Survey Research Methods (2018)
Vol. 12, No. 2, pp. 121-146
doi:10.18148/srm/2018.v12i2.7299

ISSN 1864-3361
http://www.surveymethods.org

From Strangers to Acquaintances? Interviewer Continuity and Socially Desirable Responses in Panel Surveys

Simon Kühne
University of Bielefeld
Germany

In many panel surveys that rely on face-to-face interviewing, interviewers are repeatedly allocated to the same respondents in each wave. Researchers and fieldwork agencies argue that interviewer continuity can contribute to the quality of the data collected, for instance, by reducing panel attrition. However, there is almost no empirical evidence focusing on the effects of growing familiarity between interviewers and respondents on responses and measurement error in repeated interviews. This paper focuses on questions containing socially (un)desirable answer options. It is argued that interviewer continuity promotes the development of trust, emotional closeness, and loyalty, as well as interview rapport between respondents and interviewers, and that this, in turn, increases the respondents' motivation to answer truthfully rather than in a socially desirable way. Drawing on data derived from 31 waves of an ongoing household panel study in Germany, the results show a consistent effect of interviewer continuity on response behavior: Respondents who are more familiar with their interviewers are less likely to choose answer options associated with socially desirable connotations. This study provides evidence for a rare advantageous panel conditioning effect on data quality in longitudinal studies and points to the importance of taking into account the familiarity between respondents and interviewers when investigating conditioning effects on measurement error in longitudinal studies.

Keywords: interviewer continuity; interviewer-respondent familiarity; interviewer effects; panel conditioning; social desirability; panel surveys

1 Introduction

Many ongoing large scale panel surveys in the social sciences rely on face-to-face interviewing (e.g., the Socio-Economic Panel Survey, SOEP, see Wagner, Frick, and Schupp, 2007; UK Understanding Society, see Buck and McFall, 2012). Interviewers contribute to panel data quality by maintaining high participation rates, handling complex survey instruments and questionnaires, while providing assistance and clarifications to respondents (Fowler & Mangione, 1990). This is why many researchers still rely on interviewers despite their high costs compared to other modes of data collection, such as telephone and web surveys. However, survey research also reveals that interviewers can trigger undesired response behavior and measurement error through their mere presence as well as their individual characteristics. Although numerous studies investigate these interviewer effects in the cross-sectional context (see West and Blom, 2017 for a research synthesis on the topic), little is known about the specifics of the interviewer's influence on responses in longitudinal studies. This is especially surprising since, in many ongoing panel studies, interviewers are repeatedly allocated to the same respondents. It seems reasonable to expect that multiple encounters between interviewers and respondents affect the characteristics of their relationship and social interaction, such as interpersonal trust and mutual self-disclosure. So far, the potential effects of growing familiarity between respondents and interviewers on respondents' answering behavior remain largely unknown.

This paper addresses *interviewer continuity* effects on responses in panel surveys. Whereas other studies on panel conditioning effects usually focus on the impact of the respondents' general panel participation (e.g., Warren and Halpern-Manners, 2012, Uhrig, 2012), this paper examines whether repeated encounters of respondents with the same interviewer affect their answering behavior. In this regard, the paper focuses on questions associated with socially (un)desirable answer options. It is argued that the growing familiarity between respondents and interviewers promotes the development of interpersonal trust, emotional closeness, and interview rapport, and that this, in turn, increases respondent motivation to answer truthfully rather than providing socially desirable answers.

Contact information: Simon Kühne, University of Bielefeld, Facultät für Soziologie, Postfach 100131, 333501 Bielefeld (simon.kuehne@uni-bielefeld.de)

The paper is structured as follows. Section 2 defines the key concepts and summarizes the existing literature and research gaps. The theoretical background and hypotheses are introduced in Section 3. Section 4 describes the data and methods; results are presented in Section 5. Finally, section 6 summarizes the results and discusses limitations and starting points for further research on the topic.

2 Motivation and State of Research

Many researchers and fieldwork agencies aim for interviewer continuity in their face-to-face panel studies and, consequently, try to avoid changing interviewer. This includes for instance, the Socio-Economic Panel Study (SOEP) (Wagner et al., 2007, Schräpler, 2002), Understanding Society – The UK Household Longitudinal Study (Mitchell, Collings, and Brown, 2015, p. 9), and The Household, Income and Labour Dynamics in Australia (HILDA, Watson and Wooden, 2009). The reasons for this are diverse. On the one hand, organizational or financial limitations can result in interviewer continuity. Since large-scale samples are often based on clustered sampling designs, it is usually more cost-effective to allocate interviewers who live near the primary sample clusters, thereby reducing travel costs. Thus, if neither the interviewer nor the respondent moves, allocating the interviewer living geographically closest results in interviewer continuity. On the other hand, it is argued that interviewer continuity contributes to the quality of the data collected (see e.g., Lynn and Lugtig, 2017). While there is first evidence that interviewer continuity reduces panel attrition rates (e.g., Rendtel, 1995, Lynn, Kaminska, and Goldstein, 2014, Campanelli and O'Muricheartaigh, 1999, Vassallo, Durrant, Smith, and Goldstein, 2015) and item nonresponse rates (Hajek and Schumann, 2014, Schräpler, 2004), only a few studies examine possible effects on response behavior and measurement error, such as socially desirable responding. Moreover, the few existing studies are rather inconclusive as some point to a decrease in socially desirable responses with growing interviewer continuity (Chadi, 2013, Warren and Halpern-Manners, 2012) while others point to an increase (Uhrig and Lynn, 2008, Mensch and Kandel, 1988).

2.1 Socially Desirable Responding

Socially desirable responding is one of the most often discussed sources of response bias in sensitive questions (see Krumpal, 2013 for an overview). Put simply, social desirability describes the tendency of some respondents to give positive self-descriptions in order to create a positive image of themselves rather than answering truthfully and accurately. Some respondents tend to over-report socially desirable behaviors, attitudes, and traits (e.g., charity donations, Lee and Woodlife, 2010) and under-report socially undesirable characteristics (e.g., not having voted, see Tourangeau and Yan, 2007) when confronted with questions containing (un)desirable answer options. Social desirability is shown to be especially likely in personal interviews (Tourangeau, Rips, & Rasinski, 2000), in which some respondents may seek approval from their interviewer. While survey research literature usually refers to clear social norms as the main source of desirability (Krumpal, 2013, e.g., stealing is socially undesired), the respondents own expectations and judgements, as well as anticipated interviewer expectations, may also function as a benchmark for desirability (DeMaio, 1984, p. 258).

Some scholars define social desirability as a (stable) personality characteristic (e.g., Paulhus, 1991). In this regard, respondents are thought to vary in their general proneness to engage in socially desirable answering behavior. Following Paulhus (1991), one can differentiate between the concepts of "other-deception" and "self-deception," which relate to the addressee of the social desirability behavior. Other-deception refers to the individuals' impression management behavior and need for social approval by others, for instance, the interviewer. In relation to the widely accepted model of the question-answering process (see e.g., Tourangeau et al., 2000) some respondents are expected to deliberately and consciously adjust a true answer during the process of giving an actual answer. In contrast, self-deception describes the respondent's strategies to "maintain a positive self-image, to maximize self-worth and to reduce cognitive dissonance resulting from divergence between social norms, self-perception, and self-demands, on the one hand, and reality, on the other hand" (Krumpal, 2013, p. 2030). In contrast to impression management strategies, self-deceptive response strategies are mainly thought to occur subconsciously.

Social desirability (bias) in surveys is not yet fully conceptualized and the exact determinants of its occurrence and underlying operating mechanisms are still insufficiently understood (Holtgraves, 2004). For instance, there is no final conclusion to what extent social desirability bias reflects an (conscious) adjustment of an answer in the final stage of providing an answer rather than a (subconscious) heuristic (such as "satisficing") occurring during the information retrieval and evaluation stage of generating an answer (see Holtgraves, 2004, Schaeffer, 2000). The few existing empirical studies point to an adjustment of a true answer, i.e., a response editing mechanism, rather than a (subconscious) satisficing heuristic (Kaminska and Foulsham, 2016, Kaminska and Foulsham, 2013, Holtgraves, 2004).

Furthermore, social desirability is framed as an item characteristic rather than a personality characteristic (e.g., DeMaio, 1984). From this perspective, the source of social desirability emerges from the questions and items itself. Some questions relate to topics that are perceived as sensitive by many, i.e., provoke feelings of intrusiveness, fears about the disclosure of certain information or social desirability concerns (Tourangeau and Yan, 2007, F. Kreuter, Presser, and

Tourangeau, 2008). Thus, some question topics, as well as their answer options, are generally being perceived as more sensitive and, therefore, are more likely to provoke socially desirable responding than others (e.g. Sudman and Bradburn, 1974). Socially desirable bias is shown to be present across all types of questions, including questions on self-reported behaviors (see e.g., Tourangeau and Yan, 2007, F. Kreuter et al., 2008), attitudes (see e.g., Sudman and Bradburn, 1974), as well as traits (e.g., Crowne and Marlowe, 1960).[1] Moreover, contextual factors of the survey, such as the mode and the perceived privacy of the interview situation, play an important role in the occurrence of social desirability. Less private interview situations, e.g., face-to-face interviews, are more likely to provoke socially desirable responses compared to modes without interviewer administration, like web surveys (e.g., F. Kreuter et al., 2008). Such interviewer effects are shown to be especially likely in cases where question topics are directly related to observable interviewer characteristics such as the interviewer's gender (e.g., Kane and Macaulay, 1993) or race/ethnicity (e.g., Schuman and Converse, 1971).

2.2 Social Desirability Bias in Panel Surveys

While the majority of research on social desirability bias relies on cross-sectional data, some scholars have also addressed its occurrence in the context of panel surveys and panel conditioning. The concept of "panel conditioning" or "time-in-sample bias" relates to the phenomenon that responses in later waves of a panel may be influenced by previous interviews (e.g., Sturgis, Allum, and Brunton-Smith, 2009). Usually, scholars separate between two broader categories of panel conditioning, "real changes" and "changes in reporting" (Waterton & Lievesley, 1989). Social desirability bias has been related to both of these categories. In the context of *real changes* due to panel participation, Warren and Halpern-Manners (2012, p. 502) argue that sensitive questions "may force respondents to confront the reality that their attitudes/and or behaviours conflict with what mainstream society regards as normative and appropriate" and that "respondents may react to survey questions by bringing their actual attitudes or behaviours into closer conformity with social norms." On the other hand, panel participation may also cause *changes in reporting* to sensitive questions. In this regard, it is hypothesized that respondents may feel increasingly "comfortable" (Warren & Halpern-Manners, 2012, p. 502) and "more at ease" (Waterton & Lievesley, 1989, p. 330) with the survey procedure and instruments, and consequently, feel increasingly free to answer truthfully. Moreover, the experience of an absence of negative consequences of participating in the survey may grow trust in the survey organization and the interviewer as well as "confidence given the promises of confidentiality" (Nancarrow & Cartwright, 2007, p. 574). This, in turn, likely increases the respondents' perceived privacy, thus it is expected to reduce tendencies to apply social desirability strategies.

So far, empirical evidence for panel conditioning effects on socially desirable response behavior is rare and inconclusive. Based on a longitudinal web survey, Halpern-Manners, Warren, and Torche (2014) find that questions about theft and drunk driving affect answers to the same questions in subsequent waves, but only when the surveys are close together in time. Uhrig (2012), investigating effects of panel participation on socially desirable responding in questions on respondents' reported body weight and height, only finds evidence for a decrease in social desirability bias for weight in women. Finally, Warren and Halpern-Manners (2012) provide evidence for decreasing rates of socially desirable responses to questions about social issues such as worries about the environment.

