Survey Research Methods (2019)
Vol. 13, No. 2, pp. 195-213
doi:10.18148/srm/2019.v13i2.7385

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

Doing a Time Use Survey on Smartphones Only: What Factors Predict Nonresponse at Different Stages of the Survey Process?

Anne Elevelt
Utrecht University
The Netherlands

Peter Lugtig
Utrecht University
The Netherlands

Vera Toepoel
Utrecht University
The Netherlands

The increasing use of smartphones opens up opportunities for novel ways of survey data collection, but also poses new challenges. Collecting more and different types of data means that studies can become increasingly intrusive. We risk over-asking participants, leading to nonresponse. This study documents nonresponse and nonresponse bias in a smartphone-only version of the Dutch Time Use Survey (TUS). Respondents from the Dutch LISS panel were asked to perform five sets of tasks to complete the whole TUS: 1) accept an invitation to participate in the study and install an app, 2) fill out a questionnaire on the web, 3) participate in the smartphone time use diary on their smartphone, 4) answer pop-up questions and 5) give permission to record sensor data (GPS locations and call data). Results show that 42.9% of invited panel members responded positively to the invitation to participate in a smartphone survey. However, only 28.9% of these willing panel members completed all stages of the study. Predictors of nonresponse are somewhat different at every stage. In addition, respondents who complete all smartphone tasks are different from groups who do not participate at some or any stage of the study. By using data collected in previous waves we show that nonresponse leads to nonresponse bias in estimates of time use. We conclude by discussing implications for using smartphone apps in survey research.

Keywords: Smartphone survey; nonresponse; sensor data; nonresponse bias; consent bias; time use research

1 Introduction

The increasing popularity of smartphones opens up opportunities for novel ways of data collection in survey research (e.g. Miller, 2012) that could complement and partly substitute survey questions. Unlike Internet (browser) surveys, smartphone apps enable the collection of auxiliary data, such as GPS locations or communication behavior through mobile phones (Dufau et al., 2011; Miller, 2012; Raento, Oulasvirta, & Eagle, 2009). Smartphones incorporate a large number of sensors (e.g. accelerometers, GPS, light and proximity sensors) which can be logged passively, providing a large and detailed set of measurements about respondents and their environment (Cottrill et al., 2013; Ermes, Parkka, Mantyjarvi, & Korhonen, 2008).

Most traditional surveys face declining response rates (De Leeuw, Hox, & Luiten, 2018). Respondents are increasingly reluctant to participate (Groves & Heeringa, 2006) especially when surveys are long and questions burdensome. Galesic (2006) shows that the more burdensome questions are, the less motivated respondents are to answer them. Smartphone surveys can get increasingly burdensome and intrusive as we ask respondents to share more personal information. Respondents may not be willing to share these kinds of data due to privacy concerns (Revilla, 2017). Although smartphones offer great possibilities for better measurement, we risk overasking our participants.

One type of survey research that faces this trade-off between measurement quality and response rates is time use research. Response rates in time use surveys, traditionally conducted with paper diary studies, are generally not very high (Abraham, Maitland, & Bianchi, 2006; Knulst & Van den Broek, 1999; Stoop, 2007). For example, the response rate for the Dutch Time Use Survey (TUS) ranges between 18% in 1995 and 40.3% in 2011–2012 (Cloïn et al., 2013; Statistics Netherlands (CBS), 2013; Van Ingen, Stoop, & Breedveld, 2008). Response rates for time use studies in the

Contact information: Anne Elevelt, Utrecht University, Padualaan 14, 3584CH Utrecht, The Netherlands (E-mail: a.elevelt@uu.nl)

United States are 54.6% (Abraham et al., 2006) and 45% in the United Kingdom (Fisher & Gershuny, 2013). These findings hold in other types of diary studies: An American web-based dietary study had a response rate of 10% and a diary completion rate of only 7.4% (Thompson et al., 2014).

Diary studies are burdensome. Moreover, time use data based on paper diaries suffer from measurement error and recall problems. Measurement in diary studies could potentially improve when conducted through an app (Sonck & Fernee, 2013).

The primary research objective of our study was to investigate the effect of asking intrusive questions through a smartphone app study on survey response rates, predictors of nonresponse and nonresponse bias in a smartphone TUS conducted in the Dutch probability-based LISS Panel. Panel members were asked to participate in several tasks varying in intrusiveness and burden (completing surveys, a diary, sharing sensor data and answering pop-up questions). Nonresponse bias could be introduced or accumulate at every step. We will use attributes of the participants (e.g. personality, demographics, smartphone familiarity) to predict nonresponse in these different stages. Subsequently we will examine whether this nonresponse influences our survey estimates, resulting in nonresponse bias.

2 Theoretical background

2.1 Decision making process for survey response

In order to understand why respondents do or do not participate in the separate parts of the smartphone study, we use the leverage-saliency theory (Groves, Singer, & Corning, 2000). According to this theory, respondents make a decision to participate or not with every request to participate. In making this decision, different respondents place different importance on factors of the survey request. One respondent might value the topic of a survey, another the incentive offered, or the emphasis that the advance letter puts on value for society. Negative leverage factors could be survey burden, privacy concerns, or topic difficulty. Someone's propensity to participate depends on the number of positive and negative factors perceived in the request (leverage) and the relative importance to the respondent (saliency) (Groves et al., 2000; Keusch, 2015).

