Article

Sociological Methods & Research
2023, Vol. 52(2) 909–932

DOI: 10.1177/0049124120914921
journals.sagepub.com/home/smr

Early and Late Participation during the Field Period: Response Timing in a Mixed-Mode Probability-Based Panel Survey

Tobias Gummer[1] and **Bella Struminskaya**[2]

Abstract

Reluctance of respondents to participate in surveys has long drawn the attention of survey researchers. Yet, little is known about what drives a respondent's decision to answer the survey invitation early or late during the field period. Moreover, we still lack evidence on response timing in longitudinal surveys. That is, the questions on whether response timing is a rather stable respondent characteristic and what—if anything—affects change in response timing across different interviews remain open. We relied on data from a mixed-mode general population panel survey collected between 2014 and 2016 to study the stability of response timing across 18 panel waves and factors that influence the decision to participate early or late in the field period. Our results suggest that the factors which had effects on response timing are different in the mail and web modes. Moreover, we found that experience with prior panel waves affected the respondent's decision to participate early or late. Overall, the present study advocates understanding

[1] GESIS—Leibniz Institute for the Social Sciences, Mannheim, Germany
[2] Utrecht University, The Netherlands

Corresponding Author:
Tobias Gummer, GESIS—Leibniz Institute for the Social Sciences, B2, 1 Mannheim, Baden-Wuerttemberg 68159, Germany.
Email: tobias.gummer@gesis.org

response timing as a metric variable and, consequently, the need to reflect this in modeling strategies.

Keywords

response timing, participation, field period, reluctant respondents, mixed-mode, panel surveys

Introduction

A respondent's timing with respect to responding to a survey invitation has received the attention of survey methodologists as part of the ongoing discussions on survey nonresponse and data quality (Kreuter, Müller and Trappmann 2014; Mayer and Pratt 1966; Newman 1962; Olson 2013). Previous studies have shown that early and late respondents systematically differ in demographic characteristics such as sex (Irani, Gregg, and Telg 2004; Kennickell 2008; Sigman et al. 2014), age (Bates and Creighton 2000; Kanitkar and Marlar 2015; Kruse et al. 2010), ethnic minority status (Bates and Creighton 2000; Sigman et al. 2014; Voigt, Koepsell, and Daling 2003), employment status (Bates and Creighton 2000; Kennickell 2008), and other characteristics (Díaz de Rada 2005; Green 1991; Kypri, Stephenson, and Langley 2004). Accordingly, some have made cautious suggestions that the risk of nonresponse bias may be minimized by using reminders and other additional fieldwork efforts (e.g., Díaz de Rada 2005) and follow-up surveys (e.g., Olson 2013) and that the risk of nonresponse bias generally varies across the field period (e.g., Moore, Durrant, and Smith 2018; Sigman et al. 2014; Struminskaya and Gummer 2017).

The discussions on the length of the field period and response timing are important, since they have consequences for survey management and organization. When designing and managing a survey, the decision about the optimal length of the field period has implications not only for the risk of nonresponse bias but also for the amount of effort that will be put into the fieldwork, which includes additional efforts to convert reluctant respondents at the end of the field period, contacting hard-to-reach cases, and providing the infrastructure to complete the survey during the field period (e.g., maintaining respondent hotlines, hosting web surveys, and managing the fieldwork efforts). The amount of effort that is put into fieldwork can have a severe impact on survey costs, and thus, it is in the researcher's interest to invest the most appropriate amount of effort, for instance, by choosing the optimal field length. In addition, if respondents differ in their characteristics, depending on when they participate during the field period, selecting the optimal field period length can be a tool to control the risk of nonresponse bias. The field period length can be selected in a way that the participation of underrepresented subgroups is enhanced, for instance, by keeping the fieldwork open until habitual late respondents have participated. However, this strategy would require an in-depth knowledge about what influences response timings during the field period.

Previous studies have focused mainly on differentiating subgroups of respondents by comparing early and late respondents on a set of sociodemographic characteristics (e.g., Bates and Creighton 2000; Irani et al. 2004; Kennickell 2008; Kruse et al. 2010; Sigman et al. 2014; Voigt et al. 2003). Determining why respondents chose to answer a survey invite early or late during the field period has not received much attention. Furthermore, only a few studies have examined response timing in longitudinal surveys (examples include Kanitkar and Marlar 2015; Rao and Pennington 2013; Ullman and Newcomb 1998). In contrast, research on panel attrition and nonresponse has generated theories of habitual survey participation behavior (e.g., Lugtig 2014). A notable study by Rao and Pennington (2013) has proposed a more longitudinal perspective that distinguishes between *habitually* early, intermediate, and late response timings. However, despite these authors' important contribution to the longitudinal perspective on response timing, we still lack empirical evidence on the stability of the early and late response timing of respondents across multiple interviews in a panel survey, and on what drives response timing. This lack of evidence is particularly unfortunate for two reasons. First, the information provided by respondents about themselves in previous interviews can be used to predict and influence their response timing. Second, longitudinal surveys depend on respondents participating multiple times (Lynn 2009) and, thus, selecting an optimal length of field periods for a sample can help to increase participation and possibly representativeness. In other words, a too short field period may impede the participation of typical late respondents and, thus, may introduce a systematic underrepresentation of these respondents in the sample. In addition, respondents' commitment to a panel can grow if their participation preferences are reflected by the length of the field period. Therefore, from the perspective of longitudinal surveys, it is important to gain knowledge about what drives early and late responses. Moreover, for the purpose of efficiently planning field period lengths, the existence of notoriously late respondents would be challenging.

The present study addressed two research questions regarding early and late participation in a survey. The first research question is as follows: How can early and late survey participation be explained, and do systematic differences exist between groups of respondents who respond early and late? We also examined whether the findings of prior studies can be replicated, such as differences in gender with early respondents more likely to be female, differences in age with early respondents being older (Bates and Creighton 2000; Kruse et al. 2010), and differences in ethnic minority status with early respondents more likely to be nonminority (Bates and Creighton 2000; Sigman et al. 2014; Voigt et al. 2003). The second research question is as follows: Is early or late responding a rather stable respondent characteristic or subject to change? In other words, are late respondents always late?

The next section describes the theoretical framework from which we derived a set of hypotheses about factors that affect the response timing in a survey so to address our first research question. We approached our second research question from an explorative perspective. In the following sections, we introduce our data and methods. After presenting the results of our study, we conclude with closing remarks and an outlook for future research opportunities.

