Predictors of nonresponse at different stages of a smartphone-only Time Use Survey

Anne Elevelt, Peter Lugtig & Vera Toepoel.

Introduction

Almost everyone has a mobile phone and more and more of such phones are capable of doing much more than calling. In 2015 78.8% of Dutch people aged 12 years or older had a mobile phone with internet access (Statistics Netherlands, 2016). The fact that people carry their smartphone with them everywhere they go offers great opportunities for social scientists interested in studying attitudes and behaviors. We can, for example, send pop-up questions in real-time, for instance to measure participants' feelings, and record sensor data. However, as nice as these new opportunities are, the questions we ask can get increasingly intrusive, and we risk over-asking participants.

The primary research objective of the current study was to investigate the effect of asking respondents to do additional tasks through mobile phones on the survey response rates and nonresponse bias. This article investigates an innovative smartphone-only Time Use Survey (TUS) (Sonck & Fernee, 2013). Throughout this paper, the term "stages" refers to the different times that respondents were asked to participate; first filling out a pre-questionnaire, followed by filling out a diary, answering pop-up questions and sharing GPS data, and finally filling out a post-questionnaire. Respondents were repeatedly asked to participate, hence nonresponse bias could be introduced at every step. More important, they were asked to participate in different activities with different properties.

According to the leverage-salience theory, with every request to participate respondents make a decision to participate or not. This indicates that respondents in this smartphone-only survey are forced to make several decisions based on different properties. Participation propensity per stage depends on the number of positive and negative factors in the request (leverage) and the relative importance to the respondent (saliency) (Groves, Singer, & Corning., 2000; Keusch, 2015). When the request to participate is different (e.g. questionnaire versus filling out a diary) the leverage and salience per request also differ. Therefore, participants may be willing to participate in one stage, but not in another. For example, some participants think privacy is very important and therefore choose not to share GPS data, whereas this may not be the case for other participants (Groves et al., 2000). This difference in willingness to participate may cause nonresponse bias and this bias may differ over different stages. We will use attributes of the participants (e.g. personality, demographics, smartphone familiarity) to predict nonresponse in these different stages.

However, in practice, it seems that once a panel member has agreed to participate in a study this decision is likely to be followed by continuous participation (Lemay, 2010). Social psychologists explain this continuous participating by noting that people possess a strong desire to be consistent within their attitudes, beliefs and deeds (Groves, Cialdini, & Couper, 1992). Therefore we also added a measure of consistency as a predictor to our model.

Fundamental, methodological knowledge about nonresponse in smartphone-only surveys and additional data collection is lacking, but very important to understand selection bias. Therefore, in this workshop I want to discuss with you; Does nonresponse become different in nature when we ask respondents to do additional tasks (e.g. complete diaries, share sensory data or answer pop-up questions)?

Data

The research process consisted of several stages, listed here in chronological order.

1. ***Willing to participate?*** Participants of the Dutch LISS panel, a representative panel of the entire Dutch population, were asked in three different surveys in 2012/2013 whether they were interested in participating in a

smartphone survey. From this pool of interested respondents, participants were randomly assigned to the sample of the actual TUS.

2. ***Pre-questionnaire.*** The selected participants started by completing a pre-questionnaire on the web.

3. ***Diary***. Participants were asked to fill out a diary in which they recorded their activities on two randomly selected days; one weekday and one weekend day. Participants without a smartphone could use a loan phone to reduce coverage error.

4. ***Pop-up questions***. On the same days as the diary, participants received pop-up questions which could be answered for ten minutes. After this ten minute interval the question disappeared from the screen and could not be answered anymore to ensure real-time feelings were measured.

5. ***Sensor Data.*** Participants were asked permission to record auxiliary data. These data included communication (call and text) data and GPS locations.

6. ***Post Questionnaire.*** Participants conducted a post-questionnaire on the web.

Background variables

***Sociodemographic characteristics*.** We used a set of sociodemographic characteristics that we treat as time-invariant: gender, age, net income, educational level and number of children.

Personality. Five personality factors were computed: *openness, conscientiousness, extraversion, agreeableness* and *neuroticism*.

Survey attitude. De Leeuw and colleagues (2010) developed a nine-item scale to measure one's general attitude towards surveys . These three factors are *survey enjoyment, survey value and survey burden.*

***Privacy*.** Two factors for privacy were computed; *trust* and *worries*. The factor *trust* covers how much participants trust different organizations to keep their personal information private. The factor *worries* covers about how worried participants are about their privacy.

***Smartphone usage/ownership*.** The factor *smartphone usage* consists of 22 questions regarding smartphone use for various online activities. *Smartphone ownership* only covers whether participants own a smartphone, as the questions about smartphone usage were asked within the pre-questionnaire and is thus not available for the entire LISS panel.

***Participation history*.** Participation history is a general measure of respondent commitment. These participants have been active within the LISS panel for several years. We calculated the surveys they participated in as a proportion of the number of invitations they received.