2.3 Interviewer Continuity and Social Desirability

In longitudinal studies, interviewer continuity is repeatedly associated with higher data quality. For instance, there is first evidence that interviewer continuity reduces panel attrition rates (e.g., Lynn et al., 2014, Campanelli and O'Muricheartaigh, 1999, Vassallo et al., 2015, Rendtel, 1995). However, few studies examine interviewer (dis)continuity effects on responses and measurement error. Hajek and Schumann (2014) and Schräpler (2004) investigate effects of interviewer changes on item nonresponse rates in income questions, revealing minor evidence for an increase in nonresponse rates when a change in interviewer occurs. Uhrig and Lynn (2008) examine effects of interviewer continuity on satisficing behavior, also finding limited evidence.

Only three studies that analyze interviewer (dis)-continuity effects on socially desirable responding in face-to-face panel surveys are known to the author. Focusing on self-reported life satisfaction, the results by Chadi (2013) point to

[1] While the sensitivity of a question is generally dependent upon its specific topic, there is evidence that questions about behaviors are less likely to be affected by socially desirable responding. As Holtgraves (2004, p. 167) notes, "But behaviors, because they allow for relatively less distortion, should result in lowered rates of socially desirable responding." This is in line with Schnell and Kreuter (2005), who show stronger interviewer effects in non-factual questions compared to factual questions, for instance, about behaviors. Furthermore, from a rational choice perspective, the provision of socially desirable (i.e., adjusted) answers rather than true answers may be associated with higher subjective costs in behavioral questions compared to attitudinal/opinion questions. While there is usually only a small chance that an adjustment (a lie) is detected when questions relate to opinions or traits, the odds are larger if questions relate to behaviors, i.e. factual information. For instance, respondents may worry about follow-up questions relating to the details of a specific behavior ("Which exact charity organization have you donated to?").

decreasing levels of life satisfaction with growing familiarity between respondent and interviewer. He argues (in line with Warren and Halpern-Manners, 2012 and Waterton and Lievesley, 1989), that respondents gain trust in their interviewers and, consequently, provide more honest answers as familiarity with an interviewer grows. In contrast, Uhrig and Lynn (2008) reveal an *increase* in social desirability bias for 6 out of 8 items tested. In line with this, Mensch and Kandel (1988) find lower reports of drug usage behavior with growing interviewer continuity. In this regard, the authors speculate that, "interviewer familiarity increases salience of normative standards and that participants respond not only in terms of their past familiarity but also in terms of their subjective expectations regarding the probability of a future encounter with the interviewer" (Mensch & Kandel, 1988, p. 100).[2]

To sum up, empirical evidence on the impact of interviewer changes and continuity on socially desirable responding is sparse and inconclusive. More generally, it is still unknown whether interviewer-related error (variance and bias) in longitudinal studies is rather stable or whether panel surveys are associated with changing levels of interviewer effects over time (Lynn and Lugtig, 2017, Sturgis et al., 2009). The three existing studies on the topic either rely on very few panel waves (Uhrig and Lynn, 2008, Mensch and Kandel, 1988) or isolated items (Chadi, 2013). Furthermore, past contributions provide only limited theoretical background in order to explain potential familiarity effects. Finally, the existing studies neglect potential confounding factors, for instance, due to the respondents' general panel experience as well as panel attrition.

This paper contributes to the existing literature in multiple ways. First, the analyses are based upon a large scale panel survey consisting of many waves, respondents, and interviewers, as well as a variety of questions with socially desirable connotations. This not only increases statistical power and the generalizability of the analyses, but it also allows for an investigation of potential long-term interviewer (dis)continuity effects. Second, the paper generally contributes to research on panel conditioning effects. As respondents' panel experience and their familiarity with their interviewer are usually highly correlated in ongoing panel surveys, inconsistencies in past findings on panel conditioning may be in part due to confounding interviewer continuity effects. Finally, the paper provides practical implications for fieldwork agencies regarding interviewer allocation in ongoing panel surveys.

3 Theoretical Background and Hypotheses

Interviewers and respondents form a social relationship during personal (face-to-face) interviews. Their relationship and social interaction has specific features (see Fowler and Mangione, 1990, Sudman and Bradburn, 1974, p. 5, Podell, 1955). In the cross-sectional context, respondents and interviewers are usually complete strangers to each other. Their social relationship may be described as professional rather than intimate and private, comparable, for instance, to the relationship between customers and salesmen or patients and doctors. In contrast, in many panel studies, the relationship between interviewers and respondents is ongoing and stable as interviewers are re-allocated to the same respondents in each wave of the panel. Just like in any other type of ongoing social relationship, characteristics of the relationship and social interaction between interviewer and respondent are expected to be subject to change. Over the course of multiple encounters, it seems reasonable to expect changes in the level of interpersonal trust, intimacy, mutual self-disclosure, or sympathy between respondent and interviewers. Hence, while usually complete strangers in their first encounter, interviewers and respondents get to know each other and may gradually move from being strangers to being acquaintances over time. These potential changes in characteristics of the social relationship and interaction between interviewer and respondent can be related to the occurrence of socially desirable responding.

There are several plausible accounts for *declining* levels of social desirability bias by growing familiarity between interviewer and respondent:

- Interviewer continuity likely contributes to the development of interpersonal trust and increases the respondents confidence in the quality and integrity of the fieldwork agency (e.g., Schräpler, 2004, Halpern-Manners et al., 2014). Respondents who trust their interviewer (and the study) are expected to give more honest, i.e., less socially desirable, answers (Waterton & Lievesley, 1989).
- Interviewer continuity likely contributes to the establishment of *rapport* between the interviewer and respondent. Rapport may be understood as a "range of positive psychological features of an interaction, including a situated sense of connection or affiliation between interactional partners, comfort, willingness to disclose or share sensitive information, motivation to please, and empathy" (Garbarski, Schaeffer, and Dykema, 2016, p. 1; see also Podell, 1955 and Hyman, 1951). Good rapport is associated with more honest answers to sensitive questions as respondents feel more willing to disclose sensitive/undesirable information (Holbrook, Green, and Krosnick, 2003, Sun, 2014, Sudman and Bradburn, 1974).
- Over repeated encounters, some respondents may feel emotionally closer and increasingly loyal toward their interviewers. Consequently, they may be more prone to act as a "good respondent," which generally includes answering survey questions truthfully (DeMaio, 1984, p. 258).

[2] Analyzing computer-assisted telephone interviews, Lipps (2007) finds no significant effect of interviewer continuity on socially desirable responding.

• Being interviewed by an acquainted interviewer contributes to the overall familiarity with the interview situation. Respondents may therefore feel less stressed, less nervous, and more comfortable during the interview, thusly being more willing to give more honest answers to sensitive questions.

• Research in social psychology on people lying in everyday life indicates that lies are less likely in interactions described as more intimate and more pleasant (DePaulo, Kashy, Kirkendol, Wyer, & Epstein, 1996). This seems plausible for the social interaction between respondent and interviewer as well. Over multiple encounters, respondents and interviewers may establish a more intimate and pleasant social relationship as they get to know each other better and develop mutual trust.

• As respondents and interviewers become increasingly familiar with each other, interviewers may increasingly deviate from a standardized interviewing behavior with chronological question-answer sequences toward a less formal and more conversational interviewing style, as it may appear more natural given an established social relationship. Dijkstra (1987) examine the impact of personal vs. formal interviewing style on accuracy and social desirability bias and find less socially desirable response behavior in the personal and less formal style (see also Van der Zouwen, Dijkstra, and Smit, 1991, Schober and Conrad, 1997 and West, Conrad, Kreuter, and Mittereder, 2016).

Building upon this, the main research hypothesis states: *Growing interviewer continuity decreases socially desirable response behavior.* In this regard, interviewer continuity is expected to provoke changes in response behavior rather than a change in actual attitudes, behaviors or traits. In other words, respondents who (initially) tend to adjust their true answers toward social norms are expected to answer more honestly as familiarity with their interviewer increases.

Consequently, a change in interviewer is expected to *promote* the occurrence of socially desirable responses:

• First, this is likely due to a lower level of trust toward a new interviewer compared to an already established social relationship between the respondent and a former interviewer.

• Second, the experience of an interviewer change itself may increase socially desirable responding. For instance, in the context of the Socio-Economic Panel (SOEP), respondents are informed to be (probably) re-contacted and re-interviewed by the same interviewer in the upcoming panel wave. Thus, experiencing a change in interviewer may also lower the perceived trust and emotional attachment toward the study and the fieldwork agency more generally; this may, consequently, cause respondents to be less prone to share sensitive or socially undesirable information.

• Finally, a change in interviewer may lower the respondents' perception of a study's professionalism and scientific seriousness, thereby decreasing the respondents' motivation to answer honestly. While survey research indicates that interviewer changes increase panel attrition rates (e.g., Lynn et al., 2014), it seems plausible to expect some respondents not only to feel less motivated to participate, but also less motivated to provide accurate and honest answers when experiencing a change in interviewer.

On the other hand, scholars also provide arguments for a contrary effect, in which growing interviewer continuity may *increase* socially desirable responding:

• While a certain level of rapport between the interviewer and respondent is associated with a higher willingness of respondents to disclose and share sensitive information (Garbarski et al., 2016), very high levels of rapport may yield the opposite effect. As Hyman (1951, p. 214) notes, "Friendliness is important, but a certain degree of businesslike formality may be superior to maximum rapport. At some point on the continuum of increasing rapport, friendliness may pass over into intimacy. Then one is no longer a stranger, and the respondent may prefer not to hurt the interviewer's feelings, or may be eager to the interviewer's opinion." This is in line with research on honesty in other types of professional relationships, such as between doctors/therapists and patients (e.g., Quilliam, 2011). Thus, with increasing levels of familiarity and rapport between the respondent and interviewer, some respondents may also experience a growing need for the interviewer's approval. In this regard, social approval by acquainted persons may be perceived as more valuable compared to social approval by complete strangers. Thus, respondents may increasingly aim to appear in a good light, for instance, by giving socially desirable responses (Sun, 2014, Mensch and Kandel, 1988, Weiss, 1968).

• In order to minimize interviewer effects, such as social desirability bias, fieldwork organizations usually train their interviewers to act in a professional and neutral way and to not express personal views, opinions or preferences (Groves et al., 2009, p. 305, Fowler and Mangione, 1990, p. 33). It seems plausible to expect interviewers will find it difficult to not express personal views or maintain a neutral interviewing behavior if they are acquainted (or befriended) with a respondent.

Thus, comparatively high levels of familiarity between respondents and interviewers may also induce a contrary effect. Investigating potential non-linear effects in the models also allows addressing this hypothesis.

4 Methods

4.1 The Socio-Economic Panel Study (SOEP)

Data is derived from the Socio-Economic Panel Study (SOEP), an ongoing longitudinal survey of households in Germany (Wagner et al., 2007). Conducted annually since 1984, the study covers a variety of topics such as household

Table 1
Reasons for Interviewer Changes in the SOEP

Reasons	Frequency (person-years)	%
Respondent moved	1,441	6.9
Interviewer retired[a]	1,466	7.0
Interviewer drop-out	11,526	55.3
Interviewer temp. drop-out	6,411	30.8
Total	20,844	100.0

[a] Interviewers age 60+ ultimately dropping out of staff.

composition, employment and family biography, health, education, personality and attitudes. In 2014, there were 27,465 individuals in 16,110 households participating in the study (Kroh, Kühne, & Siegers, 2015). SOEP version v31.1, years 1984 to 2014 (doi: 10.5684/soep.v31.1, see Socio-Economic Panel Study, 2016) is used for this paper. SOEP microdata is available for scholarly research and statistical purposes via the Research Data Center SOEP at the German Institute for Economic Research (DIW Berlin). More information on data access is provided at https://www.diw.de/en/diw_02.c.242211.en/criteria_fdz_soep.html.

