumptions.

### 6.1 Falsification in Surveys

Face-to-face interviews are an important mode of data collection. The interviewers play a central role, since they can probe into unclear answers from the respondents, for example (Mangione et al., 1992). There is however, the risk that interviewers may falsify parts of, or the entire, interview (cp. Bredl et al., 2013). Following the definition of the AAPOR, interviewer falsification “means the intentional departure from the designed interviewer guidelines or instructions, unreported by the interviewer, which could result in the contamination of data” (AAPOR, 2003: 1). The research question this paper aims to answer is how fabricated data affects the results of theory-driven multivariate analyses, using the example of explaining political participation.

---

<sup>5</sup> Already published: Landrock, Uta (2017a): Explaining Political Participation – A Comparison of Real and Falsified Survey Data. *Statistical Journal of the IAOS* 32(3): 447-458. The final publication is available at IOS Press through <http://dx.doi.org/10.3233/SJI-160270>.

But why do interviewers falsify survey data? Falsifying interviewers want to optimize their cost-benefit balance. That means, they want to fulfill their tasks and objectives (conducting the interviews, achieving a high response rate, maximizing the payment etc.) with a minimal effort and resources (time and travel expenditures, convincing the survey respondents to participate, etc.). Thus the behavior of the interviewers follows the logic of rational choice approaches: To save effort and time, interviewers can decide to falsify (Sodeur, 2007; Winker, 2016).

In addition, the falsifying processes themselves underlie a cost-benefit calculation. On the one hand, falsifying interviewers want to reduce their cognitive effort when choosing answer categories. On the other hand, they want to prevent detection as falsifiers, which implies that they must fabricate answers which could plausibly be the true answers of the respondents (Menold et al., 2013; Menold and Kemper, 2014). To reduce their cognitive effort, falsifiers apply satisficing strategies: Satisficing means that the falsifiers do not search for the optimal answer on a certain question; they search for an acceptable answer, one which appears to be plausible on first thought (Krosnick and Alwin, 1987; Menold et al., 2013). Respondents in a real interview situation are assumed to use satisficing strategies since they also want to reduce cognitive effort when choosing answer categories (Krosnick and Alwin, 1987).

To quickly evaluate the plausibility of answers, falsifiers apply general stereotypes and implicit everyday knowledge (Reuband, 1990; Schnell, 1991). By stereotyping, falsifiers are able to reproduce the means and marginal distributions in high accordance with reality, as described later in this section (Reuband, 1990; Schnell, 1991; Schraepler and Wagner, 2003). The same stereotypes and implicit models of response behavior can also lead to a higher consistency of the falsified interviews (Schnell, 1991), particularly since the existing variations which accompany different sociodemographic characteristics are overestimated by the interviewers (Reuband, 1990). Aside from the higher consistency and the differences in the covariance structure there are further indications for falsification, for example in the meta- and para-data and also regarding different formal indicators like response sets (e.g. acquiescent responding or primacy and recency effects) as Menold and Kemper (2014) have shown.

Reliable information on the exact proportion of falsifications in surveys does not exist; different authors indicate that it does not exceed five percent (e.g., Koch, 1995; Krejsa et al., 1999; Li et al., 2009). However, there are examples of surveys where the percentage of

falsifications is much higher under certain circumstances (Kuriakose and Robbins, 2016; Koczela et al., 2015). In previous studies only small differences in the proportions and means were identified when real and falsified data were compared (Reuband, 1990; Schnell, 1991). Menold and Kemper (2014) also report large similarities between real and falsified data, especially in attitudinal and behavioral questions. However, findings on the impact of falsified data on the covariance structure of data prove that even small proportions of falsifications can contaminate data substantially (Reuband, 1990; Schnell, 1991; Schraepler and Wagner, 2003).

To explore the impact of falsified survey data on the results of data analysis, Schnell (1991) and Reuband (1990) both produced artificially falsified data in the lab, which they compare with real survey data. The real survey data in both studies were subsamples from factually existing data sets. Basic demographic information from some of the real survey respondents were provided to the study participants, who were recruited as falsifiers. The falsifiers were asked to use this demographic information to invent responses to the survey questions. Schnell (1991) used subjective political competence as a dependent variable in a regression analysis. The explaining variables were net income, self-reported social class, and self-placement on the top-bottom scale. In his regression analysis, the explained variance was higher in the falsified than in the real data. While in the falsified data all three independent variables had significant effects, in the real data only two of the three independent variables were significant (Schnell, 1991). Reuband (1990) analyzed the effects of age and gender on the attitude towards abortion. The results of his regression analysis also show that the explained variances were higher and the effects stronger in the falsified than in the real data (Reuband, 1990).

Furthermore, Schraepler and Wagner (2003) applied a regression analysis to explore the impact of falsified survey data on the results of a multivariate analysis. They used data from the German Socio Economic Panel, which included real and falsified data. The data were collected in a real survey field setting; the falsified data were afterwards identified as “real” interviewer falsifications. The dependent variable of the regression analysis was the log of gross income. The explaining variables were age and gender, the duration of training, and the working hours per week of the respondent. Schraepler and Wagner (2003) report that the inclusion of falsified data in their analysis reduces the explained variance and biases the effect sizes in different directions. Compared to the data set, which consists exclusively of real data, the effects of age and gender were overestimated and the effects of duration of training and working hours per week were underestimated (Schraepler and Wagner, 2003).

The previous studies in this field of research show that falsifications in surveys affect the results of data analyses, particularly the covariance structure, often resulting in an overestimation of correlations and of explained variance. Nevertheless, the correlations that were investigated were intuitive, based rather on ad hoc than on theoretically deduced hypotheses; for this reason, it is possible that they meet the assumptions of the stereotypes which falsifiers use in their satisficing strategies. As a consequence, one can assume that the overestimation of correlations in falsified data holds only or particularly for cases in which a falsifier would expect a relationship, for example because of stereotypes or implicit knowledge. In contrast to these above mentioned studies Landrock and Menold (2016) compared real and falsified data by applying a certain social science theory, the theory of planned behavior, and conducted theory-driven regression analyses for explaining healthy eating behavior. They came to the result that stereotypical and non-theory-driven relationships were more strongly pronounced in the falsified data while theory-driven relationships were more strongly pronounced in the real data. Apart from their work, until now, little is known about how the differences between fabricated and real data appear when theory-driven assumptions are tested. This paper intends to reduce this research gap by answering the research question whether falsifiers are able to produce multivariate results on political participation in accordance with the assumptions of established social science approaches and thus how fabricated data affects the results of theory-driven multivariate analyses. Thereby it aims to give further evidence that falsifiers are not fully able to reproduce the complex relationships of real data.

## **6.2 Analyses and results**

As we know from previous research, falsifiers are often able to invent plausible answers and produce realistic frequency distributions. They are also able to produce correlations that exist in real data, which they often even overestimate, as long as these correspond to customary stereotypes and are intuitively guessable. The research question raised here therefore is how correlations are affected by falsified data that are theoretically deduced and well-grounded in existing scientific literature. I analyze differences between real and falsified data using theory-driven multivariate analyses to identify the determinants for political participation. I decided to apply the example of explaining political participation because this is a widely researched question in social sciences with high social relevance.

### 6.2.1 Hypotheses

The question of whether falsifiers are able to produce data in accordance with the assumptions of established social science approaches is anything but evident. One can reasonably assume that falsifiers are typically not familiar with these social science approaches and do not apply them – also because this would imply investing more cognitive effort and would contradict the falsifiers' satisficing strategies. A core assumption of the following analyses therefore is that the more causal effects are rooted in social science theoretical approaches rather than in everyday knowledge, the less they can be produced by falsifiers and the less they are found in falsified data. This assumption leads to the following four hypotheses:

H1: For the theoretically predicted determinants there are more significant effects in the real than in the falsified data.

Falsifiers reproduce and eventually overestimate effects that they assume to be true, typically because they are obvious or intuitive, such as the influence of the self-placement on the left-right-dimension. In contrast, the effects predicted by elaborate social science theories, such as the influence of political efficacy, are less likely to be intuitive for falsifiers. And unless they have an academic background, the falsifying interviewers probably do not know the underlying theoretical bases of a study. Therefore, they should generally not be able to reproduce relationships that are explained by these bases and that are empirically proven in reality by the existing research literature<sup>6</sup>. Instead, they apply implicit stereotypes and everyday knowledge about the attitudes and behavior of the respondents. If this assumption is correct, this should be reflected in the results of the theory-driven causal analysis. Consequently, H1 states that there are more significant effects for the theoretically predicted determinants in the real than in the falsified data.

H2: For real sociodemographic information known to falsifying interviewers, there are more significant effects in the falsified than in the real data.

As described above, previous research has found that falsifiers overestimate the existing sociodemographic differences between the respondents (Reuband, 1990). Furthermore, they have only little information about the respondents and their living conditions. Therefore, the falsifiers strongly rely on sociodemographic information that they happen to know or that is

---

<sup>6</sup> It seems to be plausible that more experienced falsifiers are able to guess these less obvious relationships without knowing the social science theory behind.

provided to them and give a particularly high relevance to such information<sup>7</sup>. Consequently, H2 states that within sociodemographic information known to falsifiers, there are more significant variables in the falsified than in the real data.

H3: For further correlates found in research, there are more significant effects in the falsified than in the real data.

Beyond the ground of strict theoretical deduction, there are further correlates found in research which enter the bases of scientific knowledge and are typically also used in subsequent research for formulating hypotheses and designing multivariate regression models. In the case of research on political participation, there are such independent variables which have been proven in previous empirical studies to be relevant under certain circumstances (cp. section 6.2.3). They mostly have an influence in specific contexts, depending on the concrete form of the analyzed political participation. These influences are more obvious and principally guessable by laymen as they were by experts. Based on their laymen's theories and everyday knowledge, the falsifiers invent answers and construct stereotypical relationships with the likewise fabricated attitudes and behaviors. The real respondents instead answer more inconsistently, perhaps due to their more complex life conditions and the occurrence of satisficing. Therefore, the falsifiers overestimate the relationships and construct more significant effects in the falsified data than observable in the real data. Consequently, H3 states that there are more significant effects of the further correlates found in the falsified than in the real data.

H4: The falsifiers construct more consistent cases than real life; therefore, the explained variance is higher in the falsified than in the real data.

An invented pattern of answers to a questionnaire follows less complex principles and shows fewer inconsistencies than the empirical realities created by real life. Therefore, all in all, the falsifiers invent more consistent models of respondents' behavior than the real respondents show. This leads to more consistent regression models in the falsified than in the real data<sup>8</sup>. As a result, the explained variances, measured as adjusted R-squared, should be higher in the falsified than in the real data.

---

<sup>7</sup> Nevertheless, it seems to be plausible that the level of overestimation of sociodemographic information depends on the knowledge of the subject and the experience of the falsifier.

<sup>8</sup> The higher consistencies of the falsifiers vary on different factors, for example the length of the questionnaire.

### 6.2.2 Database and Method

Data sets with *confirmed* real and *confirmed* falsified interviews rarely exist; therefore usually experimental data are used for the empirical analysis of the effects of falsifications (cp. Winker et al., 2015). The database in this paper consists of three datasets, obtained in the summer of 2011, and is described in detail by Kemper and Menold (2014) as well as by Menold and Kemper (2014):

For the first dataset, 78 interviewers conducted 710 real face-to-face interviews. The interviewers were students at the University of Giessen. The respondents were students at the University of Giessen as well and were recruited on the campus. All interviews were audio-recorded and checked afterwards to assure that they were actually conducted and not falsified. The questionnaire contains 62 questions on attitudes and behavior and sociodemographic information, mainly adopted from the German General Social Survey (ALLBUS<sup>9</sup>) with political, economic and social topics. Payment was provided either per interview (8 Euros) or per hour (12 Euros) to the interviewers (e.g., Kemper and Menold, 2014).

For the second dataset, the same interviewers fabricated survey data in the lab so that a data set of 710 falsified interviews, corresponding to each of the  $N = 710$  real interviews, was obtained. For this task, either 9 Euros per hour or 3 Euros per falsified interview was paid. The falsifying interviewers received basic sociodemographic information about the real survey participants, for whom they were supposed to invent data and who had been interviewed by a colleague (not by themselves) in the first part of the study. The information the interviewers were given was of such a kind that a falsifying interviewer could easily have collected him- or herself via a short interview with the respondent: sex, age, studied subject, number of semesters enrolled, marital status, place of residence, living situation (with whom the respondent lives in a household), and country of origin (e.g., Kemper and Menold, 2014).

The interviewers were instructed to imagine a person with these characteristics and to fill in the questionnaire, thus falsifying the data as if they were collected in a real survey setting. The exact instruction was:

Please read the description of the person, whose interview you are supposed to falsify, carefully. Please fill in the attached questionnaire as if you had conducted a personal interview with the respondent in reality. Please place the description of the respondent

---

<sup>9</sup> <http://www.gesis.org/en/allbus/allbus-home/>



next to the questionnaire while falsifying, so that you are always aware of the characteristics of this person.

The person, whose interview is to falsify...

- is female,
- is 20 years old,
- studies teaching,
- is enrolled in her second semester at a university.
- She is unmarried, in a stable relationship,
- lives in Huettenberg, a rural village in Hesse,
- with her parents or relatives.
- Country of birth: Germany.

Additionally, in the third phase of the fieldwork, the interviewers filled in the survey questionnaire for themselves, as respondents, and also a questionnaire about their falsifying strategies. These data are stored in a third data set.

This experimental setup has strengths and weaknesses. On the one hand, it allows the collection of a lot of information about the interviewers and their falsifying processes. In comparison to a standard field setting, this is a relevant strength. A significant limitation of this experimental design is the fact that interviewers and respondents were students. As a result, some of the sociodemographic variables, like age or education, show only small variances (cp. Winker et al., 2015).

I use the two data sets with real and with falsified answers from the respondents for calculating multivariate causal analysis and comparing results to find out which differences occur with the use of not only real but also falsified data.

### **6.2.3 Political participation**

Following the definition of the political action study of Barnes and Kaase et al. (1979), political participation includes all activities of individual citizens, which are voluntary and intended to influence political decisions directly or indirectly (Kaase and Marsh, 1979: 42). Political participation has been differentiated in two forms. On the one hand, conventional political participation covers forms of participation which are related to political elections and party activities (Marsh and Kaase, 1979: 84), such as engagement in a party or participation in public discussions (cp. Koch et al., 1999). On the other hand, unconventional political

participation means forms of participation which are independent of political elections and party activities (Koch et al., 1999: 12; Kaase and Marsh, 1979: 41), such as participation in demonstrations or house squatting, which means occupying an abandoned house without permission (cp. Koch et al., 1999). However, the contrast between the two forms has diminished during the last decades since most forms of unconventional political participation used to be considered as illegitimate by large parts of society and are considered as fully legitimate today. In current literature this differentiation is mostly neglected, since it is considered as “analytically elegant, but difficult to keep up empirically” (van Deth, 2001: 203).

The question arises, which factors have an impact on the degree to which individuals engage in political participation. There are different models which aim to explain political participation (cp. Lüdemann, 2001). In the political action study, Kaase and Marsh (1979: 43) developed a “heuristic device” (Kaase and Marsh, 1979: 41), which is helpful to explain political participation, but “is not meant to be a systematic specification of causal relationships” (Kaase and Marsh, 1979: 41). Based on this instrumental concept, Opp developed a rational choice model of explaining political participation (cp. Lüdemann, 2001; Opp and Finkel, 2001) which is acknowledged and frequently used today, for example, it is the theoretical foundation for the German General Social Survey (ALLBUS, cp. Koch et al., 1999). According to Opp, the following dimensions are considered as determinants for political participation (see also Koch et al., 1999; Lüdemann, 2001): dissatisfaction with the political, economic or social situation, the perceived influence on political decisions, the existence of norms of political participation, and social incentives for political engagement.

The theoretical assumption is that the more people are dissatisfied with their political, economic or social situation, the stronger is their political participation. Recent research has provided empirical evidence for this assumption. For example, Steinbrecher (2004) proves the positive effect of political dissatisfaction on political participation. The perceived influence on political decisions is captured by the concept of political efficacy, which covers two dimensions: the internal and the external. Internal political efficacy means an individual’s perception of his or her own ability to influence political decisions. External political efficacy means the actor’s perception that the political system will react (Kaase and Marsh, 1979: 48-49). The political efficacy transforms a disposition to participate in factual participation and can therefore be considered a precondition for political participation (cp. Kaase and Marsh, 1979). The higher the political efficacy is the stronger is the political participation. Opp and

Finkel (2001), Lüdemann (2001), as well as Steinbrecher (2004) report evidence for the positive impact of political efficacy. A third influencing dimension covers the perception of participation norms; the stronger the perceived norms to participate are the stronger is the political participation (cp. Opp and Finkel, 2001). The dimension of social incentives refers to the social relations and memberships an actor has in organizations in which participation is likely to be supported, for example in political parties, citizens' action groups or other associations (see also van Deth, 2001). The theoretical expectation is that higher social integration in such supporting organizations and social networks should be positively correlated with stronger political participation. Opp and Finkel (2001) for example provide evidence for this assumption.

Aside from these influencing dimensions, many studies consider additional influencing factors. These factors mostly have an influence in specific contexts, depending on the concrete form of participation that is analyzed (for example conventional or unconventional forms or legal or illegal forms). Marsh and Kaase (1979) identify in the political action study the effects of sociodemographic variables like age, gender, and education. Opp and Finkel (2001), Westle (2001) and Steinbrecher (2004) confirm the significance of these variables: Opp and Finkel (2001) report a positive relationship between age and protest as unconventional political participation, whereas Steinbrecher (2004) identify a negative relationship. Westle (2001) shows that men participate politically more often than women do. Opp and Finkel (2001) as well as Steinbrecher (2004) report a positive relationship between education and political participation. There are further correlates, which have influences on political participation: Kaase and Marsh (1979) describe the self-placement on the left-right scale as a complementing indicator, with people classifying themselves as rather politically left being more active. The same result is reported by Lüdemann (2001). Furthermore, there is evidence for the positive influence of the self-reported social class on political participation (Marsh and Kaase, 1979). For TV consumption, Lüdemann (2001) and Schulz (2001) report a negative correlation with political participation. As one can see there is a broad variety of influencing correlates.

#### **6.2.4 Operationalization and regression model**

To operationalize political participation, which serves as the dependent variable, the questionnaire contains material adopted from the German General Social Survey (ALLBUS) 2008, which covers 12 different conventional and unconventional activities (cp. Wasmer et

al., 2010). Just as in the original ALLBUS 2008 item-scale, not only the factual behavior in the past is obtained for each activity, but also the intention to participate politically in the future (idem).

The exact wording of the two questions regarding the intention to participate and the factual behavior in the past is:

If you wanted to have political influence or to make your point of view felt on an issue which was important to you: Which of the possibilities listed on these cards would you use? Which of them would you consider? Please name the corresponding letters.

[letters A to M, see below]

Which of these things have you actually already done, what have you already taken part in? Please name the corresponding letters.

- A. Express your opinion to friends and acquaintances and at work
- B. Vote at elections
- C. Take part in public discussions at meetings
- D. Participate in a citizens' action group
- E. Voluntary work for a political party
- F. Take part in an unauthorized demonstration
- G. Take part in an authorized demonstration
- H. Not vote at elections out of protest
- J. Out of protest, vote for a party other than your party of choice
- K. Sign a petition
- L. Boycott or buy goods for political, ethical or environmental reasons
- M. Take part in an online protest campaign

Marsh and Kaase (1979), and following them also Steinbrecher (2004), Westle (2001), Schulz (2001) and other researchers, use the readiness to participate politically as an indicator for political participation. In the work presented in this paper, it is operationalized accordingly. One reason for including readiness in the operationalization is the assumption that factual political participation in the past supports and determines the current readiness for political participation (cp. Westle, 2001). Another reason refers to the chronological order of cause and

effect: the interview measures factual participation in the past and other information, such as attitudes or sociodemographic characteristics, at the time of the interview. This makes it difficult to consider the current attitudes or other interview information as influencing factors for political participation, which occurred prior to the measurement of these potential causes (cp. Steinbrecher, 2004). Therefore it seems convincing to use readiness or intention to participate as indicators for political participation and as dependent variables.

To identify the latent structure of the political activities and to generate a suitable indicator for political participation, I conducted an exploratory factor analysis using the twelve forms of past political participation as well as readiness for political participation (cp. van Deth, 2001; Lüdemann, 2001; Steinbrecher, 2004; and others). As a result, I identified four factors of past political activities and also four factors of readiness for political participation. The most reliable factor, with a Cronbach's Alpha of 0.6, covers the readiness for political participation, consisting of the readiness for participation in a citizens' action group (item C), the readiness for participation in public discussions (item D), and the readiness for engagement in a party (item E). I calculated an additive index as an indicator for political participation and dependent variable for the further analysis. Steinbrecher (2004) identified in his analysis of the ALLBUS data almost the same factor<sup>10</sup>, which he calls *party activities*.

The explaining variables are organized in three blocks. The first block covers the theoretically expected determining factors for political participation as described by Opp and Finkel (2001, see above), the second block consists of sociodemographic and control variables provided to the falsifiers, and the third block refers to the further situationally relevant correlates found in research.

As described before, following Opp's explanation for political participation, there are four determining factors: political efficacy, dissatisfaction, norms of participation, and social incentives (cp. Opp and Finkel, 2001).

In the data set, political efficacy is differentiated in two forms, measured with two items each. The internal political efficacy is measured as in agreement with the following items from ALLBUS 1998 (cp. Koch et al., 1999):

---

<sup>10</sup> Steinbrecher's factor additionally covers the item "supporting a candidate", which is not used in ALLBUS 2008 and neither in my questionnaire.

I would have the confidence to take on an active role in a group concerned with political issues.

Politics is so complicated that somebody like me can't understand what's going on at all.  
[Reversed item]

The external political efficacy covers the agreement with the following items from ALLBUS 1998 (idem):

Politicians don't care much about what people like me think. [Reversed item]

In general, politicians try to represent the people's interests.

As independent variables for the analysis, the means of both items were each calculated (Koch et al., 1999; Wasmer et al., 2010).

Following the operationalization of the ALLBUS 1998, dissatisfaction is obtained as political and economic dissatisfaction. To measure political dissatisfaction, the mean of the agreement with three statements is used (Koch et al., 1999):

Only when differences in income and social status are large enough is there any incentive for personal achievement.

Differences in social position between people are acceptable because they basically reflect what one has made of the chances one had.

I consider the social differences in this country to be just on the whole.

The economic situation is measured just as in ALLBUS 1998, in accordance with two items (Koch et al., 1999). Here I did not calculate an index because of the low value of Cronbach's Alpha (.26):

How would you generally rate the current economic situation in Germany?

And your own current financial situation?

Information on the norms of participation and social incentives are not available in the used data set.

The second block of independent variables covers sociodemographic information which is provided to the falsifiers. These variables are known to have effects on political participation (Marsh and Kaase, 1979). Furthermore, this information is used by the falsifier to invent the

survey data. I included age and gender, but not education, which was also provided, because all respondents are students.

In the third block of independent variables, I included indicators which were reported as situationally relevant for political participation; these variables are self-placement on the left-right scale (Kaase and Marsh, 1979), self-reported social class (idem), and TV consumption (Lüdemann, 2001; Schulz, 2001). In ALLBUS 2008, the attractiveness of the respondent, as reported by the interviewer, is assumed to have a positive effect on political participation, because psychological research has shown that attractive people consider themselves to be more influential than others (Wasmer et al., 2010). Thus, I also included this variable as situationally relevant.

These independent variables were included in a linear OLS regression analysis, which was calculated for real and for falsified data identically, to compare the results of the real and the falsified data.

### **6.2.5 Results**

The following table (Table 6-1) shows the results of the regression analysis for the real and the falsified data.

Model 1, includes the theoretically expected determining factors for political participation, which are less obvious for falsifiers. In the real data, three of the five determining factors show significant effects. The correlations of internal political efficacy and political dissatisfaction are positive, as expected. In contrast to the theoretical assumptions, there is a negative influence of dissatisfaction with an individual's own economic situation on political participation. In addition, Westle (2001) and Opp and Finkel (2001) cannot confirm any (positive) effects of dissatisfaction. Opp and Finkel (2001) argue in their analysis that the respondents assume that they cannot reduce their dissatisfaction by participating politically, which seems to apply in this study as well. Furthermore, in this study the respondents are students. It doesn't seem plausible to assume that students, who are dissatisfied with their own economic situation, would be willing to participate politically in the forms of participation that are analyzed here (cp. section 6.2.4). It seems more convincing that these students would concentrate on the success of their studies to change their economic situation as soon as possible. Consequently, they would be less likely to willingly participate politically, as compared to students who are satisfied with their economic situation. This effect, however, is not in line with the theoretically-developed hypotheses. Aside from the three effects

described, there are no significant effects in the first block: external political efficacy and dissatisfaction with the general economic situation in Germany do not influence political participation.

Falsified data from model 1 contained only one significant variable: internal political efficacy. The other four theoretically expected determining factors for political participation remain without significant effects. This finding is generally in line with the first hypothesis, stating that for the theoretically predicted determinants, which are less obvious for falsifiers, there are more significant effects in the real than in the falsified data. Still, the question arises as to how the falsifiers are able to reproduce the correlation between internal political efficacy and political participation. This effect is even stronger in the falsified data than it is in the real data. One possible answer is that the two statements which measure internal political efficacy<sup>11</sup> are relatively coarse and are therefore, for the falsifiers obviously related to political participation, without scientific knowledge. The adjusted R-squared of .21 in the falsified data is higher than the value of .16 in the real data. Although only one independent variable is significant in the falsified data, the falsified model is more consistent than the model for the real data. The falsifiers only recognize the relevance of one theoretically predicted determinant, the internal political efficacy, but they overestimate its influence, similarly as it was predicted in hypotheses 2 and 3 for the determinants that are not grounded on theoretical approaches.

Model 2 additionally includes the control variables provided to falsifiers. These are the sociodemographic information of the real respondents regarding age and gender that are accurate also in the data set of falsified data, so that possible differences in the results can only occur due to the dependent variable. In the real data, age has no effect on political participation, presumably because of the low variance within the age range of this sample of students. Gender has a significant influence on political participation: In this sample, women participate politically more frequently than men.

Results in the falsified models are quite similar: The effect of gender is also provable and follows the same direction. Age is not significant in the falsified data either, presumably for the same reason as in the real data. These results do not support the assumption that falsifiers overestimate the sociodemographic information they happen to have, which

---

<sup>11</sup> The two statements are: "I would have the confidence to take on an active role in a group concerned with political issues" and "Politics are so complicated that somebody like me can't understand what's going on at all" (see also section operationalization).



contradicts the second hypothesis. One can conclude, however, that falsifiers use the information about gender to falsify the questionnaires. The question arises of how the falsifiers are able to guess the influence of gender, and also in the correct direction, with women being more engaged than men, which seems stereotypically contradictory. One explanation may be that the falsifiers have experience conducting real interviews prior to falsifying their data, which may have given them a good impression regarding gender differences in the sample. This is a weakness of the data set which might corrupt our findings. The adjusted R-squared is still higher in model 2 in the falsified data, with two significant independent variables, than in the real data, with four significant independent variables. This difference is still due to the variables of the first block, particularly due to the impact of internal political efficacy.

The final model 3 includes additional correlates found in research, which are reported as situationally relevant for political participation and more obvious for falsifiers. In the real data, none of the further correlates show any significant effect; for the “party activities,” as Steinbrecher (2004) calls the forms of political participation used here for operationalization, these variables do not make a difference. Furthermore, in model 3, dissatisfaction with the individual's own economic situation loses its significant influence. Internal political efficacy and political dissatisfaction remain the only two significant determining factors for political participation.

In the falsified data there is a significant effect of self-placement on the left-right-dimension: The falsifiers assume that self-placement on the left-right-dimension has a significant influence on political participation, with those defining themselves as rather left being more engaged. This is in line with the third hypothesis, stating that for further correlates, there are more significant effects in the falsified than in the real data. The adjusted R-squared value in the falsified data slightly increases to .23 in model 3 and is still higher in comparison to the real data. Thus, the falsifiers' model of political participation is more consistent than in reality.