Most research testing the leverage-saliency theory use experiments to vary aspects of the leverage and saliency in the survey request explicitly. In our study, leverage and saliency were varied more naturally as respondents were asked to perform tasks that are different in nature. For example, worries about privacy may influence the willingness to share GPS data, whereas pop-up questions that interrupt daily life may annoy some participants (e.g. those who are busy). The diary study in our app is the most time-consuming part of the study, so respondents who are sensitive to burden may dropout in this task. Because the nature of the tasks in our study differ, participants may be willing to participate in one task of the TUS, but not in another. This difference in willingness to participate per task may then cause nonresponse bias to vary per task as well.

The leverage-saliency model was developed with a one-time decision in mind, such as in cross-sectional surveys. In many smartphone studies, participants have to make multiple decisions to participate in different tasks which are not independent. From longitudinal surveys, we know that once a panel member has agreed to participate in a study, this decision is likely to be followed by continuous participation (Lemay, 2010). Similarly, once a respondent has not participated in a task, he or she may be more likely to not participate in subsequent tasks either. We can incorporate this longitudinal feature by including prior decision as a separate factor in our leverage-saliency model.

2.2 Time Use Surveys

The self-completed time use diary is considered to be the most reliable and accurate data collection instrument to obtain information on the activity patterns of participants (Michelson, 2005). Many European time use surveys follow the Harmonized European Time Use Survey (HETUS) guidelines (Eurostat, 2009). Time diaries are also used in other fields, for example to measure physical activity (e.g. Bouchard's Physical Activity Record, BAR; see Bouchard et al., 1983) or dietary intake (e.g. Automated Self-Administered 24-hour, ASA24 Dietary Assessment Tool; see Subar et al., 2012), but how the diary is designed varies between studies. Most time diary studies cover the full 24 hours of a day and divide the day into 10 minute timeslots. This works as a cognitive cue and reduces omissions due to forgetfulness (Belli, Shay, & Stafford, 2001). Time diaries also allow the collection of contextual information such as whom the respondent was with, and typically make a distinction between main and side activities. Time use diaries sometimes only span a short time (e.g. one day only), but ideally cover a longer period so that infrequent activities are also captured (Gershuny, 2012).

Time use diaries present challenges for data collection. First, diaries are burdensome to complete, often resulting in response rates that are lower than those of one-time questionnaire-based surveys. Second, if respondents do not complete the diary regularly throughout the day, recall problems may arise resulting in less accurate data. Third, the administration costs are high because the manual coding and entering of data from paper is very labor intensive (Minnen et al., 2014).

Conducting time use research on a smartphone could create a more user-friendly and less burdensome instrument compared to the traditional paper-based TUS. Respondents can complete the diary any time of the day, as long as they have their smartphone with them. In contrast to the paper diary which is usually left at home and filled out at the end of the day, using a smartphone app diary makes it possible to remind respondents to fill out their diary several times per day. This may help to reduce the recall-problem (Lai et al., 2010). In addition, smartphones enable the collection of auxiliary data, such as GPS locations or communication behaviors (Raento et al., 2009), which can reduce the number of questions we need to ask. Finally, a smartphone app can significantly reduce the time and costs of data processing. An app enables the use of pre-coded categories of time use, avoiding coding efforts after data collection has been completed.

There are also some potential pitfalls of using smartphones for diary studies. Coverage error may lead to bias when certain groups or members of the population who do not own a smartphone are automatically excluded. Participants may further be unwilling or insufficiently able to use an app, leading to nonresponse and nonresponse bias. A pilot study of a smartphone TUS by Chatzitheochari et al. (2018) showed promising results regarding response and data quality. 97 cohort members of the UK Millennium Cohort Study were invited for the pilot and could self-select into the web (28%), or a smartphone version (41%) of the TUS. The paper diary (20%) was only offered to participants without a personal computer or smartphone, or who refused to use the web and smartphone modes. There was a nonresponse rate of 11%. Mode choice was similar by gender and household income. Results show that the completion rate for the smartphone (48% on day 1 and 30% on day 2) and web version (33% and 30%) were slightly lower than the paper version (63%). Comparisons of the measurement quality across modes found that there were fewer item-missings in the smartphone app mode and more contextual data (location, and who the respondent was with). Due to a very small sample size and a non-randomized mixed mode design, the results from Chatzitheochari et al. (2018) can however only be treated as indicative.

2.3 Analytical Framework and Hypotheses

Apart from factors specific to the prediction of nonresponse in our smartphone TUS, there are also general predictors of nonresponse relevant to our study. Demographic characteristics such as gender, age, educational level, occupation, ethnicity, household status and size, urbanicity, and marital status have been shown to generally correlate with response propensity in surveys (Fan & Yan, 2010; Groves, Cialdini, & Couper, 1992). Certain demographic variables decrease the contact likelihood, such as urbanicity, living alone, and living without children (Abraham et al., 2006; Groves, 2006). Other demographic characteristics tend to increase refusal rates, such as ethnicity, educational level and age (Lipps, 2009; Lugtig, 2014; Van Ingen et al., 2008). Often there is no clear theoretical link between sociodemographic variables and nonresponse (Fan & Yan, 2010). They can however be important to include as predictors in order to compare studies, and to investigate nonresponse bias.