Data collection in the SOEP is largely based upon personal, face-to-face interviews. Since 1998, the SOEP has been gradually replacing paper and pencil personal interviewing (PAPI) with computer-assisted personal interviewing (CAPI) as the predominant mode of data collection. For instance, in wave 2014, 68% of the ~28,000 individual interviews were interviewer-administered, with 91% of those being conducted via CAPI. Since this paper addresses interviewer continuity effects, only personal interviews are included in the analyses.

Panel stability is comparatively high in the SOEP with over 90% successful re-interviews between two consecutive panel waves (see Kroh et al., 2015) and some respondents having participated for more than three decades. On average, respondents participate 7.9 times in the SOEP (SD = 6.7, Median = 6). In 2014, the average panel experience, i.e., the total number of waves a respondent has already participated in the panel, amounts to 9.2 years (SD = 8.6, Median = 5, Minimum = 1, Maximum = 31). About 25% of all respondents in 2014 have participated for at least 15 waves.

The number of encounters (= waves) of a respondent with a single interviewer is one of the key explanatory variables. The SOEP fieldwork agency (Kantar Public) aims to maintain an ongoing relationship between interviewers and respondents. Consequently, some respondents have known their interviewers for more than thirty years. In the wave 2014, the average respondent has known his/her interviewer for about 5 years (SD = 4.9, Median = 3, Minimum = 1, Maximum = 31).

Across all waves 1984 to 2014, 23% of all respondents experienced at least one interviewer change. In 2014, about 9% of all participating respondents experienced a change in interviewer. The reasons for interviewer changes are diverse (see Table 1). In most cases, interviewer changes occur due to structural changes in the interviewer staff. In over 60% of the cases, interviewers drop-out of the staff because they quit, are being dismissed, or retire.[3] In about 31% of the cases, interviewers temporarily drop out of the staff (e.g., because they work in other studies) or are re-allocated to another area or another subset of SOEP households. Only about 7% of all interviewer changes are because the household moved.

4.2 Identifying Interviewer Conditioning Effects

In principle, different approaches are suitable for identifying interviewer continuity effects on response behavior in panel surveys. In accordance with existing research on panel conditioning effects, one may identify three main strategies (see Warren & Halpern-Manners, 2012, 506f).

Ideally, an experimental design in which interviewers are randomly allocated to respondents over time is implemented. However, for practical and financial reasons, large scale face-to-face panel studies rarely implement such "interpenetrated" survey designs. A rare exception is found in Uhrig and Lynn (2008), which investigates interviewer continuity effects based on a partially interpenetrated design in the context of three waves within the UK Understanding Society Innovation Panel.

Second, scholars have compared panel survey subpopulations, for instance, a new sample being interviewed for the first time to an already existing panel population (e.g., Toepoel, Das, and van Soest, 2009). Within the given context one may compare responses from longitudinal panel members with greater familiarity with their interviewers to panel members of new refreshment samples experiencing their first encounter with an interviewer. The main advantage here is that this approach can be implemented quite easily in many large ongoing panel surveys, such as the US General Social Survey or the SOEP (Warren & Halpern-Manners, 2012, p. 510). However, differences between the compared subpanel populations may be due to confounding factors such as nonresponse bias, other conditioning effects, or varying questionnaire content (Sturgis et al., 2009, p. 116). A further disadvantage of this approach is that the focus is on aggregated differences between different populations in time, i.e. means or shares, rather than investigating intra-individual changes in response behavior over time (Bergmann, 2015).

[3] No information is available about the exact reason why individual interviewers permanently drop-out of the staff. Thus, a proxy is used to identify the cases in which a drop-out most likely reflects an interviewer's retirement. More precisely, a drop-out was defined as retirement if an interviewer aged 60 or older ultimately drops out of the staff.

Consequently, different types of conditioning effects and potential heterogeneous effects may overlap or cancel each other out.

Third, researchers may focus on changes in responses over time in a (single) panel population (see e.g., Sturgis et al., 2009, Bergmann, 2015). In line with Chadi (2013) and Mensch and Kandel (1988), I use this approach in order to investigate interviewer continuity effects in the Socio-Economic Panel. The main advantage of this method, in comparison to the analysis of aggregated differences, is the possibility to model intra-individual variation over time. Nonetheless, this approach comes along with some disadvantages as well. First, since some respondents drop out of the panel in every wave, estimates may be biased due to panel attrition. Second, interviewer changes occur regularly, yet still comparatively rarely in the SOEP. Thus, respondents' familiarity with the interviewer is usually highly correlated with their general panel experience, likely causing estimates to be associated with greater uncertainty. Finally, interviewer changes occur not completely random. Consequently, observed changes in reporting may be not due to interviewer discontinuity effects but rather confounding factors that cause interviewer changes in the first place. These three potential issues are specifically addressed in the statistical modelling of the continuity effects in this paper.

4.3 Identifying Socially Desirable Responding

Identifying measurement error due to social desirability bias is generally challenging, with a variety of strategies applied in the existing literature. At the respondent level, social desirability, or lie, scales have been developed in order to identify a respondent's general tendency to give socially desirable answers (e.g., the Balanced Inventory of Desirable Responding (BIDR), Paulhus, 1991). On an item level, scholars compare responses to questions with varying levels of socially desirable connotations (e.g., Sudman and Bradburn, 1974). Furthermore, responses are compared to external validation data or a "gold standard" (e.g., Kroh, 2005, F. Kreuter et al., 2008, Bauman and Dent, 1982). However, external validation data is rare and for characteristics such as attitudes and opinions, no external standards exist. Thus, survey researchers usually aim to manipulate the contextual factors potentially related to the occurrence of socially desirable responding, such as the perceived privacy or anonymity of the interview situation (e.g., Warren and Halpern-Manners, 2012, Lipps, 2007, A. Jäckle, Roberts, and Lynn, 2006).

For this study, I make use of a similar approach, focusing on the impact of interviewer (dis)-continuity on response behavior in questions associated with socially desirable answer options. How does the familiarity between respondent and interviewer affect the likelihood of respondents to choose socially (un)desirable answer options?

Four criteria are used to select suitable items for this research. First, questions and items needed to include clear socially desirable or undesirable answer options. Second, questions and items were supposed to be relevant for many research applications as well as be regularly asked in existing social surveys relying on face-to-face/telephone interviewing. As a consequence, this paper does not focus on very highly sensitive items only rarely asked in personal interviews (e.g., questions on sexual practices), but rather on items that are actually part of many existing (panel) survey questionnaires. Third, the items were supposed to have shown their susceptibility to social desirability bias in past studies already.[4] And fourth, as I use a panel analysis approach focusing on within-person variation over time, items needed to be a regular part of the SOEP questionnaire, if not annually, at least across many panel waves. As interviewer changes occur fairly irregularly, this not only increases the statistical power of the analysis, but also allows for an investigation of long-term continuity effects.

There are several questions in the SOEP that match all criteria (see Table 2). While the items are of moderate sensitivity only, their answer options can be easily classified into likely socially desirable and undesirable. Thus, respondents may use them to apply impression management strategies, such as adjusting a response into a desirable direction, in order to appear in a good light in front of the interviewer. In each case, other studies already provide evidence for their susceptibility for socially desirable response behavior (see Column 6 in Table 2). To simplify the interpretation and comparison of models, all selected items were transformed into binary indicators taking the value 1 if an answer option reflects a socially desirable characteristic, and 0 otherwise.[5]

Life Satisfaction. A number of studies reveal that reported life satisfaction decreases with ongoing panel participation (e.g., Wunder, Wiencierz, Schwarze, and Küchenhoff, 2013, D'Ambrosio and Frick, 2012). Being satisfied with one's own life may be perceived as a desirable trait to report since it reflects happiness, success and strength; characteristics that are generally socially approved in many societies. In this regard, scholars argue that growing trust toward the study and interviewers increases the respondents willingness to admit lower levels of life satisfaction in panel survey interviews (Chadi, 2013). Moreover, by using SOEP data, the results by Chadi (2013) reveal an increase in reported life satisfaction

[4] Ideally, empirical data regarding the levels of 'sensitiveness' or 'desirability' of questions and answer options is available (e.g., Holbrook et al., 2003, Sudman and Bradburn, 1974). However, similar to the majority of existing studies, such data is not available in the context of this study.

[5] Descriptive statistics for the non-recoded items are displayed in Table A1 in the Appendix. As the items strongly vary in their topics, and as a consequence most likely in their level of perceived sensitivity, items were not combined into an overall index. In addition, Table A3 in the Appendix contains information about the stability and variation of responses in the selected items over time.

Table 2
Selected Items and Socially Desirable Connotations in the SOEP

Variable	No. of Waves	Answer options coded as ...		Binary indicator		Past Studies
		Socially Desirable	Neutral/Undesirable	Mean	Std. Dev.	
Life satisfaction	31	completely satisfied (9/10)	8–10 (completely dissatisfied)	0.18	0.39	Chadi (2013)
Interest in politics	30	Very interested	Somewhat Interested, Not interested	0.08	0.27	Lipps (2007)
Worries about ...						Warren and Halpern-Manners (2012)
Immigration	16	Not worried at all	A little worried Very worried	0.28	0.45	
Xenophobia	23	Very worried	A little worried Not worried at all	0.24	0.43	
Environment	31	Very worried	A little worried Not worried at all	0.33	0.47	
Peace	31	Very worried	A little worried Not worried at all	0.40	0.49	
Crime	21	Very worried	A little worried Not worried at all	0.47	0.50	

for respondents experiencing a change in interviewer. Thus, growing levels of trust toward an interviewer in an already established relationship may increase the likelihood that respondents admit that they are not perfectly satisfied with their life.

Interest in Politics. SOEP participants are asked annually whether they are interested in politics. It seems reasonable that being interested in public issues and democracy is likely to be perceived as a desirable trait. Furthermore, to many, showing interest in politics may reflect characteristics such as being well-informed, educated, or intellectual. Thus, some respondents may state that they are interested in politics even though they are not. For instance, research indicates that respondents regularly over-report their habits of reading newspapers and watching the news (Prior, 2009).

Worries about Social Issues. From its start, SOEP respondents have been asked to report their worries toward a variety of social and political issues, including, among others, immigration and crime. Some of these items incorporate a clear socially desirable connotation. For example, it seems plausible that being worried about the environment, peace, and crime is perceived as a desirable trait by many, since it reflects a caring, minding, and empathetic personality or mindset. In addition, and particularly in the German context, it is likely desirable to not think negatively about immigrants and foreigners. Thus, being worried about xenophobia and *not* being worried about immigration likely reflect desirable opinions in large parts of the German society. Warren and Halpern-Manners (2012), by using SOEP data, provide first empirical results pointing to diminishing social desirability bias in these items with growing panel experience of respondents. For this analysis, I examine socially desirable answering behavior in self-reported worries toward immigration, xenophobia, the environment, peace, and crime.

4.4 Analysis and Models

A subpopulation of the SOEP was used for this analysis. First, only personal interviews with adults including both computer-assisted (CAPI) and paper-and-pencil (PAPI) interviewer assisted interviews are taken into account. Second, as this paper focuses on interviewer (dis)continuity effects, respondents who never experienced a change in interviewer were excluded from analysis. For this group of respondents, interviewer continuity effects could have not been separated from general panel experience effects. This was true for about 77% of all SOEP respondents who have participated in personal interviews. The final sample for analysis consists of 15,666 individual respondents with a total of 152,437 observations (person-years) between 1984 and 2014. As not all selected questions were asked in every wave, sample sizes vary across models and range from 12,633 for the worries about immigration to 15,666 for the self-reported life satisfaction.

Logistic panel fixed effects regressions analysis is used to assess the effects of interviewer changes and interviewer-respondent familiarity on socially desirable responding. Using respondent-level fixed effects models allows to assess differences in socially desirable responding by changes of

interviewer-familiarity within respondents over time. Moreover, fixed effects models automatically control for all time-invariant variables and potential confounders ("unobserved heterogeneity").