**Table 6-1: OLS regression analysis – determinants of political participation in real and in falsified data**

| Included variables  | Real data    |              |              | Falsified data |              |              |
|---|--------------|--------------|--------------|----------------|--------------|--------------|
|   | Model 1      | Model 2      | Model 3      | Model 1        | Model 2      | Model 3      |
|   | Beta (sign.) | Beta (sign.) | Beta (sign.) | Beta (sign.)   | Beta (sign.) | Beta (sign.) |
| <i>Determining factors for participation:</i>                         |              |              |              |                |              |              |
| Internal political efficacy   | .38 (.000)   | .41 (.000)   | .40 (.000)   | .47 (.000)     | .49 (.000)   | .46 (.000)   |
| External political efficacy   | .02 (.636)   | .01 (.797)   | .00 (.926)   | -.01 (.734)    | -.02 (.612)  | -.01 (.722)  |
| Dissatisfaction with the general economic situation in Germany        | .00 (.949)   | -.01 (.791)  | .00 (.999)   | .04 (.355)     | .04 (.387)   | .03 (.469)   |
| Dissatisfaction with individual economic situation                    | -.08 (.029)  | -.08 (.033)  | -.07 (.067)  | -.01 (.900)    | .00 (.979)   | -.01 (.859)  |
| Political dissatisfaction   | .13 (.000)   | .12 (.001)   | .11 (.006)   | .06 (.097)     | .06 (.128)   | .01 (.727)   |
| <i>Control variables provided to falsifiers:</i>                      |              |              |              |                |              |              |
| Age   |              | -.02 (.515)  | -.02 (.518)  |                | .01 (.860)   | .01 (.797)   |
| Gender = female (reference: male)                                     |              | .08 (.033)   | .08 (.034)   |                | .09 (.023)   | .08 (.048)   |
| <i>Further correlates found in research:</i>                          |              |              |              |                |              |              |
| Self-placement on the left-right-dimension (1 = right, 10 = left)     |              |              | .00 (.942)   |                |              | .13 (.002)   |
| Self-reported social class (reference: upper middle class and higher) |              |              | -.04 (.271)  |                |              | -.04 (.387)  |
| TV-consumption (min/day)  |              |              | -.06 (.100)  |                |              | -.03 (.370)  |
| Attractiveness of respondent, reported by the interviewer             |              |              | -.03 (.414)  |                |              | .01 (.801)   |
| Adj. R <sup>2</sup> (sign.)   | .16 (.000)   | .17 (.000)   | .17 (.000)   | .21 (.000)     | .22 (.000)   | .23 (.000)   |

To summarize the results of the regression analysis I will review the hypotheses:

H1 states: *For the theoretically predicted determinants, there are more significant effects in the real than in the falsified data.*

This hypothesis is confirmed. In all three models, there are more significant determining factors in the real than in the falsified data. The final model 3 shows two significant effects in the real data, namely effects of internal political efficacy and of political dissatisfaction. In model 3, regarding the falsified data, only internal political efficacy has a significant effect: The falsifiers do not guess the impact of political dissatisfaction. So, there is some support for H1. On the other hand, the support for H1 could be stronger: There is still one significant effect in the falsified data, even if it is an influence that seems intuitively guessable; and, there are only a few significant effects of the determining factors for political participation in the real data. As previously mentioned, other researchers cannot confirm any (positive) impact of dissatisfaction (cp. Opp and Finkel, 2001; Westle, 2001). The same applies to external political efficacy (Steinbrecher, 2004). In conclusion, the results presented here for the real data are in accordance with research on political participation.

H2 states: *For real sociodemographic information known to falsifying interviewers, there are more significant effects in the falsified than in the real data.*

This hypothesis is not confirmed. In each of the three models in both subsamples, gender has a significant effect on political participation but age does not. This contradicts the hypothetical assumptions of more significant effects in the falsified than in the real data. One reason probably is that age only has a small variance in this student sample. The second reason may be that the falsifiers have gained an idea about the actual gender differences in the sample before they falsified via experience conducting real interviews. So, it is very possible that the lack of support for H2 is due to weaknesses in the data set and its methodological design.

H3 states: *For further correlates found in research, there are more significant effects in the falsified than in the real data.*

This hypothesis is confirmed. One can see that there is one variable, self-placement on the left-right-dimension, with a significant effect in the falsified data, but none in the real data. This is in line with H3. However, again, there could be stronger support. It can be assumed

that falsifiers do not have many applicable stereotypes regarding readiness for political participation. This may have to do with the dependent variable: Readiness is a pretty vague construct. Also, it could be that more unconventional forms of political participation, for example participating in demonstrations, would activate more stereotypes in the expected way, in contrast to the engagement in a party or participating in a citizens' action group or participating in public discussions, which are considered in this analysis.

H4 states: *The falsifiers construct more consistent cases than in real life; therefore, the explained variance should be higher in the falsified than in the real data.*

This hypothesis is confirmed. The results regarding H4 meet the expectations of a higher explained variance in the falsified than in the real data. The falsifiers seem to invent more consistent cases in all three models.

The corrected R-squared value of .17 is not very high in the real data. This corresponds to the findings of Steinbrecher (2004), who analyses different forms of readiness for political participation and reports corrected R-squared values between .10 and .21 in his analyses of ALLBUS data. Steinbrecher (2004) argues that these results indicate that political participation depends highly on the context.

It can be summarized that three of the four hypotheses find support, even if the support is not very strong. Only H2 is left without confirmation.

### **6.3 Conclusion and discussion**

The present work started out from the observation that falsifying interviewers are, on the one hand, able to invent plausible answers to survey questions. On the other hand, previous studies show that interviewer falsifications in surveys affect the results of data analyses, mostly leading to an overestimation, partly also to an underestimation of influences. This paper was inspired by the assumption that an overestimation of influences would occur for variables to which an influence seemed intuitively plausible, based on implicit laymen's theories and stereotypes, whereas an underestimation would occur for variables whose influence seemed less obvious. The latter can be expected for effects predicted by elaborate social science theory approaches. The paper therefore intended to answer the question of whether or not falsifiers are able to produce data in accordance with the assumptions of established social science approaches. Explaining political participation is a suitable example

for investigating this question, because there is a well-established theoretical approach with proven survey questions and measurements (for example in the 1998 and 2008 ALLBUS questionnaires). Furthermore, political participation is a highly relevant topic in the social sciences.

As described in the results above, there are differences in the models between the real and the falsified data: The explained variances are higher in the regression models of the falsified data. Furthermore, correlations are present in the falsified data that cannot be proven in the real data, which supports the assumption that falsifiers use stereotypes or implicit knowledge for inventing realistic answers to interview questions. Finally, the falsifiers were not able to reproduce both effects of the theoretically predicted determinants for political participation, in particular the effect of political dissatisfaction on the readiness for political participation. As one can see, the falsifiers are not fully able to produce data in accordance with the assumptions of established social science approaches. These findings underline that good results of multivariate analysis, in terms of strong significant effects and high shares of explained variance, do not necessarily imply good data quality; they might as well be a hint for falsifications. Therefore it is crucially important for empirical research to put much effort in avoiding falsifications. And it is important to identify falsifications, for example using the formal indicators investigated by Menold and Kemper (2014). Comparing subsamples with suspicious and with unsuspecting cases based on theory-driven multivariate analysis could be an additional strategy in examining a data set for falsifications.

Despite the fact that our hypotheses do find support, there are fewer differences between real and falsified data than one could expect. This probably has four reasons: First, the falsifying interviewers were familiar with the typical responses of the interviewees because they had conducted real interviews before they falsified. That is also the case in real fieldwork, where real interviewers who falsify may have conducted true interviews beforehand. Second, the respondents, as well as the interviewers, were students, leading to the possible bias that the interviewers are acquainted with the thinking and the habits of the interviewed population. Third, there are different social science theories that are more or less “intuitive” for laymen and therefore easier or harder to reproduce for falsifiers. The theoretically predicted explaining variables for political participation are, partly, intuitive, in particular the effect of internal political efficacy. Fourth, there are dependent variables that are more or less suitable for applying stereotypes. Aside from the intuitive explaining variables mentioned above, there are only a few stereotypes related to the readiness for political

activities. One may find better-suited dependent variables to activate stereotypes and to make the differences between real and falsified data visible, for example, explaining activities like participation in demonstrations.

Nevertheless, the main finding of this paper is that falsifiers overestimate the influence of stereotypical causes and they are less able to reproduce theoretically-induced relationships. Thus this paper validates the results of Landrock and Menold's previous work (2016) and gives another piece of evidence that complex relationships in the real data are difficult for falsifiers to reconstruct.

## 7. How Interviewer Effects Differ in Real and Falsified Survey Data: Using Multilevel Analysis to Identify Interviewer Falsifications<sup>12</sup>

**Abstract:** In face-to-face interviews, interviewers can have an important positive influence on the quality of survey data, but they can also introduce interviewer effects. What is even more problematic is that interviewers may decide to falsify all or parts of interviews. The question that the present article seeks to answer is whether the interviewer effects found in falsified data are similar to those found in real data, or whether interviewer effects are larger and more diverse in falsified data and may thus be used as an indicator for data contamination by interviewer falsifications. To investigate this question, experimental data were used from controlled real interviews, interviews falsified by the same interviewers, and questionnaires completed by these interviewers themselves as respondents. Intraclass correlations and multilevel regression models were applied, and interviewer effects in the real survey data were compared with those in the falsified data. No evidence of interviewer effects was found in the real data. By contrast, interviewer effects were found in the falsified data. In particular, there was a significant association between the interviewers' own responses and the falsified responses to the same questions in the questionnaire. Thus, to detect interviewer falsifications, I recommend that researchers should also get the interviewers to complete the questionnaire and check datasets or suspicious cases for interviewer effects.

### 7.1 Introduction

Face-to-face interviews are an important mode of data collection in empirical social research. It is used in many major studies, for example the European Values Study (EVS),<sup>13</sup> the U.S. General Social Survey (GSS),<sup>14</sup> and the Programme for the International Assessment of Adult Competencies (PIAAC).<sup>15</sup> Interviewers can have a major influence on the quality of survey data. On the one hand, they can improve data quality, for example by helping the respondent to understand the survey questions correctly (Mangione et al., 1992). On the other hand, there is the risk of interviewer effects, that is, distortions of survey responses due to the presence of

---

<sup>12</sup> Already published: Landrock, Uta (2017b): How interviewer effects differ in real and falsified survey data. Using multilevel analysis to identify interviewer falsifications. *mda* 11(2): 165-190. The final publication is available at GESIS through <http://dx.doi.org/10.12758/mda.2017.03>.

<sup>13</sup> <http://www.europeanvaluesstudy.eu/>

<sup>14</sup> <http://gss.norc.org/>

<sup>15</sup> <http://www.oecd.org/skills/piaac/>

an interviewer. Interviewer effects can cause biased data and affect substantive findings (Beullens and Loosveldt, 2016; Groves and Magilavy, 1986). They occur when the respondent's answer depends not only on the intended stimulus of the question but also on the interview situation and the interviewer (Bogner and Landrock, 2016; Schanz, 1981). In the case of interviewer effects, certain interviewer behaviors (e.g., reading pace or suggestiveness) or characteristics (e.g., experience, age, gender, or education) may influence the response behavior of the respondent (Beullens and Loosveldt, 2016; Haunberger, 2006; Mangione et al., 1992). Interviewer effects therefore constitute response bias (see Groves and Magilavy, 1986), where the reported values of the respondent systematically deviate from the true values.

In this context, it is important to know whether some types of questions are more susceptible to interviewer effects than others (Mangione et al., 1992). Research on interviewer effects has yielded a large number of findings in this regard (for an overview, see Bogner and Landrock, 2016). According to Haunberger (2006), for example, difficult and sensitive questions, attitudinal questions, and open-ended questions are particularly prone to interviewer effects. Haunberger (2006) showed that, in the case of difficult questions, the gender and education of the interviewers may have an influence on responses, for example, to income-related questions. The probability that the respondent will refuse to answer such questions is reported to be higher in the case of female or highly educated interviewers (Bogner and Landrock, 2016; Haunberger, 2006). Regarding attitudinal questions, research findings are ambiguous. Whereas Liu and Stainback (2013) identified interviewer gender effects on responses to attitudinal questions, Groves and Magilavy (1986) did not find evidence of such an influence on attitudinal questions compared to factual questions. Haunberger (2006) suggested that interviewer age and education may influence responses to open-ended questions and that these questions are therefore susceptible to interviewer effects (Mangione et al., 1992). By contrast, Groves and Magilavy (1986) reported that open-ended questions were not inherently more susceptible to interviewer effects than closed questions. However, in the case of open questions that ask respondents to mention several entities, for example "What do you think are the most important problems facing the country?," the authors suggested that the likelihood that the respondent would mention a second entity might depend on the interviewer's probing behavior, and that "the differential behaviors that determine whether a second mention is given also might influence substantive responses on the second mention" (Groves and Magilavy, 1986, p. 260). In summary, therefore, research



findings show that difficult, attitudinal, and open-ended questions are susceptible to interviewer effects.

These findings provide evidence that the perceptible sociodemographic characteristics of the interviewer – namely gender, age, and education – are relevant to the occurrence of interviewer effects (Haunberger, 2006; Liu and Stainback, 2013; West and Blom, 2016). Olson and Bilgen (2011) reported that larger interviewer effects occurred with respect to acquiescence in the case of experienced interviewers than in the case of inexperienced interviewers. West and Blom (2016) described the influence of certain personality traits of the interviewers that may affect response behavior. Moreover, research findings suggest that the relation between interviewers' and respondents' characteristics may result in interviewer effects: Schanz (1981) analyzed the relevance of interaction effects and described positive correlations between the answers of the interviewer and the answers of the respondent to the same survey questions. One possible explanation for this positive correlation is that the respondent reacts to the non-verbally expressed attitudes of the interviewer (Schanz, 1981; West and Blom, 2016). Thus, interviewer effects may also depend on the content of the question and the interaction of the attitudes of the interviewers and the respondents (Schanz, 1981).

In face-to-face interviews, not only may interviewer effects occur, but interviewers may even decide to falsify all or parts of interviews. This is the most extreme and problematic form of influence that an interviewer can exert. Falsifications may severely bias the results of analyses and lead to incorrect results (Landrock, 2017a; Reuband, 1990; Schnell, 1991; Schraepfer and Wagner, 2003). A reliable strategy for identifying falsifications would therefore be extremely valuable to ensure high quality in interviewer-based survey research. However, research has shown that, based on univariate distributions (Menold and Kemper, 2014; Reuband, 1990; Schnell, 1991) and multivariate correlations (Landrock, 2017a), falsified and real data appear to be quite similar and that the existence of falsifications in data is thus not readily noticeable. Given that the falsification of interviews may be considered to be an extreme form of interviewer effect, statistically testing for interviewer effects might provide a more effective indicator for identifying falsifications. This paper therefore analyzes and compares interviewer effects in real survey data and in data falsified by interviewers. Using experimental data, the aim is to determine whether similar interviewer effects occur in falsified data and in real data or whether interviewer effects are larger and more diverse in

falsified data and may thus be used as an indicator for data contamination by interviewer falsifications (see Winker et al., 2015).

In falsified interviews, by definition, no interaction takes place between the respondent and the interviewer. Therefore, it may seem implausible to assume that interviewer effects occur in a dataset comprised of falsified data. However, in falsified interviews, interviewers obviously have a direct influence on the data reported as answers by the respondent. Yet, they have only a little information about the respondent. Consequently, the fabrication of plausible responses depends very strongly on the falsifier. Thus, interviewer effects – or, more precisely, “falsifier effects” – can be expected.

Different falsifiers may falsify the respondents’ answers in different ways. It is conceivable that certain socioeconomic, demographic, or psychological characteristics of the falsifiers may find their way into the data they falsify. Both the falsifiers’ perceptions of social reality and their falsifications are influenced by personal characteristics. Therefore, the interviewers’ characteristics should be significant explanatory variables in a dataset that is contaminated by interviewer falsifications. Moreover, I assume that interviewer effects are more pronounced in falsified than in real survey data (see Winker et al., 2015).

In the research presented in this paper, a number of variables that are known to be generally susceptible to interviewer effects are analyzed as dependent variables with the aim of determining (a) the degree to which interviewer effects occur in real and in falsified data and (b) whether there are differences between the interviewer effects in real and in falsified survey data.

## **7.2 Hypotheses**

To contribute to research on interviewer effects, to knowledge of interviewer falsifications and their impact on data quality, and to potential strategies for identifying contaminated data, the following two general hypotheses will be tested:

H1: Interviewer effects occur both in real and in falsified data.

As falsifying interviewers have only a little information about the respondent, they must draw on their personal experience of social reality in order to fabricate plausible answers to survey questions. Thus, interviewer effects may occur not only in real survey data but also in falsified survey data (see Winker et al., 2015).

H2: The interviewer effects in falsified data are larger than in real data.

I assume that sociodemographic or psychological characteristics of interviewers are more likely to find their way into falsified survey data than into real data.

Regarding the interviewer characteristics that may cause interviewer effects or influence the way in which an interviewer falsifies, explanatory variables will be analyzed that can theoretically be expected to be susceptible to interviewer effects. The following more specific hypotheses will be tested on real data and on falsified data:

H3a: The core sociodemographic characteristics of the interviewers affect the reported responses.

As reported by West and Blom (2016), Haunberger (2006), Mangione et al. (1992), and Liu and Stainback (2013), sociodemographic characteristics of the interviewer – in particular gender, age, and education – may lead to interviewer effects. I further expect that income, as an indicator of socioeconomic background, may also cause interviewer effects.

H3b: The magnitude of interviewer effects depends on the interviewer's experience.

Olson and Bilgen (2011) found that experienced interviewers caused larger interviewer effects than inexperienced interviewers. Hypothesis H3b will test whether this finding is replicated in the present study.

H3c: Associations exist between the behaviors and attitudes of interviewers and the reported behaviors and attitudes of the respondents they interview.

Following Schanz (1981), I assume that associations will be found between the answers of the interviewers and the answers of the respondents to the same survey question – in other words, that the interviewer's response to the same survey question affects the response reported by the respondent.

H3d: The occurrence and magnitude of interviewer effects depends on the personality traits of the interviewer.

Both West and Blom (2016) and Winker et al. (2015) found evidence that suggested that the personality traits of the interviewer may lead to interviewer effects. West and Blom (2016) reported an effect of interviewers' extraversion and self-confidence. Accordingly, I assume

that interviewers with higher levels of extraversion produce larger interviewer effects than introverted interviewers. By contrast, more conscientious interviewers should produce smaller interviewer effects than interviewers with a lower level of conscientiousness. With regard to self-confidence, I assume that interviewers with a higher level of perceived self-efficacy perform better, and therefore produce smaller interviewer effects, than interviewers with a lower level of perceived self-efficacy.

H3e: The magnitude of interviewer effects depends on the interviewer payment scheme used (payment per completed interview vs. payment per hour).

In their study of interviewer effects in real and falsified interviews, Winker et al. (2015) found that the payment scheme (i.e., the type of monetary compensation) applied had an impact on the collected data and therefore on the quality of a survey. I assume that interviewers who are paid per completed interview produce larger interviewer effects than interviewers paid per hour. Winker et al. (2015) also found correlations between the payment scheme and political participation (operationalized as the number of political activities mentioned by the respondent). For the real data, the authors showed that payment per hour was associated with a higher number of political activities mentioned. It would appear that payment per hour leads to more complete data and thus to higher data quality. Hypothesis H3e will test the assumption that interviewers who are paid per completed interview produce larger interviewer effects than interviewers who are paid per hour.

### **7.3 Data Base and Methods**

Due to the virtual non-existence of datasets with proven falsified interviews, experimental data were used to analyze falsified data and their differences to real data (see Winker et al., 2015). My data base comprised three datasets. The data were collected at the University of Giessen, Germany in summer 2011 in the framework of the research project IFiS – Identification of Falsifications in Surveys (see also Menold and Kemper, 2014; Winker et al., 2015).

In the first step, 78 interviewers conducted 710 real face-to-face interviews. The questionnaire consisted of 62 questions, which were taken mainly from the 1998 German General Social Survey (ALLBUS) questionnaire.<sup>16</sup> Besides sociodemographic questions, the

---

<sup>16</sup> <http://www.gesis.org/en/allbus/allbus-home/>

questionnaire comprised attitudinal and behavioral items on social, political, and economic topics. The average interview duration was 30 minutes. Both the respondents and the interviewers were students at the University of Giessen. The interviewers themselves selected the respondents on the university campus without any quota restrictions and interviewed them. The audio-recorded interviews were checked to make sure that they had been conducted correctly. Half of the interviewers were paid per completed interview (8 Euros), the other half were paid per hour (12 Euros). Prior to data collection, an interviewer training session was conducted, in the course of which the interviewers were familiarized with the research design and the questionnaire.

For the second dataset, 710 interviews were fabricated. For this purpose, the same interviewers who had conducted the real interviews were requested to fabricate survey data in the lab. Hence, for each real interview, a corresponding fabricated interview was obtained. Compensation was allocated either per interview (3 Euros per falsified interview) or per hour (9 Euros per hour). The falsifying interviewers were given details of the sociodemographic characteristics of the persons whose interviews they were to fabricate. These persons were real survey participants, who had been interviewed previously by another student interviewer. The information provided included the respondent's gender, age, subject studied, number of semesters enrolled, marital status, place of residence, living situation (i.e., the person or persons with whom the respondent lived in a household), and country of origin. In the case of a genuine (i.e., uninstructed) falsification in an actual fieldwork setting, the falsifying interviewer could easily have obtained this information by briefly interviewing the respondent. The falsifiers were requested to imagine the described person and to complete the questionnaire, thus fabricating the data as if they had been collected in a real survey fieldwork setting.

The exact instructions for falsifying an interview were:

Please read carefully the description of the person whose interview you are to falsify. Please complete the attached questionnaire as if you had really conducted a personal interview with the respondent. During falsification, please place the description of the respondent next to the questionnaire, so that you are always aware of the characteristics of that person.

The person whose interview you are to falsify...

- is female,
- is 20 years old,
- studies teaching,
- is enrolled in her second semester at a university.
- She is unmarried, in a steady relationship,
- lives in Huettenberg, a rural village in Hesse,
- with her parents or relatives.
- Country of birth: Germany.

As a last step, the interviewers themselves, as respondents, completed the same questionnaire that they had previously used for interviewing and falsifying. These self-administered interviews generated the third dataset.

This experimental setup has strengths, but it also has weaknesses. One weakness is that the respondents and interviewers were students and that core sociodemographic characteristics, such as age and education, therefore displayed only small variance (see Winker et al., 2015). The major strength of the experimental setup, compared to a standard field setting, was the possibility of collecting more information about the interviewers and their falsifying processes. Because they were surveyed with the same questionnaire as the proper respondents, the dataset includes not only information about respondents and fictitious respondents but also about the interviewers. This offers great potential for analyzing interviewer effects.

There are several possible approaches to investigating interviewer effects. Schanz (1981) analyzed the influences of interviewer characteristics on the response behavior of the participants by estimating multiple regression analyses. First, he included substantive explanatory variables; then he added interviewer variables. Mangione et al. (1992) and Groves and Magilavy (1986) measured interviewer effects by intraclass correlation. The intraclass correlation expresses the proportion of the item variance that is attributable to the interviewer (Mangione et al., 1992). In the absence of interviewer effects, the value of the intraclass correlation should be zero or close to zero (Beullens and Loosveldt, 2016). Olson and Bilgen (2011) estimated multilevel regression analyses with respondent characteristics such as age and education on the respondent level (individual level) and interviewer

characteristics such as age, education, and experience on the interviewer level (contextual level).

At first glance, it would appear to be useful to estimate ordinary least squares (OLS) regressions. However, especially when it comes to analyzing interviewer effects, it makes sense to assume that – as expressed in the above-mentioned hypotheses – the observations of the respondents (i.e., the individual interviews) are probably not independent from the interviewers. Therefore, the model assumptions of OLS regressions are not met. Rather, the data are organized hierarchically, and multilevel regression analyses are thus more appropriate (Hox, 1995). The respondents represent the individual level, and the interviewers represent the group or contextual level.

To investigate the impact of interviewer characteristics on substantive findings, intraclass correlations were also estimated and multilevel regression analyses were conducted. To answer the research question as to what influence interviewers have on the data and findings and whether there are differences between real and falsified data in this respect, identical multilevel regression models were estimated separately with real and with falsified data. Thus, to determine what differences occur, the respective results – in particular, the effects of the various independent variables – were compared. This approach also allowed the identification of interviewer effects on substantive findings.

## 7.4 Operationalization and Multilevel Regression Model

Table 7-1 gives an overview of the dependent and independent variables used. These variables are explained in more detail in the following sections.

**Table 7-1: Overview of variables used to analyze interviewer effects**

| Dependent Variables     | Independent Variables on the Individual (Respondent) Level  | Independent Variables on the Contextual (Interviewer) Level  |
|-------------------------|---|--|
| Income                  | Age<br>Living situation   | Payment scheme<br>Interviewer's gender<br>Interviewer's income<br>Interviewer's response to the same questions of the questionnaire<br>Interviewer's experience<br>Interviewer's extraversion<br>Interviewer's conscientiousness<br>Interviewer's level of perceived self-efficacy |
| Political participation | Gender<br>Internal political efficacy<br>Political dissatisfaction<br>Extremism   |  |
| Political anomy         | Economic dissatisfaction<br>External political efficacy   |  |
| Healthy eating behavior | Intention<br>Perceived behavioral control<br>TV consumption<br>Body mass index<br>Doing sports<br>Preference for healthy desserts |  |

### 7.4.1 Dependent Variables on the Individual Level

One aim of the present study was to analyze a number of dependent variables that I considered to be particularly susceptible to interviewer effects, namely (a) income, as a sensitive (and open-ended) factual question; (b) political participation, as a behavioral question; (c) political anomy, as an attitudinal question; and (d) healthy eating behavior, as an additional behavioral question.

*Income* was measured with the question: “How much money is at your disposal on average per month, during the current semester?”

*Political participation* was measured using a list of twelve political activities. The wording in the questionnaire was:

If you wanted to have political influence or to make your point of view felt on an issue that was important to you: Which of the possibilities listed on these cards would you use? Which of them would you consider? Please name the corresponding letters.

In a previous study, I analyzed the effects of falsified data on the results of multivariate theory-driven OLS regression analyses, using the explanation of political participation as an example (Landrock, 2017a). To investigate interviewer effects in the present study, the same



dependent and independent variables were applied in a multilevel regression. Factor analysis revealed that the factor *party-political activities* was an appropriate indicator for political participation. An additive index was calculated as a dependent variable measuring political participation. It consisted of the following three items:

- Participation in public discussions at meetings (factor loading: 0.701).
- Participation in a citizens' action group (factor loading 0.697).
- Voluntary work for a political party (factor loading 0.776).

*Political anomy* was measured with a scale consisting of four items that were summarized into an index that served as a third dependent variable (ZA & ZUMA, 2014). The items were:

- In spite of what some people say, the situation of the average man is getting worse, not better.
- It's hardly fair to bring a child into the world with the way things look for the future.
- Most public officials are not really interested in the problems of the average man.
- Most people don't really care what happens to the next fellow.

*Healthy eating behavior* was measured with the question: "On how many days per week do you eat healthy?" to analyze interviewer effects. I have used this variable in the past to explore the impact of falsifications on substantial findings in social science research on the basis of the theory of planned behavior (Landrock and Menold, 2016).

#### **7.4.2 Independent Variables on the Individual Level**

To implement multilevel regression models, statistically significant explanatory variables on the individual level were identified by estimating OLS regressions. These individual-level independent variables were included in the multilevel regression analyses presented in what follows. Given that my research interest here was to estimate interviewer effects, these variables may be considered as control variables.

For *income* as a dependent variable, the statistically significant explanatory variable on the respondent level – besides age – was the living situation, which was measured with the question: "Where are you living during the current semester?" This variable was dichotomized: The option "living with parents or relatives" was coded as 1; other options were coded as 0. The effect of age on income was positive. Regarding the living situation, the analysis revealed that students who lived with their parents or relatives reported less income than students who did not.

For *political participation*, the statistically significant explanatory variables on the respondent level were internal political efficacy, political dissatisfaction, extremism (captured with the left–right scale), and (female) gender. The means of the individual items were calculated for both internal political efficacy and political dissatisfaction; all items were adapted from the ALLBUS 1998 questionnaire (see Koch et al., 1999).

The items used to measure *internal political efficacy* were:

- I would have the confidence to take on an active role in a group concerned with political issues.
- Politics is so complicated that somebody like me can't understand what's going on at all. (Reverse-scored item)

*Political dissatisfaction* was measured with the following three items:

- Only when differences in income and social status are large enough is there any incentive for personal achievement.
- Differences in social position between people are acceptable because they basically reflect what one has made of the chances one had.
- I consider the social differences in this country to be just on the whole.

To measure extremism, the left–right scale from the ALLBUS 1998 questionnaire was used:

Many people use the terms “left” and “right” when they want to describe different political views. Here we have a scale which runs from left to right.

Thinking of your own political views, where would you place these on this scale?

To operationalize extremism (see Lüdemann, 2001), the original 10-point rating scale (with the value 1 on the left end of the scale and the value 10 on the right end of the scale) was recoded in such a way that the original values between 1 and 10 were assigned the new values between 5 and -5. These new values were then squared, thereby yielding a measurement for extremism where the value 1 stands for a very small degree of extremism and the value 25 for a very high degree of extremism (integrating both the left and the right ends of the left–right scale). All of these variables, except extremism, were found to have significant positive effects in the real data. As extremism had a significant positive effect in the falsified data, this independent variable was nonetheless included in the analysis of interviewer effects (Landrock, 2017a).