In longitudinal studies, socio-psychological variables are thought to be more closely related to why participants keep participating or drop out from a study. For example, respondents who are more "agreeable" on the Big Five personality scale are more cooperative, whereas "conscientious" people tend to be more reliable and determined (Costa & McCrae, 1992). Both should lead to a higher commitment to the survey, and have been associated with lower dropout rates (Lugtig, 2014; Richter, Körtner, & Saßenroth, 2014). In contrast, people with high levels of "extraversion" are reported to become easily bored and distracted (Costa & McCrae, 1992), leading to dropout. "Openness" also seems to have a robust effect on response propensity as people high in openness are considered to be more interested in new experiences and intellectually curious (Richter et al., 2014; Salthouse, 2014).

Respondents' survey attitude is also an important indicator for survey commitment (De Leeuw et al., 2010; Stocké, 2006). Respondents with a positive survey attitude—who think surveys are important and enjoy answering them—are less likely to attrite (Stocké, 2006). These respondents may place less importance on the burden of the survey, and more on the value of the survey.

Other predictors that we will use in this study are specific to our smartphone TUS: we expect that respondents who use their smartphone frequently are more willing to use this device for survey completion (De Bruijne & Wijnant, 2014; Mavletova, 2013) and that privacy concerns will prevent respondents from sharing GPS data.

In our analytical models to study nonresponse we will include socio-demographic, socio-psychological variables, survey attitudes and specific predictors for each task in order to study whether the correlates of nonresponse differ across the different tasks of the TUS. Following the leverage-saliency model, we expect the correlates to differ per task as the aspects of the request are also different. For example, worries about privacy may particularly influence the willingness to share GPS data, whereas busyness may be mainly related to interruptive pop-up questions, and survey burden to the most time-consuming part, the time use diary. Apart from looking at correlates of nonresponse, we will also look at bias in estimates of time use. As the smartphone TUS was conducted within an existing panel, we had some basic information on time use available for almost all sample members. We will test whether biases on these variables cancel each other out over the different stages, or reinforce each other.

3 Methods

3.1 Sample

In this study we used data from the LISS (Longitudinal Internet Studies for the Social sciences) panel, administered by CentERdata (Tilburg University, The Netherlands). The LISS panel started in 2007 and is the principal component of the project Measurement and Experimentation in the Social Sciences (MESS). The LISS panel consists of about 8000 individuals who complete online questionnaires every month. These questionnaires cover a large variety of domains including work, income, housing, time use, political views, values and personality. For more information about the LISS panel, see, Scherpenzeel and Das (2010).

The panel is based on a simple random sample of households drawn from the Dutch population register by Statistics Netherlands and aims to be representative of the Dutch population (Scherpenzeel, 2011; Scherpenzeel & Das, 2010). After the first sample was drawn in 2007, a refreshment sample was recruited between June and December 2009. Non-Internet households that could otherwise not participate are provided with a computer and Internet connection. Using the response metrics of Callegaro and Disogra (2008) the initial recruitment rate for the LISS panel was 63% and the profile rate 48% (Scherpenzeel, 2009). Retention is about 90% a year (Toepoel, 2013). See Appendix A for the demographical composition of the LISS panel. In 2012 and 2013, the smartphone TUS was administered to the LISS panel.

3.2 The Smartphone Time Use Survey

To study nonresponse and nonresponse bias, we examine response patterns in the Dutch Time Use Survey (TUS), developed and coordinated by the Netherlands Institute for Social Research (NL: Sociaal en Cultureel Planbureau [SCP]). In 2012 the SCP first conducted a small-scale test of their TUS through an app on a smartphone (Sonck & Fernee, 2013), after which the app was adapted and administered to the LISS panel.

Respondents in the TUS without a smartphone, or who preferred using a phone provided by LISS, could borrow one from the LISS panel: 52% of the participants in our study (n = 1120) used a borrowed phone. Respondents received an incentive of € 15 for every hour of participating. The research process of the TUS in the LISS panel consisted of several stages, listed here in chronological order:

Willing to participate Participants of the Dutch LISS panel were asked in three different surveys of the LISS panel—conducted in August 2012, March 2013 and July 2013—whether they were interested in participating in future smartphone surveys. Those who said yes to any of these three requests were considered for the TUS. See Appendix A for the demographics of the TUS sample.

Time Use Survey[1] To minimize possible seasonal influences on the time use data, data were collected for an entire year. Data collection started in September 2012. After this, each month a different batch of 176 panel members was invited (Sonck & Fernee, 2013). This resulted in a sample of 2154 participants in September 2013. People in this sample were invited for every stage of the TUS, even if they did not participate in the prior stage(s).

Pre-questionnaire Participants started by completing a pre-questionnaire on the web that mimicked the normally-used paper diary TUS. This pre-questionnaire contained more than 200 questions on various topics, like smartphone use, feeling in control about one's life, reasons to work certain hours, social support, child care, hobbies and household composition.