Within the logistic fixed-effect model, the probability of a respondent i at time (wave) t to choose an answer option with socially desirable connotations is a logistic function (see e.g., Allison, 2009, Roßmann, 2015):

$$\log\left(\frac{p_{it}}{1-p_{it}}\right) = \alpha_i + \beta_1 X_{1_{it}} + \ldots + \beta_n X_{n_{it}} + \epsilon_{it}, \quad t = 1, 2, \ldots, T \tag{1}$$

of an individual specific (fixed) intercept α_i, a set of time-variant independent variables $X_{1_{it}} ... X_{n_{it}}$ and coefficients $\beta_1 .. \beta_n$ as well as a residual error term ϵ_{it}.

A single model is estimated for each of the items. The binary indicators (see Table 2) are used as dependent variables, taking the value 1 if an answer option reflects a socially desirable connotation and 0 otherwise.

Two sets of models are estimated. In the first set, the focus is on effects of the occurrence of interviewer changes (yes vs. no) on responses. Are respondents more likely to choose a desirable answer option in cases where they experience a change in interviewer? In a second set, effects of different levels of interviewer-respondent familiarity on responses are estimated: Are respondents less likely to choose desirable answer options with their growing familiarity with an interviewer over time?

In order to address potential time-variant confounders, a set of control variables are included in all models. The most relevant rivalling account is the general panel experience of respondents, which is also highly correlated with the familiarity of respondents and interviewers (r=0.71). Therefore, the respondents panel experience – the total number of waves a respondent has already participated in the SOEP in a given year – is included in the models. In this regard, another possible confounder relates to overall changes of attitudes and traits in the underlying population over time, and – although unlikely – these changes being correlated with interviewer continuity. Thus, wave-specific dummies are added in order to control for these potential period effects. Furthermore, estimates may be confounded by panel attrition. For instance, respondents who are generally less motivated to participate in the survey may also be less motivated to provide honest and accurate answers. Using fixed-effect models, all time-constant respondent characteristics possibly influencing both panel participation and socially desirable response behavior are controlled for automatically. In addition, the respondents' prospective number of years of participation is included in the models. This measure captures whether socially desirable responding for given levels of interviewer-respondent familiarity is lower in respondents who will refuse to participate in the following wave of the survey compared to those who will participate in the survey for several more years. At a respondent level, an obvious potential confounder is the occurrence of household moves. Moves may not only initiate a change in interviewer, but also affect respondents' attitudes and opinions toward political and social issues, as these are likely to be correlated with geographical/area characteristics. A dummy variable for household moves in a given wave (yes vs. no) is therefore included into the models. Finally, changes in reporting may be due to systematic differences between the new and the previous interviewer rather than changing levels of familiarity. For instance, Annette Jäckle, Lynn, Sinibaldi, and Tipping (2013) find that experienced and skilled interviewers achieve higher co-operation rates. Thus, the interviewers' work experience as well as their response rate and number of contact attempts in each given wave are included as control variables. Descriptive statistics for all control variables are displayed in Table A2 in the Appendix.

5 Results

5.1 Interviewer Changes and Socially Desirable Responding

First, the effect of the occurrence of an interviewer change (yes vs. no) on the likelihood of choosing socially desirable answer options is estimated using multivariate panel fixed effects logistic regressions for each of the 7 items. Table 3 displays odds ratios as well as their corresponding significance levels and standard errors. The respondents' general panel experience is collapsed into broader categories in order to observe potential non-linear effects and declining marginal growth rates. As additional controls, wave fixed effects, the prospective panel participation, household moves (y/n), as well as interviewer experience, wave-specific response-rate, and mean household contacts within a wave are included into the models. For the sake of clarity, coefficients for these controls are not displayed in this table (see Table A4 in the Appendix).

In line with the hypothesis, the results reveal a significant increase in socially desirable responses in cases where respondents experience an interviewer change for 6 out of the 7 items tested. Respondents experiencing an interviewer change in a given wave were more likely to report a very high life satisfaction (9/10 out of 0-10 scale, OR = 1.10 , $p < .001$), being very interested in politics (OR = 1.30, $p < .001$), and being worried about xenophobia (OR = 1.12, $p < .001$), the environment (OR = 1.21, $p < .001$), peace (OR = 1.26, $p < .001$) and crime (OR = 1.12, $p < .001$). The results may be interpreted as reflecting a decrease in the respondents' trust toward the study and the interviewer, resulting in a reduced willingness to disclose socially undesirable opinions and traits. Moreover, as respondents may experience an increasing need to appear in a good light in front of the new and unknown interviewer, they are more likely to

Table 3
Effects of Interviewer Changes on Socially Desirable Response Propensities

	Model 1 Life Sat.	Model 2 Politics	Model 3 Immigr.	Model 4 Xenoph.	Model 5 Environm.	Model 6 Peace	Model 7 Crime
Interviewer Change (Ref.: No)							
Yes	1.10***	1.30***	1.01	1.12***	1.21***	1.26***	1.12***
	(0.03)	(0.06)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
Panel Experience (Ref.: 1st year)							
2 years	0.81***	1.04	1.05	0.76***	0.83***	0.83***	0.79***
	(0.04)	(0.09)	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)
3-4 years	0.65***	0.88	0.93	0.80***	0.86***	0.79***	0.80***
	(0.03)	(0.07)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)
5-9 years	0.62***	0.88	0.94	0.73***	0.82***	0.76***	0.87**
	(0.03)	(0.07)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)
10+ years	0.61***	0.93	0.89	0.83**	0.87**	0.70***	0.79***
	(0.04)	(0.09)	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)
$n_{resp.} \times t$	91,916	38,792	58,097	73,919	115,020	123,336	86,953
$n_{resp.}$	8,152	3,104	6,204	7,285	9,496	10,476	8,112

Panel Fixed Effects Logistic Regressions, Odds Ratios (Standard Errors in Parentheses). Controls: Wave fixed effects; Prospective panel participation; Move (y/n); Interviewer experience, response-rate, and mean household contacts within wave.
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

choose socially desirable answer options. No effect was observed for the question relating to the respondents' worries about immigration.

Interpreting the observed effect sizes (odds ratios), the results reveal a substantial impact of interviewer changes on response behavior. For instance, respondents experiencing an interviewer change are associated with a 30% increase in reporting an interest in politics compared to waves in which they do not experience a change in interviewer. Specifically, this strong effect appears to be even more relevant given the fact that political interest is shown to be a highly stable trait over the life-cycle (see Table A3 in the Appendix; see also Prior, 2010). More generally, the observed results clearly point to a large potential impact of contextual characteristics of the survey interview situation on actual response behavior.

While interviewer changes increase the tendency of respondents to provide desirable answer options, a contrary effect is observed for the respondents' general panel experience. In line with the existing literature on the topic, respondents are less likely to provide desirable answers in their further participation after wave one. However, no significant effect is observed for the interest in politics and, again, for worries about immigration. Interestingly, effect sizes do not substantially increase with the ongoing panel participation. Thus, respondents seem to change their response behavior mainly during their second panel participation.

5.2 Interviewer Continuity and Socially Desirable Responding

While the previous models estimates the effect of a change in interviewer between two consecutive waves, the next set of models focuses on the effect of growing interviewer continuity, i.e., the growing familiarity with individual interviewers over time. This analysis allows observing long-term, as well as non-linear, familiarity effects. If respondents indeed gain trust in their interviewers and feel increasingly comfortable to answer truthfully, socially desirable responding should decrease with interviewer continuity.

Table 4 displays the results of seven multivariate panel fixed effects logistic regressions. The same set of controls are included in each model. Again, for brevity, coefficients for these controls are not displayed in this table (see Table A5 in the Appendix). To simplify interpretation and to capture possible non-linear effects, both interviewer continuity and the respondents' panel experience are collapsed into broader categories.

In line with the hypothesis, the results reveal significant negative effects of interviewer-respondent familiarity on the occurrence of socially desirable responding, again, in 6 out of the 7 items tested. Respondents who had a chance to become familiar with their interviewer are less likely to choose desirable answer options compared to their first encounter with a given interviewer. This is true for the respondents' satisfaction with their life, their interest in politics, as well as

Table 4
Effects of Interviewer Continuity on Socially Desirable Response Propensities

	Model 1 Life Sat.	Model 2 Politics	Model 3 Immigr.	Model 4 Xenoph.	Model 5 Environm.	Model 6 Peace	Model 7 Crime
Interviewer Continuity (Ref.: 1st year)							
2 years	0.97	0.85***	0.98	0.93*	0.88***	0.89***	0.89***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)
3-4 years	0.98	0.79***	1.06	0.88***	0.80***	0.81***	0.88***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)
5-9 years	0.88***	0.69***	0.99	0.77***	0.80***	0.74***	0.90**
	(0.03)	(0.04)	(0.04)	(0.04)	(0.02)	(0.02)	(0.03)
10+ years	0.82***	0.68***	0.93	0.81***	0.75***	0.69***	0.94
	(0.04)	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)	(0.04)
Panel Experience (Ref.: 1st year)							
2 years	0.85***	1.24*	1.07	0.82***	0.94	0.96	0.88*
	(0.04)	(0.11)	(0.07)	(0.05)	(0.05)	(0.04)	(0.05)
3-4 years	0.68***	1.10	0.91	0.89*	1.04	0.96	0.91
	(0.03)	(0.09)	(0.05)	(0.05)	(0.05)	().04	(0.04)
5-9 years	0.67***	1.16	0.93	0.85**	1.00	0.96	0.98
	(0.03)	(0.10)	(0.05)	(0.04)	(0.05)	(0.04)	(0.05)
10+ years	0.67***	1.24*	0.91	0.97	1.07	0.89*	0.89*
	(0.04)	(0.12)	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)
$n_{resp.} \times t$	91,916	38,792	58,097	73,919	115,016	123,336	86,953
$n_{resp.}$	8,152	3,104	6,204	7,285	9,495	10,476	8,112

Panel Fixed Effects Logistic Regressions, Odds Ratios (Standard Errors in Parentheses). Controls: Wave fixed effects; Prospective panel participation; Move (y/n); Interviewer experience, response-rate, and mean household contacts within wave.
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

worries toward xenophobia, environment, peace, and crime.

The strongest effects, again, are observed for the question about the respondents' interest in politics. The more familiar a respondent is with his or her interviewer, the less likely he/she reports being highly interested in politics. A similar monotonically increasing pattern is also observed for Models 4, 5, and 6 relating to worries toward xenophobia, the environment, and peace. In other words, respondents seem to be increasingly honest in their answers, i.e., more likely to provide socially *un*desirable answers, the more familiar they are with their interviewers. Relating this results to the observed effects in the previous models on interviewer changes, it indeed seems to be familiarity with the interviewer that affects the answering behavior rather than potential confounding factors, such as household moves simultaneously causing interviewer changes *and* affecting response behavior.

While a consistent and increasing familiarity effect is observed, Models 2, 4, 5, and 6 also reveal declining marginal growth rates of the effect. In other words, and this seems plausible, growing familiarity with the interviewer at some point in time does not seem to *further* affect response behavior in socially desirable questions. No significant continuity effect was observed, again, on the respondents' worries toward immigration (Model 3). In Model 1, relating to the respondents' reported life satisfaction, all coefficients point to a decrease in socially desirable response behavior, however, only the two top category coefficients (5-9 years and 10+ years) are statistically significant.

Contrary to what was expected, the effect of the general panel experience of respondents is rather inconsistent and in cases where it reaches statistical significance, even weaker compared to the interviewer continuity (Model 4 and 7). An exception is the strong negative effect of panel experience on life satisfaction, matching the existing literature (see Chadi, 2013).