For the dependent variable *political anomy*, two statistically significant explanatory variables, economic dissatisfaction and external political efficacy were identified. *Economic*

*dissatisfaction* was measured with the question: “How would you generally rate the current economic situation in Germany?”

*External political efficacy* was measured with two items:

- Politicians don’t care much about what people like me think. (Reverse-scored item)
- In general, politicians try to represent the people’s interests.

Here, too, all items were adapted from the ALLBUS 1998 questionnaire. To operationalize external political efficacy, the means of the items were calculated (see Koch et al., 1999). Economic dissatisfaction was found to have a positive influence on political anomy, whereas external political efficacy had a negative effect.

To analyze interviewer effects on reported *healthy eating behavior*, a model based on the theory of planned behavior was adopted, which I applied in previous research on the impact of falsified data on substantive findings (Landrock and Menold, 2016).

The statistically significant independent variables for explaining healthy eating behavior on the individual level are the intention to eat healthily, perceived behavioral control, TV consumption, body mass index, doing sports, and preferring healthy desserts. The intention to eat healthily and perceived behavioral control were measured with two items each. These items were used to calculate an index for intention and for perceived behavioral control:

- In future I will eat healthy at least four days a week. (Intention)
- In the coming weeks I will eat healthy at least four days a week. (Intention)
- It is possible for me to eat healthy at least four days a week. (Perceived behavioral control)
- It is completely in my own hands to eat healthy at least four days a week. (Perceived behavioral control)

The questionnaire included the following question on TV consumption:

Thinking about the days when you watch TV, how long on average do you watch TV on these days – I mean in hours and minutes?

Body mass index was calculated on the basis of the self-reported height and weight of respondents. The variable *doing sports* was dichotomized; respondents were asked to answer an open-ended question about which sports they took part in at least occasionally. A list of 12 desserts was used to find out whether the respondents preferred healthy desserts. The variable *preference for healthy desserts* was dichotomized. Healthy desserts (fruit curd, fruit salad, or

yoghurt) were coded as 1; unhealthy desserts (mousse au chocolat, tiramisu, chocolate pudding, or pancakes) as 0.

As theory-driven explanatory variables, the intention to eat healthily and perceived behavioral control were found to have positive effects on reported healthy eating behavior. TV consumption and body mass index had negative effects, whereas doing sports and preferring healthy desserts showed positive effects, at least in the falsified data.

### **7.4.3 Independent Variables on the Contextual Level**

One aim of the present study was to identify interviewer characteristics on the contextual level that are linked to interviewer effects. The independent variables on the interviewer level that were tested are variables that are known to generally cause interviewer effects (see hypotheses in section 7.2 above). These variables are the *payment scheme* (payment per hour vs. payment per completed interview), the *interviewer's gender and income*, the *interviewer's response to the same question of the questionnaire*, and the *interviewer's experience*. Interviewers' personality traits were also tested, in particular *extraversion*, *conscientiousness*, and *perceived self-efficacy*, as they were considered relevant for analyzing interviewer effects.

First, the payment scheme was analyzed to determine whether the fact that an interviewer was paid per completed interview or per hour made a difference for the collected data, and therefore for the data quality. Winker et al. (2015) reported such an influence of the payment scheme on formal, non-content-related meta-indicators, for example non-differentiation. The payment scheme was varied in the research design: One half of the interviewers were paid per hour, the other half were paid per completed interview (see also section 7.3 above).

Many authors have described the core sociodemographic characteristics, namely gender, age, and education, as factors influencing interviewer effects (see Haunberger, 2006; Liu and Stainback, 2013). To my knowledge, researchers usually obtain only this basic information about interviewers from the fieldwork agencies, so that further interviewer characteristics typically cannot be analyzed. In the present study, I included the effects of the interviewers' gender as collected with the questionnaire completed by the interviewers themselves as respondents. Regarding age and education, the data show only small variances because all the interviewers were students and they were therefore very similar with respect to age and education. Instead, I considered the income of the interviewers, assuming that, in the

case of the student population of interviewers, income would be an appropriate indicator for the socioeconomic background of an interviewer, which might lead to interviewer effects.

As mentioned above, the interviewers themselves also completed the survey questionnaire as respondents. Thus it was possible to include as an independent variable their responses to the same questions that the respondents were also asked. The interviewers' responses were included as an explanatory variable on the contextual level in order to test whether there were positive correlations between the respondents' answers and the interviewers' answers. Schanz (1981) reported positive correlations between the attitudinal and behavioral characteristics of interviewers and respondents.

A further relevant factor for the occurrence of interviewer effects is interviewer experience (Olson and Bilgen, 2011). The question used to measure this variable was whether the interviewer had ever conducted interviews before participating in the present study. The variable was dichotomized into interviewers with experience and interviewers without experience.

The questionnaire also included scales to measure the personality traits of the interviewers. To analyze the effects of the interviewers' personality traits on the respondents' responses, these traits were included in the multilevel analyses on the contextual level. Perceived self-efficacy was measured as agreement with the following three items (Beierlein et al., 2014) using a seven-point rating scale:

- I can rely on my own abilities in difficult situations.
- I am able to solve most problems on my own.
- I can usually solve even challenging and complex tasks well.

Afterwards, the means of the items were calculated.

To measure extraversion and conscientiousness, the ten-item Big Five Inventory (BFI-10; Rammstedt et al., 2014) with a five-point rating scale was used:

I see myself as someone who...

- ...is reserved (Extraversion, reverse-scored item)
- ...is outgoing, sociable (Extraversion)
- ...tends to be lazy (Conscientiousness, reverse-scored item)
- ...does a thorough job (Conscientiousness)

For these variables, too, the means of each item were calculated.

#### 7.4.4 Multilevel Regression Model

To test the hypotheses and to investigate whether the interviewers' characteristics influenced the respondents' answers (e.g., reported income), separate identical multilevel regression models were developed for the real and the falsified data. The statistical software Stata 12 was used to conduct the multilevel analyses. First, a null model without an independent variable and without the contextual level was estimated in order to assess the goodness of fit of the baseline model on the basis of log likelihood, or deviance (Hox, 1995). Second, to estimate interviewer-level variance the contextual level was included in the random-intercept-only model (RIOM) in order to be able to answer questions such as whether the income reported by the respondent depended on the interviewer – in other words, whether the incomes of the respondents varied across interviewers. To this end, the intraclass correlation (ICC), which measures interviewer-level variance, was calculated. In the third step, the random-intercept model (RIM) was estimated. This model considers the influence of the individual respondent-level explanatory variables and controls for the contextual level. By including the interviewer-level explanatory variables of the contextual level (intercept-as-outcome model), direct effects of certain interviewer characteristics on respondents' responses were estimated. Thus, it could be determined, for example, whether the income reported by the respondents depended on the interviewers' gender. The results of the intercept-as-outcome model are shown in detail in Tables 4 and 5 (section 7.5.2).<sup>17</sup>

The likelihood-ratio test and McFadden's R-squared values were used to assess the goodness of fit of the model. With the likelihood-ratio tests, it was assessed, first, whether the multilevel approach was more appropriate than an OLS regression and, second, whether the estimated model extension (i.e., the reduction of deviance) was significant. McFadden's R-squared assesses model fit by comparing the log likelihood of the null model (i.e., the model without dependent variables and contextual level) with the log likelihood of the estimated model. According to Langer (2010: 756), values between 0.2 and 0.4 are excellent.

The dependent variables to be analyzed were required to be metric variables. Prior to the analyses, the independent variables were modified: The independent metric variables were grand-mean centered; the independent nominal variables were dichotomized and coded into binary variables.

---

<sup>17</sup> As an extension of the intercept-as-outcome models, the slope-as-outcome models were also estimated; they were not significant.

## 7.5 Results

### 7.5.1 Interviewer Effects in Real Data

First, interviewer effects in the real data were analyzed. Table 7-2 shows the random-intercept-only model (RIOM) for all of the dependent variables.<sup>18</sup> The intraclass correlations varied between 0.017 and 0.067, which means that between 1.7 percent and 6.7 percent of the total variance is accounted for by the contextual level (i.e., the interviewer level). These interviewer effects are very small. Only healthy eating behavior, with an ICC of 0.067, showed slightly increased interviewer effects (see Groves and Magilavy, 1986; Mangione et al., 1992). The likelihood-ratio test measures the significance of the models and indicates whether a multilevel model is more suitable than an OLS regression model. Regarding the dependent variables income and political participation, the RIOMs were not significant, which means that multilevel models were not appropriate and OLS regressions should be estimated instead. Regarding political anomy and healthy eating behavior, the RIOMs were significant; multilevel models could thus be preferred over OLS models. In the next step, the individual respondent-level variables were included in the model, and the random-intercept model (RIM) was developed. In the case of political anomy and healthy eating behavior as dependent variables, these models were not significant. Thus it can be assumed that interviewer effects scarcely exist in the real data.

**Table 7-2: Interviewer effects in the real data (random-intercept-only models, RIOMs)**

| RIOMs                             | Dependent Variables       |  |                                    |                                   |
|-----------------------------------|---------------------------|--|------------------------------------|-----------------------------------|
|                                   | Income<br>$\sigma^2$ (SE) | Political Particip.<br>$\sigma^2$ (SE) | Political Anomy<br>$\sigma^2$ (SE) | Healthy Eating<br>$\sigma^2$ (SE) |
| Resid. variance<br>(respondents)  | 143206.6<br>(8553.957)    | 0.131<br>(0.007)                       | 1.183<br>(0.071)                   | 2.933<br>(0.165)                  |
| Resid. variance<br>(interviewers) | 3660.958<br>(3674.375)    | 0.002<br>(0.003)                       | 0.063<br>(0.034)                   | 0.210<br>(0.087)                  |
| ICC                               | 0.025                     | 0.017                                  | 0.050                              | 0.067                             |
| LR test (p)                       | 0.1356                    | 0.1834                                 | 0.0114                             | 0.0007                            |
| N                                 | 644                       | 710                                    | 623                                | 710                               |

<sup>18</sup> Regarding political anomy, it should be mentioned that there were a large number of missing values, due, in particular, to the item “Most public officials are not really interested in the problems of the average man” (56 missing values).

### 7.5.2 Interviewer Effects in Falsified data

In the second step, interviewer effects in the falsified data were analyzed accordingly.<sup>19</sup> Table 7-3 shows the results of the RIOMs. The likelihood-ratio tests indicated that the models for all dependent variables were significant, which implies that the multilevel approach was more appropriate than the OLS regression approach. With values between 0.17 and 0.21, the intraclass correlations were much higher than in the real data, which means that the contextual level explained between 17 and 21% of the total variance. These strong interviewer effects indicate that individual characteristics, attitudes, and behaviors of the interviewers found their way into the falsified data. Thus, interviewer effects in the falsified data were further analyzed in order to determine which interviewer characteristics, attitudes or behaviors were particularly associated with interviewer effects.

**Table 7-3: Interviewer effects in the falsified data (random-intercept-only models, RIOMs)**

| RIOMs                             | Dependent Variables       |  |                                    |                                   |
|-----------------------------------|---------------------------|--|------------------------------------|-----------------------------------|
|                                   | Income<br>$\sigma^2$ (SE) | Political Particip.<br>$\sigma^2$ (SE) | Political Anomy<br>$\sigma^2$ (SE) | Healthy Eating<br>$\sigma^2$ (SE) |
| Resid. variance<br>(respondents)  | 30678.33<br>(1887.241)    | 0.102<br>(0.006)                       | 1.125<br>(0.065)                   | 1.869<br>(0.105)                  |
| Resid. variance<br>(interviewers) | 7913.874<br>(1964.437)    | 0.020<br>(0.005)                       | 0.271<br>(0.065)                   | 0.506<br>(0.115)                  |
| ICC                               | 0.205                     | 0.165                                  | 0.194                              | 0.213                             |
| LR test (p)                       | 0.0000                    | 0.0000                                 | 0.0000                             | 0.0000                            |
| N                                 | 606                       | 708                                    | 681                                | 710                               |

In the third step, the RIOM was extended by including the respondent characteristics on the individual level (RIM, not shown here). Afterwards, the interviewer characteristics on the contextual level were included, thus developing the intercept-as-outcome model (IOM), which estimates the direct effects of the independent variables on the interviewer level. The further extensions of the IOM were not significant for any of the dependent variables. Therefore, the random-intercept, random-slope models with cross-level interactions could not be estimated. Table 7-4 shows the results of the final IOM for the dependent variables income and political participation.

<sup>19</sup> In the falsified data, there were a large number of missing cases in the case of income . I assume that the question is difficult to falsify and that the falsifiers therefore preferred to report item nonresponse.



As can be seen from Table 7-4, the models fit well: The likelihood-ratio test indicated that both the models themselves and the model extensions to IOMs were significant. The McFadden R-squared values of 0.16 and 0.64 were at least very reasonable.

**Table 7-4: Results of ML regression in the falsified data (intercept-as-outcome models, IOMs)**

| IOMs                                       | Dependent Variables |          |                         |        |
|--|---------------------|----------|-------------------------|--------|
|  | Income              |          | Political Participation |        |
| Fixed Part                                 | Coeff.              | SE       | Coeff.                  | SE     |
| <b>Constant</b>                            | 725.907 ***         | 23.732   | 0.266 ***               | 0.036  |
| <b>Respondent level</b>                    |                     |          |                         |        |
| Age  | 10.381 ***          | 2.345    | -                       | -      |
| Living with parents / relatives (ref.: no) | -176.879 ***        | 21.467   | -                       | -      |
| Internal political efficacy                | -                   | -        | 0.128 ***               | 0.011  |
| Political dissatisfaction                  | -                   | -        | 0.034 +                 | 0.019  |
| Gender = female (ref.: male)               | -                   | -        | 0.035 +                 | 0.019  |
| Extremism                                  | -                   | -        | 0.017 ***               | 0.003  |
| <b>Interviewer level</b>                   |                     |          |                         |        |
| Payment per hour (ref.: per interview)     | 2.435               | 23.428   | -0.025                  | 0.035  |
| Gender = female (ref.: male)               | -51.359 +           | 26.539   | 0.086 *                 | 0.039  |
| Income                                     | -                   | -        | 0.000                   | 0.000  |
| Interviewer's answer                       | 0.114 *             | 0.053    | 0.259 ***               | 0.052  |
| Experience (ref.: no)                      | -4.696              | 29.644   | -0.034                  | 0.044  |
| Extraversion                               | -1.050              | 14.651   | 0.017                   | 0.022  |
| Conscientiousness                          | 17.575              | 15.002   | 0.022                   | 0.022  |
| Perceived self-efficacy                    | 2.372               | 12.341   | -0.013                  | 0.019  |
| <b>Random Part</b>                         |                     |          |                         |        |
|  | $\sigma^2$          | SE       | $\sigma^2$              | SE     |
| Respondents' residual variance             | 26933.240           | 1797.859 | 0.074                   | 0.005  |
| Interviewers' residual variance            | 4784.561            | 1509.125 | 0.010                   | 0.003  |
| <b>Model fit</b>                           |                     |          |                         |        |
| Log likelihood                             | -3392.254           |          | -92.393                 |        |
| N  |                     | 516      |                         | 579    |
| LR test (p)                                |                     | 0.0000   |                         | 0.0000 |
| LR test model extens. (p)                  |                     | 0.0000   |                         | 0.0000 |
| McFadden's R <sup>2</sup>                  |                     | 0.1641   |                         | 0.6433 |

Notes: \*\*\* p<0.001; \*\* p<0.01; \* p<0.05; + p<0.10

The results show that all individual variables on the respondent level were significant, at least at the ten percent level, which is not surprising as they already proved to have significant influence in the previously performed OLS regressions. However, for the analysis of interviewer effects, the more relevant results were found on the contextual level. Significant

effects on the dependent variables were not found for the payment scheme, the interviewers' personality traits, or the interviewers' experience. The interviewers' income had no significant effect on reported political participation. However, for income and political participation as dependent variables, significant effects of the interviewers' gender and their answers to the same survey questions could be identified.

Female falsifying interviewers tended to report lower incomes and higher values for political participation of the respondents than did male falsifying interviewers. Evidence was found that the gender of the interviewer tended to affect reported income and political participation in the case of the falsified data. It was also found that the interviewers' answers to the same questions had a positive effect on the reported respondents' answers. Thus, there were positive correlations between the falsifiers' attitudes and behaviors and the falsified reported attitudes and behaviors of the respondents. Presumably, the interviewers used their own income and political participation as a knowledge base for what a realistic income and political participation level might be for the interviews they were falsifying.

The models estimated for political anomy and healthy eating behavior as dependent variables yielded very similar results (Table 7-5). In both cases, the interviewers' answers to the same questions had a positive effect on the falsified reported answers of the respondents. In the case of healthy eating behavior as a dependent variable, the interviewers' gender affected the reported falsified response. Male falsifiers reported higher values for healthy eating. Thus, an impact of the attitudes and behaviors of the falsifying interviewers on all four analyzed variables could be identified.

**Table 7-5: Results of ML regression in the falsified data (intercept-as-outcome models, IOMs)**

| IOMs                                      | Dependent Variables |       |                |       |
|---|---------------------|-------|----------------|-------|
|   | Political Anomy     |       | Healthy Eating |       |
| Fixed Part                                | Coeff.              | SE    | Coeff.         | SE    |
| <b>Constant</b>                           | 1.691 ***           | 0.130 | 4.580 ***      | 0.140 |
| <b>Respondent level</b>                   |                     |       |                |       |
| External political efficacy               | -0.544 ***          | 0.045 | -              | -     |
| Economic dissatisfaction                  | 0.091               | 0.079 | -              | -     |
| Intention                                 | -                   | -     | 0.353 ***      | 0.032 |
| Perceived behavioral control              | -                   | -     | 0.359 ***      | 0.046 |
| TV consumption                            | -                   | -     | -0.003 **      | 0.001 |
| Doing sports (ref.: no)                   | -                   | -     | 0.117 +        | 0.070 |
| Preference for health desserts (ref.: no) | -                   | -     | 0.005          | 0.010 |
| BMI                                       | -                   | -     | -0.100 ***     | 0.018 |
| <b>Interviewer level</b>                  |                     |       |                |       |
| Payment per hour (ref.: per interview)    | 0.041               | 0.126 | -0.089         | 0.133 |
| Gender = female (ref.: male)              | -0.229              | 0.148 | -0.341 *       | 0.147 |
| Income                                    | 0.000               | 0.000 | 0.000          | 0.000 |
| Interviewer's answer                      | 0.195 ***           | 0.054 | 0.160 ***      | 0.039 |
| Experience (ref.: no)                     | -0.026              | 0.156 | 0.020          | 0.163 |
| Extraversion                              | 0.085               | 0.079 | 0.123          | 0.083 |
| Conscientiousness                         | 0.032               | 0.081 | 0.130          | 0.083 |
| Perceived self-efficacy                   | -0.079              | 0.067 | -0.079         | 0.070 |
| <b>Random Part</b>                        |                     |       |                |       |
|   | $\sigma^2$          | SE    | $\sigma^2$     | SE    |
| Respondents' resid. variance              | 0.896               | 0.057 | 0.998          | 0.063 |
| Interviewers' resid. variance             | 0.133               | 0.042 | 0.143          | 0.047 |
| <b>Model fit</b>                          |                     |       |                |       |
| Log likelihood                            | -797.383            |       | -827.605       |       |
| N   | 565                 |       | 565            |       |
| LR test (p)                               | 0.0000              |       | 0.0000         |       |
| LR test model extension (p)               | 0.0000              |       | 0.0000         |       |
| McFadden's R <sup>2</sup>                 | 0.2613              |       | 0.3703         |       |

Notes: \*\*\* p<0.001; \*\* p<0.01; \* p<0.05; + p<0.10

### 7.5.3 Summary and Review of Hypotheses

First, I will review the two general hypotheses:

H1: Interviewer effects occur both in real and in falsified data.

This hypothesis cannot be confirmed. Interviewer effects were identified in the falsified data but not in the real data.

H2: The interviewer effects in falsified data are larger than in real data.

This hypothesis can be clearly confirmed. Large interviewer effects occurred in the falsified data, whereas interviewer effects could not be identified in the real data.

Next, I will review the more specific hypotheses regarding characteristics of the interviewers that may cause interviewer effects:

H3a: The core sociodemographic characteristics of the interviewers affect the reported responses.

As no effects of the core sociodemographic characteristics of the interviewers were measurable in the real data, this hypothesis must be rejected for the real data. With regard to the falsified data, the analysis of the effect of the interviewers' gender on the dependent variables revealed that female falsifiers reported lower income, higher political participation, and lower values for healthy eating behavior than did their male counterparts. The interviewers' age and education were too homogeneous to be tested. With the exception of income as a dependent variable (see H3c), the interviewers' income does not appear to have affected the falsified responses. Accordingly, for the falsified data, the hypothesis can be confirmed with respect to gender.

H3b: The magnitude of interviewer effects depends on the interviewer's experience.

This hypothesis could not be confirmed for the real or the falsified data: No effect of interviewer experience on any of the dependent variables was found.

H3c: Associations exist between the behaviors and attitudes of interviewers and the reported behaviors and attitudes of the respondents they interview.

This hypothesis cannot be confirmed for the real data, where no interviewer effects were found. However, strong evidence was found in support of the hypothesis in the falsified data:

For all four dependent variables, significant positive correlations were found between the interviewers' answers as respondents and the falsified answers to the same survey questions.

H3d: The occurrence and magnitude of interviewer effects depends on the personality traits of the interviewer.

This hypothesis cannot be confirmed for the real data or for the falsified data. No effects of the personality traits on the dependent variables could be identified either in the real data or the falsified data.

H3e: The magnitude of interviewer effects depends on the interviewer payment scheme used (payment per completed interview vs. payment per hour).

This hypothesis cannot be confirmed for the real data or for the falsified data. Although previous research (see Winker et al., 2015) has shown that the payment scheme used (payment per completed interview vs. payment per hour) generally has an impact on the collected data, the present analyses did not detect effects of the payment scheme.

In summary, it can be stated that no interviewer effects of any kind were found in the real data. In the falsified data, the occurrence and magnitude of interviewer effects does not appear to have depended on the interviewers' experience or personality traits, or on the payment scheme used. However, effects of the interviewers' gender were found on the falsified reported income, political participation, and eating behavior of respondents. Furthermore, the interviewers' own attitudes and behaviors were correlated with the falsified reported attitudes and behaviors of the respondents. Thus, the falsifiers' attitudes and behaviors found their way into the falsified data and influenced the data reported as answers of the respondents.

## **7.6 Conclusions and Recommendations**

The findings of the present study suggest that interviewer effects are clearly stronger in falsified data than in real data: The real data, derived from actual conducted interviews, does not appear to be contaminated by interviewer effects at all. This can be taken as an indication of high data quality, which may be due to the fact that the real interviews were audio-recorded and the fieldwork was intensively monitored. By contrast, very strong interviewer effects were measured in the falsified dataset. This suggests that the process of falsifying leads to a

pronounced impact of the falsifiers' sociodemographic characteristics, attitudes, and behaviors on the data reported as answers of the respondents.

However, the interviewer effects (or, more precisely, "falsifier effects") identified in the falsified data were smaller than expected. One reason for this may be that both the respondents and the interviewers were students. Therefore, the falsifiers were familiar with the respondents' social reality and were able to give realistic answers – which reduced the magnitude of the interviewer effects. (This may also be a reason for the absence of interviewer effects in the real data.) A second reason why interviewer effects in the falsified data were smaller than expected may be that, despite the fact that the dependent variables used were empirically shown to be susceptible to interviewer effects, more appropriate dependent variables could possibly have been found to analyze interviewer effects.

The fact that neither the payment scheme nor the interviewers' experience caused interviewer effects is surprising because current findings in the literature suggest that they should have. Winker et al. (2015) found that the payment scheme had an impact on formal, non-content-related meta-indicators such as non-differentiation. However, the present study analyzed content-related dependent variables. A further reason why the payment scheme did not have the hypothesized influence could be that the instructed falsifiers in the present experimental study had an intrinsic motivation to participate in the study and were therefore less frustrated by payment per completed interview than an interviewer in a real fieldwork setting might have been. Moreover, the interviewers in the present study selected the respondents on the university campus and interviewed them themselves. In a real fieldwork setting, the interviewers must contact certain predefined target persons, which may be time-consuming. In such a case it would appear plausible that the payment scheme would make a difference and that payment per hour might enhance motivation to contact the predefined target person. The lack of support for the hypothesized influence of interviewer experience might be due to the fact that the students who stated that they had conducted interviews before were still less experienced than the experienced interviewers in the studies in which interviewer effects have been found.

One limitation of the present study is the fact that the respondents and interviewers were students and that core sociodemographic characteristics, such as age and education, displayed only small variance. Moreover, in a real fieldwork setting, it would hardly be possible to implement an experimental approach such as that employed here. Nonetheless, I assume that the present results are generalizable, not least because interviewers in social

science research and market research are often students. However, further research will be needed to confirm the generalizability of my results to real survey settings.

A number of recommendations can be derived from the present findings. First, researchers conducting interviewer-based surveys should collect as much information about the interviewers as possible and feasible (see Bogner and Landrock, 2016; Winker et al., 2015). In particular, as the present study shows, interviewer responses to the same questions that the respondents are asked are highly suitable for detecting interviewer effects in the case of falsified interviews. The interviewers could be requested to complete the survey questionnaire as part of interviewer training, for example. This would have at least two positive effects: First, the interviewers would familiarize themselves with the questionnaire, as a preparation for conducting the interviews; second, the researchers could get to know the interviewers.

A further recommendation that can be derived from the findings of the present study is that researchers using interviewer-based data should check the data for interviewer effects, especially if they suspect that falsifications may have occurred. Falsification checking should be implemented at least by calculating intraclass correlations or conducting multilevel analyses as presented in this paper. This can be done for the entire dataset or only for suspicious cases – provided, of course, more than one interviewer is involved. If a large share of the variance is explained by interviewer-level variables, this may be an indication of contamination of the dataset by interviewer falsifications. In light of the fact that neither bivariate nor multivariate correlational analyses have proved effective in unambiguously establishing the existence of falsifications, the assessment strategies presented here may be very valuable for improving the quality and accuracy of survey data.

## 8. Investigating Interviewer Falsifications – A Quasi-experimental Design<sup>20</sup>

**Abstract:** This paper discusses the strengths and limitations of experimentally-orientated research in Sociology. In principal, results from experiments cannot be generalized. Nevertheless they allow controlling for certain conditions that are impossible to measure in a survey. This paper gives an example: It presents findings from a quasi-experimental research on interviewer falsifications: Interviewers conducted real standardized interviews and subsequently falsified corresponding interviews in the lab. This enables a comparison of real and falsified survey data. Our research design has the limitation that there is no proof that actual falsifiers in an actual survey fieldwork environment would have falsified interviews in the same way as did the participants of our study. However, only this quasi-experimental approach allows us to know for sure which interviews are falsified.

### 8.1 Research Question

The research question this paper aims to discuss is whether findings from experimentally orientated research are generalizable, as well as what the strengths and limitations are. These questions are addressed theoretically, referring to existing methodological literature, and discussed on the basis of empirical findings, using an example for a quasi-experimental setup of a methodological social science research project. The project “IFiS – Identification of Falsifications in Surveys”, which was conducted by two research teams at GESIS Mannheim and at the University of Giessen, aims to compare real and falsified survey data and to enable researchers to identify interviewer falsifications. Referring to this example, the question posed here is: Can experimental designs in general relevantly contribute to social science knowledge by asking whether our experimental setup in particular, as an example, is appropriate to analyze differences between real and falsified survey data. Or more concretely: Are our findings on interviewer falsifications generalizable?

With respect to the fact that face-to-face interviews are widely used for data collection in empirical social science research, the question arises how interviewer falsifications can be

---

<sup>20</sup> Already published: Landrock, Uta (2017c): Investigating Interviewer Falsifications – A Quasi-experimental Design. *Bulletin of Sociological Methodology* 136(1): 5-20. Copyright © 2017 (Uta Landrock). Reprinted by permission of SAGE Publications. The final publication is available at SAGE through <http://dx.doi.org/10.1177/0759106317725640>.



empirically studied. The employment of interviewers generally has a positive influence on data quality (Mangione et al., 1992; van Meter, 2005), but always bears the risk that interviewers deviate from their instructions and interviewer guidelines intentionally and falsify parts of an interview or the entire interview (Bredl et al., 2013; van der Zouwen, 2006). This may contaminate the data and thus affect the data quality (AAPOR, 2003). A problem connected to this risk is that it is hardly possible to determine a suspicious interview unambiguously as falsified or not (Menold et al., 2013). Re-interviewing as a common mean to detect falsifying interviewers (Biemer and Stokes, 1989) is costly and time consuming and therefore impractical for each collected interview (Menold et al., 2013). The same is true for contacting and interviewing the interviewers subsequently to clarify data anomalies (Pagès et al., 2006). Therefore the question arises how to efficiently detect interviewer falsifications and thus, how real and falsified data differ. The research on interviewer falsifications, however, must deal with the obstacle that data sets with confirmed falsified interviews rarely exist in social science fieldwork (cp. Winker et al., 2015). Therefore our research project IFiS applies an experimental setup which allows us to know undoubtedly whether a certain interview is falsified or conducted correctly. Thus we are able to identify and analyze differences between real and falsified survey data. Furthermore we get to know how falsifiers proceed while falsifying. Finally, we can support empirically researching social scientists in identifying interviewer falsifications.