Diary About a week after the invitation to the pre-questionnaire, participants were asked to download an app, and complete a diary in which they recorded their activities on two randomly selected days; one weekday and one weekend day. The activities were predefined from a list of 41 categories following HETUS guidelines (Eurostat, 2009). Participants had to complete the full 24 hours (from 04:00 am to 04:00 am the next day) using ten-minute intervals. See Figure 1 for a screenshot of the app and time use diary. The left panel shows the screen where activities were reported by a hypothetical respondent, the middle panel shows an overview of the respondent's set of recorded activities, the right panel shows the questions that were asked using Experience Sampling (see 4. Pop-up questions below). The app was available for iOS and Android users.

When participants failed to complete one of the two days they were assigned to, they were invited to participate on a third day. This third day was exactly one week after the first weekday. We coded participants as respondents in this task when they filled out the diary for at least one day.

Pop-up questions On the same days as the diary, participants received six pop-up questions which were sent at random times of the day between 8:00 am and 10:00 pm. The pop-up question would show up on the screen for ten minutes. After this ten minute interval the question disappeared and could not be answered anymore to ensure real-time feelings were measured. These pop-up questions asked respondents either about their emotional state or smartphone use in the past hour. See Figure 1 for a screenshot of the pop-up questions. We coded participants as respondents in this task when they answered at least one pop-up question.

Sensor Data Participants were asked permission to passively record additional data through the app. These data included communication data (number of incoming and outgoing calls and text messages) and GPS locations. When downloading the app, participants were asked permission for this passive data collection. By default, the GPS tracker was turned on, but respondents could turn this off any time. We could not directly observe who turned off the GPS tracker and when this happened; we can only observe the number of GPS data points available per participant. The distribution of GPS data points showed a clear peak around 576–600 GPS data points for a two-day period implying that under normal conditions one location measurement was taken every five minutes. Participants with fewer than 576 data points were assumed to have turned off their GPS tracker or phone at some point during the study and were treated as nonrespondents in this stage. We performed sensitivity analyses treating everyone with at least 278 GPS point as respondents, but found no differences with the results we present below.

Post Questionnaire After all smartphone tasks were completed, respondents completed another questionnaire on the web. This post-questionnaire had the same design as the pre-questionnaire and contained about 180 questions on various topics, like Internet use, use of social networks, and family life.

3.3 Instruments

We used a variety of background variables from previous waves of the LISS panel to predict nonresponse. As most of these variables are measured annually, we used the data from respondents that were recorded closest before the start of the TUS.

Sociodemographic characteristics We used a set of sociodemographic characteristics: gender, age, net income, highest level of education (7 categories), and number of children living with the respondent.

Personality Five personality factors were computed: openness, conscientiousness, extraversion, agreeableness and neuroticism. These five factors are based on the Big Five, a taxonomy for describing the basic dimensions of personality (Costa & McCrae, 1992). We used the self-rating Big Five Questionnaire (Goldberg et al., 2006). This questionnaire consists of fifty items on which respondents must rate how they apply to them on a five-point scale. See Appendix B for all question wordings and results of our factor analyses.

Survey attitude The LISS panel contained nine questions about one's general attitude towards surveys. These items asked the participants for example whether they think surveys are important for society, and exhaustive to answer. Three factors for survey attitude were computed; survey enjoyment, survey value and survey burden (De Leeuw et al., 2010).

Privacy Two factors regarding privacy concerns were computed; trust and worries. The factor trust covers three questions about how much participants trust different organizations to keep their personal information private. The factor worries covers two questions about how worried participants are about their privacy. These questions were part of a survey conducted in July and August 2008. For that reason, data were not available for all respondents.

Smartphone use The factor smartphone use is based on questions in the pre-questionnaire of the TUS. Participants were asked whether they used their smartphone for 22 internet activities, for example for watching television, surfing the web and sending tweets. We calculated a factor score based on these 22 activities, see Appendix B. Participants were also asked to report for themselves how often they used mobile Internet on their phone. The correlation between this measure and the factor 'smartphone use' was $r = 0.733$.

Participation history We calculated the proportion of surveys panel members participated in relative to the number of invitations they received over the course of participation in the LISS Panel.