Theoretical Framework: Response Timing

Most of the previous studies about response timing have used empirical categorizations of early, intermediate, or late respondents. For example, Kreuter et al. (2014:456) used the number of call attempts as a measure of operational effort and divided the sample into quintiles. Rao and Pennington (2013) identified respondents as *early* if they participated before the first reminder, *intermediate* if they participated before the second reminder, and *late* if a third reminder was necessary. Similarly, Ullman and Newcomb (1998) defined *early* and *late respondents* by the number of reminders needed until participation. Other studies such as Voigt et al. (2003) relied on time intervals to classify their respondents as *early, intermediate,* or *late*. In practice, after receiving a survey invite, the decision to participate can be made at any time as long as a potential respondent has access to the survey (e.g., the mail questionnaire rests on the desk; the invitation e-mail is in the inbox; and the survey is active). We argue that the empirically driven classification of response timing into a few classes does not reflect the practical reality, since participation can occur at any point in time at which access to the survey is possible. Respondents make a decision when to participate, which can range from participating at the field start to participating at the end of the field period.

Based on our expectation that respondents differ in their response timing, we assumed that the timing is a function of factors that describe the respondents' contactability and cooperativeness. Respondents make a decision when to participate in a survey every time they have an opportunity to participate (i.e., each time respondents look at a mail questionnaire [MAIL] or at an invitation e-mail [WEB]). Following this reasoning, we drew on two theories of survey participation as a theoretical framework from which we derived hypotheses on response timing. However, explaining (non-)response differs from explaining response timing. Thus, when addressing response timing, we focused on people who ultimately respond and not on nonrespondents.

First, in their Social Exchange Theory, Dillman, Smyth, and Christian (2014) argued that respondents compare the costs against the rewards of participating in a survey after adjusting for how much they actually trust receiving this reward. The reward is not necessarily monetary or material in nature, since respondents can deem their survey participation as being useful for society (Dillman et al. 2014). Costs can include the amount of time that survey participation will require, weighted by the amount of time available to the respondent. Second, Leverage Saliency Theory provides a theoretical framework for respondent attributes that facilitate cooperativeness in survey participation (Groves, Singer, and Corning 2000). Depending on the saliency of these attributes (e.g., importance of the survey topic) to respondents, their likelihood of participation changes. Both frameworks conceptualize survey participation from a cost-benefit perspective, which is the basis for integrating them into one model and applying their logic to response timing decisions. By combing these two frameworks—similar to what Gummer (2019) suggested for explaining nonresponse—we expected that participating early during the field period is a function of the costs and rewards that a respondent associates with immediately cooperating with a survey invitation. The more personally enjoyable and socially beneficial respondents evaluate their immediate cooperation, the more likely will be their early response, since the rewards should be high. The more costly respondents deem their immediate cooperation, the more likely they are to postpone their participation. How respondents interpret costs and saliency also can depend on the exact situation in which they are making a decision about participating.

Previous Research and Hypotheses

With respect to a panel survey, respondents have prior experience with previous waves of the panel on which they can base their decision to continue their participation (e.g., Schnell 1997:141-46). For instance, interviewer observations of respondents being hostile, restless, or impatient were correlated negatively with their participation in the next panel wave (Olsen 2005), and respondents who perceived the questionnaires as difficult or too personal were found to be more likely to attrite (Struminskaya 2014:hapter 4). We argue that respondents will incorporate these prior experiences into their calculations of the costs and rewards of immediately completing a survey or postponing its completion. Some researchers have shown that survey value and survey enjoyment positively affect participation in a panel survey (De Leeuw, Hox, and Rosche 2017). Thus, our hypotheses regarding our first research question are as follows:

Hypothesis 1 (H1): The more important for science respondents perceive their latest completed panel wave, the quicker they will respond to a pending survey invite.

Hypothesis 2 (H2): The more interesting respondents perceive their latest completed panel wave, the quicker they will respond to a pending survey invite.

Hypothesis 3 (H3): The more diverse respondents perceive their latest completed panel wave, the quicker they will respond to a pending survey invite.

Furthermore, the perceived burden of completing a survey (De Leeuw et al. 2017; Groves 1998; Groves et al. 2002) and privacy concerns (Bates, Dahlhamer, and Singer 2008) can negatively impact survey participation. Accordingly, we hypothesized that:

Hypothesis 4 (H4): The more too personal respondents perceived their latest completed panel wave, the slower they will respond to a pending survey invite.

Hypothesis 5 (H5): The longer their latest completed panel wave was perceived by respondents, the slower they will respond to a pending survey invite.

Hypothesis 6 (H6): The more difficult their latest completed panel wave was perceived by respondents, the slower they will respond to a pending survey invite.

With respect to self-administered mixed-mode surveys (web and mail modes), we expected the decision to respond early or later to differ between the modes. Gummer and Daikeler (2018) have argued that being able to perceive the questionnaire not only visually on a screen (i.e., as in web mode) but also haptically (i.e., by touch) enables a more intense processing associated with the survey and facilitates memorizing information related to the questionnaire characteristics. Accordingly, with respect to the mail mode, these authors found that the length, difficulty, and diversity of questionnaires impacted the decision of respondents to participate in later waves of a panel. With respect to the web mode, these authors reported that questionnaire length and privacy concerns had significant effects on participation in a panel survey. Consequently, we assumed that the effects of prior experiences on a respondent's decision to immediately act on a survey invite or postpone participation will differ between the web and mail modes.

In addition to testing our hypotheses regarding our first research question on the impact of prior survey experience, we set out to replicate the findings of previous studies. Since most previous studies did not derive their independent variables from a theoretical framework comparable to ours, we will not present hypotheses but will establish theoretical links where appropriate. For example, Kennickell (2008:493) has argued that a group of younger respondents "may contain relatively many people who are busy with the advancement of their careers" while "older respondents are more likely to be at home" and, thus, age proxies for the time available for respondents to complete surveys. With respect to our theoretical framework, this argument assumes a relationship between age and contactability. However, it seems sensible to test whether an age effect on response timing is linear or whether it diminishes over time as individuals get settled and, in older years, retire (i.e., ∩-shaped relationship). Also, previous research has reported that early respondents are less likely to be part of an ethnic minority, for instance, as defined by citizenship (Bates and Creighton 2000; Kanitkar and Marlar 2015; Kruse et al. 2010; Rao and Pennington 2013; Sigman et al. 2014). With respect to our theoretical framework, this finding might be explained as follows: contributing to the benefit of society should be more valuable to members of an ethnic majority as a way to help maintain the status-quo that favors them, compared to ethnic minorities (e.g., van Tubergen, Maas, and Flap 2004). Consequently, immediate participation might be less likely for ethnic minorities compared to members of the ethnic majority. The research findings on other sociodemographic characteristics such as sex (e.g., Irani et al. 2004; Sigman et al. 2014) and education (e.g., Rao and Pennington 2013; Ullman and Newcomb 1998; Voigt et al. 2003) remain inconclusive and demand further testing. Consequently, we include these characteristics in our analysis without specifying prior expectations.