To answer the research question regarding the generalizability of experimentally-orientated research we describe in the following sections the IFiS project with its experimental setup and data base as well as our findings. Afterwards we present the strengths and limitations of experimental and quasi-experimental research designs. In the final section we discuss whether our findings on interviewer falsifications are generalizable.

## **8.2 State of Research on Interviewer Falsifications**

Precise statements about the exact numbers of falsifications in surveys are hardly possible; however, different authors report that they do not exceed five percent (Koch, 1995; Krejsa et al., 1999; Li et al., 2009). Some authors describe even higher proportions, for example in surveys without broad field controls (Bredl et al., 2012). Either way, the occurrence of falsifications and its consequences for the collected data call for research on the effects of interviewer falsifications.

The question arises of why interviewers decide to falsify survey data. Interviewers

may decide to falsify because they want to optimize their cost-benefit balance. Thus, they reduce time and effort, particularly when their objective is difficult, for example reaching a certain difficult-to-reach target person and obtaining participation (Sodeur, 2007; Turner et al., 2002; Winker, 2016).

Previous research has indicated that real and falsified data differ only slightly with respect to proportions and means of the responses, which means that there are large similarities between real and falsified survey data (Reuband, 1990; Schnell, 1991). Nevertheless, these studies also show that falsifications have a relevant influence on the covariance structure of data and may lead to biased results of multivariate analyses (Reuband, 1990; Schnell, 1991; Schraepler and Wagner, 2003).

Schraepler and Wagner (2003) estimated an OLS regression on the log gross income to analyze the impact of falsified data on the results. The authors used data from the German Socio Economic Panel. This data consists of real and falsified data which were obtained in a real survey field setting and afterwards identified as interviewer falsifications. The inclusion of these falsified data led to less explained variances and to biases of the effects of independent variables in different directions. The effects of some explanatory variables, namely age and gender, were overestimated, while the effects of others, i.e. the duration of training and the working hours per week, were underestimated (Schraepler and Wagner, 2003).

In contrast, the research of Reuband (1990) as well as of Schnell (1991) ascertained that falsifications increased the explained variances. In both studies artificially-produced falsified data were used. Reuband (1990) as well as Schnell (1991) instructed their study participants to use the provided factual sociodemographic information of real survey respondents to invent responses to the survey questionnaire. Reuband analyzed the attitude toward abortion and reports that the effects of the included independent variables were larger in the falsified than in the real data (Reuband, 1990). The results described by Schnell (1991) were very similar. Schnell analyzed subjective political competence. In his analyses there are also more significant effects of explanatory variables in the falsified than in the real data (Schnell, 1991).

There is evidence in both studies that falsifiers, at least under certain conditions, may produce more consistent models of behavior than may occur in real life (Reuband, 1990; Schnell, 1991). The falsifiers overestimate the variation related to sociodemographic characteristics. This very likely reflects that they rely on the provided sociodemographic

information, which is the only information about the respondents that they have. The falsifiers apply general stereotypes and implicit everyday knowledge, such as “men earn more than women do,”. This can, in the case of falsified data, lead to significant relationships with variables which are insignificant when analyzing real data (Reuband, 1990; Schnell, 1991). Thus, the regression models of the falsified data show more consistencies (Reuband, 1990; Schnell, 1991).

To summarize the state of research on interviewer falsifications: Falsifying interviewers are able to reproduce the proportions and means of actual responses to survey questions. Falsifications impact the covariance structure of data and may bias substantial findings. Furthermore, falsifiers tend to overestimate the relevance of the sociodemographic information they have about the respondents for the dependent variables; and they apply stereotypes and implicit knowledge in order to guess the direction of the effects. This may lead to more consistent models of explaining behavior in falsified interview data than expectable in real life (cp. Landrock and Menold, 2016).

### **8.3 The IFiS Project - Experimental Setup and Data Base**

The main goal of the IFiS project was the research on consequences and identification of interviewer falsifications in survey data. For this reason it was crucial to analyze the differences between undoubtedly real and undoubtedly falsified survey data. As previously mentioned, survey data with confirmed falsified interviews rarely exist (Winker et al., 2015). Similarly it is difficult to claim with certainty that a particular survey interview is *not* falsified. Therefore, we had to implement an experimental setup which allowed us to obtain evidentially falsified as well as evidentially real survey data (see also Menold and Kemper, 2014).

First, we recruited 78 interviewers. These interviewers as well as the respondents were students at the University of Giessen. The interviewers recruited the respondents themselves without any quota restriction on the University campus and conducted 710 face-to-face interviews in a real fieldwork setting. All interviews were audio-recorded and checked afterwards to make sure that the interviews were actually conducted and performed correctly. Each interviewer conducted nine interviews on average.

Second, the interviewers took over the role of falsifying interviewers and were instructed to falsify interviews in the lab. For this purpose the respondents were randomly assigned to the falsifying interviewers. We checked that no interviewer falsified an interview

of a survey respondent whom he or she had previously interviewed. For the falsification task, the interviewers received basic sociodemographic information of the real survey respondents whose interviews they were supposed to falsify. The given information were of that kind a falsifying interviewer in a real fieldwork setting could also have obtained easily, for example through a short interview with the assigned respondent without going through the complete questionnaire. Thus, for each real interview, a corresponding falsified interview was obtained to make the closest comparison to each other as possible. The exact instruction for the falsifications and the given information (illustrated by one exemplary case) were:

Please read the description of the person, whose interview you are supposed to falsify, carefully. Please fill in the attached questionnaire as if you had conducted a personal interview with the respondent in reality. Please place the description of the respondent next to the questionnaire while falsifying, so that you are always aware of the characteristics of this person.

The person, whose interview is to falsify...

- is female,
- is 20 years old,
- studies teaching,
- is enrolled in her second semester at a university.
- She is unmarried, in a stable relationship,
- lives in Huettenberg, a rural village in Hesse,
- with her parents or relatives.
- Country of birth: Germany.

Third, the interviewers were surveyed using an additional self-administered interviewer questionnaire, about *how* they conducted the falsifications and which strategies they applied. Additionally, the interviewers filled in the actual survey questionnaire for themselves.

The survey questionnaire contained 62 questions which were mainly adopted from the ALLBUS 2008 and 2010, the German General Social survey (GESIS, 2013; GESIS, 2015). It covered questions on political, social, and economic attitudes, as well as behaviors, personality traits, and sociodemographic information. The average duration of one interview was 30 minutes.

The interviewers were randomly assigned to two different payment schemes: half of

the interviewers was paid per completed interview while the other half was paid per hour. The payment per hour was 12 Euros in the case of conducting the real interviews and 9 Euros in the case of falsifying interviews. The payment per interview was 8 Euros per real interview and 3 Euros per falsified interview.

#### **8.4 Main Findings of the IFiS Project**

With respect to the state of the research, there are at least two reasons why we would expect differences between real and falsified data (cp. Menold et al., 2013). Firstly, as mentioned above, falsifying interviewers may overestimate the impact of the respondents' sociodemographic characteristics on the responses respondents would give in an interview. They may, on the base of the sociodemographic information they have about a respondent, apply stereotypes about the respondent's possible attitudes and behaviors when answering the questionnaire in place of the real respondent (Schnell, 1991). This may lead to stronger and more significant effects and to more consistent regression models. Secondly, real respondents and falsifying interviewers differ in the mode of answering survey questions. Real respondents tend to satisficing response behaviors, as described by Krosnick and Alwin (1987). They want to reduce their cognitive effort; therefore they tend to give a satisfactory, plausible answer, which may not necessarily be the optimal, perfectly precise answer (Krosnick and Alwin, 1987). Falsifiers, on the one hand, follow a similar motivation. Just as the respondents, they want to reduce effort and time so they decide to falsify. Therefore falsifying interviewers also tend to satisfice; their satisficing processes may even be stronger than those of real respondents (Menold et al., 2013). On the other hand, falsifying interviewers have to avoid being detected. Therefore they spend much effort to invent plausible answers. One can assume that the effort of a falsifying interviewer in inventing a plausible answer is higher than the effort of a respondent to give an acceptable answer. These differences should be visible when comparing real and falsified data (Menold et al., 2013).

As already reported by Reuband (1990) and Schnell (1991), as well as in the IFiS project, we ascertain that in falsified data the proportions and means are very similar to the proportions and means in the real data. Menold and Kemper (2014) identified only slight differences with respect to the distributions, particularly for attitudinal and behavioral questions.

In order to identify and measure satisficing strategies, we developed formal indicators which inform about the level of response sets (Menold et al., 2013; Kemper and Menold,

2014; Menold and Kemper, 2014). The assumption was that falsifying interviewers, who wanted to avoid being caught, would show less satisficing than real survey respondents, who wanted to reduce their cognitive effort. For this assumption we find strong support. We could identify significant differences between the response sets of real and those of falsified survey data. Amongst others, the real data show higher levels of extreme response style and acquiescence response style as well as stronger primacy and recency effects (Menold et al., 2013). These results indicate that falsifiers respond more carefully than real respondents (Kemper and Menold, 2014) and may help to identify falsifications.

In order to assess differences regarding the covariance structure, we developed the approach of estimating theory-driven models; this way we were able to research whether falsifiers are also able to construct complex models of behavior in accordance with social reality (Landrock and Menold, 2016; Landrock, 2017a). We estimated theory-driven OLS regressions and compared the results of the falsified with those of the real data. We applied two established theories of social science research. Firstly, applying the theory of planned behavior (Ajzen, 1991; Ajzen and Krebs, 1994), we developed a model to explain healthy eating behavior (Landrock and Menold, 2016). Secondly, we adopted a rational choice model of explaining political participation (cp. Kaase and Marsh, 1979; Opp and Finkel, 2001) and, again, compared the results of falsified and of real data (Landrock, 2017a).

Our hypothesis had been that falsifiers underestimate theoretically predicted correlations and overestimate the effects of sociodemographic information given in the instructions as well as stereotypical relations. In the case of applying the theory of planned behavior (see Table 8-1) as well as in the case of explaining political participation (see Table 8-2), we can report that falsifying interviewers are indeed not able to apply social science theories to invent plausible answers. While in the real data the theoretically predicted relations can mainly be verified, this is not the case for the falsified interviews. We can also confirm in both applications that falsifiers overestimate stereotypical relations. Our further expectation that falsifiers would construct more consistent models of behavior than the respondents in real life can partly be confirmed. It is true for explaining political participation. The assumption that falsifiers overestimate the effects of sociodemographic variables is true for applying the theory of planned behavior to predict healthy eating behavior. In the case of explaining political participation, we can find no differences between real and falsified data in that respect. Nevertheless we can conclude that the estimation of complex theory-driven models makes differences regarding the covariance structure between real and falsified data

visible (Landrock and Menold, 2016; Landrock, 2017a).

**Table 8-1: Overview over independent variables that are significant either in the real or in the falsified data; dependent variable: healthy food consumption**

| Included variables   | Real data  | Falsified data |
|--|------------|----------------|
| <i>Theory-driven variables:</i>  |            |                |
| Attitude   | -          | (-)            |
| Subjective norm  | +          | (n.s.)         |
| Perceived behavioral control   | +          | +              |
| Intention  | +          | +              |
| <i>Sociodemographic variables:</i>   |            |                |
| Gender = female (reference: male)  | (n.s.)     | (+)            |
| Living with parents (reference: all others)  | (n.s.)     | +              |
| Being single (reference: having a partner)   | (n.s.)     | +              |
| University as the center of interests and activities<br>(ref.: university not as center) | +          | (+)            |
| <i>Stereotypical relations:</i>  |            |                |
| TV consumption in min/day  | (n.s.)     | -              |
| Doing sports (reference: no)   | (n.s.)     | +              |
| Favorite dessert is healthy (reference: unhealthy)                                       | (n.s.)     | +              |
| BMI  | (n.s.)     | -              |
| Adj. R <sup>2</sup> (sign.)  | .50 (.000) | .48 (.000)     |

+ positive significant effect (p<.05); - negative significant effect (p<.05);  
 (+) positive effect p<.10; (-) negative effect (p<.10); (n.s.) not significant  
 Source: Landrock and Menold, 2016: 310 (condensed table)

**Table 8-2: Overview over independent variables that are significant either in the real or in the falsified data; dependent variable: political participation**

| Included variables   | Real data  | Falsified data |
|--|------------|----------------|
| <i>Theory-driven variables:</i>                                      |            |                |
| Internal political efficacy  | +          | +              |
| Dissatisfaction with individual economic situation                   | (-)        | (n.s.)         |
| Political dissatisfaction  | +          | (n.s.)         |
| <i>Sociodemographic variables:</i>                                   |            |                |
| Gender = female (reference: male)                                    | +          | +              |
| <i>Stereotypical relations:</i>                                      |            |                |
| Self-placement on the left-right-dimension<br>(1 = right, 10 = left) | (n.s.)     | +              |
| Adj. R <sup>2</sup> (sign.)  | .17 (.000) | .23 (.000)     |

+ positive significant effect (p<.05); - negative significant effect (p<.05);  
 (+) positive effect p<.10; (-) negative effect (p<.10); (n.s.) not significant  
 Source: Landrock, 2017a: 454 (condensed table)

For our final research question we aimed to identify which impact the interviewer and respectively the falsifier, has on the respondents' responses in terms of statistically measurable interviewer effects (Winker et al., 2015; Landrock, 2017b). With respect to existing research on interviewer effects, we assumed that in real data interviewer effects may occur which may affect substantive findings (cp. Groves and Magilavy, 1986). Nevertheless we expected stronger interviewer effects in the falsified data than in the real data (Winker et al., 2015; Landrock, 2017b) since by inventing plausible answers the falsifying interviewers directly influence the reported responses. Therefore these responses should severely depend on the falsifying interviewers with their attitudes, behaviors, and their social reality (Landrock, 2017b).

Winker et al. (2015) analyzed the effects of interviewer characteristics on several formal meta-indicators, which had been used before to compare the response sets and satisficing strategies between real and falsified data (Menold et al., 2013). The results (see Table 8-3) show that, for example, rounding and non-differentiation are not affected by interviewer characteristics in the real data, while in the falsified data interviewer effects on primacy, non-differentiation, acquiescent response style, and rounding can be found (Winker et al., 2015).

**Table 8-3: Overview over the significant effects of interviewer' characteristics on formal meta-indicators in the real and in the falsified data**

|                                       | Effects in<br>real data | Effects in<br>falsified data (p<.05)               |
|---------------------------------------|-------------------------|--|
| Interviewer's gender                  | (no effect)             | Primacy  |
| Payment scheme                        | (no effect)             | Non-differentiation                                |
| Interviewer's attitude towards future | (no effect)             | Non-differentiation,<br>acquiescent response style |
| Interviewer's attitude towards rules  | (no effect)             | Rounding, non-differentiation                      |
| Interviewer's awareness               | (no effect)             | Non-differentiation                                |

Source: Winker et al., 2015: 429-431 (condensed table)

We also analyzed interviewer effects on substantive findings and for that purpose estimated different multi-level models. We analyzed the income of the respondents and their anomy (i.e. the perceived lack of norms and rules) as well as the healthy eating behavior and political participation as dependent variables (Landrock, 2017b). For all four examples, we find no interviewer effects in the real data, but strong interviewer effects in the falsified data (see Table 8-4).



**Table 8-4: Interviewer effects (i.e., intraclass correlations ICC) in the real and in the falsified data, dependent variables: income, political participation, anomy, healthy eating behavior**

| ICC            | Income | Political participation | Anomy | Eating healthy |
|----------------|--------|-------------------------|-------|----------------|
| Real data      | .025   | .017                    | .050  | .067           |
| Falsified data | .205   | .165                    | .194  | .213           |

Source: Landrock, 2017b: 19-20 (condensed table)

For three of the four dependent variables, the interviewers' gender influences the reported responses in the case of falsified data. Again, in the case of falsified data, we find in all four models significant effects and robust correlations between the answer of the fictitious respondents and the interviewers' own answer to the same question in the survey questionnaire (see Table 8-5). This means that the reported falsified income, anomy, healthy eating behavior, and political participation, are strongly influenced by the income, anomy, healthy eating behavior and political participation of the falsifying interviewer (Landrock, 2017b). Against the background of strong interviewer effects in the falsified data we can recommend empirically researching social scientists to collect as much information as possible about their interviewers, since this information may help to identify falsifications (Winker et al., 2015; Landrock, 2017b).

**Table 8-5: Results of ML regression in the falsified data, overview over significant independent variables on the interviewer level; dependent variables: income, political participation, anomy, eating healthy**

|                         | Interviewer's gender = female<br>(ref. male) | Interviewers' answer<br>to the same question |
|-------------------------|--|--|
| Income                  | (-)  | +  |
| Political participation | +  | +  |
| Anomy                   | (n.s.)                                       | +  |
| Eating healthy          | -  | +  |

+ positive significant effect ( $p < .05$ ); - negative significant effect ( $p < .05$ );

(+) positive effect  $p < .10$ ; (-) negative effect ( $p < .10$ ); (n.s.) not significant

Source: Landrock (2017b: 21-22, 24-25), (condensed table)

## 8.5 Experimental and Quasi-experimental Research Designs

### 8.5.1 The Rationale of Experimental Research

One major disadvantage of non-experimental social science research, in particular of cross-sectional studies, is the fact that it is not possible to control for all relevant or potentially confounding variables. These non-controlled variables may be correlated with explaining

variables, and they may influence the outcome variables, i.e. the dependent variable a researcher is interested in (Angrist and Pischke, 2015). Thus, in non-experimental research, selection biases may occur<sup>21</sup>: Differences, for example, in the means of the outcome variable are not caused by the identified explaining variable, but by unobserved differences between the groups, regardless of the treatment (Angrist and Pischke, 2015).

The elimination of this selection bias is the main advantage of experimental approaches over survey research and a way in which experimental designs can specifically contribute to social science knowledge. The key element of experimental research design is, aside from manipulation of one condition, random assignment (Angrist and Pischke, 2015; Diekmann, 2010). Randomized assignment to treatment and control groups eliminates selection biases (Angrist and Pischke, 2015), because treatment and control groups differ only in the type of treatment and are similar otherwise (Angrist and Pischke, 2015). Randomization leads to similar groups because the randomly assigned individuals come from the same population. Thus, experiments assure a high internal validity because through randomization the researcher can control for selection biases, which may be caused by unobserved differences (Diekmann, 2010).

However, there are also drawbacks regarding the implementation of randomized experimental designs. The treatment and control group will only be similar if the groups are large enough (Angrist and Pischke, 2015; Diekmann, 2010). Difficulties arise if the number of cases remains small, because in that case randomization may not lead to similarity of the treatment and control groups (Diekmann, 2010).

The implementation of experiments and random assignment requires much effort and there may be cases where randomization is impossible, for example due to practical or ethical issues<sup>22</sup> (Diekmann, 2010). One may consider that randomization is not the only way to identify causal relations and to make comparisons possible (Angrist and Pischke, 2015). In any randomized and unrandomized setup, it is useful to check for balanced answers to questions, whether participants of the treatment and control groups are in fact similar (Angrist and Pischke, 2015). All in all, there are research questions where randomization is neither suitable nor required. In such cases, quasi-experiments – i.e. experiments without

---

<sup>21</sup> Angrist and Pischke (2015) described the selection bias on the example of health insurance. People with health insurance are healthier than people without. However, there might be a causal effect of health insurance on health, but there are also unobserved differences like education, employment and income, which are related to health insurance status as well as to health (Angrist and Pischke, 2015).

<sup>22</sup> For example the non-treatment in medical experiments touches ethical issues (cp. Diekmann, 2010).

randomization – may be appropriate (Diekmann, 2010).

### **8.5.2 Quasi-experimental Designs**

As described above, random assignments enable experimental control as well as internal validity, because selection biases cannot occur since the groups are similar (Shadish et al., 2002). Nevertheless randomization “is just one part of experimental design” (Shadish et al., 2002: 256), which may be “conceptually irrelevant” (Shadish et al., 2002: 256) and difficult to implement.

There are some further limitations concerning the feasibility of a large experiment, in particular in cases like the IFiS project where the intended and conducted analyses required a large number of cases. If randomization cannot adequately be implemented, a quasi-experimental controlled design may be more appropriate (Shadish et al., 2002).

Quasi-experimental controlled designs are suitable when the researcher knows that the treatment works and that the effects or changes are caused by treatment. Then quasi-experimental controlled designs can give evidence regarding treatment effects (Shadish et al., 2002). Although quasi-experimental designs cannot completely control for all unobserved differences “alternative explanations can sometimes be made implausible” (Shadish et al., 2002: 252). The application of quasi-experimentation may be suitable in cases which require following the experimental rationale of manipulation without randomized assignment (Diekmann, 2010).

### **8.5.3 External Validity and Generalizability**

One central issue concerning experimental research is the question of external validity, i.e. whether the results of experimental studies can be transferred to social reality (Diekmann, 2010; Levitt and List, 2007). There are some differences between experimental settings and social reality, namely moral considerations of the participants, the processes of self-selection, and their knowledge of being monitored (Levitt and List, 2007). This means that the participants are volunteers and they know that they are being observed by researchers (Levitt and List, 2007). An additional, and presumably in the case of the IFiS project the most serious aspect, is that the participants assume that there is a legitimate purpose of the experiment which may influence decisions (Levitt and List, 2007).

The generalizability of laboratory experiments is further limited by the fact that alternative actions are artificially restricted and that it is not possible to control completely for

all contextual factors of a participant (Levitt and List, 2007). In contrast, field experiments as well as quasi-experimental approaches require a theory-driven research design to conduct the study adequately and interpret the results reasonably (Wolbring and Keuschnigg, 2015).

Further problems of generalizability refer to the fact that the situation is abstract and not comparable with real life (Franzen and Pointner, 2013). Additionally, the participants of experiments are usually students (Franzen and Pointner, 2013), which is also true for the IFiS project. All these issues indicate that the results of laboratory or field experiments as well as quasi-experimental designs are not always generalizable. However, there is a strong recommendation for a theory-driven implementation of experimental designs to enable transferring the attained results into social reality:

“Theory is the tool that permits us to take results from one environment to predict in another.”  
(Levitt and List, 2007: 170)

## **8.6 Benefits of Quasi-experimental Controlled Design Implemented by IFiS**

Coming from the methodological principles of experimental research, the optimal experimental design to conduct the IFiS project would have consisted of three stages (see Angrist and Pischke, 2015): Firstly, we would have defined target persons and recruited these respondents to fill in the questionnaire as a reference. Secondly, we would have recruited interviewers and randomly assigned them to the control group of real interviewers or to the treatment group of falsifying interviewers. There would have been comprehensive field control procedures implemented to ensure that the interviews were obtained correctly and that the real interviewers were honest. The treatment group of falsifying interviewers would have fabricated the interviews in the lab as done in the actual implemented experimental setup. Finally, we would have collected the data and compared and analyzed the responses reported by real and falsifying interviewers.

There are many good reasons why we did not implement an experimental design with randomized assignment: Firstly, our intended multivariate analyses required a large amount of data. That implied that we had to obtain the largest number of cases as possible to compare real and falsified data comprehensively. Thus, it would have been quite counterproductive to split our group of interviewers. The two randomly assigned groups would not have met the requirements of being large enough to be similar with respect to unobserved differences (cp. Angrist and Pischke, 2015) as well as to conduct the intended multivariate analyses.

We would also have had to face the problem of repeated measurements of the survey

respondents, which may have led to maturing effects (Imbens and Rubin, 2015).

For these important reasons we decided to implement a design where in each step the same subjects were observed. Firstly, all interviewers conducted real interviews; secondly, all interviewers conducted falsified interviews. Thus, applying a sort of “within-subject design” (Franzen and Pointner, 2013), we were able to control for unobserved differences and therefore eliminate selection biases through an approach where each interviewer conducted real and falsified interviews.

We have to consider that randomized experiments have “many practical problems” (Shadish et al., 2002: 277), which limit randomization (Shadish et al., 2002). In the case of our study randomization could not adequately be implemented because of the requirement of the largest possible number of cases.

When randomization is not feasible, quasi-experimental controlled designs, or quasi-experiments, can be implemented instead. Quasi-experimental approaches are suitable particularly when the researcher knows that the treatment works and that the effects or observed changes are in fact caused by the treatment. That in turn requires a theory-driven implementation of study design (Franzen and Pointner, 2013; Wolbring and Keuschnigg, 2015) which enables quasi-experimental controlled designs to give evidence about the treatment effects. One can say that quasi-experimental designs are suitable when “alternative explanations can (...) be made implausible” (Shadish et al., 2002: 252; see above). In the IFiS project we applied a quasi-experimental design, i.e. an experiment without randomization, which follows an experimental rationale.

Although we gain important benefit with the chosen quasi-experimental setup, there are some limitations and problems regarding the generalizability of findings. Firstly, like in all laboratory experiments, our participants have volunteered to take part in the experiment, i.e. their decision to take part was intentional, which also means that they were aware of the fact that they were under observation (Franzen and Pointner, 2013). In the framework of the IFiS project one may assume that student volunteers are particularly interested in science and research and therefore, for example, more capable than others to apply social science theories while falsifying (Levitt and List, 2007). Furthermore, experimental and quasi-experimental designs cannot control completely for all contextual factors of a participant (Franzen and Pointner, 2013). The context of the participant, in the case of conducting real interviews in IFiS, includes for example the interviewer’s gender or his/her attitudes; these may influence the real interviews and cause interviewer effects. In the case of falsifications, probably the

perception of social norms or own experiences may influence the falsified data (cp. Levitt and List, 2007).

Experimental and quasi-experimental designs artificially restrict the possible actions of the participants (Franzen and Pointner, 2013). In a real fieldwork setting the interviewers would not be restricted to the two options of being either honest or falsifying entire interviews; they would also be able to choose alternative actions and, for example, decide to conduct partial falsifications. A further limitation is that participants in experimental and quasi-experimental studies assume that there is a legitimate purpose for the experience, which may influence decisions (Franzen and Pointner, 2013). In the case of the IFiS project, instructed falsifications in the lab were intended and legitimated, which is definitely not the case in real life. It is, for example, not completely clear, whether the falsifiers have invented the same responses as “real” falsifiers would have, who would have needed to avoid being detected. We tried to solve this problem by implementing a lottery game, where three of those falsifiers, who were not detected by us, were paid 100 Euros (Menold and Kemper, 2014).

Additionally, as mentioned before, all interviewers conducted both, real and falsified interviews, which eliminates selection bias. Regarding this repeated measurement, we must consider that we cannot control for changes which may have occurred in the meantime (cp. Angrist and Pischke, 2015). The falsifiers became familiar with the real respondents and therefore may have been able to reproduce typical responses. However, this is not problematic, since in a real fieldwork setting falsifying interviewers may also conduct honest interviews before beginning to falsify.

## **8.7 Conclusion**

This paper discusses the potentials and limitations of experimentally orientated research in Sociology – using the example of the IFiS project on identification of interviewer falsifications. The research question was whether our quasi-experimental setup is appropriate to analyze differences between real and falsified survey data. Or more concrete: Whether our findings on interviewer falsifications are generalizable.

The starting point was that research on interviewer falsifications in survey data is nearly impossible without experimental manipulation, as it is hard to find confirmed falsifications and confirmed real interviews from a real fieldwork setting. Furthermore, in the case of IFiS, randomization was not feasible because of practical reasons in obtaining the largest possible number of cases.

To conclude, there are advantages, but also limitations of quasi-experimental research on the example of interviewer falsifications in surveys. With respect to limitations, one can find reasons which may lead to the assumption that the results are not generalizable. Firstly, the experimental setup, particularly the instructed falsifications in the lab, generates an abstract situation which is not comparable with real life (cp. Franzen and Pointner, 2013). This also implies that it is not possible to control for all potentially relevant social conditions. Another reason which may limit generalizability can be found in the non-representativeness of the sample of student participants (cp. Franzen and Pointner, 2013). In our study, this is true for the interviewers as well as for the interviewees. Nevertheless, one may assume that in a real fieldwork setting at least interviewers are often students. Secondly, there is no proof for the assumption that actual falsifiers in real survey fieldwork would falsify interviews in the same way as the participants in our study. However, there is also no reason to believe that this assumption is wrong. Thirdly, the results of experimental setups are not always generalizable, but the strength of the IFiS project lies in the theory-driven approach to research on interviewer falsifications. Theoretically reasoned experimental designs allow transferring the results out of the laboratory and into social reality (cp. Levitt and List, 2007).