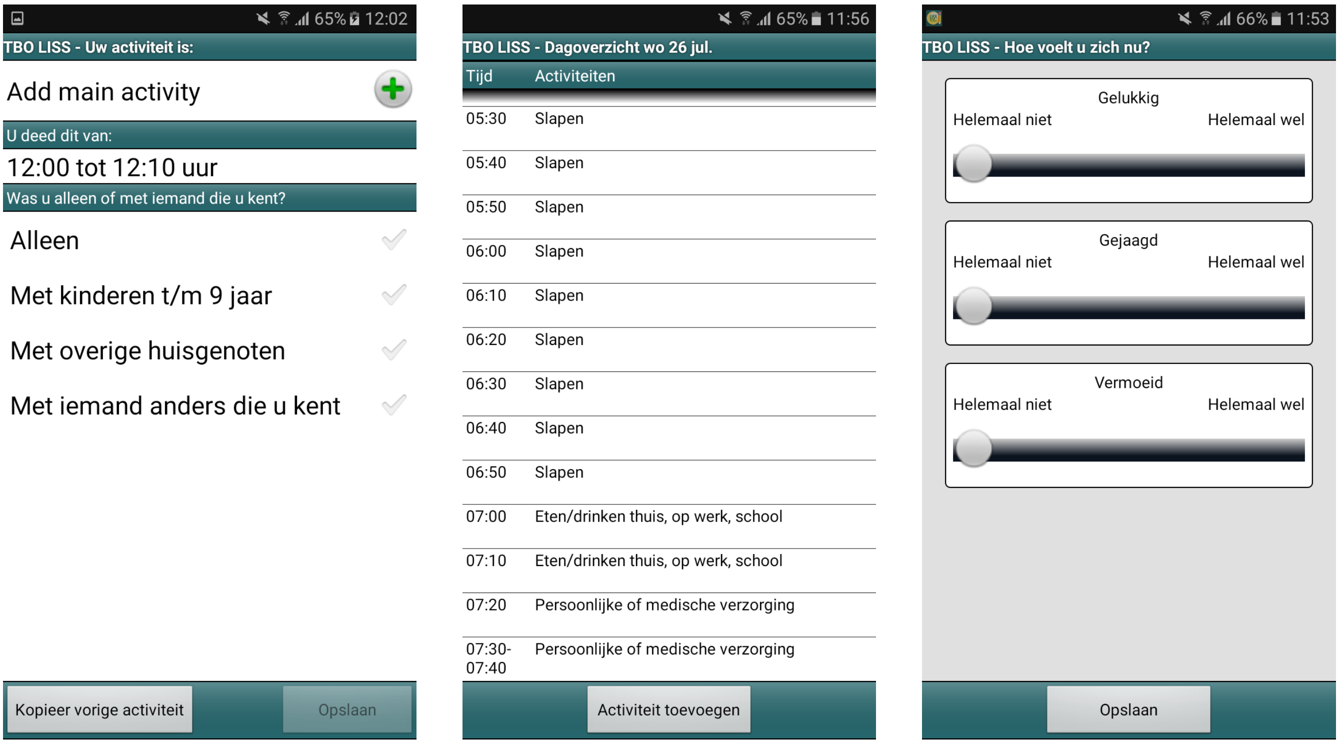
Prior decision Prior decision is a stage-specific measure of continuous participation. We added participation (0 = no, 1 = yes) in the previous stage of the smartphone TUS as a predictor for the subsequent stages. For example, participation in the pre-questionnaire is the prior decision for participation in the TUS diary.

3.4 Missing data

Missing data on covariates were imputed using the EM-algorithm in the Missing Values Analysis module in SPSS 24.0 (IBM Corp., 2016). 14.87% of the cases were complete. There were 77 different missing data patterns. By far the largest group had missing data only on variables that made up the factor score "privacy". 44.5 % of the cases had a missing value on privacy. Sociodemographic variables were available for everyone, except for income (5.7%), urbanization (1.1%) and educational level (1.0%).

3.5 Analyses

Our first, descriptive objective was to see how many participants participate in every stage, how many drop out and



(a) The screen where activities were reported. TUS LISS—Your activity is: add main activity. You did this between: 12:00 and 12:10. Were you alone of with someone you know? Alone/ With children up to 9 years old / With other family members / With someone else you know. The buttons below in the screen show "Copy previous activity"; and "Save".

(b) The day overview of one set of recorded activities. TUS LISS—Day overview Wednesday 26 July. Time & Activities: 05:30 Sleeping, 05:40 Sleeping, 05:50 Sleeping, 06:00 Sleeping, 06:10 Sleeping, 06:20 Sleeping, 06:30 Sleeping, 06:40 Sleeping, 06:50 Sleeping, 07:00 Eating/Drinking at home, at work/school, 07:10 Eating/Drinking at home, at work/school, 07:20 Personal or Medical care, 7:30–7:40 Personal or Medical care. The button below in the screen shows "Add activity".

(c) The three pop-up questions, asking: TUS LISS—How do you feel at this moment? Happy, Rushed, Tired. The scale labels are: Not at all—Extremely. The button below in the screen shows "save".

Figure 1. Screenshots of the TUS app

how many return. Second, we predicted nonresponse in every stage of the smartphone TUS, using the covariates described in the section above. We ran multivariate, logistic regression models with response (0 = nonresponse, 1 = response) as the dependent variable. We also included participation history and prior decision in our logistic regression analyses hierarchically to investigate the effect of continuous participation. Third, we investigated nonresponse bias to discover if and how nonresponse influences the survey estimates.

All models were estimated using R 3.4.1 (R Core Team, 2017). We conducted several multivariate logistic regression analyses and calculated Average Marginal Effects (AME) with the R-Package "mfx" (Fernihough, 2014). 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). We report AME instead of odds-ratios because odds-ratios reflect unobserved heterogeneity (Mood, 2010). Unobserved heterogeneity 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. AME's are not affected by unobserved heterogeneity and can thus be compared across models and groups. We transformed the AME's into percentages in the tables we present, to make them easier to interpret.

4 Results

Figure 2 shows the responses at different stages of the TUS. Only participants who said that they were willing to participate were invited for stage 1, the pre-questionnaire. The reported numbers in the squares represent the participants who start in that specific stage.

At three moments in 2012 and 2013, 7296 participants of the LISS panel received the question whether they would be willing to participate in a smartphone survey. Participants who answered they were not willing to participate were treated as refusals, while participants who did not answer this question at all were treated as noncontacts. Following the AAPOR 2006 guidelines, we observe a noncontact rate of 16.0% (n = 1168), and a refusal rate of 41.1% (n = 2996). 42.9% of all respondents (n = 3132) said they would be willing to participate in our smartphone study (Callegaro & Disogra, 2008).