Data

To investigate our research questions, we used data from the GESIS (2017) Panel, which is a mixed-mode general population panel in Germany (cf. Bosnjak et al. 2018). The panel was recruited in 2013 using face-to-face interviews based on a probability-based register sample. The American Association for Public Opinion Research (2016) Response Rate 5 to this initial interview was 38.6 percent. 4,888 respondents consented to be re-interviewed and were invited to participate in the first panel wave when the panel went fully operational in 2014. Since 2014, the panelists have been interviewed six times per year (i.e., every two months). These interviews were conducted via web or mail questionnaires. At each wave, web participants received a mailed invitation letter and an e-mail inviting them to participate in the survey. Up to two reminder e-mails were sent to web respondents. Mail respondents received a mailed invitation letter, a paper-based questionnaire, and a prepaid return envelope. In both instances, the mailed invitation letters included a 5€ unconditional incentive.

During the recruitment interview, the interviewers slightly pushed the respondents who regularly used the Internet to participate in the web mode, whereas those who did not use the Internet or were not comfortable with the web as a survey mode could opt for the mail mode (in 2014 62 percent web, 38 percent mail). Apart from being pushed slightly by the interviewers to choose the web option, respondents were free to select their preferred mode of participation. In addition, despite not being encouraged to do so, respondents were free to change their mode of participation in between the waves. With respect to the data we used, mode changes were very rare ($N = 3$). As we differentiated between response timing in web and mail mode (see below), we excluded these three respondents from our analyses.

For the present study, we selected all waves that were conducted between 2014 and 2016—a total of 18 waves. Overall, participation in the panel remained at a rather high level, and in the 18th wave, 69.7 percent of the initial panelists were still active (3,408). Selecting 18 waves of interviews gave us a data set of 66,372 spells (i.e., person waves).

The dependent variable in our analyses was days after each wave's field start until a respondent participated in an interview, either via a web or via mail mode. This information was available for the web interviews as time stamps that were automatically collected by the survey software. For the offline interviews, each mail questionnaire featured a textbox at the end of the questionnaire in which respondents were asked to provide their date of participation.

Table 1. Descriptive Statistics.

	Web		Mail		Total				
	M	*SD*	*M*	*SD*	*M*	*SD*	Min	Max	*N*
Dependent variable									
Days until participation									
Web mode	7.825	7.732	—	—	7.825	7.732	1	55	2,770
Mail mode	—	—	7.761	8.388	7.761	8.388	2	53	1,497
Independent variables									
Sex: female	0.492	0.500	0.569	0.495	0.520	0.500	0	1	4,291
Age	42.888	14.175	51.165	13.364	45.815	14.445	18	70	4,278
Education									
Low	0.134	0.341	0.374	0.484	0.219	0.414	0	1	4,291
Intermediate	0.336	0.472	0.390	0.488	0.355	0.479	0	1	4,246
High	0.529	0.499	0.232	0.422	0.423	0.494	0	1	4,246
Ethnic majority	0.953	0.212	0.956	0.205	0.954	0.209	0	1	4,289
Last survey									
Important for science	2.500	0.841	2.542	0.922	2.515	0.870	0	4	4,231
Interesting	2.679	0.783	2.700	0.813	2.686	0.794	0	4	4,193
Diverse	2.725	0.805	2.783	0.826	2.745	0.813	0	4	4,133
Too personal	0.784	0.765	0.783	0.851	0.783	0.796	0	4	4,180
Long	0.897	0.787	0.863	0.826	0.885	0.801	0	4	4,199
Difficult	0.406	0.615	0.501	0.736	0.439	0.661	0	4	4,197

Note. All figures are given for wave 1 of the panel. *M* = mean; *SD* = standard deviation; min = minimum; max = maximum.

To test our hypotheses and replicate previous research, we created a set of independent variables. Descriptive statistics for all the independent variables are provided in Table 1. Each respondent's sex was measured in the initial face-to-face interview (0 = *male*, 1 = *female*). At the time of the panel recruitment, we computed each respondent's age by drawing on their birth year as a metric variable. Furthermore, we recoded education levels into *low*, *intermediate*, and *high*. To assess whether a respondent was part of the ethnic majority, we drew on their citizenship and differentiated between Germans and non-Germans (0 = *minority*, 1 = *majority*).

In addition to these time-invariant measures, we operationalized the experience that respondents had with their last interview in the panel as a set of time-variant variables. Each wave of the panel featured questions concerning an evaluation of the survey that enabled us to assess how respondents perceived the survey—whether it was interesting, difficult, important for science, long, and diverse—and whether the questions were too personal. We measured each of these items on a scale from 0 (*not at all*) to 4 (*very*).

Method

Our dependent variable (days until participation) was a count variable and followed a Poisson distribution (Appendix Figure A1, which can be found at http://smr.sagepub.com/supplemental/). However, testing the distribution showed signs of overdispersion. To cope with this overdispersion, we fitted mixed-effects negative binomial regressions. These models are described in more detail in Hox, Moerbeek, and van de Schoot (2017:chapter 7) and Cameron and Trivedi (2005:chapters 20 and 23). In our case, mixed-effects models were necessary to account for the panel data structure in which the interviews were nested within respondents (i.e., multiple interviews per respondent). Consequently, our mixed models incorporated equations on two levels: the level of the respondents and the level of the panel waves. This approach is similar to what Allison (2009:chapter 4) suggested concerning the analysis of panel data with negative binomial regressions.

Using panel data enables a calculation of fixed effects estimators that control for time-invariant unobserved heterogeneity (Allison 2009). To take advantage of these beneficial properties of our data, we used a hybrid approach to predict our estimators as suggested by Allison (2009:23-25, 65-68). In contrast to a conventional fixed effects regression approach, a hybrid approach includes fixed estimates for time-variant predictors and also enables the prediction of time-invariant estimators. For all the time-variant variables, we calculated their deviations from the person-specific means. Thus, we included person-specific means and deviations in the regressions. Finally, the coefficients for the deviations of the time-variant variables can be interpreted as fixed effects coefficients.