Nevertheless, our project has meaningful strengths. Our data enable comparisons between real and falsified survey data and thus to find ways and methods which help to identify interviewer falsifications. Finally, only experimental or quasi-experimental approaches allow controlling for the number of falsifications as well as knowing precisely which interviews are confirmed falsified and which were confirmed real.

On the base of the knowledge gained from the described quasi-experimental study, we continued our work on interviewer falsifications and conducted the next stage of the project as a field experiment in the strict sense of the word to obtain “real” falsifications in a real fieldwork setting.

## 9. Summary of results

### 9.1 First application: Explaining healthy food consumption

Publication 1:

Landrock, Uta and Menold, Natalja (2016): Validation of Theoretical Assumptions with Real and Falsified Survey Data. *Statistical Journal of the IAOS* 32(3): 305-312.

In this paper, with Natalja Menold as co-author, I analyze differences between real and falsified data with respect to theory-driven multivariate regression models. As a first of two examples for a realistic research endeavor in social sciences, I examine the determinants that explain healthy eating behavior. As an underlying theoretical approach for predicting self-reported healthy eating behavior, the theory of planned behavior (Ajzen, 1991) was applied. Attitudes towards this behavior, subjective norms, perceived behavioral control, and intention to eat healthy are assumed to be predicting variables. The regression model includes three blocks of independent variables: theory-driven explaining variables as mentioned above, sociodemographic variables, and variables regarding stereotypical relations as ad hoc predictors. OLS regressions are calculated separately for real and falsified data, afterwards the results are compared (Table 9-1).

In the following the hypotheses are reviewed. The hypotheses H1.1, H1.2, and H1.3 can be confirmed:

H1.1: For the theoretically predicted determinants there are more significant effects in the real than in the falsified data.

Although in the falsified data intention and perceived behavioral control show significant effects ( $p < .05$ ), there are more significant effects of theory-driven explaining variables in the real data. Attitude, subjective norm, perceived behavioral control, and intention significantly affect the healthy food consumption, as predicted by the theory of planned behavior.

H1.2: For real sociodemographic information known to falsifying interviewers, there are more significant effects in the falsified than in the real data.

All four sociodemographic variables were significant at least at the 10% level in the case of the falsified data. In the real data only one variable is significant.



H1.3: For further correlates (based on common stereotypes), there are more significant effects in the falsified than in the real data.

In the real data no stereotypical correlation can be identified. By contrast, in the fabricated data the falsifiers invent significant correlations for all four stereotypical independent variables.

Hypothesis H1.4 cannot be confirmed:

H1.4: The falsifiers construct more consistent cases than real life; therefore, the explained variance is higher in the falsified than in the real data.

In the real data the overall explained variance is slightly higher than in the falsified data. The findings of Reuband (1990) and Schnell (1991) cannot be replicated. One possible explanation may be that the question whether explained variances are higher in the real or in the falsified data depends on the dependent variable: The better a variable can be explained by social science theories and the less it is linked to existing stereotypes the more likely it is that the explained variance of regression models is higher in the real than in the falsified data. Furthermore, this result corresponds with the findings of Schraepler and Wagner (2003), who also report that the consideration of falsified data reduces the explained variances. In the case of healthy eating, high ratios of explained variance are achieved in the real data (adjusted R-squared of .50) because the theory of planned behavior is very well suited for explaining healthy eating. In the case of falsified data high ratios of explained variance are achieved (adjusted R-squared of .48) because falsifiers here evidently draw on strong stereotypes. Consequently, there are virtually no differences in the explained variances. Thus, one may conclude that the explained variance per se is not a good indicator to distinguish between real and falsified data. For this purpose one should, instead, compare which explaining variables are statistically significant.

**Table 9-1: Overview over independent variables that are significant either in the real or in the falsified data; dependent variable: healthy food consumption**

| Included variables   | Real data  | Falsified data |
|--|------------|----------------|
| <i>Theory-driven variables:</i>  |            |                |
| Attitude   | -          | (-)            |
| Subjective norm  | +          | (n.s.)         |
| Perceived behavioral control   | +          | +              |
| Intention  | +          | +              |
| <i>Sociodemographic variables:</i>   |            |                |
| Gender = female (reference: male)  | (n.s.)     | (+)            |
| Living with parents (reference: all others)  | (n.s.)     | +              |
| Being single (reference: having a partner)   | (n.s.)     | +              |
| University as the center of interests and activities<br>(ref.: university not as center) | +          | (+)            |
| <i>Stereotypical relations:</i>  |            |                |
| TV consumption in min/day  | (n.s.)     | -              |
| Doing sports (reference: no)   | (n.s.)     | +              |
| Favorite dessert is healthy (reference: unhealthy)                                       | (n.s.)     | +              |
| BMI  | (n.s.)     | -              |
| Adj. R <sup>2</sup> (sign.)  | .50 (.000) | .48 (.000)     |

+ positive significant effect (p<.05); - negative significant effect (p<.05);

(+) positive effect p<.10; (-) negative effect (p<.10); (n.s.) not significant

Source: Landrock, 2017c; Landrock and Menold, 2016: 310 (condensed table)

The results show that falsifiers rather do not apply social science theories. Instead, they seem to rely on stereotypes in order to invent as plausible data as possible. Furthermore they strongly rely on the sociodemographic information that they have been given (and that in a real fieldwork setting could easily be collected by a falsifying interviewer in a short interview). Falsifiers overestimate the influence of sociodemographic background and (other) stereotypical variables and underestimate the predicting effects of theory-driven variables – these relations seem to be difficult to reproduce by falsifying interviewers. This is a new finding in the research on interviewer falsifications, which may help to better understand the impact of falsified data on the results of substantial social science research.

The research design in this paper could have been altered in a few ways: The theory of planned behavior could have been applied to predict the intention (rather than the behavior) by attitude, subjective norm, and perceived behavioral control. Healthy eating behavior, in turn, could have been explained by intention and perceived behavioral control only, as

suggested by Ajzen (1991), instead of including attitude, subjective norm, and perceived behavioral control as well as intention in one single model. Also, one could have used the entire dataset and not only a subsample of interviewers. It remains to future research to test whether the presented results are robust enough to hold up if the research design is varied in these ways.

## 9.2 Second application: Explaining Political Participation

Publication 2:

Landrock, Uta (2017a): Explaining Political Participation – A Comparison of Real and Falsified Survey Data. *Statistical Journal of the IAOS* 32(3): 447-458.

The second paper, just as the first paper, aims to answer the question how real and falsified data differ with respect to theory-driven multivariate analyses and, accordingly, what effects falsified data have on substantial social science research. In the first publication, the explanation of healthy food consumption was taken as an example for a research endeavor – a behavior for which presumably comparably many stereotypes exist (even with respect to the theoretically predicted variables) and everyday knowledge can be applied rather easily. By contrast, in the second publication a considerably more complex and abstract correlation is examined, where falsifying interviewers presumably have to spend higher effort: I take the example of explaining political participation for which a scientifically accepted model with established operationalization and measurements (see Barnes et al., 1979) was applied. This second example for a research endeavor that actual social sciences are concerned with also serves to increase the reliability and generalizability of the findings. Therefore the comparison of the findings of the first two papers is of particular interest.

The corresponding OLS regression in the second paper, explaining political participation, follows the rationale of the model for healthy eating behavior (see previous section) to enable comparisons between the two examples: In a first block of independent variables, the theory-driven predicting variables are included. In this case, these are based on the explanatory model of Barnes and Kaase and colleagues (1979). The second block consists of sociodemographic variables, known to the falsifiers, and the third one of stereotypical variables (Table 9-2).

**Table 9-2: Overview over independent variables that are significant either in the real or in the falsified data; dependent variable: political participation**

| Included variables   | Real data  | Falsified data |
|--|------------|----------------|
| <i>Theory-driven variables:</i>                                      |            |                |
| Internal political efficacy  | +          | +              |
| Dissatisfaction with individual economic situation                   | (-)        | (n.s.)         |
| Political dissatisfaction  | +          | (n.s.)         |
| <i>Sociodemographic variables:</i>                                   |            |                |
| Gender = female (reference: male)                                    | +          | +              |
| <i>Further correlates found in research:</i>                         |            |                |
| Self-placement on the left-right-dimension<br>(1 = right, 10 = left) | (n.s.)     | +              |
| Adj. R <sup>2</sup> (sign.)  | .17 (.000) | .23 (.000)     |

+ positive significant effect (p<.05); - negative significant effect (p<.05);  
 (+) positive effect p<.10; (-) negative effect (p<.10); (n.s.) not significant  
 Source: Landrock, 2017c; Landrock, 2017a: 454 (condensed table)

With respect to the hypotheses, the following results can be reported: The hypotheses H1.1, H1.3, and H1.4 can be confirmed.

H1.1: For the theoretically predicted determinants there are more significant effects in the real than in the falsified data.

In the falsified data only internal political efficacy has a significant effect. In the real data, three independent variables – *internal political efficacy*, *dissatisfaction with individual economic situation*, and *political dissatisfaction* – are statistically significant, at least at the 10% level.

H1.3: For further correlates (found in research), there are more significant effects in the falsified than in the real data.

In the falsified data, the effect of further correlates found in research, namely *self-placement on the left-right-dimension*, is significant – this is not the case in the real data.

H1.4: The falsifiers construct more consistent cases than real life; therefore, the explained variance is higher in the falsified than in the real data.

The adjusted R-squared reaches 0.23 for the falsified data, but only 0.17 in the real data. With respect to the first paper and the interpretation of the unconfirmed hypothesis there, it has to be noted that it cannot be generalized that explained variances are higher in falsified than in real data. However, political participation seems to be a dependent variable for which the

assumption of higher explained variances in falsified data is true. This may be interpreted in a way that for political participation the existing stereotypes and implicit knowledge of falsifiers are stronger in relation to the ability of social science theories to actually explain the real phenomenon. In the real data where one could have expected a strong influence of the theory-driven variables the reached adjusted R-squared value is lower than in the falsified data where stereotypes and implicit knowledge gain relevance.

By contrast, hypothesis H1.2 cannot be confirmed:

H1.2: For real sociodemographic information known to falsifying interviewers, there are more significant effects in the falsified than in the real data.

Gender as sociodemographic variable is significant in both datasets. In the case of falsified data such effect was expected because it is known that falsifiers overestimate the effect of sociodemographic information. For the real dataset the effect of gender was shown in previous research on political participation (see Marsh and Kaase, 1979). Considering the fact that in this case the influence of gender is real, the data provide little opportunity for an overestimation of the sociodemographic background and a confirmation of H1.2. Probably the falsifying interviewers learned about the effect of gender in the real interviews that they had conducted before they falsified interviews. This may have contributed to diminish the differences between real and falsified data.

It needs to be concluded that the predicted differences between real and falsified data can only be confirmed if there is a suitable selection of independent variables in the models. In the case of political participation particularly the variable gender is not suitable since it is meaningful in real and in falsified data.

**Table 9-3: Overview over the tested hypotheses regarding the dependent variables: healthy eating and political participation**

|                                  | Healthy eating       | Political participation |
|----------------------------------|----------------------|-------------------------|
| H1: theory-driven variables      | Confirmed            | Confirmed               |
| H2: sociodemographic information | Confirmed            | <i>Not confirmed</i>    |
| H3: further correlates           | Confirmed            | Confirmed               |
| H4: explained variances          | <i>Not confirmed</i> | Confirmed               |

The results show, like in the case of the explanation of healthy eating and the application of the theory of planned behavior, that falsifiers are not fully able to invent models of explaining behavior that are consistent with social science theories (see overview in Table 9-3). In

contrast to the case of healthy eating, in the example of explaining political participation the falsifiers do not overestimate the influence of sociodemographic characteristics, namely gender. That is foremost due to the relevance of gender in the real data, which limits the possibility of any overestimation. The only other sociodemographic variable that had been tested in the models is age, which remains statistically insignificant in both datasets, probably because of the very limited variation within the sample of student participants. Nevertheless, this paper replicates the main result of the first publication and confirms that falsifying interviewers overestimate stereotypical relationships and underestimate the influence of the theory-driven explaining determinants. Complex correlations that do not correspond to stereotypes and everyday knowledge are difficult to falsify.

### 9.3 Interviewer effects

Publication 3:

Landrock, Uta (2017b): How interviewer effects differ in real and falsified survey data. Using multilevel analysis to identify interviewer falsifications. *mda* 11(2): 165-190.

This paper aims to answer how real and falsified survey data differ with respect to interviewer effects. Interviewer effects occur when respondents do not only respond to survey questions, but also to characteristics and behaviors of interviewers (see Bogner and Landrock, 2016). Consequently, interviewer effects may cause biased data and influence substantive social science research (Beullens and Loosveldt, 2016). Thus, they may affect data quality.

For falsified responses, interviewer effects in the sense of its definition cannot occur since an interaction of interviewer and respondent has never occurred. Nevertheless, the falsifying interviewer has a direct and obviously very strong impact on the reported responses. Therefore, one can expect that interviewer effects – in the sense of a statistical correlation between interviewer and reported answers – may be particularly strong in falsified data.

The question is how interviewers bring in their life in the real as well as in the falsified survey data in comparison. To provide an answer to this question intraclass correlations are calculated and multilevel regression analyses are applied on the example of four dependent variables, which can be considered as particularly prone to interviewer effects (at least in real data). The dependent variables are income, political participation, anomy, and healthy eating. *Income* is an example of a sensitive, open-ended, factual question. *Political participation* and

*healthy eating* are behavioral questions. *Anomy* is an application for an attitudinal question. Real and falsified data are analyzed separately, and then the results are compared (Table 9-4 and 9-5).

**Table 9-4: Interviewer effects (i.e., intraclass correlations ICC) in the real and in the falsified data, dependent variables: income, political participation, anomy, healthy eating behavior**

| ICC            | Income | Political participation | Anomy | Healthy eating |
|----------------|--------|-------------------------|-------|----------------|
| Real data      | .025   | .017                    | .050  | .067           |
| Falsified data | .205   | .165                    | .194  | .213           |

Source: Landrock, 2017c; Landrock, 2017b: 16, 17 (condensed table)

**Table 9-5: Results of multilevel regression analyses in the falsified data, overview over significant independent variables on the interviewer level; dependent variables: income, political participation, anomy, eating healthy**

|                         | Interviewer's gender = female<br>(ref. male) | Interviewers' answer<br>to the same question |
|-------------------------|--|--|
| Income                  | (-)  | +  |
| Political participation | +  | +  |
| Anomy                   | (n.s.)                                       | +  |
| Eating healthy          | -  | +  |

+ positive significant effect (p<.05); - negative significant effect (p<.05);  
(+) positive effect p<.10; (-) negative effect (p<.10); (n.s.) not significant  
Source: Landrock, 2017c; Landrock, 2017b: 18, 20 (condensed table)

With respect to the general hypotheses I find the following:

H2.1: Interviewer effects occur both in real and in falsified data.

This hypothesis cannot be confirmed, because interviewer effects were identified in the falsified data only, but not in the real data (see ICCs in Table 9-4). Despite the fact that interviewer effects generally can occur in survey research this is not the case in the interviews conducted in the IFiS project – maybe because the student interviewers were particularly intrinsically motivated to fulfil the interviewer guidelines accurately (when asked to do so) and because the interviews were audio-recorded so that there was a high level of social control.

H2.2: The interviewer effects in falsified data are larger than in real data.

This hypothesis can be evidently confirmed (see ICCs in Table 9-4). In the falsified data strong interviewer effects are found. In the real data no interviewer effects occur at all. This

supports the assumption that the direct influence of a falsifying interviewer on the reported answers creates particularly strong interviewer effects in the sense of statistical correlations between the interviewer and the reported answers.

Which interviewer characteristics find their way into the data? Regarding this question the same hypotheses were formulated for real and for falsified data equally. Given that, in the real data, no interviewer effects occurred at all none of these hypotheses can be confirmed for the real data. For the falsified data, some of the hypotheses can be confirmed and others cannot (see Table 9-5). Hypotheses H2.3b, H2.3d, and H2.3e cannot be confirmed:

H2.3b: The magnitude of interviewer effects depends on the interviewer's experience.

The effect of the interviewer's experience is statistically insignificant.

H2.3d: The occurrence and magnitude of interviewer effects depends on the personality traits of the interviewer.

The data at hand allowed testing the influence of extraversion, conscientiousness, and perceived self-efficacy. The effects of the personality traits of the interviewer are statistically insignificant.

H2.3e: The magnitude of interviewer effects depends on the interviewer payment scheme used (payment per completed interview vs. payment per hour).

The effect of the payment scheme is statistically insignificant.

Neither in the real nor in the falsified data an effect of any of the mentioned characteristics of the interviewer or of the interview situation, namely experience, personality traits, and payment scheme, on any of the dependent variables is found. Possibly, this can be explained by the fact that these hypotheses are inspired by the state of research on interviewer effects in real data, whereas the statistical influence of falsifiers on the reported answers might follow different mechanisms and rules. The payment scheme, for example, may be a condition affecting the interviewer's motivation to invest time and energy in the accurate realization of interviewer instructions (when conducting real interviews). It may also affect the likelihood that an interviewer decides to falsify, instead of conducting real interviews. But maybe it does not affect the way in which a falsifier falsifies an interview. For both groups of interviewers,



interviewers paid per hour and interviewers paid per completed interview, it is of crucial importance not to be detected – independent from the applied payment scheme. Thus, both have the same motivation to invest time and effort in the invention of false responses. The same may be true in respect to the interviewer's experience and the interviewer's personality traits.

Hypotheses H2.3a and H2.3c do find empirical support:

H2.3a: The core sociodemographic characteristics of the interviewers affect the reported responses.

For the real data, no effect of the core sociodemographic characteristics is found. By contrast, in the falsified data, I identify an effect of the interviewer's gender on the reported income, political participation, and healthy food consumption. The reported political participation is higher if falsifying interviewers are female; the reported income is higher and the eating behavior is healthier if falsifying interviewers are male. Therefore, in the case of falsified data, this hypothesis can be confirmed with respect to gender. Female and male falsifiers seem to perceive social realities differently and have different stereotypes which they translate into their invented responses.

H2.3c: Associations exist between the behaviors and attitudes of interviewers and the reported behaviors and attitudes of the respondents they interview.

This hypothesis has to be rejected for the real data (see above) but finds strong support in the falsified data. Significant positive correlations between the invented respondents' answers and the interviewers' answers to same survey questions can be shown for all four dependent variables. Obviously falsifying interviewers take their own social reality and their own attitudes and behaviors as an inspiration for inventing plausible responses.

To sum up: In the real data no interviewer effects occur at all – a finding that can be taken as an indication for good data quality in the real interviews conducted in the IFiS project. By contrast, in the fabricated data the falsifiers' characteristics, attitudes and behaviors find their way into the data and affect the fictitious answers of the respondents. These influences are very strong. Thus, the assumption is confirmed that due to the direct strong impact of falsifying interviewers on the reported responses, interviewer effects are particularly strong in falsified data. Among the falsifiers' characteristics, attitudes and behaviors, particularly the

falsifier's gender as well as his or her answer to the same question in the questionnaire have strong influence on the reported answers. By these insights, this publication contributes substantially to the research on interviewer falsifications. It also may offer a key for identifying datasets that are substantially contaminated by interviewer falsifications, that is, strong statistically significant interviewer effects can serve as an indication.

#### **9.4 Methodological Considerations regarding the quasi-experimental design**

Publication 4:

Landrock, Uta (2017c): Investigating Interviewer Falsifications – A Quasi-experimental Design. *Bulletin of Sociological Methodology* 136(1): 5-20.

In this paper the potentials and limitations of quasi-experimental setups are discussed in general, using the example of the IFiS research project. The question this publication aims to answer is, whether the reported results are generalizable. In other word: Is the applied quasi-experimental design appropriate to analyze differences between real and falsified survey data at all? One important aspect lies in the fact that it is nearly impossible to find datasets with confirmed real and confirmed falsified survey data (see Winker et al., 2015). This means there are no good alternatives for experimental or quasi-experimental methods in order to research differences between real and falsified survey data. The main argument for choosing a quasi-experimental rather than an experimental design is that in the frame of the IFiS project the largest possible number of cases is needed to realize the statistical analyses. Therefore randomization does not seem appropriate.

There are some reasons to doubt that the reported results are generalizable. One reason may be that the experimental setup creates an artificial situation that is not comparable to a real fieldwork setting, in which falsifying implies deviating from the interviewer instruction, instead of following the instruction to falsify. Another argument could be that the sample is too homogeneous because the participants are students. Additionally, there is no evidence that actual falsifiers would fabricate interviews in the same way as done in the IFiS project. On the other hand, there is also no reason to assume that this should not be the case. And the fact that all interviewers are students could also occur in a real fieldwork setting in actual survey research. The main argument for considering the data based on quasi-experimental research as

generalizable is the theory-driven approach of the IFiS project which allows transferring the results from the lab into social reality (see Levitt and List, 2007).

Accordingly, one can conclude that the quasi-experimental setup of the IFiS project is appropriate for research on interviewer falsifications, or at least the most appropriate approach that is feasible. The collected data allow comparing real and falsified survey data and thus offer the opportunity to search for strategies for identifying interviewer falsifications.

## 10. Conclusion

This PhD thesis researches differences between real and falsified survey data and aims to answer two questions: Which effects do falsified data have on substantial findings from theory-driven multivariate analyses? And: In what ways do interviewer effects differ in falsified data, compared to real data? Thus, it contributes to the research on interviewer falsifications.

The estimated theory-driven OLS regressions give evidence that falsifiers are not able to produce data in full accordance with the predictions of social science theories. The *theory-driven* explaining variables have – as predicted – more significant effects in the real and fewer in the falsified data. By contrast, *further correlates* that either refer to common stereotypes or that were found to be relevant under certain conditions have more significant effects in the falsified than in the real data. These findings may be considered as new insights in the research on interviewer falsification.

The same principally seems to be true for *sociodemographic* characteristics, even if the assumption of more significant effects of sociodemographic variables in the falsified than in the real data cannot be confirmed in both papers. However, falsifying interviewers obviously utilize satisficing strategies, general stereotypes and implicit or explicit everyday knowledge in order to invent plausible patterns of answers, and yet minimize the invested time and effort. Insofar, interviewers try to reach given goals with a minimum of invested resources. This also confirms the underlying theoretical assumptions that the behavior of interviewers – when deciding to conduct a real interview or to falsify and when deciding on how to falsify – can be explained by rational choice models.

This difference in *which* types of independent variables – theory-driven vs. further correlates vs. sociodemographic – turn out to be significant and have strong effects seems to

be decisive for the question which effects falsified data have on the results of content-related, theory-driven multivariate research. And, in principal, it could provide a key for identifying interviewer falsifications.

However, in a real research setting it will be hard or even impossible to evaluate whether the effects of the theoretically predicted explaining variables are weak enough and the effects of the stereotypical and sociodemographic variables strong enough to assume that interviewer falsifications have occurred. The difference *can* be detected in direct comparison of confirmed real and confirmed falsified data. By contrast, without a reference which effects exactly have to be expected in a dataset that is not contaminated by falsified interviews at all, the effects in a contaminated dataset will very likely not appear to be suspicious. Therefore it is doubtful whether the differences between real and falsified data, identified in the here presented research, may provide a *reliable* key for identifying interviewer falsifications in practice.

A further reason why the presented differences between real and falsified data (regarding which independent variables have significant effects) are not a practical key for the identification of falsifications is that they do not occur reliably enough. Their occurrence depends on rather technical aspects, as is shown by the second paper. In the case of explaining political participation (unlike in the case of explaining healthy eating behavior), the assumption that falsifiers overestimate the effects of sociodemographic variables cannot be confirmed. In the sample consisting of students only, the respondents' age turns out to be too homogeneous to be significant in the real or in the falsified data. Gender, instead, turns out to be significant in the real data already and therefore does not provide a further occasion in the falsified data to overestimate any sociodemographic influence. Accordingly in this case, the selection of independent sociodemographic variables does not allow confirming the hypothesis that the effect of sociodemographic characteristics is overestimated by falsifiers. Such may happen in other cases also. That means that the described differences between real and falsified data with respect to which independent variables have significant effects principally in fact *do* exist. However, it is not guaranteed that they can be measured in practice and, thus, can be applied to identify interviewer falsifications.

Regarding overall explained variances of regression analyses, also ambiguous findings have to be reported. In the case of explaining healthy eating behavior, the assumption of higher explained variances in the falsified data is not confirmed. In the case of explaining political participation, it is. The fact that the overall explained variances are not generally

higher in falsified data may be due to the fact that this assumption needs to be differentiated. The question, whether explained variances are higher in the real or in the falsified data probably depends on the dependent variable: With real data, a regression analysis will reach a high adjusted R-squared value if the dependent variable is well explained by existing social science theories and their operationalization in the dataset. With falsified data, a regression analysis will reach a high adjusted R-squared value if the dependent variable is associated with strong stereotypes or experiences in everyday knowledge, providing inspiration for the falsifiers. The better a variable can be explained by social science theories and the less it is linked to existing stereotypes the more likely it is that the explained variance of regression models is higher in the real than in the falsified data – and vice-versa: The weaker a variable is explained by social science theories and the more it is linked to existing stereotypes the more likely it is that the explained variance of regression models is higher in the falsified than in the real data. Obviously, healthy eating behavior is relatively well explained by social science theory in the real data and, at the same time, associated to strong stereotypes in the falsified data, so that in real and in falsified data high adjusted R-squared values are achieved. In the case of explaining political participation, in the real data the theoretically guided explanations are weaker than the linkage to stereotypes in the falsified data. Therefore, in the case of political participation the explained variances are higher in the falsified than in the real data.

All in all, the differences between real and falsified data turn out to be more subtle than one might have thought. Even the differences that could be theoretically predicted and empirically confirmed are not particularly profound: Even in falsified data, some of the theoretically predicted explaining variables are significant; even in real data there are significant effects of sociodemographic variables and further correlates. The question arises, why the falsifying interviewers are able to reproduce several aspects of social reality. First, in my view, this has to do with the fact that social sciences are based on successfully operating empirical research. Therefore, many data are available, substantial findings are empirically proven, and they are disseminated into the public. The results from social science research are published not only in scientific journals but also in mass media – in newspapers, TV documentations and talk shows – and find their way into public knowledge and general education. (The public discussions about the results of the PISA study may serve as an example for the dissemination of scientific knowledge into general knowledge.) Falsifiers are exposed to this information, even without a social science education, and remember a certain

share of it. Thus, through the extent of social science knowledge in society, the falsifiers are enabled to produce data in high accordance with social reality, which is comprehensively researched by empirical social scientists. A second reason may be that not all predictions of social science theories are far away from everyday knowledge and stereotypes; some seem intuitive and can be guessed by non-scientists. Third, the falsifiers have all conducted real interviews before falsifying and gained knowledge about the interviewees through this experience. A fourth explanation for the unexpectedly small differences between real and falsified data may lie in the fact that the interviewers as well as the respondents are students and therefore the interviewers are familiar with the living conditions of their respondents. The last two issues may be considered as limitations of the research design. These and other potential limitations will be discussed below.

Another implication of the described findings is that seemingly good results of multivariate analyses, referring to high proportions of explained variances or strong significant effects, do not necessarily stand for good data quality. Such findings can also be an indication for the occurrence of falsifications. Thus, if there are any doubts regarding the data, it can be recommended to separate suspicious from unsuspecting cases and to conduct theory-driven analyses with both subsamples separately. If such differences as the here reported ones occur, this may be a hint for possible interviewer falsifications. Of particular interest in that case is which subsample has more significant effects in which type of explaining variables (theory-driven, further correlates, or sociodemographic). Fewer significant effects among the theory-driven variables and more significant effects among the other explaining variables may point towards the occurrence of falsifications.

The results of multilevel regression analyses give evidence for the existence of interviewer effects in the falsified data, but not in real data. However, these “falsifier effects” are, in terms of occurrence and magnitude with respect to certain interviewers characteristics, smaller than one could have assumed. Nevertheless, the measures of intraclass correlations point out clear and statistically significant differences regarding interviewer effects. In the falsified data on average one fifth of the total variance is explained by the contextual level – that means by interviewers characteristics, attitudes and behaviors.

This finding firstly provides a further answer to the question how falsified data differ from real data. Secondly, it may provide a much more promising key for detecting interviewer falsifications: A high (two digit number) percentage of variance explained on the interviewer

level in a multilevel analysis may be taken as indicator for a high contamination of a dataset or subsample. The same may be true for a particularly strong influence of the interviewer's gender or the interviewer's answer to the same question in the questionnaire – if that information is available.

Thus, in order to detect interviewer falsifications in survey research, I recommend to, firstly, collect as much information as possible about the interviewers (see Bogner and Landrock, 2016; Winker et al., 2015). This recommendation is derived particularly from the result of strong correlations between the answers of the interviewer and the respondents' response to the same survey question. As part of the interviewer training, the interviewers should be requested to fill in the survey questionnaire themselves. This actually helps the interviewers to get to know the questionnaire better; at the same time, it helps the researchers to get to know their interviewers and to be able to check for correlations between the answers of interviewers and their respondents later on.