In Figure 2, the numbers in the arrows represent the response probabilities for the subsequent stages, conditional on whether the respondent participated in the preceding stage. The arrows between stages represent consistent participation. For example, 85.6% of the participants who complete the pre-questionnaire also complete the diary, and 74.0% of the participants who complete the diary share GPS data. The arrows that do not connect boxes represent participants returning after missing a previous stage. For example, 32.0% of the nonrespondents in the pre-questionnaire fill out the diary, and 0.2% of the nonrespondents in the diary share GPS data.

The flowchart (Figure 2) shows that at every stage participants drop out. This results in a smaller number of participants at every subsequent stage. An exception forms the stage of GPS sharing, which has the lowest number of participants (n = 1193). A relatively large group misses one stage, but then returns to complete the next stage. This is especially so for the diary and post-questionnaire.

4.1 Willing to participate

In order to identify who is willing to participate in the smartphone survey, we ran a multivariate logistic regression model. In this analysis we excluded the noncontacts[2] at the invitation stage since we are uncertain whether they would be willing to participate. Table 1 shows the result of this multivariate logistic regression model (e.g. 0 = not willing, 1 = willing). Age, educational level, extraversion, conscientiousness, openness, survey value, enjoyment and burden, worries about privacy, and smartphone ownership are all significant predictors of being willing to participate in smartphone studies. For example, if people own a smartphone, their probability of participating increases by 21.5% conditional on the other covariates in the model. Furthermore, a one year increase in participant age decreases the probability to participate by 0.55%. Participation history, added to the model to investigate the additional effect of respondents' prior commitment to LISS panel, does not have a significant effect. The results of the model without controlling for participation history are reported in Appendix C.

Table 1
Average Marginal Effects for Predicting Willingness to Participate

	AME	Std. Err.
Sociodemographics		
Gender	−1.84	1.57
Age	−0.55***	0.06
Educational level	3.74***	0.51
Number of children	1.04	0.69
Income	−0.02	0.02
Personality		
Neuroticism	0.07	0.76
Extraversion	−2.08**	0.76
Agreeableness	−0.91	0.84
Conscientiousness	−4.41***	0.77
Openness	3.42***	0.77
Survey Attitude		
Survey value	3.81***	1.15
Survey enjoyment	10.96***	1.21
Survey burden	−2.62*	1.17
Privacy		
Trust	1.41	0.95
Worries	−3.69***	0.85
Smartphone Use		
Smartphone Ownership	21.51***	1.88
Continuous Participation		
Participation History	4.33	3.97
Nagelkerke R^2	0.19	

$* \, p < 0.05$ $** \, p < 0.01$ $*** \, p < 0.001$

4.2 Participation in the Time Use Survey

To assess who does and who does not participate in the different stages of the smartphone TUS, we ran separate multivariate logistic regression analyses per stage (0 = not willing, 1 = willing). To investigate the effect of continuous participation and to control for effects of the previous stage, we added prior decision and participation history next to sociodemographic, socio-psychological and smartphone specific predictors in our model. The results of these final multivariate logistic regression analyses are presented in Table 2.

[2] When comparing the socio-demographic characteristics of the contacts and noncontacts it appeared that noncontacts are on average younger and more likely to live in a household with children.

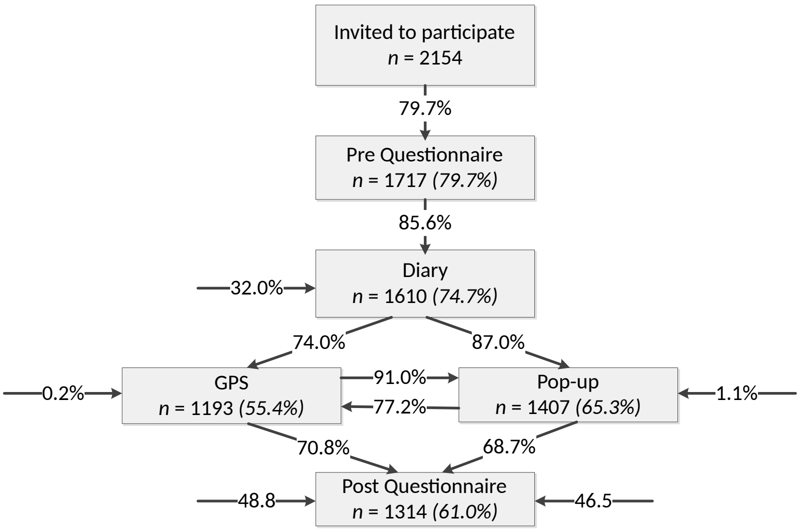


Figure 2. Flowchart of response behavior in the smartphone TUS. The percentages indicate the response probability per stage, dependent on participation (arrow from one stage to another) or no participation (arrow from the outside to the square) in the preceding stage.

The results of the models without controlling for continuous participation are reported in Appendix C.

First, we look at the sociodemographic predictors across all stages. Sociodemographic variables have no effect on participating in the pre-questionnaire. Willingness to participate in the smartphone parts of the study declines with age. The willingness to fill out the diary increases 1.54% per educational level. All other effects of sociodemographic variables across the stages are nonsignificant and/or small.