To sum up, the results point to the fact that – at least in the context of socially desirable responding – it seems to be the familiarity with the interviewer, rather than the familiarity with the survey, that causes changes in reporting, i.e., conditioning effects, in panel surveys. Therefore, estimates of panel conditioning effects on social desirability bias in past studies are potentially confounded by the familiarity between the respondent and the interviewer. This points to the necessity of taking into account the continuity of the relation-

ship between respondents and interviewers when investigating panel conditioning effects in longitudinal studies.

5.3 Further Robustness Checks

Scholars note that satisficing behavior is a potential heuristic process underlying social desirability bias (Holtgraves, 2004). Some respondents may apply shortcutting strategies such as choosing the same answer option for many items (straightlining) or simply choosing the first answer option (acquiescence behavior) in order to reduce the cognitive burden of responding to survey questions (see Krosnick, 1991). Moreover, respondents may generally change their response style over time, for instance, by increasingly avoiding extreme answer options. Although the results of Kaminska and Foulsham (2016) and Kaminska and Foulsham (2013) do not provide evidence for satisficing heuristics underlying social desirability bias, a set of control indicators that relate to potential satisficing response behavior (and other undesired response behaviors) are added to the models.

First, an indicator for straightlining response behavior was included, reflecting whether respondents tend to choose the same/similar answer option for many consecutive items or not. In this regard, responses to a total of eight sets of items were analyzed covering topics such as health, personality and subjective-well being. I coded straightlining as the relative absence of variation in these sets of items. A threshold was used so that about 25% of all respondents were assigned to the straightlining group (= 1) while the rest was not (= 0). Hence, the indicator does not reflect a strict straightlining behavior (all answers identical) but the relative tendency of respondents to choose similar answer options for many items. In addition, a categorical response style indicator was constructed that reflect whether a respondent generally tends to a) choose extreme answer options (i.e., at the beginning and end of scales), b) answer options centred around middle categories, or c) none of the above. Again, a large number of responses to item scales were analyzed for each individual and wave. More precisely, the indicator reflects how individual respondents score across the scales compared to the sample average latent factors across the respective scales. Finally, an individual's overall item nonresponse rate in a given wave was included.

Adding these controls into the models does not substantially change the results (see Table A6 in the Appendix).

Moreover, the analyses were replicated using items that are comparable in type and form but yield no or less socially desirable connotations. More precisely, familiarity effects are estimated on the respondents' reported satisfaction with their health, worries about their own health, worries about their own economic situation, and worries about the general economic situation. No significant effects of interviewer (dis)continuity on responses are observed for any of these items (see Table A7 in the Appendix).

6 Discussion

This paper investigates interviewer continuity effects on socially desirable responding in panel surveys. While existing research mainly focuses on the (beneficial) effects of interviewer continuity on participation rates, the present study expands knowledge on (dis)continuity effects on actual response behavior and measurement error. The study draws on data from 31 waves of the German Socio-Economic Panel (SOEP), including a large number of respondents and interviewers as well as a variety of questions associated with socially desirable connotations. By additionally controlling for potential confounding factors, the unique database allows for an in-depth analysis of short-term and long-term interviewer continuity effects in longitudinal studies.

The results indicate that respondents provide more honest, i.e., less socially desirable, answers with the growing familiarity with their interviewers. Based on respondent-level fixed effects models, a consistent effect is observed for questions relating to traits and opinions, such as self-reported life satisfaction, interest in politics, as well as worries about immigration, xenophobia, peace, the environment, and crime. In line with this, the occurrence of an interviewer change is associated with an increase in answers with socially desirable connotations. Interpreting these results, as respondents gain trust with their interviewer and the study, they may feel increasingly comfortable during their interview and, consequently, are motivated to answer more truthfully. Thus, repeatedly allocating the same interviewer to respondents in longitudinal surveys – the default procedure in many ongoing panel surveys – may not only maintain high panel participation rates but also contribute to the quality of the data collected by reducing respondents' tendencies to provide socially desirable responses.

While existing research on conditioning effects in panel surveys almost exclusively reveals time-in-panel effects that threaten the quality of the data collected, this paper provides evidence for one of the rarely documented advantageous (interviewer) conditioning effects on data quality in panel surveys (see also Kroh, Winter, and Schupp, 2016). Surprisingly, the effect of the respondents' general panel experience on socially desirable responding remains unsystematic and (often) not statistically significant once the interviewer-respondent familiarity is included in the models. This points strongly to the importance of controlling for the familiarity between respondents and interviewers when investigating panel conditioning effects on measurement error in panel surveys. Thus, observed panel conditioning effects in past studies may not be due to growing panel experience of respondents but due to correlated familiarity between respondents and interviewers.

This study also faces limitations and leaves room for further research. First, empirical data on the level of social desirability connotations is not available within the SOEP con-

text and items were selected based on results in past studies. Given the questions' topics and contextual social norms, the selected items are of moderate sensitivity only. Thus, other continuity effects may be observed for highly sensitive items, for instance, in questions about drug abuse or criminal activities. Even though highly sensitive items are only rarely implemented in face-to-face interviewing modes, further research should incorporate information on social desirability connotations and investigate whether effects vary over different levels of item sensitivity.

Second, interviewer changes are not entirely random in the SOEP, thereby potentially threatening the generalizability of the results. While some reasons for interviewer changes, such as the retirement of the interviewer, are expected to be comparatively randomly distributed over participating households, others may not (interviewers quit, households move). In this regard, however, replicating the analyses using only interviewer changes due to retirement and temporary drop-outs show highly comparable results. Nonetheless, and although I control for household moves and the interviewers' work experience and performance, the results may not be entirely generalizable to the underlying population since respondents experiencing a change in interviewer may still be – in principle – systematically different from those who do not. Thus, restricting the sample to respondents who have experienced a change in interviewer potentially introduces (nonresponse) bias into the familiarity effect estimates. Respondents who decide to no longer participate in the panel due to a change in interviewer may be systematically different from those who continue to take part. For instance, those who decide to drop out of the panel may be particularly in need of a stable relationship with their interviewer in order to provide honest answers to questions perceived as sensitive. For these respondents, effects are expected to be even stronger and, consequently, the given approach would underestimate the size of the familiarity effect.

Finally, changes in answers over time may not only be due to changes in reporting – the focus of this paper – but due to real changes as well. Even though real changes in behavior, attitudes/opinions or traits due to growing familiarity with the interviewer seem implausible, the given research design does not allow for empirically testing this.

This paper provides first comprehensive evidence for interviewer continuity effects on response quality in panel surveys. Clearly, more research is needed to unequivocally identify the exact causal processes underlying the observed effects. For instance: What kind of respondent and interviewer characteristics (e.g., personality traits) as well as their interactions contribute to the occurrence of interviewer continuity effects? Which exact underlying mechanisms associated with interviewer changes affect respondents answering behavior? Future research may make use audio- and video-recordings for behavior and interaction coding purposes. This would allow evaluating whether interviewers are more likely to apply conversational interviewing behavior (rather than standardized behavior) as familiarity with their respondents grows. Moreover, future studies may collect data on the respondents' and interviewers' interpersonal perceptions and ratings of the social relationship in order to observe changes over time. Finally, further research should investigate familiarity effects on other types of items associated with socially desirable connotations such as questions about sensitive behaviors.

For currently ongoing panel surveys, the results point to the fact that allocating the same interviewers to households has a positive effect not only on participation propensities, but also on measurement and responses. Thus, sending the same interviewers not only reduces survey costs, but also likely contributes to the quality of the data collected.

Acknowledgements

The author would like to especially thank Martin Kroh and Zeina Mneimneh for their feedback on early versions of this paper. The majority of work on this paper was conducted during the author's time as a research associate at the Socio-Economic Panel Study (SOEP) at DIW Berlin. This research was supported by the RECLINK project funded by the Leibniz Association. The author gratefully acknowledges support from the SOEP and the Berlin Graduate School of Social Sciences (BGSS) at Humboldt-Universität zu Berlin.

References

Allison, P. D. (2009). *Fixed effects regression models*. New Jersey, NJ: Wiley.

Bauman, K. & Dent, C. (1982). Influence of an objective measure on self-reports of behavior. *Journal of Applied Psychology*, *67*(5), 623–628.

Bergmann, M. (2015). *Panel Conditioning. Wirkungsmechanismen und Konsequenzen wiederholter Befragungen*. Baden-Baden: Nomos.

Buck, N. & McFall, S. (2012). Understanding society: design overview. *Longitudinal and Life Course Studies*, *3*(1), 5–17.

Campanelli, P. & O'Muricheartaigh, C. (1999). Interviewers, interviewer continuity, and panel survey nonresponse. *Quality & Quantity*, *33*, 59–76.

Chadi, A. (2013). The role of interviewer encounters in panel responses on life satisfaction. *Economics Letters*, *121*(3), 550–554.

Crowne, D. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *J Consult Clin Psychol*, *24*, 349–354.

D'Ambrosio, C. & Frick, J. (2012). Individual wellbeing in a dynamic perspective. *Economica*, *79*, 284–302.

DeMaio, T. (1984). Social desirability and survey measurement. A review. In C. Turner & E. Martin (Eds.), *Surveying subjective phenomena (vol. 2)* (pp. 257–281). New York: Russell Sage Foundation.

DePaulo, B., Kashy, D., Kirkendol, S., Wyer, M., & Epstein, J. (1996). Lying in everyday life. *Journal of Personality and Social Psychology*, *70*(5), 979–995.

Dijkstra, W. (1987). Interviewing style and respondent behavior. An experimental study of the survey-interview. *Sociological Methods & Research*, *16*(2), 309–334.

Fowler, F. J. & Mangione, T. W. (1990). *Standardized survey interviewing: minimizing interviewer-related error*. Newbury Park, CA: Sage.

Garbarski, D., Schaeffer, N. C., & Dykema, J. (2016). Interviewing practices, conversational practices, and rapport: responsiveness and engagement in the standardized survey interview. *Sociological Methodolgy*, *46*(1), 1–38.

Groves, R. M., Fowler Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Hoboken, NJ: Wiley.

Hajek, K. & Schumann, N. (2014). *Continuity trumps? The impact of interviewer change on item nonresponse.* Handout prepared for the Panel Survey Methodology Workshop 2014, Ann Arbor, Michigan, U.S.

Halpern-Manners, A., Warren, J., & Torche, F. (2014). Panel conditioning in a longitudinal study of illicit behaviors. *Public Opinion Quarterly*, *78*(3), 565–590.

Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability sample with long questionnaires. Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly*, *67*, 79–125.

Holtgraves, T. (2004). Social desirability and self-reports: testing models of socially desirable responding. *Personality and Social Psychology Bulletin*, *30*, 161–172.

Hyman, H. (1951). Interviewing as a scientific procedure. In D. Lerner & H. Lasswell (Eds.), *The policy sciences*. Stanford, CA: Stanford University Press.

Jäckle, A. [A.], Roberts, C., & Lynn, P. (2006). Telephone versus face-to-face interviewing: mode effects on data quality and likely causes. report on phase ii of the ess-gallup mixed mode methodology project. ISER Working Paper No. 2006-41. ISER Institute for Social & Economic Research, Essex UK.

Jäckle, A. [Annette], Lynn, P., Sinibaldi, J., & Tipping, S. (2013). The effect of interviewer experience, attitudes, personality and skills on respondent co-operation with face-to-face surveys. *Survey Research Methods*, *7*(1), 1–15.

Kaminska, O. & Foulsham, T. (2013). Understanding sources of social desirability bias in different modes: evidence from eye-tracking. ISER Working Paper No. 2013-04. Institute for Social & Economic Research (ISER), Essex UK.

Kaminska, O. & Foulsham, T. (2016). Eye-tracking social desirability bias. *Bulletin de Méthodologie Sociologique*, *130*, 73–89.