The second recommendation is to use this information about the interviewers to estimate multilevel regression analyses, in order to examine interviewer effects. At least intraclass correlations should be calculated. These analyses can be conducted for the entire dataset or for suspicious subsamples of the dataset (if more than one interviewer is considered as suspicious). A high share of variance explained by the contextual level as well as a strong and significant effect of interviewer-level variables can be indications of contamination by interviewer falsifications.

Given that multivariate analyses of statistical correlations have not found to be *unambiguously* able to differ between real and falsified data, the assessment of interviewer effects may be a valuable supplementary strategy to identify interviewer falsifications and hence to assure a high data quality. Still, the best strategy for ensuring a high data quality is certainly providing conditions that make it unlikely that interviewers decide to falsify at all: That is, interviewers should rather be paid per working hours than per completed interviews; the access to the field should be as easy as possible; the questionnaire should be easy to understand and to use; interviews should be short; the content of the questions should be interesting enough so that respondents have an interest in participating.

The implemented research design has few limitations that have to be mentioned and considered in future research. First, the described analyses are based on falsifications of entire interviews. In a real fieldwork setting, partial falsifications are probably the most prevalent

form of interviewer falsifications and therefore assumingly more relevant than falsifying entire interviews. Second, the reported findings are derived of datasets with 100 percent real and 100 percent falsified interviews. By contrast, in a real fieldwork setting the proportion of falsifications is probably much lower. On the one hand, this is fortunate since it means that the damage created by falsifications is limited. On the other hand, it implies that falsifications are harder to detect. A third limitation that needs to be considered is that in the quasi-experimental setup the interviewers conducted real interviews in the first stage of data collection and afterwards falsified the data in the lab. Consequently, the falsifying interviewers had experienced typical response patterns and were familiar with the expectable responses. However, this circumstance is also realistic for a real fieldwork setting because there an interviewer who decides to falsify would typically also conduct honest interviews before falsifying. Fourth, and perhaps most important of all, the study participants, interviewers as well as respondents, were all students. This implies that the interviewers and respondents are similar to each other. Therefore, the interviewers are familiar with the lifestyles, living conditions, and thoughts of their respondents. For the falsified data, that might improve the closeness to social reality regarding significant effects as well as explained variances. This could lead to an underestimation of the differences between the results for real and falsified data. In turn, one could expect larger differences between the models if not only students were employed as interviewers and if a representative sample would be examined. A fifth limitation of the research design, related to the previous one, is that the student population displays only small variances concerning age and education. That limits the possibilities to analyze the effects of these core sociodemographic variables. Further research will be needed that might replicate the reported findings without being affected by the mentioned limitations.

The subject of interviewer falsifications will not lose relevance since some of the major survey programs will continue working with interviewers. Therefore, the next steps of research on interviewer falsifications should be to investigate whether, on the base of the reported findings – particularly with respect to interviewer effects –, it will be possible to identify factual falsifications that occurred in real fieldwork settings of surveys. In a first step, one could try to identify confirmed falsifications, for example, in the SHARE data (Schuller and Bergmann, 2017). Next, highly suspicious datasets, like the German GGS (Ruckdeschel et al., 2016) could be investigated by searching for interviewer effects in order to identify



presumably falsified cases or interviewers who have presumably falsified (many) interviews. Such probably contaminated data may be used to enhance the knowledge about their effects on content-related, theory-driven multivariate research.

To deepen the understanding of interviewer falsifications and to answer my research question on the influence of falsifications on findings of substantial social science research more comprehensively a further step could be to conduct an additional experiment. This subsequent experiment should base on a representative sample of respondents. Thus, one would achieve higher variances concerning age and education. Furthermore, one could avoid that student interviewers are familiar with the lives and thoughts of their respondents. The most important advantage would be that, in contrast to the here described experimental setup, the interviewers have to achieve and recruit certain predefined target persons. Consequently, this procedure would match stronger with the factual tasks, and difficulties, of interviewers in a real fieldwork setting. (To remember, in the IFiS design the interviewers recruited the respondents themselves and without any quota or other restriction on the campus of their own university.)

In a further experiment, one could also think of implementing experimental randomization: The assignment of the student participants to the treatment or control group – that means to be a falsifying or an honest interviewer – could be randomized. But, as ascertained in the fourth manuscript, one should also consider that under certain circumstances, randomization may not be necessary, as long as theory-driven approaches were applied when implementing quasi-experimental setups. Additionally, one could employ not only student interviewers, but a more “representative” group of interviewers. Another fruitful approach could be to widen the focus on partial falsifications that are probably more relevant than falsifications of entire interviews.

Although still quite a number of gaps in the research on interviewer falsifications have to be admitted, there are already a few steps taken.

## References

- AAPOR American Association of Public Opinion Research (2003). Interviewer Falsification in Survey Research: Current Best Methods for Prevention, Detection and Repair of Its Effects. Published online: [https://www.aapor.org/AAPOR\\_Main/media/MainSiteFiles/falsification.pdf](https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/falsification.pdf). Also published (2004) in Survey Research 35(1), 1-5, <http://www.srl.uic.edu/Publist/Newsletter/2004/04v35n1.pdf>.
- Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes* 50(2), 179-211.
- Ajzen, I., & Krebs, D. (1994). Attitude Theory and Measurement: Implications for Survey Research. In: Borg, I., & Mohler, P. (Eds.), *Trends and Perspectives in Empirical Social Research*. Berlin: de Gruyter, 250-265.
- Angrist, J. D., & Pischke, J.-S. (2015). *Mastering Metrics. The Path from Cause to Effect*. Princeton: Princeton University Press, 1-33.
- Barnes, S. H., Kaase, M., Allerbeck, K. R., Farah, B. G., Heunks, F., Inglehart, R., Jennings, M. K., Klingemann, D., Marsh, A., & Rosenmayr, L. (Eds.) (1979). *Political Action. Mass Participation in Five Western Democracies*. Beverly Hills (CA): Sage.
- Beierlein, C., Kovaleva, A., Kemper, C. J., & Rammstedt, B. (2014). Allgemeine Selbstwirksamkeit Kurzskala (ASKU). Zusammenstellung sozialwissenschaftlicher Items und Skalen. doi: 10.6102/zis35. Retrieved from [http://zis.gesis.org/skala/Beierlein-Kovaleva-Kemper-Rammstedt-Allgemeine-Selbstwirksamkeit-Kurzskala-\(ASKU\)](http://zis.gesis.org/skala/Beierlein-Kovaleva-Kemper-Rammstedt-Allgemeine-Selbstwirksamkeit-Kurzskala-(ASKU)).
- Beullens, K., & Loosveldt, G. (2016). Interviewer effects in the European Social Survey. *Survey Research Methods* 10(2), 103-118. doi: 10.18148/srm/2016.v10i2.6261.
- Biemer, P. P., & Stokes, S. L. (1989). The Optimal Design of Quality Control Samples to Detect Interviewer Cheating. *Journal of Official Statistics* 5(1), 23-39.
- Billari, F. C., Philipov, D., & Testa, M.R. (2009). Attitudes, norms and perceived behavioural control: explaining fertility intentions in Bulgaria. *European Journal of Population / Revue européenne de Démographie* 25(4), 439-465. doi: 10.1007/s10680-009-9187-9.
- Bogner, K., & Landrock, U. (2016). Response biases in standardised surveys. *GESIS Survey Guidelines*. Mannheim: GESIS – Leibniz Institute for the Social Sciences. doi: 10.15465/gesis-sg\_en\_016.

- Bredl, S., Storfinger, N., & Menold, N. (2013). A Literature Review of Methods to Detect Fabricated Survey Data. In: Winker, P., Menold, N., & Porst, R. (Eds.), *Interviewers' Deviations in Surveys*. Frankfurt: Peter Lang, 3-24.
- Bredl, S., Winker, P., & Kötschau, K. (2012). A Statistical Approach to Detect Interviewer Falsification of Survey Data. *Survey Methodology* 38(1), 1-10.
- Diekmann, A. (2010). *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen*. Reinbek bei Hamburg: Rowohlt.
- EVS (2015). European Values Study Longitudinal Data File 1981-2008 (EVS 1981-2008) – Restricted Use File. GESIS Datenarchiv, Köln. ZA5174 Datenfile Version 1.0.0. doi: 10.4232/1.5174.
- Franzen, A., & Pointner, S. (2013). The External Validity of Giving in the Dictator Game. *Experimental Economics* 16(2), 155-169. doi: 10.1007/s10683-012-9337-5.
- GESIS – Leibniz Institute for the Social Sciences (2013). German General Social Survey – ALLBUS 2010. Köln: GESIS Datenarchiv. ZA4612 Version 1.0.1. doi: 10.4232/1.11782.
- GESIS – Leibniz Institute for the Social Sciences (2015). German General Social Survey – ALLBUS 2008. Köln: GESIS Datenarchiv. ZA4602 Version 1.0.0. doi: 10.4232/1.11690.
- GESIS – Leibniz Institute for the Social Sciences (2017). Allgemeine Bevölkerungsumfrage der Sozialwissenschaften ALLBUS 2016. GESIS Datenarchiv, Köln. ZA5250 Datenfile Version 2.1.0. doi: 10.4232/1.12796.
- Groves, R. M., & Magilavy, L. J. (1986). Measuring and Explaining Interviewer Effects in Centralized Telephone Surveys. *Public Opinion Quarterly* 50(2), 251-266. doi: 10.1086/268979.
- Haunberger, S. (2006). Das standardisierte Interview als soziale Interaktion: Interviewer-effekte in der Umfrageforschung. *ZA-Information* 58, 23-46.
- Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology* 47, 237-271. doi: 10.1146/annurev.psych.47.1.237.
- Hox, J. J. (1995). *Applied multilevel analysis*. Amsterdam: TT-Publikaties.
- Imbens, G. W., & Rubin, D. R. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences – An Introduction*. New York (NY): Cambridge University Press.
- Japéc, L. (2006). Quality Issues in Interview Surveys – Some contributions. *Bulletin de Méthodologie Sociologique* 90, 26-42.

- Kaase, M., & Marsh, A. (1979). Political Action. A Theoretical Perspective. In: Barnes, S. H., Kaase, M., Allerbeck, K. R., Farah, B. G., Heunks, F., Inglehart, R., Jennings, M. K., Klingemann, D., Marsh, A., & Rosenmayr, L. (Eds.), *Political Action. Mass Participation in Five Western Democracies*. Beverly Hills (CA): Sage, 27-56.
- Kemper, C. J., & Menold, N. (2014). Nuisance or Remedy? The Utility of Stylistic Responding as an Indicator of Data Fabrication in Surveys. *Methodology* 10(3), 92-99. doi: 10.1027/1614-2241/a000078.
- Koch, A. (1995). Gefälschte Interviews: Ergebnisse der Interviewerkontrolle beim ALLBUS 1994. *ZUMA-Nachrichten* 36, 89-105.
- Koch, A., Kurz, K., Mahr-George, H., & Wasmer, M. (1999). Konzeption und Durchführung der "Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften" (ALLBUS) 1998. Mannheim: ZUMA-Arbeitsbericht 1999/02, <http://nbn-resolving.de/urn:nbn:de:0168-ssoar-200413>.
- Koczela, S., Furlong, C., McCarthy, J., & Mushtaq, A. (2015). Curbstoning and beyond: Confronting data fabrication in survey research. *Statistical Journal of the IAOS* 31(3), 413-422. doi: 10.3233/SJI-150917.
- Krejsa, E. A., Davis, M. C., & Hill, J. M. (1999). Evaluation of the Quality Assurance Falsification Interview used in the Census 2000 Dress Rehearsal. *Proceedings of the American Statistical Association (Survey Research Methods Section)*, 635-640.
- Krosnick, J. A., & Alwin, D. (1987). An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement. *Public Opinion Quarterly* 51(2), 201-219.
- Krosnick, J. A., Narayan, S. S., & Smith, W. R. (1996). Satisficing in Surveys: Initial Evidence. In: Braverman, M. T., & Slater, J. K. (Eds.), *Advances in survey research*. San Francisco: Jossey-Bass, 29-44.
- Kuriakose, N., & Robbins, M. (2016). Don't get duped: Fraud through duplication in public opinion surveys. *Statistical Journal of the IAOS* 32(3), 283-291. doi: 10.3233/SJI-160978.
- Landrock, U. (2017a). Explaining Political Participation: A Comparison of Real and Falsified Survey Data. *Statistical Journal of the IAOS* 32(3), 447-458. doi: 10.3233/SJI-160270.
- Landrock, U. (2017b). How interviewer effects differ in real and falsified survey data. Using multilevel analysis to identify interviewer falsifications. *methods, data, analyses* 11(2), 163-187. doi: 10.12758/mda.2017.03.

- Landrock, U. (2017c). Investigating Interviewer Falsifications – A Quasi-experimental Design. *Bulletin of Sociological Methodology* 136(1), 5-20. doi: 10.1177/0759106317725640.
- Landrock, U., & Menold, N. (2016). Validation of Theoretical Assumptions with Real and Falsified Survey Data. *Statistical Journal of the IAOS* 32(3), 305-312. doi: 10.3233/SJI-161020.
- Langer, W. (2010). Mehrebenenanalyse mit Querschnittsdaten. In: Wolf, C., & Best, H. (Eds.), *Handbuch der sozialwissenschaftlichen Datenanalyse*. Wiesbaden: VS Verlag für Sozialwissenschaften, 741-774.
- Levitt, S. D., & List, J. A. (2007). What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?. *Journal of Economic Perspectives* 21(2), 153-174. doi: 10.1257/jep.21.2.153.
- Li, J., Brick, J. M., Tran, B., & Singer, P. (2009). Using Statistical Models for Sample Design of a Reinterview Program. *Proceedings of the American Statistical Association (Survey Research Methods Section)*, 4681-4695.
- Liu, M., & Stainback, K. (2013). Interviewer gender effects on survey responses to marriage related questions. *Public Opinion Quarterly* 77(2), 606-618. doi: 10.1093/poq/nft019.
- Lüdemann, C. (2001). Politische Partizipation, Anreize und Ressourcen. Ein Test verschiedener Handlungsmodelle und Anschlusstheorien am ALLBUS 1998. In: Koch, A., Schmidt, P., & Wasmer, M. (Eds.), *Politische Partizipation in der Bundesrepublik Deutschland. Empirische Befunde und theoretische Erklärungen*. Opladen: Leske und Budrich, 43-71.
- Mangione, T. W., Fowler, F. J., & Louis, T. A. (1992). Question Characteristics and Interviewer Effects. *Journal of Official Statistics* 8(3), 293-307.
- Marsh, A., & Kaase, M. (1979). Measuring Political Action. In: Barnes, S. H., Kaase, M., Allerbeck, K. R., Farah, B. G., Heunks, F., Inglehart, R., Jennings, M. K., Klingemann, D., Marsh, A., & Rosenmayr, L. (Eds.), *Political Action. Mass Participation in Five Western Democracies*. Beverly Hills (CA): Sage, 57-96.
- Menold, N., & Kemper, C. J. (2014). How do Real and Falsified Data Differ? Psychology of Survey Response as a Source of Falsification Indicators in Face-to-face Surveys. *International Journal of Public Opinion Research* 26(1), 41-65. doi: 10.1093/ijpor/edt017.

- Menold, N., Winker, P., Storfinger, N., & Kemper, C. J., (2013). A Method for Ex-post-identification of Falsifications in Survey Data. In: Winker, P., Menold, N., & Porst, R. (Eds.), *Interviewers' Deviations in Surveys*. Frankfurt: Peter Lang, 25-48.
- Moskowitz, G. B. (2005). *Social cognition. Understanding self and others*. New York: Guilford Press.
- Olson, K., & Bilgen, I. (2011). The role of interviewer experience on acquiescence. *Public Opinion Quarterly* 75(1), 99-114. doi: 10.1093/poq/nfq067.
- Opp, K. D., & Finkel, S. E. (2001). Politischer Protest, Rationalität und Lebensstile. Eine empirische Überprüfung alternativer Erklärungsmodelle. In: Koch, A., Schmidt, P., & Wasmer, M. (Eds.), *Politische Partizipation in der Bundesrepublik Deutschland. Empirische Befunde und theoretische Erklärungen*. Opladen: Leske und Budrich, 73-108.
- Pagès, F., Tribel, M., & Bonis-Charancle, A. (2006). Indélicatesse et manque de rigueur dans les sondages. Enquête auprès des professionnels du sondage. *BMS* 89, 31-48.
- Prapavessis, H., Gaston, A., & DeJesus, S. (2015). The Theory of Planned Behavior as a model for understanding sedentary behavior. *Psychology of Sport and Exercise* 19, 23-32. doi: 10.1016/j.psychsport.2015.02.001.
- Rammstedt, B., Kemper, C. J., Klein, M. C., Beierlein, C., & Kovaleva, A. (2014). Big Five Inventory (BFI-10). Zusammenstellung sozialwissenschaftlicher Items und Skalen. doi:10.6102/zis76. Retrieved from [http://zis.gesis.org/skala/Rammstedt-Kemper-Klein-Beierlein-Kovaleva-Big-Five-Inventory-\(BFI-10\)](http://zis.gesis.org/skala/Rammstedt-Kemper-Klein-Beierlein-Kovaleva-Big-Five-Inventory-(BFI-10)).
- Rammstedt, B., Martin, S., Zabal, A., Konradt, I., Maehler, D., Perry, A., Massing, N., Ackermann-Piek, D., & Helmschrott, S. (2016). Programme for the International Assessment of Adult Competencies (PIAAC), Germany – Reduzierte Version. GESIS Datenarchiv, Köln. ZA5845 Datenfile Version 2.2.0. doi: 10.4232/1.12660.
- Reuband, K. H. (1990). Interviews, die keine sind. „Erfolge“ und „Mißerfolge“ beim Fälschen von Interviews. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 42(4), 706-733.
- Ruckdeschel, K., Sauer, L., & Naderi, R. (2016). Reliability of retrospective event histories within the German Generations and Gender Survey. The role of interviewer and survey design factors. *Demographic Research* 34(11), 321-358.

- Schanz, V. (1981). Interviewereffekte: zusammenfassende Darstellung von Arbeiten, die im Rahmen zweier von ZUMA betreuter Projekte entstanden sind. *ZUMA Nachrichten* 9, 36-46.
- Schnell, R. (1991). Der Einfluss gefälschter Interviews auf Survey-Ergebnisse. *Zeitschrift für Soziologie* 20(1), 25-35.
- Schraepfer, J. P., & Wagner, G. G. (2003). Identification, Characteristics and Impact of Faked Interviews in Surveys. An Analysis by Means of Genuine Fakes in the Raw Data of SOEP. IZA Discussion Paper No. 969. Forschungsinstitut zur Zukunft der Arbeit.
- Schuller, K., & Bergmann, M. (2017). Identifying fake interviews in a cross-national panel study (SHARE). Presentation at the 7th Conference of the European Survey Research Association (ESRA) in Lisbon, Portugal, from 17th to 21st July 2017, session: "Deviations and Fraud in Surveys - the Impact of Motivation and Incentives 2", July, 20th 2017.
- Schulz, W. (2001). Politische Mobilisierung durch Mediennutzung? Beziehungen zwischen Kommunikationsverhalten, politischer Kompetenz und Partizipationsbereitschaft. In: Koch, A., Schmidt, P., & Wasmer, M. (Eds.), *Politische Partizipation in der Bundesrepublik Deutschland. Empirische Befunde und theoretische Erklärungen*. Opladen: Leske und Budrich, 170-194.
- Schwarz, N. (2007). Cognitive Aspects of Survey Methodology. *Applied Cognitive Psychology* 21(2), 277-287. doi: 10.1002/acp.1340.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Belmont (CA): Wadsworth Cengage Learning.
- Sodeur, W. (2007). Entscheidungsspielräume von Interviewern bei der Wahrscheinlichkeitsauswahl. *Methoden – Daten – Analysen* 1(2), 107-130.
- Steinbrecher M. (2004). *Politische Partizipation in Deutschland*. Baden-Baden: Nomos.
- Storfinger, N., & Winker, P. (2013). Assessing the Performance of Clustering Methods in Falsification Identification using Bootstrap. In: Winker, P., Menold, N., & Porst, R. (Eds.), *Interviewers' Deviations in Surveys*. Frankfurt: Peter Lang, 49-65.
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Turner, C. F., Gribbe, J. N., Al-Tayyip, A. A., & Chromy, J. R. (2002). *Falsification in Epidemiologic Surveys – Detection and Remediation (Prepublication Draft)*.

- Technical Papers on Health and Behavior Measurement No. 53. Research Triangle Institute. Washington (DC).
- van der Zouwen, J. (2006). Interviewer and Survey Researcher. Mutual Dependencies. *BMS* 89, 49-64.
- van Deth, J. W. (2001). Soziale und politische Beteiligung: Alternativen, Ergänzungen oder Zwillinge? In: Koch, A., Schmidt, P., & Wasmer, M. (Eds.), *Politische Partizipation in der Bundesrepublik Deutschland. Empirische Befunde und theoretische Erklärungen*. Opladen: Leske und Budrich, 196-219.
- van Meter, K. (2005). Studying Survey Interviewers: A Call for Research and an Overview. *BMS* 88, 51-71.
- Wasmer, M., Scholz, E., & Blohm, M. (2010). Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 2008. *GESIS-Technical Reports* 2010, 2010/04.
- West, B. T., & Blom, A. G. (2016). Explaining interviewer effects: A research synthesis. *Journal of Survey Statistics and Methodology* 0, 1-37. doi: 10.1093/jssam/smw024.
- Westle, B. (2001). Politische Partizipation und Geschlecht. In: Koch, A., Schmidt, P., & Wasmer, M. (Eds.), *Politische Partizipation in der Bundesrepublik Deutschland. Empirische Befunde und theoretische Erklärungen*. Opladen: Leske und Budrich, 132-168.
- Winker, P. (2016). Assuring the Quality of Survey Data: Incentives, Detection and Documentation of Deviant Behavior. *Statistical Journal of the IAOS* 32(3), 295-303. doi: 10.3233/SJI-161012.
- Winker, P., Kruse, K. W., Menold, N., & Landrock, U. (2015). Interviewer Effects in Real and Falsified Interviews: Results from a Large Scale Experiment. *Statistical Journal of the IAOS* 31(3), 423-34. doi: 10.3233/SJI-150908.
- Wolbring, T., & Keuschnigg, M. (2015). Feldexperimente in den Sozialwissenschaften. Grundlagen, Herausforderungen, Beispiele. In: Keuschnigg, M., & Wolbring, T. (Eds.), *Experimente in den Sozialwissenschaften (Soziale Welt Special Issue 22)*. Baden-Baden: Nomos, 219-245.
- Zentralarchiv für empirische Sozialforschung (ZA) & Zentrum für Umfragen, Methoden und Analysen (ZUMA) e.V. (2014). Anomie (ALLBUS). Zusammenstellung sozialwissenschaftlicher Items und Skalen. doi: 10.6102/zis58. Retrieved from [http://zis.gesis.org/skala/ZA-ZUMA-Anomie-\(ALLBUS\)](http://zis.gesis.org/skala/ZA-ZUMA-Anomie-(ALLBUS)).



## **Appendix**

Codeplan

Spezifizierung der Koauthorschaften

Erklärung

## Codeplan Real

Für alle Fragen gilt:

1. Wenn eine Frage zutrifft aber ausgelassen wurde mit -9 vercoden (auch einzelne Items)
2. -9 für Auslassungen gilt, solange sie bei einer Variable nicht gesondert deklariert worden sind. Wenn die Auslassungen bei einer Variable im Codeplan deklariert sind, benutzen Sie die deklarierten Werte und nicht mehr -9.
3. Wenn bei einer Frage eine Antwort doppelt bzw. mehrfach angekreuzt wurde, obwohl keine Mehrfachnennung möglich war, ist diese Frage nicht verwertbar und wird mit -8 vercodet.

## Hauptstudie – Dokument 1

Interviewer ID:

\_ \_ \_

*InterviewerID*

Liebe Kommilitonin, lieber Kommilitone,

GESIS, das Leibniz-Institut für Sozialwissenschaften in Mannheim, und die Justus Liebig-Universität Gießen (JLU) führen zur Zeit eine Umfrage unter Studierenden der Universität Gießen und Mannheim durch, um etwas über deren wirtschaftliche Situation, ihre politischen Einstellungen und ihr Freizeitverhalten zu erfahren. Ich arbeite derzeit als InterviewerIn für GESIS/JLU und möchte Dich bitten, an der Befragung teilzunehmen. Wir werden etwa 30 bis 35 Minuten brauchen.

Alle Angaben, die Du machst, werden natürlich streng vertraulich behandelt, d.h. die Auswertung der Befragung erfolgt ohne Namen oder Adressen. Damit ist sichergestellt, dass Deine Angaben nicht mit Deiner Person in Verbindung gebracht werden.

Um die Auswertung zu erleichtern, müssen wir das Interview allerdings aufzeichnen. Dazu müsstest Du zunächst diese Erklärung lesen und unterschreiben.

--> *Erklärung und Stift übergeben!*

Bei Nachfragen: Diese Erklärung enthält die Regeln, die eingehalten werden müssen, um den geltenden Datenschutzbestimmungen gerecht zu werden. Bei Fragen kannst Du Dich gerne an diese Adresse wenden:

--> *ggfs. Kontaktdaten weitergeben: Dr. Christoph Kemper,*

*Email: christoph.kemper @gesis.org*

Wenn Du Interesse an den Ergebnissen der Untersuchung hast, werden dir diese zur Verfügung gestellt. Dazu benötigen wir Deine Email-Adresse.

*Falls Zielperson Interesse an Ergebnissen der Untersuchung hat, bitte Email-Adresse notieren!*

*✍:* \_\_\_\_\_

*Wurde die Erklärung von der Zielperson unterschrieben?*

- Ja → Weiter mit nächster Seite  
(Wenn nicht unterschrieben, bitte neue Befragungsperson suchen)

## Hauptstudie – Dokument 1

Danke für die Bereitschaft, an dieser Befragung teilzunehmen.

Wir können gleich anfangen. Ich muss nur noch das Datum und die Uhrzeit eintragen.  
Einen Moment bitte!

--> Bitte das Aufzeichnungsgerät einschalten.

--> Die Interviewer ID und Befragtennummer bitte LAUT und DEUTLICH sagen!

--> F000A: Die Befragungsperson sollte Ihre Antwort nicht einsehen können!

### F 00A\_T

--> Beginn des Interviews eintragen! Bitte Datum und Uhrzeit angeben!

**Datum** ✎: FOOA\_TDA. FOOA\_ID22011 (TT.MM.JJJJ)

**Uhrzeit** ✎: FOOA\_TZ1:FOOA\_TZ2Uhr

KEINE ANGABE  
= 99

**Befragten Nr.** ✎: \_\_\_ \_\_\_ BefragtenNR

### F000A

F000A

--> Nur für den Interviewer / die Interviewerin!

Wie attraktiv oder unattraktiv ist der / die Befragte?

Bitte entscheiden Sie spontan, gemäß Ihrem ersten Eindruck.

- 11 ○ 11 = Attraktiv  
10 ○ 10  
9 ○ 09  
8 ○ 08  
7 ○ 07  
6 ○ 06  
5 ○ 05  
4 ○ 04  
3 ○ 03  
2 ○ 02  
1 ○ 01 = Unattraktiv

KEINE MARKIERUNG = -9

## Hauptstudie – Dokument 1

Wir möchten uns in diesem Interview zunächst mit den Themen Politik und Gesellschaft beschäftigen.

### F 01

Ich werde Dir nun einige Aussagen vorlesen.  
Sag mir bitte zu jeder einzelnen dieser Aussagen, ob Du persönlich -

- > derselben Meinung bist oder
- > ob Du anderer Meinung bist.

--> *Aussagen bitte vorlesen!*

|  | Bin derselben Meinung   | Bin anderer Meinung     | Weiß nicht              | K A                     |       |
|--|-------------------------|-------------------------|-------------------------|-------------------------|-------|
| Egal, was manche Leute sagen: Die Situation der einfachen Leute wird nicht besser, sondern schlechter.   | 1 <input type="radio"/> | 2 <input type="radio"/> | 8 <input type="radio"/> | 9 <input type="radio"/> | FOI_a |
| So wie die Zukunft aussieht, kann man es kaum noch verantworten, Kinder auf die Welt zu bringen.         | 1 <input type="radio"/> | 2 <input type="radio"/> | 8 <input type="radio"/> | 9 <input type="radio"/> | FOI_b |
| Die meisten Politiker interessieren sich in Wirklichkeit gar nicht für die Probleme der einfachen Leute. | 1 <input type="radio"/> | 2 <input type="radio"/> | 8 <input type="radio"/> | 9 <input type="radio"/> | FOI_c |
| Die meisten Leute kümmern sich in Wirklichkeit gar nicht darum, was mit ihren Mitmenschen geschieht.     | 1 <input type="radio"/> | 2 <input type="radio"/> | 8 <input type="radio"/> | 9 <input type="radio"/> | FOI_d |

## Hauptstudie – Dokument 1

## F 02

--> Liste 02 vorlegen!