For the socio-psychological variables we find no large effects. Willingness to participate in the pre-questionnaire is higher for respondents who are more conscientious. With every increase of one standard deviation in the factor score conscientiousness the probability to participate increases by 1.95%. Willingness to share GPS data is larger for respondents who are more introvert.

Furthermore, smartphone use significantly predicts sharing GPS data. Participants who use their phone more often are more likely to be willing to share their GPS data. No variables related to survey attitude or privacy have an effect in any of the stages.

Finally, we look at the effects of continuous participation. Respondents prior commitment to the LISS panel, measured by participation history, increases respondents' willingness to fill out the pre-questionnaire or diary. Participating in the diary seems a strong predictor for the two subsequent phases. Participants for the diary have a 86.69% higher probability of being willing to answer pop-up questions, and 72.97% to share GPS data. Prior decision does not have a large effect on the post-questionnaire, but filling out the diary increases the willingness with 19.54% and sharing GPS data with 8.86%.

The explanatory power of our final model is particularly high for the smartphone parts of the study. The high explained variance for the pop-up (0.706) and GPS stage (0.554), in combination with the large effect of participating in the diary, implies that (non)response in these stages is mostly conditional on the previous stage.

4.3 Nonresponse bias

Apart from looking at the characteristics of those who respond and not respond, the final goal of this paper was to see how nonresponse matters for substantive statistics. The goal of the TUS is to estimate time use. To investigate the bias in these estimates we compared the time use of respondents and nonrespondents. We derived our estimates of time use for both groups from the "Social Integration and Leisure" and "Work and Schooling" questionnaires of the LISS panel. The "Social Integration and Leisure" study was conducted in February/March, and "Work and Schooling" in April/May 2012, roughly a year before the TUS. Respondents' time use could have changed in the meantime, but we find it safe to assume that over the whole sample on average no shifts oc-

Table 2
Average Marginal Effects for Participants' Willingness to Participate in the Different Stages.

	Pre Questionnaire		Diary		Pop-up		GPS		Post Questionnaire	
	AME	Std. Err.	AME	Std. Err.	AME	Std. Err.	AME	Std. Err.	AME	Std. Err.
Sociodemographics										
Gender	−1.14	2.02	−3.10	2.27	−4.24	4.42	4.44	3.00	1.25	2.58
Age	−0.06	0.07	−0.41***	0.08	−0.81***	0.16	−0.06	0.11	0.08	0.09
Educational level	−0.46	0.69	1.54*	0.78	−0.32	1.46	−0.52	1.00	1.05	0.88
Number of kids	0.28	0.76	0.83	0.90	0.16	1.74	−1.65	1.13	−0.63	0.99
Income	0.19	0.11	0.12	0.12	0.42	0.25	0.21	0.17	0.10	0.13
Participation History	33.14***	4.29	13.04*	5.34	10.87	11.83	6.90	7.21	8.32	6.25
Personality										
Neuroticism	0.36	0.95	0.44	1.07	−0.92	1.99	−2.02	1.38	0.59	1.21
Extraversion	−1.67	0.92	−1.15	1.05	0.37	1.96	−4.05**	1.38	−0.79	1.17
Agreeableness	1.05	1.01	1.63	1.12	−0.61	2.16	−1.12	1.47	1.44	1.28
Conscientiousness	1.95*	0.91	1.43	1.05	−1.02	1.99	2.37	1.34	0.82	1.19
Openness	−1.13	0.93	−1.47	1.05	−0.45	1.97	0.89	1.33	−0.78	1.19
Survey Attitude										
Survey value	−0.09	1.46	2.21	1.60	3.94	3.13	1.11	2.12	0.15	1.84
Survey enjoyment	−0.36	1.50	−1.91	1.67	−7.40*	3.17	−3.94	2.18	0.90	1.89
Survey burden	−0.54	1.52	−2.01	1.69	−0.19	3.08	−1.00	2.16	−0.67	1.95
Privacy										
Trust	0.53	1.17	1.40	1.30	−1.09	2.37	1.47	1.65	0.92	1.48
Worries	−0.69	1.03	1.94	1.11	0.42	2.07	−0.02	1.45	0.48	1.28
Smartphone Use										
Smartphone use	−0.18	1.09	0.97	1.34	3.26	2.33	4.25**	1.57	−2.65	1.41
Continuous Participation										
Participation History	33.14***	4.29	13.04*	5.34	10.87	11.83	6.90	7.21	8.32	6.25
PD[a]: Pre-Questionnaire	-	-	53.84***	2.54	10.45	6.37	10.89**	3.71	1.45	3.14
PD: Diary	-	-	-	-	86.69***	1.05	72.97***	1.23	19.54***	4.56
PD: Pop-up	-	-	-	-	-	-	-	-	3.15	3.86
PD: GPS	-	-	-	-	-	-	-	-	8.86**	2.93
Nagelkerke R^2	0.077		0.326		0.706		0.554		0.120	

[a] PD = "Prior decision"
* $p < 0.05$ ** $p < 0.01$ *** $p < 0.001$

curred. We excluded respondents for whom data were missing $n = 618$). For our analyses of nonresponse bias, we looked at four distinctive groups of respondents:

1. Respondents who said they were not willing to participate ($n = 2601$).
2. Respondents who said they were willing to participate, but never did ($n = 130$).
3. Respondents who participated in the TUS, but only in the pre- and/or post-Web questionnaire, not in any smartphone parts (pop-up, diary, GPS) ($n = 278$).
4. Respondents who participated in all tasks ($n = 622$).