Following the reasoning by Gummer and Daikeler (2018) that participation processes likely differ between the web and mail modes, we fitted two regressions, one for each mode. Thus, we treated the web and the mail mode as "mode systems" (Biemer and Lyberg 2003:208; Struminskaya, De Leeuw, and Kaczmirek 2015), that is, as entire data collection processes designed around a specific mode.

As described previously, respondents self-selected into modes during the recruitment interview. This process potentially could produce biased results if the respondents systematically differ between the modes. To control for potential self-selection bias, we applied propensity score weighting

(Rosenbaum and Rubin 1953). In the first step, we fitted a logistic regression with the participation mode (0 = *mail*, 1 = *web*) as the dependent variable, and a set of independent variables from the recruitment interview that we expected to predict affinity toward the web mode (see Table A1 in the Appendix, which can be found at http://smr.sagepub.com/supplemental/). In the second step, we predicted propensity scores based on this model. The propensity scores represented the likelihood of respondents to participate via the web mode. In the third step, we used the propensity scores to calculate weighting factors (Horvitz and Thompson 1952). Based on the variables used in the regression models, these weights assign more importance to cases that are underrepresented in the respective mixed-effects negative binomial regressions. In other words, in our model explaining response timing in the mail mode, respondents who were more likely to participate in the web mode were given more importance, whereas respondents who were more likely to participate in the mail mode were given less importance. Similarly, in the model explaining response timing in the web mode, likely mail respondents were weighted up, whereas likely web respondents were weighted down. We included the weighting factors as a weight on the respondent level in each of our mixed models, since they correct whether the respondent decided to participate via mail or web mode. In our final data set, for each given respondent, the participation mode did not vary between waves (see above).

In addition, panel surveys suffer from attrition, which may include wave nonresponse, spells of multiple nonresponse episodes, and permanent dropout (Lugtig 2014). If respondents attrite systematically, this could be a further source of bias. To also account for potential attrition bias, we again applied propensity score weighting. For each mode, we fitted logistic regressions for each wave with participation in the wave as the dependent variable (0 = *no*, 1 = *yes*) and the evaluation of the prior survey and sociodemographic variables (sex, age, education) as the independent variables (cf. Gummer and Roßmann 2019). We then predicted the respondents' response propensity for the given wave and used the inverted propensity score to calculate the weighting factors. These attrition weights resulted in likely attritors being weighted up, whereas likely stayers were weighted down. In our mixed-effects models, we included the attrition weights on the wave level as the weighting factors that represented each respondent's participation likelihood for a given wave and thus varied between the waves.

To investigate the stability of response timing across waves, we grouped respondents into intervals of days until their participation in each wave. These intervals corresponded to different possible definitions of *fast* and *slow respondents*. For instance, when defining all respondents as being part of the *fastest* that answered within 10 days, all respondents that participated later were considered *slow*. Following this reasoning, we created a set of variables that indicated whether a respondent was part of the fastest group who participated within 5 days, 10 days, 15 days, . . . 60 days or part of the slowest respondents who participated after 5 days, 10 days, . . . 60 days, respectively.[1] We then used these dichotomous variables to calculate the transition probabilities between different response timings. For instance, we assessed the average likelihood of the respondents who participated within 10 days to show the same response timing again or to become part of the slower respondents who needed more time than 10 days to participate and vice versa. Again, to control for the possible effects of self-selection into modes and panel attrition, we applied both propensity score weights when calculating the transition probabilities.

Results

In this section, we address our first research question on the differences between early and late respondents. Table 2 details the results of our regression models. Significant tests on the overdispersion parameter log (α) and the variance component σ_u^2 illustrate that it was appropriate, first, to choose a negative binomial model to account for overdispersion and, second, to represent the panel structure in a mixed-effects model.

With regard to the evaluation of prior panel waves and to our hypotheses on the perceived importance of the survey (H1), we found a significant effect in the web mode with respondents taking longer, but not in the mail mode, since no effect was found. For H2, we found that respondents were more inclined to respond early to a web survey when they had deemed the prior survey more interesting compared to respondents who reported a lower interest in the survey. While this finding partially supports H2, we found the effect of interest to be present only for web participants. We had to reject H3, since we found that the diversity of the latest completed panel survey had no effect on response timing. With respect to the web mode, we found that perceiving the prior survey as too personal affected time until participation, which partially supported H4. However, in the mail mode, perceiving the prior survey as too personal had no significant effect on response timing, which contradicted the expectation of H4. This result differed from our findings on the effect of the prior survey's perceived length. In line with our expectations (H5), if respondents perceived prior surveys as rather lengthy, they were less likely to quickly respond to a pending survey invitation. Contrary to our assumption regarding the differences between the mail and

Table 2. Results of Mixed-Effects Negative Binomial Regressions on Days until Survey Participation.

	Web Coeff. (*SE*)	Mail Coeff. (*SE*)
Last survey		
Important for science	0.022*** (0.006)	−0.013 (0.010)
Interesting	−0.036*** (0.007)	−0.019 (0.012)
Diverse	−0.003 (0.007)	−0.004 (0.011)
Too personal	0.015** (0.005)	−0.017 (0.009)
Long	0.029*** (0.005)	0.026** (0.009)
Difficult	−0.006 (0.006)	0.015 (0.010)
Age	0.047*** (0.007)	0.030*** (0.009)
Age^2	−0.001*** (0.000)	−0.000*** (0.000)
Ethnic majority	−0.149* (0.068)	−0.148 (0.096)
Education		
Low	Ref.	Ref.
Intermediate	0.104** (0.040)	0.048 (0.048)
High	−0.001 (0.038)	0.164** (0.052)
Sex: female	0.146*** (0.029)	0.047 (0.040)
Intercept	1.916*** (0.198)	1.614*** (0.246)
Log (α)	−1.000*** (0.010)	−0.776*** (0.014)
σ^2 (u)	0.497*** (0.015)	0.478*** (0.021)
N (spells)	39,365	17,694
N (waves)	18	18

Note. *p < .05. **p < .01. ***p < .001.

web modes, this effect occurred in both modes. With respect to the effect of a prior survey's difficulty (H6), we found no significant effects for both modes and, thus, had to reject this hypothesis.