Auf dieser Liste stehen einige Meinungen, die man gelegentlich hört.  
Sag mir bitte zu jeder Meinung, ob Du ihr -

- > überhaupt nicht zustimmst (=1),
- > voll und ganz zustimmst (=6).

Du kannst Deine Zustimmung zu der Aussage auch abstimmen und eine Zahl zwischen 1 und 6 nennen.

|  | Stimme überhaupt nicht zu |                       |                       |                       |                       |                       | Stimme voll und ganz zu | Weiß nicht            | KA                    |       |
|--|---------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------------------|-----------------------|-----------------------|-------|
|  | 1                         | 2                     | 3                     | 4                     | 5                     | 6                     |                         |                       |                       |       |
| A Die Politiker kümmern sich nicht viel darum, was Leute wie ich denken.                                       | <input type="radio"/>     | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>   | <input type="radio"/> | <input type="radio"/> | FO2_a |
|  | 1                         | 2                     | 3                     | 4                     | 5                     | 6                     | 8                       | 9                     |                       |       |
| B Ich traue mir zu, in einer Gruppe, die sich mit politischen Fragen befasst, eine aktive Rolle zu übernehmen. | <input type="radio"/>     | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>   | <input type="radio"/> | <input type="radio"/> | FO2_b |
|  | 1                         | 2                     | 3                     | 4                     | 5                     | 6                     | 8                       | 9                     |                       |       |
| C Die ganze Politik ist so kompliziert, dass jemand wie ich gar nicht versteht, was vorgeht.                   | <input type="radio"/>     | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>   | <input type="radio"/> | <input type="radio"/> | FO2_c |
|  | 1                         | 2                     | 3                     | 4                     | 5                     | 6                     | 8                       | 9                     |                       |       |
| D Die Politiker bemühen sich im Allgemeinen darum, die Interessen der Bevölkerung zu vertreten.                | <input type="radio"/>     | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>   | <input type="radio"/> | <input type="radio"/> | FO2_d |
|  | 1                         | 2                     | 3                     | 4                     | 5                     | 6                     | 8                       | 9                     |                       |       |
| E Im Allgemeinen weiß ich eher wenig über Politik.   | <input type="radio"/>     | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>   | <input type="radio"/> | <input type="radio"/> | FO2_e |
|  | 1                         | 2                     | 3                     | 4                     | 5                     | 6                     | 8                       | 9                     |                       |       |
| F Die Durchschnittsbürger verstehen nur wenig von Politik.   | <input type="radio"/>     | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/>   | <input type="radio"/> | <input type="radio"/> | FO2_f |
|  | 1                         | 2                     | 3                     | 4                     | 5                     | 6                     | 8                       | 9                     |                       |       |

## Hauptstudie – Dokument 1

F 03 F03

--&gt; Liste 03 vorlegen!

Viele Leute verwenden die Begriffe „links“ und „rechts“, wenn es darum geht, unterschiedliche politische Einstellungen zu kennzeichnen.

Wir haben hier einen Maßstab, der von links nach rechts verläuft.

Wenn Du an Deine eigenen politischen Ansichten denkst, wo würdest Du diese Ansichten auf dieser Skala einstufen?

Entscheide Dich bitte für eines der Kästchen und nenne mir den darunter stehenden Buchstaben.

| Links |     |     |     |     |     |     |     |     | Rechts |  |
|-------|-----|-----|-----|-----|-----|-----|-----|-----|--------|--|
| 1 ○   | 2 ○ | 3 ○ | 4 ○ | 5 ○ | 6 ○ | 7 ○ | 8 ○ | 9 ○ | 10 ○   |  |
| F     | A   | M   | O   | G   | Z   | E   | Y   | I   | P      |  |
| 33 ○  | KA  |     |     |     |     |     |     |     |        |  |

F 04 F04

Würdest Du mir bitte sagen, was Du mit dem Begriff „links“ verbindest?

--> Antworten bitte genau notieren!

☒: F04 = TEXT

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

33 ○ KA

F 05 F05

Und würdest Du mir bitte sagen, was Du mit dem Begriff „rechts“ verbindest?

--> Antworten bitte genau notieren!

☒: F05 = TEXT

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

33 ○ KA

## Hauptstudie – Dokument 1

## F 06 A

--> Kartensatz mischen und übergeben!

Wenn Du politisch in einer Sache, die Dir wichtig ist, Einfluss nehmen, Deinen Standpunkt zur Geltung bringen willst:

Welche der Möglichkeiten auf diesen Karten würdest Du dann nutzen, was davon käme für Dich in Frage?

Bitte nenne mir die entsprechenden Kennbuchstaben.

--> Alle genannten Kennbuchstaben markieren!

- FOG A\_a  A Seine Meinung sagen, im Bekanntenkreis und am Arbeitsplatz  
 FOG A\_b  B Sich an Wahlen beteiligen  
 FOG A\_c  C Sich in Versammlungen an öffentlichen Diskussionen beteiligen  
 FOG A\_d  D Mitarbeit in einer Bürgerinitiative  
 FOG A\_e  E In einer Partei aktiv mitarbeiten  
 FOG A\_f  F Teilnahme an einer nicht genehmigten Demonstration  
 FOG A\_g  G Teilnahme an einer genehmigten Demonstration  
 FOG A\_h  H Sich aus Protest nicht an Wahlen beteiligen  
 FOG A\_j  J Aus Protest einmal eine andere Partei wählen als die, der man nahesteht  
 FOG A\_k  K Beteiligung an einer Unterschriftensammlung  
 FOG A\_l  L Aus politischen, ethischen oder Umweltgründen Waren boykottieren oder kaufen  
 FOG A\_m  M Sich an einer Online-Protestaktion beteiligen

6  Nichts davon

9  KA

0 = NICHT GENANNT

1 = GENANNT

## F 06 B

--> Alle Karten erneut mischen und übergeben.

Was davon hast Du selbst schon gemacht, woran warst Du schon einmal beteiligt?

Bitte nenne mir die entsprechenden Kennbuchstaben.

--> Alle genannten Kennbuchstaben markieren!

- FOG B\_a  A Seine Meinung sagen, im Bekanntenkreis und am Arbeitsplatz  
 FOG B\_b  B Sich an Wahlen beteiligen  
 FOG B\_c  C Sich in Versammlungen an öffentlichen Diskussionen beteiligen  
 FOG B\_d  D Mitarbeit in einer Bürgerinitiative  
 FOG B\_e  E In einer Partei aktiv mitarbeiten  
 FOG B\_f  F Teilnahme an einer nicht genehmigten Demonstration  
 FOG B\_g  G Teilnahme an einer genehmigten Demonstration  
 FOG B\_h  H Sich aus Protest nicht an Wahlen beteiligen  
 FOG B\_j  J Aus Protest einmal eine andere Partei wählen als die, der man nahesteht  
 FOG B\_k  K Beteiligung an einer Unterschriftensammlung  
 FOG B\_l  L Aus politischen, ethischen oder Umweltgründen Waren boykottieren oder kaufen  
 FOG B\_m  M Sich an einer Online-Protestaktion beteiligen

6  Nichts davon

9  KA

0 = NICHT GENANNT

1 = GENANNT



Hauptstudie – Dokument 1

F 07 **FO7**

--> Liste 07 vorlegen!

Wenn am nächsten Sonntag Bundestagswahl wäre, welche Partei würdest Du dann mit Deiner ZWEITSTIMME wählen?

- 1  CDU/CSU
  - 2  SPD
  - 3  FDP
  - 4  Bündnis 90 / Die Grünen
  - 6  Die Linke
  - 30  Andere Partei, und zwar : FOZLS
  - 31  Würde nicht wählen
  - 37  Angabe verweigert
  - 38  Weiß nicht
  - 36  Nicht wahlberechtigt, da keine deutsche Staatsbürgerschaft
  - 39  KA
- 5 = REPUBLIKANER**  
**20 = NPD**

F 08 **FO8**

--> Liste 08 vorlegen!

Hier ist eine Liste mit verschiedenen Auffassungen darüber, wie es in Deutschland mit den sozialen Unterschieden tatsächlich aussieht und wie es sein sollte. Bitte gehe die Aussagen der Reihe nach durch und sage mir, ob Du der jeweiligen Auffassung -

- > überhaupt nicht zustimmst,
- > eher nicht zustimmst,
- > eher zustimmst oder
- > voll zustimmst.

|   | Stimme überhaupt nicht zu | Stimme eher nicht zu    | Stimme eher zu          | Stimme voll zu          | Weiß nicht              | KA                      |              |
|---|---------------------------|-------------------------|-------------------------|-------------------------|-------------------------|-------------------------|--------------|
| A Nur wenn die Unterschiede im Einkommen und im sozialen Ansehen groß genug sind, gibt es auch einen Anreiz für persönliche Leistungen.                 | 1 <input type="radio"/>   | 2 <input type="radio"/> | 3 <input type="radio"/> | 4 <input type="radio"/> | 8 <input type="radio"/> | 9 <input type="radio"/> | <b>FO8_a</b> |
| B Die Rangunterschiede zwischen den Menschen sind akzeptabel, weil sie im wesentlichen ausdrücken, was man aus den Chancen, die man hatte, gemacht hat. | 1 <input type="radio"/>   | 2 <input type="radio"/> | 3 <input type="radio"/> | 4 <input type="radio"/> | 8 <input type="radio"/> | 9 <input type="radio"/> | <b>FO8_b</b> |
| C Ich finde die sozialen Unterschiede in unserem Land im Großen und Ganzen gerecht.   | 1 <input type="radio"/>   | 2 <input type="radio"/> | 3 <input type="radio"/> | 4 <input type="radio"/> | 8 <input type="radio"/> | 9 <input type="radio"/> | <b>FO8_c</b> |

## Hauptstudie – Dokument 1

F 09 F09

--&gt; Liste 09 vorlegen!

Über die Aufgaben der Frau in der Familie und bei der Kindererziehung gibt es verschiedene Meinungen.

Bitte sag mir nun zu jeder Aussage auf dieser Liste, ob Du ihr -

- > überhaupt nicht zustimmst,
- > eher nicht zustimmst,
- > eher zustimmst oder
- > voll und ganz zustimmst.

|  | Stimme<br>überhaupt<br>nicht zu | Stimme<br>eher nicht<br>zu | Stimme<br>eher zu | Stimme<br>voll und<br>ganz zu | Weiß<br>nicht | KA  |       |
|--|---------------------------------|----------------------------|-------------------|-------------------------------|---------------|-----|-------|
| A Eine berufstätige Mutter kann ein genauso herzliches und vertrauensvolles Verhältnis zu ihren Kindern finden wie eine Mutter, die nicht berufstätig ist.   | 1 ○                             | 2 ○                        | 3 ○               | 4 ○                           | 8 ○           | 9 ○ | F09_a |
| B Für eine Frau ist es wichtiger, ihrem Mann bei seiner Karriere zu helfen, als selbst Karriere zu machen.   | 1 ○                             | 2 ○                        | 3 ○               | 4 ○                           | 8 ○           | 9 ○ | F09_b |
| C Ein Kleinkind wird sicherlich darunter leiden, wenn seine Mutter berufstätig ist.  | 1 ○                             | 2 ○                        | 3 ○               | 4 ○                           | 8 ○           | 9 ○ | F09_c |
| D Es ist für alle Beteiligten viel besser, wenn der Mann voll im Berufsleben steht und die Frau zu Hause bleibt und sich um den Haushalt und die Kinder kümmert.   | 1 ○                             | 2 ○                        | 3 ○               | 4 ○                           | 8 ○           | 9 ○ | F09_d |
| E Es ist für ein Kind sogar gut, wenn seine Mutter berufstätig ist und sich nicht nur auf den Haushalt konzentriert.   | 1 ○                             | 2 ○                        | 3 ○               | 4 ○                           | 8 ○           | 9 ○ | F09_e |
| F Eine verheiratete Frau sollte auf eine Berufs-tätigkeit verzichten, wenn es nur eine begrenzte Anzahl von Arbeitsplätzen gibt, und wenn ihr Mann in der Lage ist, für den Unterhalt der Familie zu sorgen. | 1 ○                             | 2 ○                        | 3 ○               | 4 ○                           | 8 ○           | 9 ○ | F09_f |

## Hauptstudie – Dokument 1

F 10 F10

Es wird heute viel über die verschiedenen Bevölkerungsschichten gesprochen.  
Welcher Schicht rechnest Du Dich selbst eher zu -

--> *Vorgaben bitte vorlesen! Nur eine Nennung möglich!*

- 1  >der Unterschicht,
- 2  >der Arbeiterschicht,
- 3  >der Mittelschicht,
- 4  >der oberen Mittelschicht oder
- 5  >der Oberschicht?

- 
- 6  Keiner dieser Schichten
  - 7  Einstufung abgelehnt
  - 8  Weiß nicht
  - 9  KA

F 11 F11

Im Vergleich dazu, wie andere hier in Deutschland leben:  
Glaubst Du, dass Du Deinen -

--> *Vorgaben bitte vorlesen! Nur eine Nennung möglich!*

- 1  >gerechten Anteil erhält,
- 2  >mehr als Deinen gerechten Anteil,
- 3  >etwas weniger oder
- 4  >sehr viel weniger?

- 
- 8  Weiß nicht
  - 9  KA

## Hauptstudie – Dokument 1

Mit dem Umfrageprojekt möchten wir erfahren, wie die Menschen in Deutschland verschiedene Situationen des Alltags erleben und wie sie auf diese Situationen reagieren. Im kommenden Abschnitt des Interviews soll es daher darum gehen, wie Du Dich selbst siehst und einschätzt.

F 12

F12

--> Liste 12 vorlegen!

Inwieweit treffen die folgenden Aussagen auf Dich zu?  
Bitte nenne mir die Kategorie von dieser Liste hier, die am ehesten Deiner Einschätzung entspricht.

--> Bitte machen Sie in JEDER Zeile eine Markierung!

|  | Trifft überhaupt nicht zu | Trifft eher nicht zu | Weder noch | Eher zutreffend | Trifft voll und ganz zu | Kann ich nicht sagen |       |
|--|---------------------------|----------------------|------------|-----------------|-------------------------|----------------------|-------|
| A Ich bin eher zurückhaltend, reserviert.                                | 1 ○                       | 2 ○                  | 3 ○        | 4 ○             | 5 ○                     | 8 ○                  | F12-a |
| B Ich schenke anderen leicht Vertrauen, glaube an das Gute im Menschen.  | 1 ○                       | 2 ○                  | 3 ○        | 4 ○             | 5 ○                     | 8 ○                  | F12-b |
| C Ich erledige Aufgaben gründlich.                                       | 1 ○                       | 2 ○                  | 3 ○        | 4 ○             | 5 ○                     | 8 ○                  | F12-c |
| D Ich bin entspannt, lasse mich durch Stress nicht aus der Ruhe bringen. | 1 ○                       | 2 ○                  | 3 ○        | 4 ○             | 5 ○                     | 8 ○                  | F12-d |
| E Ich habe eine aktive Vorstellungskraft, bin phantasievoll.             | 1 ○                       | 2 ○                  | 3 ○        | 4 ○             | 5 ○                     | 8 ○                  | F12-e |
| F Ich gehe aus mir heraus, bin gesellig.                                 | 1 ○                       | 2 ○                  | 3 ○        | 4 ○             | 5 ○                     | 8 ○                  | F12-f |
| G Ich neige dazu, andere zu kritisieren.                                 | 1 ○                       | 2 ○                  | 3 ○        | 4 ○             | 5 ○                     | 8 ○                  | F12-g |
| H Ich bin bequem, neige zur Faulheit.                                    | 1 ○                       | 2 ○                  | 3 ○        | 4 ○             | 5 ○                     | 8 ○                  | F12-h |
| I Ich werde leicht nervös und unsicher.                                  | 1 ○                       | 2 ○                  | 3 ○        | 4 ○             | 5 ○                     | 8 ○                  | F12-i |
| J Ich habe nur wenig künstlerisches Interesse.                           | 1 ○                       | 2 ○                  | 3 ○        | 4 ○             | 5 ○                     | 8 ○                  | F12-j |

KEINE ANGABE = 9

## Hauptstudie – Dokument 1

F 13

F13

--&gt; Liste 13 vorlegen!

Bitte beschreibe dich selbst mit Hilfe der folgenden Aussagen.

Diese Aussagen können auf Dich persönlich mehr oder weniger zutreffen.

Bitte nenne mir die Kategorie von dieser Liste hier, die am ehesten Deiner Einschätzung entspricht.

--&gt; Bitte machen Sie in JEDER Zeile eine Markierung!

|  | Trifft gar<br>nicht zu | Trifft<br>wenig zu | Trifft<br>etwas zu | Trifft<br>ziemlich<br>zu | Trifft voll<br>und<br>ganz zu |       |
|--|------------------------|--------------------|--------------------|--------------------------|-------------------------------|-------|
| A Ich akzeptiere immer alle Meinungen, auch wenn ich eine andere habe.                       | <u>1</u> 1             | 2 2                | 3 3                | 4 4                      | 5 5                           | F13_a |
| B Es ist schon mal vorgekommen, dass ich jemanden ausgenutzt habe.                           | <u>1</u> 1             | 2 2                | 3 3                | 4 4                      | 5 5                           | F13_b |
| C Auch wenn ich selbst gestresst bin, behandle ich andere immer freundlich und zuvorkommend. | <u>1</u> 1             | 2 2                | 3 3                | 4 4                      | 5 5                           | F13_c |
| D Gelegentlich rede ich schlecht über andere hinter deren Rücken.                            | 1                      | 2                  | <del>3</del>       | 4                        | 5                             | F13_d |
| E Im Streit bleibe ich stets sachlich und objektiv.  | 1                      | 2                  | <del>3</del>       | 4                        | 5                             | F13_e |
| F Manchmal helfe ich jemandem nur, wenn ich eine Gegenleistung erwarten kann.                | 1                      | 2                  | <del>3</del>       | 4                        | 5                             | F13_f |
| G Ich habe schon mal Müll einfach in die Landschaft oder auf die Straße geworfen.            | 1                      | 2                  | <del>3</del>       | 4                        | 5                             | F13_g |
| H Wenn ich mich mit jemandem unterhalte, höre ich ihm immer aufmerksam zu.                   | <u>1</u> 1             | 2 2                | 3 3                | 4 4                      | 5 5                           | F13_h |

## Hauptstudie – Dokument 1

F 14

F14

--&gt; Liste 14 vorlegen!

Bitte gib bei jeder der folgenden Aussagen an, inwieweit diese auf Dich persönlich zutrifft. Bitte nenne mir die Kategorie von dieser Liste hier, die am ehesten Deiner Einschätzung entspricht.

--&gt; Bitte machen Sie in JEDER Zeile eine Markierung!

|   | Trifft<br>gar<br>nicht<br>zu |   |   |   |   |   |   |  | Trifft<br>voll<br>und<br>ganz<br>zu |       |
|---|------------------------------|---|---|---|---|---|---|--|-------------------------------------|-------|
| A In schwierigen Situationen kann ich mich auf meine Fähigkeiten verlassen.                     | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     | F14_a |
|   | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     |       |
| B Die meisten Probleme kann ich aus eigener Kraft gut meistern.                                 | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     | F14_b |
|   | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     |       |
| C Auch anstrengende und komplizierte Aufgaben kann ich in der Regel gut lösen.                  | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     | F14_c |
|   | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     |       |
| D Ich kann auch mit Unerwartetem gut zurechtkommen.   | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     | F14_d |
|   | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     |       |
| E Auch wenn ich auf Widerstände treffe, gelingt es mir normalerweise, meine Ziele zu verfolgen. | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     | F14_e |
|   | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     |       |
| F Wenn etwas unerwartet auf mich zukommt, weiß ich normalerweise, wie ich mich verhalten soll.  | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     | F14_f |
|   | 1                            | 2 | 3 | 4 | 5 | 6 | 7 |  |                                     |       |

## Hauptstudie – Dokument 1

Die kommenden Fragen thematisieren wie Du Deine Freizeit verbringst, z.B. auch Dein Gesundheitsverhalten.

### F 15

Welche Zeitungen oder Zeitschriften liest Du zumindest gelegentlich, online oder in gedruckter Form?

--> Antwort bitte genau notieren!



F15\_1 = NOTIEREN DER ZEITSCHRIFTEN TEXT

:

F15\_2 ( 1 = ANGABE IN F15\_1 )  
 9 = KEINE ZEITUNGEN

9  Keine Zeitungen/Zeitschriften

### F 16

Wenn Du einmal an die Tage denkst, an denen Du fern siehst bzw. Dir Fernsehsendungen über einen Computer ansiehst:

Wie lange - ich meine in Stunden und Minuten – siehst Du da im Durchschnitt pro Tag fern?

--> Wenn Zeitspanne angegeben wird, nach Durchschnittswerten pro Tag fragen.  
 In Stunden und Minuten notieren!

Durchschnittliche Fernsehdauer pro Tag:

: F16\_h Stunde(n) F16\_m Minuten

9999  KA 96 = SIEHT KEIN TV

## Hauptstudie – Dokument 1

F 17 *F17*--> *Liste 17 vorlegen!*

Im Folgenden findest Du eine Liste von Wörtern. Die Wörter können Dir persönlich mehr oder weniger bekannt sein. Bitte gib bei jedem Wort an, wie bekannt es Dir ist.

Wenn Du z.B. ein Wort „gar nicht“ kennst, dann wähle bitte die „1“.

Wenn Du ein Wort „sehr gut“ kennst, dann wähle bitte die „7“.

Du kannst Deine Angabe auch abstufen, indem Du eine Antwort zwischen „1“ und „7“ wählst.

--> **Bitte Wörter (A bis S) nicht vorlesen!**

--> *Bitte machen Sie in JEDER Zeile eine Markierung!*

|   |           | Ich kenne<br>das Wort<br>gar nicht |            |               |            |            |            | Ich kenne<br>das Wort<br>sehr gut |              |
|---|-----------|------------------------------------|------------|---------------|------------|------------|------------|-----------------------------------|--------------|
|   |           | 1                                  | 2          | 3             | 4          | 5          | 6          | 7                                 |              |
| A | Mahnung   | <i>1</i> ○                         | <i>2</i> ○ | <i>3</i> ○    | <i>4</i> ○ | <i>5</i> ○ | <i>6</i> ○ | <i>7</i> ○                        | <i>F17_a</i> |
| B | Eichmaß   | <i>1</i> ○                         | <i>2</i> ○ | <i>3</i> ○    | <i>4</i> ○ | <i>5</i> ○ | <i>6</i> ○ | <i>7</i> ○                        | <i>F17_b</i> |
| C | Holozän   | <i>1</i> ○                         | <i>2</i> ○ | <i>3</i> ○    | <i>4</i> ○ | <i>5</i> ○ | <i>6</i> ○ | <i>7</i> ○                        | <i>F17_c</i> |
| D | Offerte   | <i>1</i> ○                         | <i>2</i> ○ | <i>3</i> ○    | <i>4</i> ○ | <i>5</i> ○ | <i>6</i> ○ | <i>7</i> ○                        | <i>F17_d</i> |
| E | Habitat   | <i>1</i> ○                         | <i>2</i> ○ | <i>3</i> ○    | <i>4</i> ○ | <i>5</i> ○ | <i>6</i> ○ | <i>7</i> ○                        | <i>F17_e</i> |
| F | Altenuse  | <i>1</i> ○                         | <i>2</i> ○ | <i>3</i> ○    | <i>4</i> ○ | <i>5</i> ○ | <i>6</i> ○ | <i>7</i> ○                        | <i>F17_f</i> |
| G | Erosion   | ○                                  | ○          | <i>usw.</i> ○ | ○          | ○          | ○          | ○                                 | <i>usw.</i>  |
| H | Platine   | ○                                  | ○          | ○             | ○          | ○          | ○          | ○                                 |              |
| J | Halali    | ○                                  | ○          | ○             | ○          | ○          | ○          | ○                                 |              |
| K | Fazit     | ○                                  | ○          | ○             | ○          | ○          | ○          | ○                                 |              |
| L | Kürschner | ○                                  | ○          | ○             | ○          | ○          | ○          | ○                                 |              |
| M | Triasmus  | ○                                  | ○          | ○             | ○          | ○          | ○          | ○                                 |              |
| N | Sparta    | ○                                  | ○          | ○             | ○          | ○          | ○          | ○                                 |              |
| O | Hybris    | ○                                  | ○          | ○             | ○          | ○          | ○          | ○                                 |              |
| P | Enklavie  | ○                                  | ○          | ○             | ○          | ○          | ○          | ○                                 |              |
| Q | Galaxie   | ○                                  | ○          | ○             | ○          | ○          | ○          | ○                                 |              |
| R | Fagott    | ○                                  | ○          | ○             | ○          | ○          | ○          | ○                                 |              |
| S | Sinus     | ○                                  | ○          | ○             | ○          | ○          | ○          | ○                                 |              |



## Hauptstudie – Dokument 1

--> Liste 18 vorlegen und bis F 20 liegen lassen!

Nun wollen wir ein kurzes Quiz machen. Das Quiz hat 7 Fragen zu unterschiedlichen Themengebieten. Wie man es aus TV-Quiz-Sendungen kennt, hat jede Frage vier Antwortalternativen. Deine Aufgabe besteht darin, aus vier vorgegebenen Antwortalternativen die Richtige herauszufinden. Wähle bitte bei jeder Frage die Antwort aus, die Du für richtig hältst. Nenne mir bitte den Buchstaben, der die korrekte Lösung kennzeichnet.

--> Bitte machen Sie zu JEDER Frage eine Markierung!

### F 18 F18

Auf einem bekannten Gemälde von Dalí werden "zerfließende Uhren" dargestellt. Welcher Stilrichtung ist dieses Gemälde zuzuordnen?

- 1  A Naturalismus  
 2  B Impressionismus  
 3  C Surrealismus  
 4  D Romantik  
 \_\_\_\_\_  
 8  Weiß ich nicht

### F 19 F19

Nach der „Völkerschlacht bei Leipzig“...

- 1  A ...musste sich Napoleon aus Deutschland zurückziehen.  
 2  B ...rückten die alliierten Truppen nach Berlin vor.  
 3  C ...musste Kaiser Wilhelm II. abdanken.  
 4  D ...wurde Karl der Große zum Kaiser gekrönt.  
 \_\_\_\_\_  
 8  Weiß ich nicht

### F 20 F20

Welche Krankheit zerstört das zentrale Nervensystem?

- 1  A Rheuma  
 2  B Multiple Sklerose  
 3  C Osteoporose  
 4  D Anorexie  
 \_\_\_\_\_  
 8  Weiß ich nicht

## Hauptstudie – Dokument 1

- > Liste 21 vorlegen und bis F 24 liegen lassen!  
 --> Bitte machen Sie zu JEDER Frage eine Markierung!

## F 21 F21

Welcher Musiker gilt als bedeutender Vertreter des Jazz?

- 1  A Johnny Cash  
 2  B Louis Armstrong  
 3  C Mick Jagger  
 4  D John Lennon  
 \_\_\_\_\_  
 8  Weiß ich nicht

## F 22 F22

Ein Terabyte sind...

- 1  A ... 1000 Megabyte.  
 2  B ... 1024 Gigabyte.  
 3  C ... 8 Gigabyte.  
 4  D ... 100000 Kilobyte.  
 \_\_\_\_\_  
 8  Weiß ich nicht

## F 23 F23

Was versteht man unter "Nihilismus"?

- 1  A Weltanschauung, die das Positive im Menschen betont  
 2  B Weltanschauung, die die Rolle der Moral betont  
 3  C Weltanschauung, die eine Sinnhaftigkeit der Welt bestreitet  
 4  D Weltanschauung, die den Erkenntnisgewinn als wichtigstes Prinzip ansieht  
 \_\_\_\_\_  
 8  Weiß ich nicht

## F 24 F24

Welche Eigenschaft kennzeichnet eine Diode?

- 1  A Eine Diode lässt den elektrischen Strom nur in einer Richtung durch.  
 2  B Eine Diode speichert elektrische Ladungen.  
 3  C Eine Diode verstärkt elektrische Signale.  
 4  D Eine Diode erzeugt ein Magnetfeld.  
 \_\_\_\_\_  
 8  Weiß ich nicht

## Hauptstudie – Dokument 1

## F 25

Welche Sportarten oder körperliche Aktivitäten übst Du zumindest gelegentlich aus?

--> Wenn keinerlei Sport oder körperliche Aktivität ausgeübt wird,  
kreuzen Sie bitte das dafür unten vorgesehene Kästchen an.

Sportarten oder körperliche Aktivitäten:

: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Ich übe keinerlei Sport oder körperliche Aktivität aus.

F25\_1 = NOTIEREN DER SPORTARTEN TEXT

F25\_2 = ( 1 = ANGABE IN F25\_1 )  
          ( 3 = KEINEN SPORT )

## Hauptstudie – Dokument 1

--> Liste 26 vorlegen und bis F 28 liegen lassen!