Prior decision. *Prior decision* is a study-specific measure of consistency. We added participation (0=no, 1=yes) in the previous stage as a predictor for the subsequent stages.

Methodological strategy

In order to predict who responds and who does not, we conducted several multivariate logistic regression analyses and calculated Average Marginal Effects (AME). Marginal effects (MFX) are the estimated probabilities that the respondent participates for a specific, marginal change in the explanatory variable, holding all other variables fixed. AME expresses the average MFX of the explanatory variable on the dependent variable (Mood, 2010). For example, an AME of -0.54 on age indicates that per year older, the probability to participate decreases by 0.54%.

We report AME instead of odds-ratios because AME are not affected by unobserved heterogeneity and therefore can be compared across models and groups. Odds-ratios reflect unobserved heterogeneity, which is the variation in the dependent variable that is caused by variables that are not observed and thus not included as predictors in the model. As this unobserved heterogeneity varies across models we cannot simply compare the effect of specific predictors at different stages when using odds-ratios.

Results

7296 participants of the LISS panel received the question whether they would be willing to participate in a smartphone survey. Following the AAPOR 2006 guidelines, we observed a noncontact rate of 16.00% (*n* = 1168), a refusal rate of 41.1% (*n* = 2996) and a completion rate of 42.9% (*n* = 3132) (Callegaro & Disogra, 2008).

See Figure 1 and Table 1 for the results within the Time Use Survey.

Points for discussion

1. Does nonresponse become different in nature when we ask respondents to do additional tasks?
2. Do participants really make different decisions based on the properties of every stage?
3. Why are participants not willing to participate?
    - *We do now know for certain why people are not willing to participate. This may be because the request is too large or because the burden of the entire research process is too high. However, this may also be due to (a lack of) technology knowledge.*

References

Callegaro, M., & DiSogra, C. (2008). Computing response metrics for online panels. *Public Opinion Quarterly*, *72*, 1008-1032.

De Leeuw, E. D., Hox, J. J., Lugtig, P., Scherpenzeel, C. V., Goritz, A., & Bartsch, S. (2010). Measuring and Comparing Survey Attitude Among New and Repeat Respondents Cross-Culturally. In Wapor annual conference. Chicago.

Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, *56*, 475-495.

Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: description and an illustration. *Public Opinion Quarterly*, *64*, 299-308.

Keusch, F. (2015). Why do people participate in Web surveys? Applying survey participation theory to Internet survey data collection. *Management Review Quarterly*, *65*, 183-216.

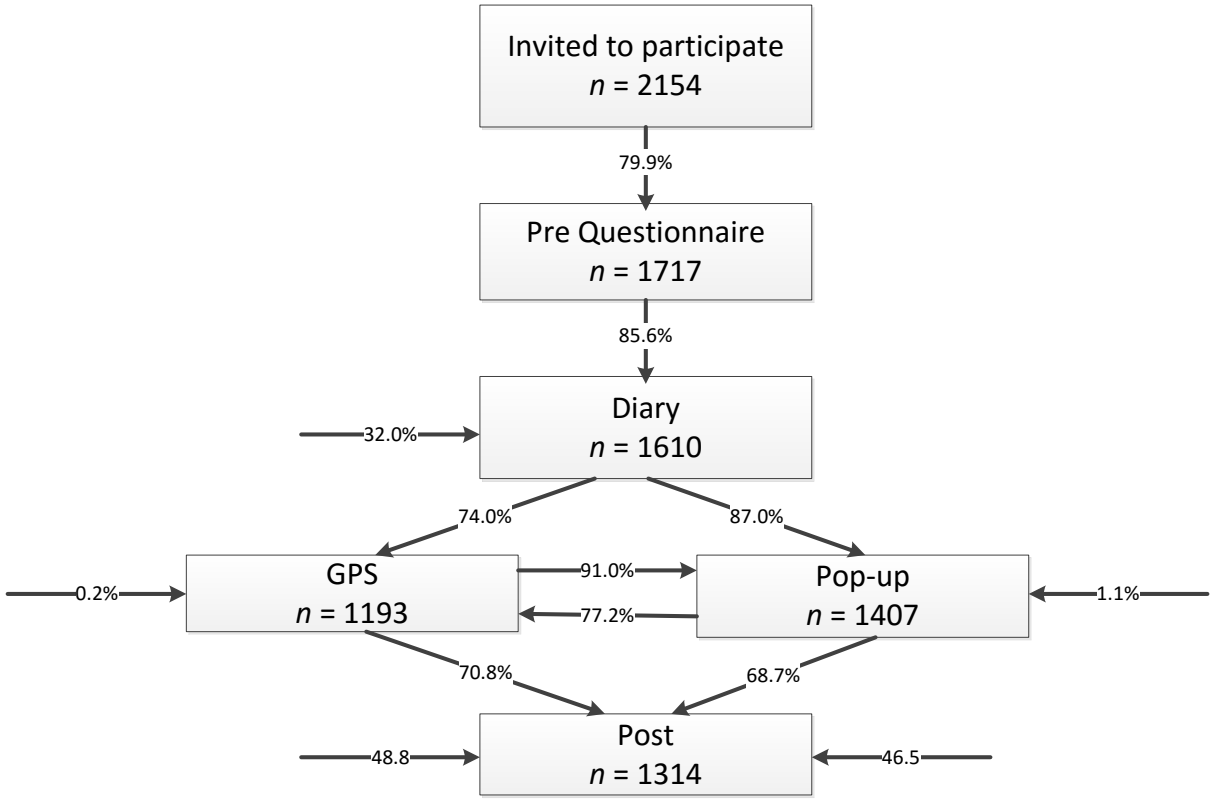
Lemay, M. 2010. ‘‘Understanding the Mechanism of Panel Attrition.’’ Unpublished Doctoral thesis, Doctor of Philosophy, University of Maryland, College Park, MD.

Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, *26*, 67-82.

Sonck, N., & Fernee, H. (2013). *Using smartphones in survey research: A multifunctional tool.* The Hague: The Netherlands Institute for Social Research.

Statistics Netherlands. (2016). *Statline Database*. Retrieved September 29, 2016, from www.cbs.statline.nl

Figure 1. Flowchart of response in the Time Use Survey.



Note: the percentages indicate the response probability per stage, dependent on participation (arrow from one stage to another) or no participation (arrow from the outside to the square) in the preceding stage.

Table 1. Results of the multivariate logistic regression analyses. We conducted a separate analysis per stage, with participation as binary outcome variable.

	Willing to participate?		**Pre Questionnaire**		**Diary**		**Pop-up**		**GPS**		**Post Questionnaire**	
	AME %	s.e.	AME %	s.e.	AME %	s.e.	AME %	s.e.	AME %	s.e.	AME %	s.e.
Gender	-1.08	1.58	-1.93	2.00	-3.70	2.22	-4.12	4.35	4.32	2.93	0.75	2.57
Age	**-0.54*** **	0.05	-0.04	0.07	**-0.34*** **	0.08	**-0.75*** **	0.16	-0.04	0.10	**0.18***	0.09
Educational level	**3.80*** **	0.53	0.16	0.71	**1.62***	0.78	-1.08	1.49	-0.80	1.03	1.20	0.91
Number of kids	0.51	0.70	0.09	0.78	1.47	0.93	1.29	1.81	-1.38	1.13	1.01	1.01
Income	0.07	0.04	0.08	0.10	-0.06	0.11	0.36	0.23	0.24	0.15	-0.06	0.12
Participation History	5.15	3.97	**32.32*** **	4.30	**14.55** **	5.32	11.32	11.79	6.55	7.16	9.98	6.22
Neuroticism	0.10	0.76	0.17	0.95	0.22	1.07	-0.98	1.99	-1.90	1.37	0.65	1.21
Extraversion	**-2.13** **	0.76	-1.75	0.93	-1.10	1.06	0.41	1.96	**-4.10** **	1.39	-0.83	1.17
Agreeableness	-0.82	0.84	1.07	1.01	1.32	1.12	-0.87	2.16	-0.74	1.46	1.43	1.28
Conscientiousness	**-4.52*** **	0.77	**1.99***	0.92	1.46	1.04	-1.15	1.99	2.29	1.34	0.89	1.19
Openness	**3.27*** **	0.77	-1.10	0.94	-1.08	1.06	0.28	1.98	1.13	1.34	-0.42	1.19
Survey value	**3.72** **	1.15	-0.22	1.47	2.22	1.60	4.07	3.13	1.13	2.12	0.10	1.84
Survey enjoyment	**11.11*** **	1.22	-0.17	1.51	-2.07	1.68	**-7.51***	3.20	-3.74	2.19	0.62	1.90
Survey burden	**-2.56***	1.17	-0.45	1.53	-1.95	1.69	-0.09	3.08	-0.75	2.15	-0.79	1.95
Trust	1.22	0.95	0.60	1.18	1.41	1.30	-1.08	2.38	1.62	1.66	0.90	1.48
Worries	**-3.55*** **	0.85	-0.66	1.04	1.84	1.11	0.41	2.07	-0.23	1.45	0.21	1.29
Smartphone usage / ownership	**21.17*** **	1.89	-0.17	1.08	1.25	1.37	3.20	2.34	**3.92***	1.54	-2.56	1.40
Prior decision: PRE					**53.96*** **	2.54	10.80	6.36	**11.04** **	1.23	1.40	3.14
Prior decision: DIARY							**86.63*** **	1.05	**72.91*** **	3.70	**19.32*** **	4.56
Prior decision: POPUP											3.36	3.87
Prior decision: GPS											**9.05** **	2.93
Nagelkerke R^2	0.191		0.077		0.327		0.705		0.554		0.119	

Note : p- values < .001 = '***' , < .01 ='**', < .05 = '*'. Rows represent background variables, columns represent stages.