Turning to the replication of previous studies, we found—in line with our expectations—that the relationship between age and days until participation was ∩-shaped. Figure 1 illustrates this nonlinear relationship in more detail by drawing on the predicted day of participation conditional on the respondent's age. With increasing age, respondents were less likely to participate early during the field period until at the age of 30 years for the web mode and 40 years for the mail mode, after which a turning point was reached and respondents again became more likely to participate early. Only partial support was found for the expected effect of belonging to the ethnic majority. With respect to the mail mode, we found no significant differences between respondents who belonged to the ethnic majority (i.e., German citizens) or

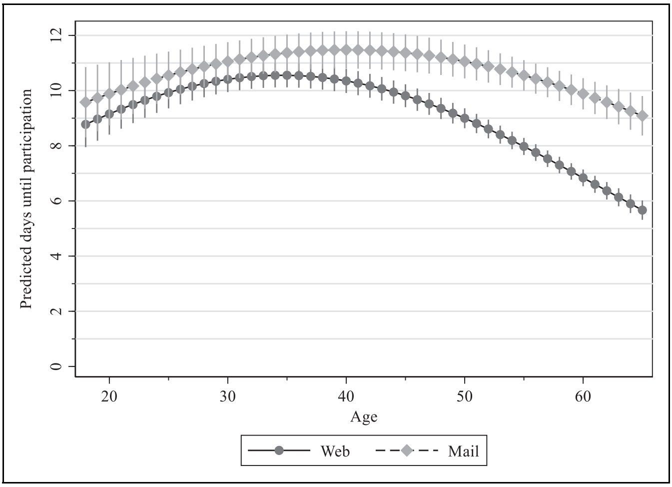

Figure 1. Predicted day of survey participation by respondent's age. 95 percent confidence intervals represented by lines.

minority. However, in line with our expectations, ethnic minorities were less likely to respond early when participating online. In addition, we found no clear effect of education for web participation, whereas panel members with a higher level of education who responded by mail were less likely to respond early compared to those with a lower education; and respondents with an intermediate level of education who responded in the web mode were less likely to respond early than respondents with a lower level of education. We also included sex as an independent variable. Surprisingly, we only found a significant positive effect of being female for those women who responded in the web mode. In our opinion, this finding may be one explanation for the ambivalent findings reported in the literature—response timing processes differ depending on survey characteristics such as the mode of response. Since most of the findings from previous research were based on single mode studies, our findings may help to explain the previous ambiguous reports on the respondent attributes that influence response timing.

To address our second research question on the stability of response timing across different waves of a panel, we computed transition probabilities between the different intervals of days until participation. Figure 2 shows an illustrative transition plot based on the transition probabilities of the fastest respondents who participated within 5 days and those who needed more than 5 days to participate using the web mode. This transition plot

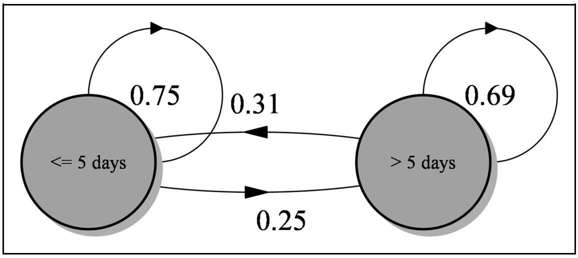

Figure 2. Transition probability plot for the slowest (participating after five days) and the fastest (participating within five days) respondents in the web mode. Arrows indicate the probability of remaining (white circles) or becoming (arrows between gray circles) part of the fastest/slowest respondents. Transition plots cover a specific group of respondents (e.g., fastest as defined by participating within five days) and their counterpart (e.g., slowest as defined by participating after five days).

details the likelihood of the future response timing for a group of respondents (those who participated within 5 days) and their counterparts (those who participated after 5 days). Two possible transitions are given for both groups: (i) remain in this state or (ii) become part of their counterpart. In this example, we found that the probability of the fastest respondents remaining in this group (.75) was 7 percentage points lower than the probability of the slower respondents to remain in the same group (.69). Consequently, the probability to transition into a respective counterpart was higher for the slowest respondents, compared to the fastest respondents. The probability of becoming part of the fastest group was .31 for the slowest of the respondents, whereas the probability of becoming part of the slowest group was .25 for the fastest.

The transition probabilities for all groups and their counterparts are presented in Figure 3, which aggregates the information that would have been included in the transition probability plots, such as in Figure 2, but that is more suitable for comparisons. For instance, with respect to the web mode, the fastest of the respondents who participated until day 20—on average—had a probability of .94 to again be part of the fastest group and a probability of .06 to become part of the slowest group who participated after 20 days. On the other hand, the respondents who participated after day 20 had a probability of .53 to become part of the fastest group and a probability of .47 to remain in the slowest group.

If response timing in later waves of a panel was independent from prior waves, we would have observed a linear pattern across the different intervals in Figure 3. In this case, the probability to participate on any given day (p_d)

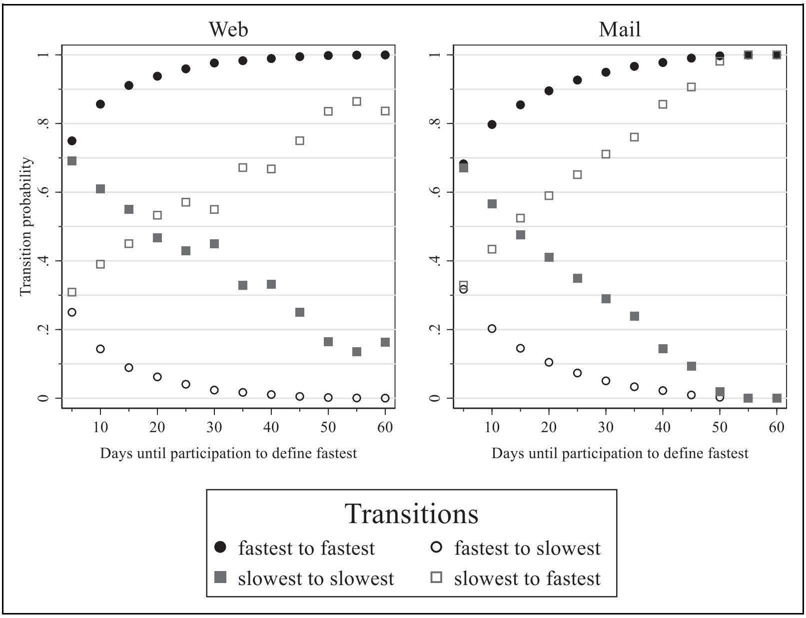

Figure 3. Transition probabilities for web and mail respondents. Dots indicate the probability of remaining in the same definition of fastest or slowest respondents (solid markers) or becoming part of the remaining respondents (hollow markers). For instance, in the web mode, the fast respondents who participated until day 10 had a probability of .86 to remain among the fastest and a probability of .14 to become part of the slower respondents, that is those who participated after day 10.