Kane, E. & Macaulay, J. (1993). Interviewer gender and gender attitudes. *Public Opinion Quarterly*, *57*, 1–28.

Kreuter, F. [F.], Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys. *Public Opinion Quarterly*, *72*(5), 847–865.

Kroh, M. (2005). Intervieweffekte bei der Erhebung des Körpergewichts in Bevölkerungsumfragen. *Gesundheitswesen*, *67*, 646–655.

Kroh, M., Kühne, S., & Siegers, R. (2015). Documentation of sample sizes and panel attrition in the german socio-economic panel (soep) (1984 until 2014). SOEP Survey Papers No. 297. Berlin: DIW Berlin.

Kroh, M., Winter, F., & Schupp, J. (2016). Using person-fit measures to assess the impact of panel conditioning on reliability. *Public Opinion Quarterly*, *80*(4), 914–942.

Krosnick, J. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*, 213–236.

Krumpal, I. (2013). Determinants of social desirability bias in sensitve surveys: a literature review. *Quality & Quantity*, *47*, 2025–2047.

Lee, Z. & Woodlife, L. (2010). Donor misreporting: conceptualizing social desirability bias in giving surveys. *Voluntas: International Journal of Voluntary and Nonprofit Organizations*, *21*(4), 569–587.

Lipps, O. (2007). Interviewer and respondent survey quality effects in a cati panel. *Bulletin de méthodologie sociologique*, *95*, 1–16.

Lynn, P., Kaminska, O., & Goldstein, H. (2014). Panel attrition: how important is interviewer continuity? *Journal of Official Statistics*, *30*(3), 443–457.

Lynn, P. & Lugtig, P. (2017). Total survey error for longitudinal surveys. In P. P. Biemer, E. D. De Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. E. Lyberg, ... B. T. West (Eds.), *Total survey error in practice: improving quality in the era of big data*. New Jersey, NJ: Wiley.

Mensch, B. S. & Kandel, D. B. (1988). Underreporting of substance use in a National Longitudinal Youth Cohort: individual and interviewer effects. *Public Opinion Quarterly*, *52*(1), 100–124.

Mitchell, M., Collings, D., & Brown, A. (2015). Factors affaffect participation in understanding society: qualitative study with panel mmember. Understanding Society Working Paper Series No. 2015-04. ISER Institute for Social and Economic Research, Essex UK.

Nancarrow, C. & Cartwright, T. (2007). Online access panels and tracking research. The conditioning issue. *International Journal of Market Research*, *49*(5), 573–595.

Paulhus, D. (1991). Measurement and control of response bias. In J. Robinson, P. Shaver, & L. Wrightman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). New York: Academic Press.

Podell, L. (1955). The structured interview as a social relationship. *Social Forces*, *34*(2), 150–155.

Prior, M. (2009). The immensely inflated news audience: assessing bias in self-reported news exposure. *Public Opinion Quarterly*, *73*(1), 130–143.

Prior, M. (2010). You've either got it or you don't? The stability of political interest over the life cycle. *The Journal of Politics*, *72*(3), 747–766.

Quilliam, S. (2011). 'the cringe report': why patients don't dare ask questions, and what we can do about that. *J Fam Plann Reprod Health Care*, *37*(2), 110–112.

Rendtel, U. (1995). *Lebenslagen im Wandel: Panelausfälle und Panelrepräsentativität*. Frankfurt/Main: Campus Verlag.

Roßmann, J. (2015). *Satisficing in Befragungen. Theorie, Messung und Erklärung* (Doctoral dissertation, Universität Mannheim, Germany).

Schaeffer, N. C. (2000). Asking questions about threatening topics: a selective overview. In A. A. Stone, J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The science of self-report: implications for research and practice*. London: Lawrence Erlbaum Associates.

Schnell, R. & Kreuter, F. [Frauke]. (2005). Separating interviewer and sampling-point effects. *Journal of Official Statistics*, *21*(3), 389–410.

Schober, M. F. & Conrad, F. G. (1997). Does conversational interviewing reduce survey measurement error? *Public Opinion Quarterly*, *61*, 576–602.

Schräpler, J.-P. (2002). Respondent behavior in panel studies – a case study for income-nonresponse by means of the german socio-economic panel (gsoep). Discussion Paper No. 299. Berlin: DIW Berlin.

Schräpler, J.-P. (2004). Respondent behavior in panel studies. *Sociological Methods & Research*, *33*(1).

Schuman, H. & Converse, J. (1971). The effects of black and white interviewers on black responses in 1968. *Public Opinion Quarterly*, *35*(1), 44–68.

Socio-Economic Panel Study. (2016). Data for years 1984-2014, version 31.1, Socio-Economic Panel Study, 2016, doi:10.5684/soep.v31.1. SOEP/DIW Berlin. https://www.diw.de/de/diw_01.c.529324.de/soep_v31_1.html.

Sturgis, P., Allum, N., & Brunton-Smith, I. (2009). Attitudes over time: the psychology of panel conditioning. In P. Lynn (Ed.), *Methodology of longitudinal surveys*. New Jersey, NJ: Wiley.

Sudman, S. & Bradburn, N. (1974). *Response effects in surveys*. Chicago: Aldine Publishing.

Sun, H. (2014). *Rapport and its impact on the disclosure of sensitive information in standardized interviews* (Doctoral dissertation, Joint Program in Survey Methodology, University of Maryland, USA).

Toepoel, V., Das, M., & van Soest, A. (2009). Relating question type to panel conditioning: comparing trained and fresh respondents. *Survey Research Methods*, *3*(2), 73–80.

Tourangeau, R., Rips, L., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.

Tourangeau, R. & Yan, T. (2007). Sensitve questions in surveys. *Psychological Bulletin*, *133*(5), 859–883.

Uhrig, N. (2012). Understanding panel conditioning: an examination of social desirability bias in self-reported height and weight in panel surveys using experimental data. *Longitudinal and Life Course Studies*, *3*(1), 120–136.

Uhrig, N. & Lynn, P. (2008). The effect of interviewer continuity on measurement error in panel surveys. Presentation. Institute for Social and Economic Research (ISER), Essex, UK. Retrieved from http://panelsurveymethods.files.wordpress.com/2013/03/2008uhriglynn.pdf

Van der Zouwen, J., Dijkstra, W., & Smit, J. H. (1991). Studying respondent–interviewer interaction: the relationship between interviewing style, interviewer behavior, and response behavior. In P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys*. New Jersey, NJ: Wiley.

Vassallo, R., Durrant, G. B., Smith, P. W., & Goldstein, H. (2015). Interviewer effects on non-response prepensitiy in longitudinal surveys: a multilevel modeling approach. *Journal of the Royal Statistical Society, Series A*, *178*(1), 83–99.

Wagner, G. G., Frick, J. R., & Schupp, J. (2007). The german socio-economic panel study (soep) – scope, evaluation and enhancements. SOEPpapers No. 1. Berlin: DIW Berlin.

Warren, J. R. & Halpern-Manners, A. (2012). Panel conditioning in longitudinal social science surveys. *Sociological Methods & Research*, *41*, 491–534.

Waterton, J. & Lievesley, D. (1989). Evidence of conditioning effects in the british social attitudes panel. In D. Kasprzyk, G. Duncan, G. Kalton, & M. Singh (Eds.), *Panel surveys* (pp. 319–339). New York: Wiley.

Watson, N. & Wooden, M. (2009). Identifying factors affecting longitudinal survey response. In P. Lynn (Ed.), *Methodology of longitudinal surveys*. Chichester (UK): Wiley.

Weiss, C. (1968). Validity of welfare mothers' interview responses. *Public Opinion Quarterly*, *32*(4), 622–633.

West, B. T. & Blom, A. G. (2017). Explaining interviewer effects: a research synthesis. *Journal of Survey Statistics and Methodology*, *5*(2), 175–211.

West, B. T., Conrad, F. G., Kreuter, F., & Mittereder, F. (2016). Can conversational interviewing improve survey response quality without increasing interviewer effects? *Journal of the Royal Statistical Society A*, *Online first. DOI: 10.1111/rssa.12255*.

Wunder, C., Wiencierz, A., Schwarze, J., & Küchenhoff, H. (2013). Well-being over the life span: semiparametrice evidence from british and german longitudinal data. *Review of Economics and Statistics*, *95*, 154–167.

Appendix
Tables

Table A1
Descriptive Statistics of Dependent Variables (Pre-Recoding) in the Respondent-Subgroup for Analysis

	Mean/ Proportion	- Std. Dev.	Min	- Max	- Person-years	Respondents
Life Satisfaction	7.10	1.78	0	10	151, 673	15, 665
Interest in Politics	2.72	0.83	1	4	142, 830	15, 505
Worries about ...						
Immigration	1.98	0.74	1	3	90, 840	12, 633
Xenophobia	1.99	0.70	1	3	103, 553	13, 258
Environment	1.79	0.64	1	3	145, 861	15, 505
Peace	1.74	0.69	1	3	145, 871	15, 505
Crime	1.63	0.67	1	3	110, 590	13, 664

Table A2
Descriptive Statistics of Explanatory Variables in the Respondent-Subgroup for Analysis

	-		Range	
	Mean/Prop.	Std. Dev.	Min	Max
Interviewer Change	0.14	0.34	0	1
Interviewer Familiarity	4.77	4.52	1	30
Panel Experience	9.07	7.07	1	31
Household Move	0.03	0.17	0	1
Interv. Response Rate = High	0.33	0.47	0	1
Interv. Contact Rate = High	0.22	0.41	0	1
Item Nonresponse Rate	0.02	0.03	0	0.70
Straighlining Behavior	0.08	0.27	0	1
Response Style	1.12	0.49	1	3

15,666 individuals with 152,437 observations.

Table A3
Stability and Variation of Responses over Time

		Observations		Between		Within	Individuals
Binary Variable	Value	n	%	n	%	%	n
Life satisfaction	0	124,194	82	15137	97	83	15,665
	1	27,475	18	8,860	55	37	-
Interest in politics	0	131,509	92	15,199	98[a]	94	15,504
	1	11,317	8	3,409	22	36	-
Worries about ...	-		-		-		
Immigration	0	65,310	72	11,309	90	78	12,619
	1	25,303	28	7,319	61	49	-
Xenophobia	0	78,348	76	12,490	94	80	13,250
	1	25,205	24	8,045	61	41	-
Environment	0	97,411	67	14,223	92	73	15,494
	1	48,443	33	10,766	70	47	-
Peace	0	86,933	60	13,857	90	67	15,494
	1	58,934	40	12,113	78	50	-
Crime	0	58,372	53	11,529	84	65	13,653
	1	52,218	47	10,236	75	60	-

Table A3 contains information about the distribution, stability, and change rates of the 7 dependent binary variables, thereby providing information about a) the general tendency of respondents to choose answer options with socially desirable connotations, and b) whether there is substantial variation within respondents over time, i.e., whether or not respondents choose desirable answer options in some waves, but not in others.

Column 3 (overall) lists the total number of observations (person-years) in which a socially desirable response (= 1) or a neutral/undesirable response (= 0) was given. For instance, in 18% of all observations, respondents have reported high levels of life satisfaction (binary indicator = 1). In addition, column 4 (between) reflects the total share of individuals, who ever provided a neutral/undesirable category (= 0) and a socially desirable category respectively (= 1). For example, 98% of all respondents in the sub-sample have ever reported a low/neutral interest in politics (binary indicator = 0), while 22% ever reported a strong interest in politics (binary indicator = 1). Finally, column 5 lists the within-person change rates for the binary outcomes, reflecting the (in)stability of responses within individuals over time. For example, respondents who ever answered that they do not worry about immigration (indicator = 0) did so in 78% of all their responses. A greater variation is observed for those respondents ever reported to worry about immigration, saying so in about 49% of their observations only. To sum up, while there are differences in response stability across items, all selected items show substantial variation in individual response behavior over time. In other words, many respondents select socially (un)desirable answer options in some waves, but not in others.