In den folgenden Fragen geht es um gesunde Ernährung.

Unter gesunder Ernährung wird verstanden z.B. ...

- ... Ausgewogenes Essen, mit viel Gemüse und Obst.
- ... Mindestens drei Mahlzeiten am Tag.
- ... Wenig „Junk Food“ (d.h. Fast Food, Chips, Süßigkeiten).
- ... Auf ausreichende Menge an Nahrung achten (d.h. nicht zu viel und nicht zu wenig).
- ... Nur essen, wenn man hungrig ist und bei Sättigungsgefühl aufhören zu essen.
- ... Wenig gesättigte Fettsäuren (z.B. tierische Fette).
- ... Zuckerhaltige Getränke (z.B. Cola, Limo) auf maximal ein Glas pro Tag reduzieren.

Bitte beantworte nun die folgenden Fragen:

--> Bitte machen Sie zu JEDER Frage eine Markierung!

## F 26 F26

An wie vielen Tagen in der Woche ernährst Du Dich gesund?

g = K.A.

An ... 0 1 2 3 4 5 6 7

|   |   |   |   |   |   |   |   |                     |
|---|---|---|---|---|---|---|---|---------------------|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Tagen in der Woche. |
|---|---|---|---|---|---|---|---|---------------------|

## F 27 F27

Wenn ich mich an mindestens 4 Tagen in der Woche gesund ernähre, wäre das

--> Bitte machen Sie in JEDER Zeile eine Markierung!

g = K.A.

|   |             |   |   |   |   |   |   |   |               |
|---|-------------|---|---|---|---|---|---|---|---------------|
| A | Schlecht    | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Gut           |
| B | Nutzlos     | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Nützlich      |
| C | Vorteilhaft | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Unvorteilhaft |

## F 28 F28

Meine Freunde würden es befürworten, wenn ich mich an mindestens 4 Tagen in der Woche gesund ernähre.

g = K.A.

|                 |   |   |   |   |   |   |   |           |
|-----------------|---|---|---|---|---|---|---|-----------|
| Stimme nicht zu | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Stimme zu |
|-----------------|---|---|---|---|---|---|---|-----------|

## Hauptstudie – Dokument 1

--> Liste 29 vorlegen und bis F 33 liegen lassen!

--> Bitte machen Sie zu JEDER Frage eine Markierung!

F 29 F29

Wenn ich mich an mindestens 4 Tagen in der Woche gesund ernähre, würden das Menschen, die mir wichtig sind, gut finden.

g = K.A.

Unwahrscheinlich

|     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|
| 1 1 | 2 2 | 3 3 | 4 4 | 5 5 | 6 6 | 7 7 |
|-----|-----|-----|-----|-----|-----|-----|

Wahrscheinlich

F 30 F30

Es ist mir möglich mich an mindestens 4 Tagen in der Woche gesund zu ernähren. g = K.A.

Trifft überhaupt  
nicht zu

|     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|
| 1 1 | 2 2 | 3 3 | 4 4 | 5 5 | 6 6 | 7 7 |
|-----|-----|-----|-----|-----|-----|-----|

Trifft voll und  
ganz zu

F 31 F31

Ob ich mich an mindestens 4 Tagen in der Woche gesund ernähre oder nicht, liegt ausschließlich in meiner Hand.

g = K.A.

Stimme nicht  
zu

|     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|
| 1 1 | 2 2 | 3 3 | 4 4 | 5 5 | 6 6 | 7 7 |
|-----|-----|-----|-----|-----|-----|-----|

Stimme zu

F 32 F32

In Zukunft werde ich mich an mindestens 4 Tagen in der Woche gesund ernähren.

Stimme nicht  
zu

|     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|
| 1 1 | 2 2 | 3 3 | 4 4 | 5 5 | 6 6 | 7 7 |
|-----|-----|-----|-----|-----|-----|-----|

Stimme zu

F 33 F33

In den kommenden Wochen werde ich mich an mindestens 4 Tagen in der Woche gesund ernähren.

g = K.A.

Werde ich auf  
jeden Fall

|     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|
| 1 1 | 2 2 | 3 3 | 4 4 | 5 5 | 6 6 | 7 7 |
|-----|-----|-----|-----|-----|-----|-----|

Werde ich auf  
keinen Fall

## Hauptstudie – Dokument 1

F 34 F34

In der folgenden Frage geht es um Ernährung.

Ich lese Dir jetzt 12 Desserts vor:

Bitte sag mir welches dieser Desserts Du am meisten magst. Bitte entscheide Dich für ein Desert.

-->Angaben vorlesen!

-->Bitte darauf achten, dass die befragte Person diese Liste nicht sehen darf!

-->Nur eine Nennung möglich!

- 1  Rote Grütze
- 2  Herrencreme
- 3  Fruchtquark
- 4  Eis
- 5  Obstsalat
- 6  Mousse au Chocolat
- 7  Sorbet
- 8  Tiramisu
- 9  Schokoladenpudding
- 10  Joghurt
- 11  Crêpes bzw. Pfannkuchen
- 12  Wackelpudding
- \_\_\_\_\_
- 39  KA

## Hauptstudie – Dokument 1

Mit dem Umfrageprojekt möchten wir feststellen, wie die wirtschaftliche Lage der Studierenden ist.

### F 35 F35

--> Liste 35 vorlegen und bis Frage 36 liegenlassen!

Wie beurteilst Du ganz allgemein die heutige wirtschaftliche Lage in Deutschland?  
Benutze für Deine Antworten bitte die Liste.

- 1  Sehr gut  
 2  Gut  
 3  Teils gut / teils schlecht  
 4  Schlecht  
 5  Sehr schlecht  
 \_\_\_\_\_  
 8  Weiß nicht  
 9  KA

### F 36 F36

--> Liste 35 liegt vor!

Und Deine eigene wirtschaftliche Lage heute?

- 1  Sehr gut  
 2  Gut  
 3  Teils gut / teils schlecht  
 4  Schlecht  
 5  Sehr schlecht  
 \_\_\_\_\_  
 8  Weiß nicht  
 9  KA

## Hauptstudie – Dokument 1

F 37 *F37*

Wie viel Geld steht Dir durchschnittlich im Monat während des aktuellen Semesters zur Verfügung? Dabei meine ich nur die Mittel für Deinen monatlichen Lebensunterhalt.

Bitte gib für jede zutreffende Finanzierungsquelle den Betrag an.

|  |  | KA                                 |
|--|--|------------------------------------|
| A Von den Eltern, vom Partner/In, von Verwandten/Bekanntem (ggf. inkl. Kindergeld).<br>Bar auf die Hand/per Überweisung auf Dein Konto.  | <i>0 = KEIN EINKOMMEN</i><br>_____ €                                 | <input type="radio"/> <i>F37_a</i> |
| B BaföG, sonstige soziale Leistungen und/oder Kredit,<br>z.B. Bildungskredit, Studienkredit der KfW Bankengruppe o.ä.  | <i>99997 = VERWEIGERT</i><br><i>99999 = KEINE ANGABE</i><br>_____ €  | <input type="radio"/> <i>F37_b</i> |
| C Eigener Verdienst aus Tätigkeiten während der Vorlesungszeit und/oder der vorlesungsfreien Zeit.<br>Bitte nenne mir den Betrag, den Du davon monatlich im Durchschnitt einsetzt. | _____ €  | <input type="radio"/> <i>F37_c</i> |
| D Andere Finanzierungsquellen.<br>Bitte benenne diese:   | <i>F37_d_1 = ART DER QUELLE</i><br><i>F37_d_2 = SUMME</i><br>_____ € | <input type="radio"/>              |
| ☞: _____<br>_____  |  |                                    |

F 38 *F38*

Wie viel Geld stand Dir im letzten Monat für Deinen Lebensunterhalt insgesamt zur Verfügung?

|   |   | KA                    |
|---|---|-----------------------|
| Im letzten Monat standen mir _____ € zur Verfügung. | <i>0 = KEIN EINKOMMEN</i><br><i>99997 = VERWEIGERT</i><br><i>99999 = K.A.</i> | <input type="radio"/> |

F 39 *F39*

Wie hoch sind Deine Ausgaben für die Miete einschließlich der Nebenkosten?  
Bitte rechne bei Deiner Angabe auch den Teil dazu, den Deine Eltern bzw. Dein Partner/Deine Partnerin bezahlen.

|  |  | KA                    |
|--|--|-----------------------|
| A Miete einschl. der Nebenkosten _____ € | <i>0 = KEINE MIETE</i><br><i>99997 = VERWEIGERT</i><br><i>99999 = K.A.</i> | <input type="radio"/> |



## Hauptstudie – Dokument 1

F 40 F40

--&gt; Liste 40 vorlegen!

Welche Zeitungen oder Zeitschriften von dieser Liste liest Du zumindest gelegentlich, online oder in gedruckter Form?

- |       |                       |   |  |
|-------|-----------------------|---|--|
| F40-a | <input type="radio"/> | Süddeutsche Zeitung                       |  |
| F40-b | <input type="radio"/> | GQ  |  |
| F40-c | <input type="radio"/> | Charm                                     |  |
| F40-d | <input type="radio"/> | Wirtschaftwoche                           |  |
| F40-e | <input type="radio"/> | Mitteldeutsches Tageblatt                 |  |
| F40-f | <input type="radio"/> | Spiegel                                   |  |
| F40-g | <input type="radio"/> | Taz                                       |  |
| F40-h | <input type="radio"/> | InStyle                                   |  |
| F40-i | <input type="radio"/> | GEO                                       |  |
| F40-j | <input type="radio"/> | Frankfurter Rundschau                     |  |
| F40-k | <input type="radio"/> | Triange                                   |  |
| F40-l | <input type="radio"/> | Cosmopolitan                              |  |
| F40-m | <input type="radio"/> | Frankfurter Allgemeine Zeitung            |  |
| F40-n | <input type="radio"/> | Trendsetter                               |  |
| F40-o | <input type="radio"/> | Neon                                      |  |
| <hr/> |                       |   |  |
| 6     | <input type="radio"/> | Keine dieser Zeitungen oder Zeitschriften |  |

0 = NICHT GENANNT

1 = GENANNT

6 = NICHTS DAVON

9 = KA

## Hauptstudie – Dokument 1

Nun möchte ich Dich noch um einige Angaben zu Deinem Studium bitten.

## F 41 F41

-->Liste 41 vorlegen!

Welches Hauptfach bzw. welche Fächer studierst Du im aktuellen Semester?

- |    |                       |    |   |                        |      |
|----|-----------------------|----|---|------------------------|------|
| 1  | <input type="radio"/> | A  | Agrarwissenschaften                       | F41-a = 1. STUDIENFACH |      |
| 2  | <input type="radio"/> | B  | Bauwesen, Architektur                     |                        |      |
| 3  | <input type="radio"/> | C  | Geowissenschaften, Physik                 | F41-b = 2.             | -11- |
| 4  | <input type="radio"/> | D  | Biologie, Chemie                          |                        |      |
| 5  | <input type="radio"/> | E  | Elektrotechnik                            | F41-c = 3.             | -11- |
| 6  | <input type="radio"/> | F  | Sprach-, Kulturwissenschaften             |                        |      |
| 7  | <input type="radio"/> | G  | Kunst, Kunstwissenschaften                | F41-d = 4.             | -11- |
| 8  | <input type="radio"/> | H  | Maschinenbau                              |                        |      |
| 9  | <input type="radio"/> | J  | Mathematik, Informatik                    | F41-e = 5.             | -11- |
| 10 | <input type="radio"/> | K  | Medizin (Human-, Tiermedizin)             |                        |      |
| 11 | <input type="radio"/> | L  | Gesundheitswissenschaften                 | F41-f = 6.             | -11- |
| 12 | <input type="radio"/> | M  | Pädagogik                                 |                        |      |
| 13 | <input type="radio"/> | N  | Psychologie                               |                        |      |
| 14 | <input type="radio"/> | O  | Rechtswissenschaften                      |                        |      |
| 15 | <input type="radio"/> | P  | Sozialwissenschaften, Sozialwesen         |                        |      |
| 16 | <input type="radio"/> | Q  | Wirtschaftswissenschaften                 |                        |      |
| 17 | <input type="radio"/> |    | Sonstiges, und zwar: <del>☒</del> : _____ | F41-g = TEXT           |      |
| 39 | <input type="radio"/> | KA |   |                        |      |

## F 42 F42

-->Liste 42 vorlegen!

Welchen Abschluss strebst Du in Deinem Studiengang an?

Bitte gib nur den Abschluss an, den Du zunächst erwerben willst.

- |    |                       |    |   |                    |      |
|----|-----------------------|----|---|--------------------|------|
| 1  | <input type="radio"/> | A  | Bachelor (nicht Lehramt)  | F42-1 = 1. NENNUNG |      |
| 2  | <input type="radio"/> | B  | Bachelor mit dem Ziel Lehramt   |                    |      |
| 3  | <input type="radio"/> | C  | Master (nicht Lehramt)  | F42-2 = 2.         | -11- |
| 4  | <input type="radio"/> | D  | Master mit dem Ziel Lehramt   | F42-3 = 3.         | -11- |
| 5  | <input type="radio"/> | E  | Fachhochschuldiplom   |                    |      |
| 6  | <input type="radio"/> | F  | Diplom einer Universität/Kunsthochschule o. ä.  |                    |      |
| 7  | <input type="radio"/> | G  | Magister  |                    |      |
| 8  | <input type="radio"/> | H  | Staatsexamen (ohne Lehramt)   |                    |      |
| 9  | <input type="radio"/> | J  | Staatsexamen für ein Lehramt  |                    |      |
| 10 | <input type="radio"/> | K  | Kirchliche Prüfung  |                    |      |
| 11 | <input type="radio"/> | L  | Promotion   |                    |      |
| 12 | <input type="radio"/> | M  | Keinen Abschluss  |                    |      |
| 13 | <input type="radio"/> |    | Anderer Abschluss (einschließlich Abschluss in Ausland),<br>und zwar <del>☒</del> : _____ | F42-5 = TEXT       |      |
| 39 | <input type="radio"/> | KA |   |                    |      |

## Hauptstudie – Dokument 1

F 43 F43

Wie viele Semester, einschließlich des aktuellen Semesters, bist Du bisher insgesamt an Hochschulen eingeschrieben?

Insgesamt heißt: Semester im derzeitigen Studiengang plus gegebenenfalls in einem anderen Studiengang absolvierte Semester, sowie Urlaubssemester, Praxissemester und Semester an ausländischen Hochschulen.

Seit \_\_\_\_ Semestern

33  KA

F 44 F44

Seit wie vielen Semestern, einschl. des aktuellen Semesters, bist Du im derzeitigen Studiengang eingeschrieben?

Seit \_\_\_\_ Semestern

33  KA

F 45 F45

--> Liste 45 vorlegen!

Welche der drei folgenden Aussagen trifft am ehesten auf Deine derzeitige Studien- und Lebenssituation zu?

- 1  A Studium und Hochschule bilden den Mittelpunkt, auf den fast alle meine Interessen und Aktivitäten ausgerichtet sind.
- 2  B Studium und Hochschule sind mir gleich wichtig wie andere Interessen und Aktivitäten außerhalb der Hochschule.
- 3  C Studium und Hochschule stehen eher im Hintergrund, weil meine Interessen und Aktivitäten außerhalb der Hochschule vorrangig sind.

9  KA

## Hauptstudie – Dokument 1

F 46 F46

--&gt;Liste 46 vorlegen!

In welchem Bundesland hast Du Deine Studienberechtigung erworben?

1  A Baden-Württemberg2  B Bayern3  C Berlin4  D Brandenburg5  E Bremen6  F Hamburg7  G Hessen8  H Mecklenburg-Vorpommern9  J Niedersachsen10  K Nordrhein-Westfalen11  L Rheinland-Pfalz12  M Saarland13  N Sachsen14  O Sachsen-Anhalt15  P Schleswig-Holstein16  Q Thüringen17  im Ausland, und zwar in:

F46\_1 = BUNDESLAND-NUMMER

F46\_2 = TEXT LAND

99  KA

F 47 F47

Hattest Du vor der Erstimmatrikulation bereits eine Berufsausbildung erfolgreich abgeschlossen?

1  Ja2  Nein9  KA

## Hauptstudie – Dokument 1

Nun folgt der letzte Teil des Interviews, in dem wir noch einige demographische Angaben von Dir erheben möchten.

F 48 **F48**

--> Geschlecht der befragten Person ohne Befragen eintragen!

- 1  Männlich  
2  Weiblich

F 49 **F49**

Wann bist Du geboren?

~~☒~~: F49\_m Monat F49\_j Jahr

KA

99 = KA MONAT

9999 = KA JAHR

F 50 **F50**

Würdest Du bitte folgende Angaben machen:

A Größe:

~~☒~~: F50\_a cm

B Gewicht:

~~☒~~: F50\_b , \_\_\_ kg

Weiß ich nicht

KA

998

999

998

999

## Hauptstudie – Dokument 1

F 51 *FS1*

--> *Liste 51 vorlegen und liegenlassen!*

Als Du 15 Jahre alt warst, welche berufliche Stellung traf damals auf Deinen Vater zu?  
Ordne es bitte nach dieser Liste ein.

--> *Bitte Kennziffer aus der Liste 51 eintragen!*

--> *Falls Vater zu dieser Zeit nicht erwerbstätig war, bitte informell ermitteln, welche Antwortvorgabe zutrifft!*

Kennziffer:

~~☒~~: *KENNZIFFER EINTRAGEN* ..... → Weiter mit F 52

- |  |                   |
|--|-------------------|
| <i>32</i> ○ Vater war zu der Zeit Rentner / Pensionär          | } Weiter mit F 53 |
| <i>33</i> ○ Vater war zu der Zeit arbeitslos                   |                   |
| <i>34</i> ○ Vater war zu der Zeit im Krieg / in Gefangenschaft |                   |
| <i>35</i> ○ Vater lebte zu der Zeit nicht mehr                 |                   |

*36* ○ Vater unbekannt ..... → Weiter mit F 55

*38* ○ Weiß nicht ..... → Weiter mit F 52

*39* ○ KA ..... → Weiter mit F 53

F 52

**FALLS Vater erwerbstätig war oder „weiß nicht“ in F 51:**

Welche berufliche Tätigkeit übte Dein VATER damals aus?  
Bitte beschreibe mir diese berufliche Tätigkeit genau.

--> *Bitte genau nachfragen:*

Hat dieser Beruf, diese Tätigkeit noch einen besonderen Namen?

~~☒~~: *F52 = TEXT BERUF*

\_\_\_\_\_

\_\_\_\_\_

*39* ○ KA

## Hauptstudie – Dokument 1

F 53 **F53**

FALLS Vater bekannt in F 51

--&gt; Liste 53 vorlegen!

Welchen allgemeinbildenden Schulabschluss hat (hatte) Dein VATER?

--&gt; Nur eine Nennung möglich!

--&gt; Nur h ö c h s t e n Schulabschluss angeben lassen!

- 1  A Schule beendet ohne Abschluss  
 2  B Volks- / Hauptschulabschluss bzw. Polytechnische Oberschule mit Abschluss 8. oder 9. Klasse  
 3  C Mittlere Reife, Realschulabschluss bzw. Polytechnische Oberschule mit Abschluss 10. Klasse  
 4  D Fachhochschulreife (Abschluss einer Fachoberschule etc.)  
 5  E Abitur bzw. Erweiterte Oberschule mit Abschluss 12. Klasse (Hochschulreife)  
 6  F Anderen Schulabschluss

- 8  Weiß nicht  
 9  KA

0 = VATER UNBEKANNT

F 54 **F54**

FALLS Vater bekannt in F 51

--&gt; Liste 54 vorlegen!

Welchen beruflichen Ausbildungsabschluss hat (hatte) Dein Vater?

Was von dieser Liste trifft / traf auf ihn zu?

Nenne mir bitte den entsprechenden Kennbuchstaben.

--&gt; Nur eine Nennung möglich

--&gt; Nur h ö c h s t e n Abschluss angeben lassen!

- 1  A Abgeschlossene gewerbliche oder landwirtschaftliche Lehre  
 2  B Abgeschlossene kaufmännische Lehre  
 3  C Fachschulabschluss (einschließlich Meister- und gleichwertiger Technikerabschluss)  
 4  D Fachhochschulabschluss (auch Abschluss einer Ingenieurschule)  
 5  E Hochschulabschluss  
 6  F Anderen beruflichen Ausbildungsabschluss, und zwar:

☒: F54\_5 = TEXT ABSCHLUSS

- 7  G Keinen beruflichen Ausbildungsabschluss

- 88  Weiß nicht  
 99  KA

0 = VATER UNBEKANNT

## Hauptstudie – Dokument 1

## F 55 A FSSa

--> Liste 55 vorlegen!

In der folgenden Frage geht es darum, wo Du wohnst. Es geht um den Wohnort während einer Arbeitswoche – Montag bis Freitag – in der Vorlesungszeit.

Wo wohnst Du während des aktuellen Semesters?

--> Nur eine Nennung möglich!

Ich wohne:

- 1  A bei den Eltern oder Verwandten  
 2  B in einer Wohngemeinschaft  
 3  C bei einer Familie als Untermieter  
 4  D in einem Studentenheim  
 5  E in einer Mietwohnung  
 6  F in einer Eigentumswohnung  
 7  G im eigenen Haus

8  KA

## F 55 B FSSb\_1

Mit wem wohnst Du während des aktuellen Semesters?

--> Nur eine Nennung möglich! --> Kategorien vorlesen!

Ich wohne:

- 1  A alleine  
 2  B mit Partner  
 3  C mit Kind  
 4  D mit Partner und Kind  
 5  E Sonstiges, und zwar ~~☒~~: FSSb\_2 = TEXT

8  KA

## F 56

Bitte bezieh Dich in der folgenden Frage auf den soeben genannten Wohnort.  
 Wie ist die Postleitzahl Deines Wohnortes?

--> Bitte genau PLZ notieren!

~~☒~~ PLZ: F56 = TEXT

KA 99999



## Hauptstudie – Dokument 1

F 57 FS7

Wie lange brauchst Du für den Weg von Deinem soeben genannten Wohnort zur Universität, d.h. von Tür zu Tür? Bitte gib die Dauer in Minuten an.

--> Bitte Angabe genau notieren!

~~☒~~: FS7 = ZAHL Minuten

999  KA

F 57 A FS7a

Bist Du in Deutschland geboren?

1  Ja → Weiter mit F 58

2  Nein

9  KA

F 57 B FS7b

In welchem Land bist Du geboren?

--> Bitte genau das Land notieren!

~~☒~~: FS7b = TEXT LAND

9999  KA

0 = IN DEUTSCHLAND GEBOREN

F 57 C FS7c

In welchem Jahr bist Du nach Deutschland gekommen?

--> Bitte genau notieren!

~~☒~~: FS7c = ZAHL (Jahr)

9999  KA

0 = IN DEUTSCHLAND GEBOREN

F 58 FS8

Ist Dein Vater in Deutschland geboren?

1  Ja

2  Nein

98  Weiß nicht

96  Vater unbekannt

## Hauptstudie – Dokument 1

F 59 F59

Ist Deine Mutter in Deutschland geboren?

- 1  Ja  
 2  Nein  
 \_\_\_\_\_  
 98  Weiß nicht  
 96  Mutter unbekannt

F 60

Was ist Deine Muttersprache?

--&gt; Bitte genau notieren!

☒: F60\_1 = TEXT 1. MUTTERSPRACHE

999  KA F60\_2 = TEXT 2. MUTTERSPRACHE

F 61 F61

--&gt; Liste 61 vorlegen!

Welchen Familienstand hast Du?

„Eingetragene“ Lebenspartnerschaft heißt, dass es sich um eine gleichgeschlechtliche Partnerschaft im Sinne des Lebenspartnerschaftsgesetzes handelt oder gehandelt hat.

--&gt; Nur eine Nennung möglich!

- 1  A Verheiratet und lebe mit meinem/r Ehepartner/-in zusammen  
 2  B In eingetragener Lebenspartnerschaft zusammenlebend (gleichgeschlechtlich)  
 3  C Verheiratet und lebe von meinem/meiner Ehepartner/-in getrennt  
 4  D Ledig  
 5  E Ledig, in einer festen Partnerschaft  
 6  F Geschieden  
 7  G Verwitwet  
 8  H Eingetragene Lebenspartnerschaft, getrennt lebend (gleichgeschlechtlich)  
 9  J Eingetragene Lebenspartnerschaft aufgehoben (gleichgeschlechtlich)  
 10  K Eingetragene/-r Lebenspartner/-in verstorben (gleichgeschlechtlich)

999  KA

## Hauptstudie – Dokument 1

F 61 F61\_b

Wärst Du bereit, unser Forschungsprojekt auch weiterhin durch Deine Mitarbeit bei einer Internet-Umfrage zu unterstützen und uns dazu Deine Email-Adresse anzugeben?

- 1  Ja, wäre bereit → Weiter mit F 62  
 2  Nein, wäre nicht bereit

F 62 F62

Bitte notieren Sie die Email-Adresse der Befragungsperson!

~~☒~~: F62 = ADRESSE AUFSCHREIBEN

--> F000A2: Die Befragungsperson sollte Ihre Antwort nicht einsehen können!

F000A2

→ Nur für den Interviewer / die Interviewerin!

Bitte beurteilen Sie noch einmal die Attraktivität des / der Befragten. Entscheiden Sie bitte wieder spontan.

--> Nur EINE Markierung möglich!

- 11  11 = Attraktiv  
 10  10  
 9  09  
 8  08  
 7  07  
 6  06  
 5  05  
 4  04  
 3  03  
 2  02  
 1  01 = Unattraktiv

KEINE MARKIERUNG = -9

F1000\_T

--> Bitte Uhrzeit eintragen!

Uhrzeit ~~☒~~: F1000\_TS : F1000\_TM Uhr

K.A. = 99

**Vielen Dank für die Teilnahme an unserer Umfrage!!!**

Der Fragebogen, den die Interviewer ausgefüllt haben, unterscheidet sich nur in der letzten Seite vom Hauptfragebogen:

---

Hauptstudie – Dokument 1

33

F1000\_T

--> Bitte Uhrzeit eintragen!

**Uhrzeit** ~~☒~~: \_\_\_\_ \_\_\_\_ : \_\_\_\_ \_\_\_\_ Uhr

FI01 **FI01**

Haben Sie vor der Tätigkeit für dieses Projekt schon irgendwann einmal Interviews durchgeführt?

- 1**  Ja → Weiter mit FI02  
**2**  Nein → ENDE!

FI02 **FI02**

Falls ja: Wie viele Interviews haben Sie bisher insgesamt etwa durchgeführt?

~~☒~~ Anzahl der Interviews eintragen: \_\_\_\_\_ Interviews

FI03

Welche Arten von Interviews haben Sie bisher schon durchgeführt?

--> Mehrfachnennungen möglich!

- 33-a**  Persönlich-mündlich konventionell  
**33-b**  Telefonisch konventionell  
**03-c**  CAPI  
**03-d**  CATI

**0 = NICHT GENANNT**

**1 = GENANNT**

**6 = KEINS DAVON**

**9 = K.A.**

**Vielen Dank für das Ausfüllen des Fragebogens!**

## Spezifizierung der Koautorschaften

Der Artikel "Validation of Theoretical Assumptions with Real and Falsified Survey Data"<sup>1</sup> wurde in Zusammenarbeit mit PD Dr. Natalja Menold erstellt. Ich bin Erstautorin des Artikels.

### Anteil Uta Landrock:

- 1) Manuskriptidee
- 2) Identifikation, Sammlung und Aufbereitung der relevanten Literatur
- 3) Erstellung der Struktur des Manuskripts
- 4) Konzeption und Durchführung der statistischen Analysen
- 5) Aufbereitung der Ergebnisse der Analysen für die Publikation
- 6) Verfassen des Manuskripts

### Anteil Dr. Natalja Menold:

- 1) Diskussion der Manuskriptidee
- 2) Erstellen des Erhebungsinstruments
- 3) Diskussion der Analyseergebnisse
- 4) Diskussion der Konzeption und Struktur des Manuskripts
- 5) Rückmeldung zum Text, Beschreibung der Datenerhebung, Ergänzungen zur Einleitung und zum Fazit

Insgesamt beträgt mein Anteil an der Publikation 75 Prozent, der Anteil von PD Dr. Natalja Menold beträgt 25 Prozent.

Mainz, den 03.08.2017

  
PD Dr. Natalja Menold

  
Uta Landrock

<sup>1</sup> Landrock, Uta and Menold, Natalja (2016): Validation of Theoretical Assumptions with Real and Falsified Survey Data. Statistical Journal of the IAOS 32(3): 305-12.

## Erklärung

Ich erkläre: Ich habe die vorgelegte Dissertation selbständig, ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der  $\text{\textcircled{S}}$ atzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis  $\text{\textcircled{S}}$  niedergelegt sind, eingehalten.

---

Ort, Datum

---

Unterschrift