would be similar for all days during the field period. In our case, we assumed that the field period would end after 60 days, which would result in $p_d = 1/60$ for any given day. For an interval until day D, the fastest respondents from a previous wave would have had a probability of $D \times p_d$ to be among the fastest again and a probability of $1 - (D \times p_d)$ to participate after day D and vice versa for the slowest respondents of a prior wave. The result would have been a linear pattern of transition probabilities across intervals, but we did not find this to be the case. An illustration of the pattern that would indicate an independence of response timing between the waves is provided in Figure A2 in the Appendix (which can be found at http://smr.sagepub.com/supplemental/).

In Figure 3, we compared the probabilities for transitions between the different groups of both modes. The more we shifted the definition of *late* *respondents* toward the lower end of the field period (i.e., by defining *late respondents* as those who participated close to the end of the field period, such as after day 50), the less stable this group became (i.e., the probability to remain part of the slowest respondents group became lower). This decrease in stability highlights two important points. The first point is the relevance of how late respondents are defined, and how such an arbitrary distinction impacts the estimates of stability and change. Second, a gap exists in the probabilities when comparing slowest to fastest: the slowest respondents were, for most of the groups and in both modes, more likely to change their state (i.e., become part of the faster counterpart) when compared to the fastest respondents. Similarly, the slowest respondents were less likely to remain part of the slowest group, compared to the fastest respondents. This finding challenges the binary nature of distinguishing between early and late respondents, which often is the position taken by researchers (e.g., Kennickell 2008; Sigman et al. 2014) when providing analyses from a longitudinal perspective. In contrast, our findings hint at the existence of a group of notoriously early respondents. This finding is particularly interesting, since late respondents increase the field duration. In other words, late respondents are those who cause additional management efforts and ultimately increase survey costs. Our findings suggest that those respondents who we defined as *late* (with different thresholds) had a high likelihood to change their response timing over additional waves.

In addition, our analysis of our first research question on determinants of response timing showed that different factors were relevant in mail and web modes with respect to the decision of when to participate. However, the patterns we found for transition probabilities between different response timings (Figure 3) progressed in similar ways. With respect to both modes, the probability of remaining within the same group decreased as the definition for *slow respondents* became more rigid. Also, the slowest respondents were less likely to remain in the slowest group and more likely to become part of the fastest group, compared to the fastest respondents' probability of remaining in the fastest group or becoming part of the slowest group. We interpret this finding as encouraging with respect to the robustness of our findings on transition probabilities. In practical terms, this means that late responding is a decision that respondents make that has a potential to change for both the mail and web modes, whereas the reasons for which respondents change their behavior—such as perceiving a prior wave questionnaire as too personal or interesting—differ between the mail and web modes.

Conclusion

In the present study, we set out to investigate what affected response timing in a mixed-mode general population panel survey, how response timing varied depending on respondent characteristics, and whether response timing was stable across different waves of the panel. We found that the response timing was different for the mail and web modes, but most importantly, that different factors affected the response timing for the web and mail modes. In line with our expectations, we found that prior experience with panel waves affected the response timing. Also, the length of the latest questionnaire affected the response timing for the mail mode, whereas the perceived length of the prior wave, its importance for science, and the perception of a prior wave as interesting or too personal had significant impacts on the response timing for the web mode. Apart from these differences in how respondents perceived their latest questionnaire in a panel, we found that the age differences of respondents impacted when they participated in a panel survey during the field period. With respect to the mail mode, we found further differences—compared to the web mode—in belonging to the ethnic majority. Our data also revealed no clear effect of education on response timing when comparing the mail and web modes. Overall, we found that response timing changed across different waves of the panel. This is an encouraging finding because it challenges the notion of the existence of a large number of notoriously late responders. Although our analyses show that certain response timings are more likely for specific respondents, timing can change over the waves or even possibly be influenced by researchers.

Our results have practical implications for survey practitioners and survey methodologists. First, in our analysis, we focused in detail on testing the effects of prior experience with surveys in the same panel study. The experience that respondents have with surveys is under the researchers' control and, thus, may be influenced to stimulate a quicker response in subsequent interviews. For instance, our findings suggest that providing respondents with shorter questionnaires may invoke a faster response, which has the potential to decrease the overall required field management efforts. This finding may be promising for other surveys that have found very pronounced differences between early and late respondents, since a responsive or adaptive design may be applied to stimulate response timing based on systematically varied prior survey experiences. Similar tools have been suggested—for example, targeted invitation letters—for tackling attrition in panel surveys (cf. Lynn 2017). Lynn (2016) has shown that when the content of invitation letters is tailored to the potential interest of a group to which a respondent belongs—such as employed, parent of young children, young, and so on—the response rates of previous wave nonrespondents and reluctant respondents increase. Given that the information about late respondents is available from prior waves, a similar approach can be implemented for web and mail panels.

Second, we found that certain groups of respondents were more likely to respond early to a survey, whereas other groups were more likely to respond late. This finding supports previous research that found that depending on the length of the field period, the risk of nonresponse bias may vary (e.g., Moore et al. 2018; Sigman et al. 2014; Struminskaya and Gummer 2017).

Third, different processes seem to determine response timing when the respondent is participating via mail or the web. Consequently, when trying to analyze whether response timing relates to respondent characteristics (e.g., when assessing data quality in a survey), we caution that the response mode is a potential confounder that needs controlling, since different factors have an impact in the mail and web modes. Our findings that highlight mode differences—even after controlling for self-selection into a survey mode—and the effects of survey design (i.e., the design of a prior survey wave in a panel) may serve as possible explanations why previous studies have reported mixed results on the effects of respondent characteristics on response timing.

Fourth, we advocate for considering response timing as metric rather than a (arbitrarily) constructed binary distinction between early and late respondents. Reducing the information level of a metric variable such as a time stamp to an ordinal variable seems counterintuitive and may hinder researchers from drawing on the full potential of their data.