Table A4
Effects of Interviewer Changes: Panel Fixed Effects Logistic Regressions, Odds Ratios (Standard Errors in Parentheses)

	Life Sat.	Politics	Immigr.	Xenoph.	Environm.	Peace	Crime
Interviewer Change (Ref.: No)							
Yes	1.10	1.30	1.01	1.12	1.21	1.26	1.12
	(0.03)	(0.06)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
Panel (Ref.: 1st year)							
2 years	0.81	1.04	1.05	0.76	0.83	0.83	0.79
	(0.04)	(0.09)	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)
3-4 years	0.65	0.88	0.93	0.80	0.86	0.79	0.80
	(0.03)	(0.07)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)
5-9 years	0.62	0.88	0.94	0.73	0.82	0.76	0.87
	(0.03)	(0.07)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)
10+ years	0.61	0.93	0.89	0.83	0.87	0.70	0.79
	(0.04)	(0.09)	(0.06)	(0.05)	(0.05)	(0.04)	(0.04)
Prospective (Ref.: Final Year)							
1 year	1.07	1.06	1.01	1.02	1.02	1.05	1.07
	(0.06)	(0.09)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)
2 years	1.02	1.09	1.01	1.13	0.97	1.06	1.04
	(0.06)	(0.10)	(0.07)	(0.07)	(0.05)	(0.05)	(0.05)
3 years	1.17	1.08	1.03	1.14	0.96	0.95	1.12
	(0.07)	(0.10)	(0.08)	(0.07)	(0.05)	(0.04)	(0.06)
4 years	1.24	1.13	1.00	1.20	0.93	1.03	1.06
	(0.08)	(0.11)	(0.08)	(0.08)	(0.05)	(0.05)	(0.06)
5-8 years	1.25	1.16	0.99	1.24	0.91	1.00	1.01
	(0.08)	(0.11)	(0.10)	(0.10)	(0.04)	(0.05)	(0.07)
9-12 years	1.31	1.19	0.99	1.55	0.85	0.99	1.01
	(0.10)	(0.14)	(0.16)	(0.18)	(0.05)	(0.06)	(0.10)
13+ years	1.37	1.21	1.04	1.75	0.81	0.87	0.97
	(0.13)	(0.19)	(0.21)	(0.27)	(0.06)	(0.06)	(0.12)
Waves (Ref.: 1984)							
1985	0.78	1.00	-	-	0.71	0.74	-
	(0.07)	(.)	-	-	(0.06)	(0.06)	-
1986	1.01	1.10	-	-	0.57	0.69	-
	(0.08)	(0.15)	-	-	(0.04)	(0.05)	-
1987	0.75	1.28	-	-	1.49	0.78	-
	(0.06)	(0.18)	-	-	(0.11)	(0.06)	-
1988	0.69	1.16	-	-	1.31	0.57	-
	(0.06)	(0.17)	-	-	(0.10)	(0.04)	-
1989	0.75	1.76	-	-	1.85	0.39	-
	(0.07)	(0.26)	-	-	(0.15)	(0.03)	-
1990	0.75	2.33	-	-	1.76	0.25	-

Continues on next page

Continues from previous page

	Life Sat.	Politics	Immigr.	Xenoph.	Environm.	Peace	Crime
	(0.07)	(0.34)	-	-	(0.14)	(0.02)	-
1991	0.68	2.93	-	-	0.87	0.59	-
	(0.06)	(0.39)	-	-	(0.07)	(0.04)	-
1992	0.60	1.35	-	1.00	1.40	0.54	-
	(0.06)	(0.20)	-	(.)	(0.11)	(0.04)	-
1993	0.63	1.33	-	2.58	0.77	0.68	-
	(0.06)	(0.20)	-	(3.04)	(0.06)	(0.05)	-
1994	0.51	1.50	-	1.54	0.40	0.67	1.00
	(0.05)	(0.23)	-	(3.96)	(0.03)	(0.05)	(.)
1995	0.53	1.18	-	1.30	0.49	0.68	0.80
	(0.05)	(0.19)	-	(1.16)	(0.04)	(0.05)	(0.05)
1996	0.52	1.36	-	1.05	0.35	0.50	0.97
	(0.05)	(0.22)	-	(0.79)	(0.03)	(0.04)	(0.06)
1997	0.45	1.02	-	1.13	0.32	0.36	1.15
	(0.05)	(0.17)	-	(0.85)	(0.03)	(0.03)	(0.07)
1998	0.45	1.43	-	1.12	0.31	0.49	0.96
	(0.04)	(0.23)	-	(0.84)	(0.03)	(0.04)	(0.06)
1999	0.48	1.30	1.00	0.92	0.17	0.50	0.78
	(0.05)	(0.22)	(.)	(0.70)	(0.02)	(0.04)	(0.05)
2000	0.40	1.40	1.57	0.85	0.20	0.30	0.64
	(0.04)	(0.23)	(0.10)	(0.64)	(0.02)	(0.02)	(0.04)
2001	0.50	1.19	2.15	1.32	0.18	0.24	0.61
	(0.05)	(0.19)	(0.14)	(0.99)	(0.02)	(0.02)	(0.04)
2002	0.36	1.43	1.61	0.78	0.14	0.54	0.55
	(0.04)	(0.26)	(0.13)	(0.59)	(0.01)	(0.05)	(0.04)
2003	0.35	1.75	2.01	0.63	0.17	1.71	0.45
	(0.04)	(0.31)	(0.17)	(0.48)	(0.02)	(0.16)	(0.04)
2004	0.29	1.46	1.43	0.63	0.17	0.51	0.43
	(0.03)	(0.27)	(0.13)	(0.48)	(0.02)	(0.05)	(0.04)
2005	0.37	1.61	0.90	1.15	0.23	0.46	0.67
	(0.04)	(0.31)	(0.09)	(0.87)	(0.02)	(0.04)	(0.06)
2006	0.37	1.76	1.20	1.05	0.23	0.59	0.50
	(0.05)	(0.36)	(0.15)	(0.80)	(0.02)	(0.06)	(0.05)
2007	0.38	1.34	1.71	0.93	0.36	0.38	0.51
	(0.05)	(0.28)	(0.22)	(0.71)	(0.04)	(0.04)	(0.06)
2008	0.39	1.28	2.34	0.78	0.23	0.31	0.39
	(0.05)	(0.27)	(0.32)	(0.60)	(0.03)	(0.03)	(0.04)
2009	0.38	1.61	3.33	0.53	0.21	0.31	0.26
	(0.05)	(0.35)	(0.48)	(0.40)	(0.02)	(0.03)	(0.03)
2010	0.42	1.38	3.31	0.64	0.23	0.30	0.35
	(0.06)	(0.32)	(0.55)	(0.49)	(0.03)	(0.03)	(0.05)
2011	0.41	2.07	2.73	0.67	0.26	0.38	0.24
	(0.06)	(0.49)	(0.50)	(0.52)	(0.03)	(0.04)	(0.03)

Continues on next page

Continues from previous page

	Life Sat.	Politics	Immigr.	Xenoph.	Environm.	Peace	Crime
2012	0.46	1.22	4.40	0.74	0.20	0.33	0.22
	(0.07)	(0.30)	(0.86)	(0.58)	(0.03)	(0.04)	(0.03)
2013	0.54	1.33	3.06	0.76	0.16	0.31	0.23
	(0.08)	(0.34)	(0.65)	(0.59)	(0.02)	(0.04)	(0.04)
2014	0.52	2.02	2.24	1.01	0.18	0.53	0.28
	(0.09)	(0.53)	(0.51)	(0.79)	(0.02)	(0.07)	(0.05)
Move (Ref.: No)							
Yes	1.19	0.96	1.06	0.98	0.96	0.97	1.09
	(0.06)	(0.09)	(0.05)	(0.05)	(0.04)	(0.04)	(0.05)
Mode (Ref.: PAPI)							
CAPI	0.84	1.19	1.03	0.99	1.08	1.05	1.07
	(0.03)	(0.06)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Interviewer (Ref.: 1st year)							
2-4 years	0.99	1.08	1.09	0.99	1.01	1.10	0.88
	(0.04)	(0.08)	(0.06)	(0.05)	(0.04)	(0.04)	(0.04)
5-8 years	0.97	1.11	1.02	1.12	0.98	1.12	0.92
	(0.04)	(0.08)	(0.06)	(0.06)	(0.04)	(0.04)	(0.04)
9-15 years	0.92	1.10	0.93	1.16	1.00	1.08	0.97
	(0.04)	(0.09)	(0.05)	(0.06)	(0.04)	(0.04)	(0.04)
16+ years	0.88	1.12	1.04	1.08	0.88	0.99	0.91
	(0.05)	(0.09)	(0.06)	(0.06)	(0.04)	(0.04)	(0.05)
Interviewer (Ref.: Low)							
High	1.01	0.96	1.08	1.01	0.93	0.92	0.88
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Interviewer (Ref.: Low)							
High	0.99	1.05	0.88	0.98	1.07	1.04	0.93
	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
N	91916	38792	58097	73919	115016	123336	86953
N_g	8152	3104	6204	7285	9495	10476	8112

Multivariate Panel Fixed Effects Regressions. Odds Ratios
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Table A5
Effects of Interviewer Familiarity: Panel Fixed Effects Logistic Regressions, Odds Ratios (Standard Errors in Parentheses)

	Life Sat.	Politics	Immigr.	Xenoph.	Environm.	Peace	Crime
Interviewer (Ref.: 1st year)							
2 years	0.97	0.85	0.98	0.93	0.88	0.89	0.89
	(0.03)	(0.04)	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)
3-4 years	0.98	0.79	1.06	0.88	0.80	0.81	0.88
	(0.03)	(0.04)	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)
5-9 years	0.88	0.69	0.99	0.77	0.80	0.74	0.90
	(0.03)	(0.04)	(0.04)	(0.03)	(0.02)	(0.02)	(0.03)
10+ years	0.82	0.68	0.93	0.81	0.75	0.69	0.94
	(0.04)	(0.05)	(0.05)	(0.04)	(0.03)	(0.03)	(0.04)
Panel (Ref.: 1st year)							
2 years	0.85	1.24	1.07	0.82	0.94	0.96	0.88
	(0.04)	(0.11)	(0.07)	(0.05)	(0.05)	(0.04)	(0.05)
3-4 years	0.68	1.10	0.91	0.89	1.04	0.96	0.91
	(0.03)	(0.09)	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)
5-9 years	0.67	1.16	0.93	0.85	1.00	0.96	0.98
	(0.03)	(0.10)	(0.06)	(0.04)	(0.05)	(0.04)	(0.05)
10+ years	0.67	1.24	0.91	0.97	1.07	0.89	0.89
	(0.04)	(0.13)	(0.07)	(0.06)	(0.06)	(0.05)	(0.05)
Prospective (Ref.: Final Year)							
1 year	1.07	1.07	1.01	1.03	1.02	1.05	1.07
	(0.06)	(0.09)	(0.06)	(0.06)	(0.05)	(0.05)	(0.05)
2 years	1.02	1.10	1.01	1.14	0.98	1.06	1.04
	(0.06)	(0.10)	(0.07)	(0.07)	(0.05)	(0.05)	(0.05)
3 years	1.17	1.10	1.03	1.14	0.97	0.95	1.12
	(0.07)	(0.10)	(0.08)	(0.07)	(0.05)	(0.04)	(0.06)
4 years	1.24	1.15	1.00	1.21	0.93	1.03	1.06
	(0.08)	(0.11)	(0.08)	(0.08)	(0.05)	(0.05)	(0.06)
5-8 years	1.25	1.17	0.99	1.25	0.91	1.00	1.01
	(0.08)	(0.11)	(0.10)	(0.10)	(0.04)	(0.05)	(0.07)
9-12 years	1.30	1.19	0.99	1.54	0.85	0.99	1.01
	(0.10)	(0.14)	(0.16)	(0.18)	(0.05)	(0.05)	(0.09)
13+ years	1.37	1.21	1.04	1.73	0.80	0.86	0.96
	(0.13)	(0.19)	(0.21)	(0.27)	(0.06)	(0.06)	(0.12)
Waves (Ref.: 1984)							
1985	0.81	1.00	-	-	0.73	0.78	-
	(0.07)	(.)	-	-	(0.06)	(0.06)	-
1986	1.03	1.06	-	-	0.56	0.70	-
	(0.09)	(0.15)	-	-	(0.04)	(0.05)	-
1987	0.76	1.25	-	-	1.49	0.80	-