Although our study has limitations, it also yields future research opportunities. First, it drew on 18 waves of a mixed-mode general population panel survey that used both a mail and web mode. Based on these data, however, we were not able to compare the process of the response timing of the two modes (mail and web) to interviewer-administered modes such as face-to-face or telephone surveys. In addition, it might be interesting to replicate our study with an experimental design in which survey modes are randomly allocated to respondents to account for self-selected mode choice in the GESIS Panel. However, we would like to note that this might limit the number of panel waves that could be conducted due to cost restrictions. Advancing our line of research in this general direction may provide useful insights that will be especially relevant for face-to-face surveys for which field duration causes significant costs.

Second, our research questions focused on factors affecting response timing and the stability of timing across several waves of a panel. Therefore, we did not attempt to research the optimal length of fieldwork periods, since this approach would exceed the scope of our study. Optimizing fieldwork periods can be done toward different indicators such as the risk of nonresponse bias, survey costs, survey outcome rates, or even the statistical power of collected data (e.g., Schouten, Peytchev, and Wagner 2018). However, the optimization of one of these indicators may have detrimental effects on the other indicators (e.g., optimizing outcome rates may increase survey costs). This potential problem makes the discussion on optimization a nontrivial problem that requires studies that approach this challenge from a different perspective, for instance, see Moore et al. (2018) with respect to the risk of nonresponse bias.

Third, we focused on data from a panel survey because it enabled us to measure respondents' experience with prior waves, draw on the beneficial statistical properties of panels, and assess the degree of stability of the response timings. However, we acknowledge that this approach may limit the generalizability of our findings, for instance, when working with purely cross-sectional surveys. Moreover, attrition may impact an analysis that draws on panel data. To alleviate the potential effects of attrition, we chose a modeling strategy that enabled us to control for time-invariant unobserved heterogeneity in parts of our independent variables and to apply propensity score weighting.

Fourth, because the reminder protocol was different for the web and mail modes, we chose to ignore the potential influence of the reminder on response timing. For the purposes of our study, we treated the web and the mail modes as "mode systems" (Biemer and Lyberg 2003:208; Struminskaya et al. 2015), that is, as entire data collection processes designed around a specific mode.

Fifth, while our study lays the groundwork for stimulating faster response timing by using adaptive and responsive designs by identifying factors that are under a researcher's control (i.e., survey design), testing this idea exceeds the scope of the present study. We see merit in advancing such approaches by means of experimental studies in which respondents in a panel receive questionnaires of different designs (e.g., varying difficulty or length) and their response timing in subsequent surveys is analyzed.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Supplemental Material

The supplemental material for this article is available online

ORCID iD

Tobias Gummer https://orcid.org/0000-0001-6469-7802

Note

1. For ease of interpretation and readability, we decided to group response timing by steps of 5 days. As a robustness check, we reran our analysis of transition probabilities with grouping by steps of 1 day. The patterns we observed in the 1-day grouping were similar to the grouping by steps of 5 days, and so the conclusions we would have drawn remained the same. We interpret the findings from our robustness check as supporting our conclusions.

References

American Association for Public Opinion Research. 2016. *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. The American Association for Public Opinion Research.

Allison, Paul D. 2009. *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage.

Bates, Nancy and Kathleen Creighton. 2000. "The Last Five Percent: What Can We Learn from Difficult/Late Interviews?" Pp. 120-25 in *Proceedings of the Section on Government Statistics*. Alexandria, Egypt: American Statistical Association.

Bates, Nancy, James Dahlhamer, and Eleanor Singer. 2008. "Privacy Concerns, Too Busy, or Just Not Interested: Using Doorstep Concerns to Predict Survey Nonresponse." *Journal of Official Statistics* 24(4):591-612.

Biemer, Paul P. and Lars E. Lyberg. 2003. *Introduction to Survey Quality*. Hoboken, NJ: Wiley.

Bosnjak, Michael, Tanja Dannwolf, Tobias Enderle, Ines Schaurer, Bella Struminskaya, Angela Tanner, and Kai W. Weyandt. 2018. "Establishing an Open Probability-Based Mixed-Mode Panel of the General Population in Germany: The GESIS Panel." *Social Science Computer Review* 36(1):103-15.

Cameron, A. Colin and Pravin K. Trivedi. 2005. *Microeconometrics. Methods and Applications*. Cambridge, UK: Cambridge University Press.

De Leeuw, Edith D., Joop Hox, and Benjamin Rosche. 2017. "Predicting Nonresponse and Attrition in a Probability-Based Online Panel." *Paper presented at the General Online Research Conference (GOR)*, Berlin, Germany.

Díaz de Rada, Vidal. 2005. "The Effect of Follow-up Mailings on the Response Rate and Response Quality in Mail Surveys." *Quality & Quantity* 39:1-18.

Dillman, Don A., Jolene D. Smyth, and Leah M. Christian. 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. Hoboken, NJ: John Wiley.

GESIS. 2017. *GESIS Panel—Standard Edition, Za5665, V19.0.0*. Cologne, Germany: GESIS Data Archive.

Green, Kathy E. 1991. "Reluctant Respondents: Differences between Early, Late, and Nonresponders to a Mail Survey." *The Journal of Experimental Education* 59(3): 268-76.

Groves, Robert M. 1998. *Survey Errors and Survey Costs*. New York: Wiley.

Groves, Robert M., Don A. Dillman, John L. Eltinge, and Roderick J. A. Little. 2002. *Survey Nonresponse*. New York: Wiley.

Groves, Robert M., Eleanor Singer, and Amy Corning. 2000. "Leverage-Saliency Theory of Survey Participation." *Public Opinion Quarterly* 64(3):299-308. doi: 10.1086/317990.

Gummer, Tobias and Jessica Daikeler. 2018. "A Note on How Prior Survey Experience with Self-Administered Panel Surveys Affects Attrition in Different Modes." *Social Science Computer Review*. Online first.

Gummer, Tobias and Joss Roßmann. 2019. "The Effects of Propensity Score Weighting on Attrition Biases in Attitudinal, Behavioral, and Socio-Demographic Variables in a Short-Term Web-Based Panel Survey." *International Journal of Social Research Methodology* 22(1):81-95.

Gummer, Tobias. 2019. "Assessing Trends and Decomposing Change in Nonresponse Bias: The Case of Bias in Cohort Distributions." *Sociological Methods & Research* 48(1):92-115.

Horvitz, Daniel G. and Donovan J. Thompson. 1952. "A Generalization of Sampling without Replacement from a Finite Universe." *Journal of the American Statistical Association* 47(260):663-85.

Hox, Joop, Mirjam Moerbeek, and Rens van de Schoot. 2017. *Multilevel Analysis. Techniques and Applications*. New York: Routledge.

Irani, Tracy A., Jon A. Gregg, and Ricky Telg. 2004. "Choosing to Use the Web: Comparing Early and Late Respondents to an Online Web-based Survey Designed to Assess It Computer Skills Perceptions of County Extension Agents." *Journal of Southern Agricultural Education Research* 54(1):168-79.

Kanitkar, Kirti and Jenny Marlar. 2015. "How Fast Can I Get Survey Results? Assessing Demographics Differences in Cumulative Daily Response Rates to Web

Panel Surveys." *Paper presented at the American Association for Public Opinion Research 70th Annual Conference (AAPOR)*, Hollywood.
Kennickell, Arthur B. 2008. "The Bitter End? The Close of the 2007 SCF Field Period." Pp. 490-97 in *Joint Statistical Meeting Proceedings, Survey Research Methods Section*. Alexandria, VA: American Statistical Association.
Kreuter, Frauke, Gerrit Müller, and Mark Trappmann. 2014. "A Note on Mechanisms Leading to Lower Data Quality of Late or Reluctant Respondents." *Sociological Methods & Research* 43(3):452-64.
Kruse, Yelena, Erlina Hendarwan, J. Michael Dennis, and Charles DiSogra. 2010. "Analysis of Late Responders to Probability-Based Web Panel Recruitment and Panel Surveys." *Paper presented at the American Association for Public Opinion Research 65th Annual Conference (AAPOR)*, Chicago.
Kypri, Kypros, Shaun Stephenson, and John Langley. 2004. "Assessment of Non-response Bias in an Internet Survey of Alcohol Use." *Alcoholism: Clinical and Experimental Research* 28(4):630-34.
Lugtig, Peter. 2014. "Panel Attrition: Separating Stayers, Fast Attriters, Gradual Attriters, and Lurkers." *Sociological Methods & Research* 43(4):699-723.
Lynn, Peter. 2009. "Methods for Longitudinal Surveys." Pp. 1-19 in *Methodology of Longitudinal Surveys*, edited by P. Lynn. Chichester: Wiley.
Lynn, Peter. 2016. "Targeted Appeals for Participation in Letters to Panel Survey Members." *Public Opinion Quarterly* 80(3):771-82.
Lynn, Peter. 2017. "From Standardised to Targeted Survey Procedures for Tackling Non-Response and Attrition." *Survey Research Methods* 11(1):93-103.
Mayer, Charles S. and Robert W. Pratt. 1966. "A Note on Nonresponse in a Mail Survey." *Public Opinion Quarterly* 30(4):637-46.
Moore, Jamie C., Gabriele B. Durrant, and Peter W. F. Smith. 2018. "Data Set Representativeness During Data Collection in Three UK Social Surveys: Generalizability and the Effects of Auxiliary Covariate Choice." *Journal of the Royal Statistical Society A* 181(1):229-48.
Newman, Sheldon W. 1962. "Differences between Early and Late Respondents to a Mailed Survey." *Journal of Advertising Research* 2(2):37-39.
Olsen, Randall J. 2005. "The Problem of Respondent Attrition: Survey Methodology Is Key." *Monthly Labor Review* 128:63-70.
Olson, Kristen. 2013. "Do Non-Response Follow-Ups Improve or Reduce Data Quality? A Review of the Existing Literature." *Journal of the Royal Statistical Society A* 176:129-45.
Rao, Kumar and Julia Pennington. 2013. "Should the Third Reminder Be Sent? The Role of Survey Response Timing on Web Survey Results." *International Journal of Market Research* 55(5):651-74.

Rosenbaum, Paul R. and Donald B. Rubin. 1953. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1):41-55.

Schnell, Rainer. 1997. *Nonresponse in Bevölkerungsumfragen: Ausmaß, Entwicklung Und Ursachen*. Opladen: Leske + Budrich.

Schouten, Barry, Andy Peytchev, and James Wagner. 2018. *Adaptive Survey Design*. Boca Raton, FL: CRC Press.

Sigman, Richard, Taylor Lewis, Naomi D. Yount, and Kimya Lee. 2014. "Does the Length of Fielding Period Matter? Examining Response Scores of Early Versus Late Responders." *Journal of Official Statistics* 30(4):651-74.

Struminskaya, Bella and Tobias Gummer. 2017. "The Link between Nonresponse Bias and the Length of the Field Period in a Mixed-Mode General Population Panel." *Paper presented at the Joint Statistical Meeting*, Baltimore.

Struminskaya, Bella, Edith D. De Leeuw, and Lars Kaczmirek. 2015. "Mode System Effects in an Online Panel Study: Comparing a Probability-Based Online Panel with Two Face-to-Face Reference Surveys." *Methods, Data, Analyses* 9(1):3-56.

Struminskaya, Bella. 2014. "Data Quality in Probability-Based Online Panels." PhD thesis, Utrecht University.

Ullman, Jodie B. and Michael D. Newcomb. 1998. "Eager, Reluctant, and Nonresponders to a Mailed Longitudinal Survey: Attitudinal and Substance Use Characteristics Differentiate Respondents." *Journal of Applied Social Psychology* 28(4):357-75.

van Tubergen, Frank, Ineke Maas, and Henk Flap. 2004. "The Economic Incorporation of Immigrants in 18 Western Societies: Origin, Destination, and Community Effects." *American Sociological Review* 69:704-27.

Voigt, Lynda F., Thomas D. Koepsell, and Janet R. Daling. 2003. "Characteristics of Telephone Survey Respondents According to Willingness to Participate." *American Journal of Epidemiology* 157(1):66-73.

Author Biographies

Tobias Gummer is a senior researcher and team leader at GESIS—Leibniz Institute for the Social Sciences. His methodological research interests include survey design, data quality, nonresponse, prevention and correction methods for biases, and analysis of social and individual change.

Bella Struminskaya is an assistant professor of methods and statistics at Utrecht University, the Netherland. She has published on the issues of data quality in online panels, mobile web surveys, and panel conditioning. Her current research focuses on nonresponse and measurement errors in panel surveys and smartphone sensor data collection.