Continues on next page

Continues from previous page

	Life Sat.	Politics	Immigr.	Xenoph.	Environm.	Peace	Crime
	(0.06)	(0.17)	-	-	(0.12)	(0.06)	-
1988	0.69	1.05	-	-	1.30	0.56	-
	(0.06)	(0.16)	-	-	(0.10)	(0.04)	-
1989	0.76	1.63	-	-	1.82	0.39	-
	(0.07)	(0.24)	-	-	(0.14)	(0.03)	-
1990	0.76	2.16	-	-	1.73	0.25	-
	(0.07)	(0.31)	-	-	(0.14)	(0.02)	-
1991	0.69	2.64	-	-	0.83	0.57	-
	(0.06)	(0.35)	-	-	(0.06)	(0.04)	-
1992	0.61	1.25	-	1.00	1.36	0.53	-
	(0.06)	(0.18)	-	(.)	(0.11)	(0.04)	-
1993	0.63	1.25	-	2.57	0.77	0.67	-
	(0.06)	(0.19)	-	(3.05)	(0.06)	(0.05)	-
1994	0.51	1.38	-	1.48	0.40	0.66	1.00
	(0.05)	(0.22)	-	(3.88)	(0.03)	(0.05)	(.)
1995	0.53	1.09	-	1.27	0.48	0.68	0.80
	(0.05)	(0.17)	-	(1.14)	(0.04)	(0.05)	(0.05)
1996	0.52	1.26	-	1.15	0.34	0.49	0.96
	(0.05)	(0.20)	-	(0.87)	(0.03)	(0.04)	(0.06)
1997	0.46	0.96	-	1.26	0.32	0.36	1.15
	(0.05)	(0.16)	-	(0.95)	(0.03)	(0.03)	(0.07)
1998	0.46	1.35	-	1.26	0.31	0.49	0.96
	(0.04)	(0.22)	-	(0.95)	(0.03)	(0.04)	(0.06)
1999	0.49	1.22	1.00	1.03	0.16	0.50	0.78
	(0.05)	(0.20)	(.)	(0.78)	(0.02)	(0.04)	(0.05)
2000	0.41	1.31	1.58	0.95	0.20	0.30	0.63
	(0.04)	(0.21)	(0.10)	(0.72)	(0.02)	(0.02)	(0.04)
2001	0.51	1.12	2.18	1.47	0.18	0.24	0.61
	(0.05)	(0.18)	(0.14)	(1.12)	(0.02)	(0.02)	(0.04)
2002	0.36	1.34	1.63	0.87	0.14	0.55	0.54
	(0.04)	(0.24)	(0.13)	(0.67)	(0.01)	(0.05)	(0.04)
2003	0.36	1.65	2.03	0.71	0.17	1.74	0.44
	(0.04)	(0.30)	(0.17)	(0.54)	(0.02)	(0.16)	(0.04)
2004	0.30	1.37	1.45	0.70	0.17	0.51	0.42
	(0.04)	(0.26)	(0.13)	(0.54)	(0.02)	(0.05)	(0.04)
2005	0.38	1.51	0.91	1.28	0.23	0.46	0.66
	(0.05)	(0.29)	(0.09)	(0.98)	(0.02)	(0.04)	(0.06)
2006	0.38	1.64	1.22	1.17	0.23	0.60	0.49
	(0.05)	(0.33)	(0.15)	(0.89)	(0.02)	(0.06)	(0.05)
2007	0.39	1.26	1.73	1.04	0.36	0.38	0.50
	(0.05)	(0.26)	(0.22)	(0.80)	(0.04)	(0.04)	(0.05)
2008	0.40	1.20	2.36	0.87	0.23	0.31	0.38
	(0.05)	(0.25)	(0.32)	(0.67)	(0.03)	(0.03)	(0.04)

Continues on next page

Continues from previous page

	Life Sat.	Politics	Immigr.	Xenoph.	Environm.	Peace	Crime
2009	0.39	1.52	3.35	0.59	0.21	0.32	0.25
	(0.05)	(0.33)	(0.48)	(0.45)	(0.02)	(0.03)	(0.03)
2010	0.43	1.30	3.35	0.72	0.24	0.31	0.34
	(0.06)	(0.30)	(0.56)	(0.55)	(0.03)	(0.03)	(0.04)
2011	0.42	1.94	2.76	0.74	0.26	0.39	0.24
	(0.06)	(0.46)	(0.50)	(0.58)	(0.03)	(0.05)	(0.03)
2012	0.47	1.16	4.45	0.82	0.20	0.34	0.22
	(0.07)	(0.28)	(0.87)	(0.64)	(0.03)	(0.04)	(0.03)
2013	0.56	1.26	3.10	0.84	0.16	0.31	0.23
	(0.09)	(0.32)	(0.65)	(0.66)	(0.02)	(0.04)	(0.04)
2014	0.54	1.91	2.28	1.12	0.18	0.54	0.28
	(0.09)	(0.50)	(0.52)	(0.88)	(0.02)	(0.07)	(0.05)
Move (Ref.: No)							
Yes	1.19	0.95	1.07	0.98	0.96	0.98	1.09
	(0.06)	(0.09)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)
CAPI	0.84	1.19	1.03	0.99	1.07	1.04	1.07
	(0.03)	(0.06)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Interviewer (Ref.: 1st year)							
2-4 years	0.95	1.03	1.07	0.98	1.00	1.05	0.88
	(0.04)	(0.07)	(0.06)	(0.05)	(0.04)	(0.04)	(0.04)
5-8 years	0.95	1.14	1.00	1.18	1.00	1.13	0.93
	(0.04)	(0.09)	(0.06)	(0.06)	(0.04)	(0.04)	(0.04)
9-15 years	0.93	1.16	0.93	1.23	1.04	1.12	0.96
	(0.05)	(0.09)	(0.06)	(0.06)	(0.04)	(0.05)	(0.05)
16+ years	0.90	1.20	1.05	1.16	0.92	1.05	0.90
	(0.05)	(0.11)	(0.06)	(0.06)	(0.04)	(0.05)	(0.05)
Interviewer (Ref.: Low)							
High	1.01	0.97	1.08	1.02	0.94	0.93	0.88
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)
Interviewer (Ref.: Low)							
High	0.99	1.05	0.88	0.98	1.07	1.04	0.93
	(0.03)	(0.04)	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)
N	91916	38792	58097	73919	115016	123336	86953
N_g	8152	3104	6204	7285	9495	10476	8112

Multivariate Panel Fixed Effects Regressions. Odds Ratios. Controls: Wave fixed effects; Prospective panel participation; Move (y/n); Interviewer experience, response-rate, and mean household contacts within wave.

This is in line with the results of Prior (2010), showing a strong stability of interest in politics over the life-cycle.

* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Table A6
Robustness Check I: Satisficing

	Life Sat.	Politics	Immigr.	Xenoph.	Environm.	Peace	Crime
Interviewer Continuity (Ref.: 1st year)							
2 years	1.00	0.82	0.93	0.81	0.85	0.81	0.83
	(0.06)	(0.07)	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
3-4 years	0.95	0.75	1.00	0.78	0.79	0.76	0.82
	(0.05)	(0.06)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)
5-9 years	0.80	0.61	0.93	0.66	0.78	0.68	0.84
	(0.05)	(0.05)	(0.05)	(0.04)	(0.04)	(0.03)	(0.04)
10+ years	0.69	0.60	0.89	0.71	0.71	0.59	0.87
	(0.05)	(0.06)	(0.06)	(0.05)	(0.04)	(0.03)	(0.05)
Panel (Ref.: 1st year)							
2 years	0.80	0.94	1.19	1.03	1.14	0.98	1.08
	(0.09)	(0.17)	(0.14)	(0.12)	(0.11)	(0.09)	(0.10)
3-4 years	0.68	1.14	0.94	0.93	1.25	0.97	1.16
	(0.06)	(0.16)	(0.09)	(0.09)	(0.11)	(0.08)	(0.10)
5-9 years	0.70	1.20	1.00	0.87	1.10	0.99	1.21
	(0.07)	(0.18)	(0.10)	(0.08)	(0.10)	(0.08)	(0.10)
10+ years	0.70	1.29	0.99	0.95	1.26	0.95	1.06
	(0.08)	(0.22)	(0.12)	(0.11)	(0.12)	(0.09)	(0.10)
Item Nonresponse (Ref.: None)							
Some	0.99	0.94	1.02	0.95	1.04	0.99	0.94
	(0.04)	(0.05)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Top 25%	1.03	0.97	1.19	0.87	0.98	0.86	0.85
	(0.05)	(0.06)	(0.05)	(0.04)	(0.04)	(0.03)	(0.03)
Straightlining (Ref.: No)							
Yes	0.78	0.94	0.93	0.93	0.91	0.97	0.92
	(0.04)	(0.07)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)
Response Style (Ref.: No style)							
Centrist	0.86	1.03	0.99	1.01	0.98	0.96	0.98
	(0.04)	(0.06)	(0.04)	(0.04)	(0.03)	(0.03)	(0.03)
Extreme	1.37	1.10	1.02	0.99	0.94	0.97	0.98
	(0.06)	(0.08)	(0.05)	(0.05)	(0.04)	(0.04)	(0.04)
N	27415	13396	29728	33166	38816	44129	44393
N_g	3878	1820	4429	4833	5432	6268	6211

Controls: Wave fixed effects; Prospective panel participation; Move (y/n); Interviewer experience response-rate, and mean household contacts within wave.
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

Table A7
Robustness Check II: Continuity Effects on Items with no/less Socially Desirable Connotations

	Immigr.	Xenoph.	Environm.	Peace
Interviewer Continuity (Ref.: 1st year)				
2 years	0.99	0.93	0.96	0.96
	(0.03)	(0.04)	(0.03)	(0.02)
3-4 years	1.04	0.99	0.95	0.98
	(0.03)	(0.05)	(0.03)	(0.03)
5-9 years	1.00	0.97	1.04	1.04
	(0.04)	(0.05)	(0.04)	(0.03)
10+ years	1.01	0.98	1.02	1.07
	(0.05)	(0.06)	(0.05)	(0.04)
Panel Experience (Ref.: 1st year)				
2 years	0.84	1.05	1.07	1.11
	(0.04)	(0.08)	(0.05)	(0.05)
3-4 years	0.70	1.03	0.93	1.00
	(0.03)	(0.07)	(0.04)	(0.04)
5-9 years	0.56	0.82	0.77	1.01
	(0.03)	(0.06)	(0.04)	(0.04)
10+ years	0.51	0.72	0.73	1.14
	(0.03)	(0.06)	(0.04)	(0.06)
N	89677	47746	94947	125058
N_g	8088	4952	8120	10626

Controls: Wave fixed effects; Prospective panel participation; Move (y/n); Interviewer experience response-rate, and mean household contacts within wave.
[a] This is in line with the results of Prior (2010), showing a strong stability of interest in politics over the life-cycle.
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$