Table 3 shows how many hours per week respondents on average spend on several activities. These activities are working, watching television, doing volunteer work, doing sports, going out (to bars, restaurants, cinema), making music, going to the theater (ballet, plays or musical) and doing creative things. In addition, we calculated the absolute and relative nonresponse bias.

An analysis of variance (ANOVA) test revealed that there are no differences between the groups in time spent on volunteer work, music, doing sports, theater or creative activities. The groups do differ significantly on time spent working ($F(3, 3627) = 18.49$, $p < 0.001$, $\eta_p^2 = 0.015$), watching television ($F(3, 3627) = 15.25$, $p < 0.001$, $\eta_p^2 = 0.012$), and going out ($F(3, 3627) = 4.60$, $p = 0.003$, $\eta_p^2 = 0.004$).[3]

[3]When we included all cases, also those including missings on some items of time use from earlier LISS surveys used to assess nonresponse bias, most results did not differ. The only differ-

The relative and absolute bias of watching television and working are also high, 3.70 (18.3%) and 4.96 hours (30.1%). The relative bias of theater and creativity are also high (47.8%), but this is due to very low prevalence of these activity in general. The absolute bias is only 0.04 and 0.08 hours, which is less than five minutes.

Post-hoc analyses reveal that only certain groups differ. First, full respondents who participated in all tasks of the TUS, appear to work significantly more hours than participants of the other three groups. The groups with (partial) nonrespondents do not differ from each other in the time spent on working. Second, full respondents watch significantly less television than the group that was not willing to participate at all (group 1) and the group that participated only in the non-smartphone parts (group 3). Finally, the full respondents (group 4) and respondents who were not willing to participate (group 1) go out less often than the other two groups (group 2 and 3).

5 Discussion and Conclusion

This paper shows how nonresponse differs over different tasks in a smartphone survey. This study is a first step into building a methodological framework for understanding nonresponse error in smartphone surveys. We know little about whether people are willing to perform smartphone-specific survey tasks and how this affects data quality.

Results show that 42.9% of the LISS panelists are willing to participate, and that from this group of willing respondents 74.7% actually complete the smartphone TUS. Only 28.9% of the willing respondents complete all the tasks of the study. The basic response rate is comparable to other, offline time use surveys, although we do not take nonresponse in the recruitment of LISS into account (Abraham et al., 2006; Van Ingen et al., 2008).

Predictors of nonresponse differ per task. However, some variables consistently predict nonresponse in every stage. Being younger, more conscientious, more open and more introvert increase the response probability of participating in every task in the smartphone TUS. These results are consistent with other, offline Time Use Studies (e.g. Abraham et al., 2006; Stoop, 2005; Van Ingen et al., 2008) and other longitudinal studies (Costa & McCrae, 1992; Lugtig, 2014; Richter et al., 2014). Smartphone ownership is an important predictor for being willing to participate in a smartphone study. Even though respondents could borrow a smartphone to participate, many participants were unwilling to do so. Haan, Lugtig, and Toepoel (2019) showed that device familiarity is an important predictor for using a particular device to complete a survey. This is confirmed in the actual TUS, where participation in some of the smartphone parts is predicted by age, rather than smartphone use. We conjecture that the respondent's attitude towards and familiarity with the specific functions of the smartphones are the main determinants of willingness to participate, rather than actual frequency of smartphone use, or age. Providing equipment is probably not enough to warrant participation in smartphone studies. In this study, the LISS panel tried to increase familiarity by showing an instruction manual and video. This may be a possible way to improve participation, but it is not very powerful as many participants did not view this video. Using interviewers might be a more powerful tool to increase device familiarity or to ensure everyone sees the video.

When panel members participated in the prior stage of the TUS, they are more likely to participate again. This finding can be explained by the foot-in-the-door technique (Cialdini, 1993), where a small initial request increases compliance with the next, larger request. Our study was rather successful in this respect, probably because it started with a regular survey respondents were used to.

When we frame our results in light of the leverage-saliency model we see that respondents who have a more positive smartphone attitude are more likely to participate in the smartphone parts of the study than in the survey parts. This suggests that respondents make a thoughtful decision to participate. Most of the variance is however explained by continuous participation, not by aspects of the survey request of that specific task. Future research could test the leverage-saliency model more extensively, by varying the survey request wording per task.

Nonresponse in itself may not be a problem, as long as it does not influence the survey estimates (Groves, 2006). However in this study, nonresponse does influence the survey estimates and therefore induces nonresponse bias. A specific group participates in the smartphone parts; this group works more and watches less TV than the average LISS panel participants. The respondents who only complete the pre- or post-questionnaires are more similar to those who do not participate in our study at all than like the full respondents. These results also replicate the results of Van Ingen et al. (2008) and Abraham et al. (2006), who also found that busy people are more likely to participate in an offline TUS. According to Stoop (2005) busy or working people are more involved in society. This involvement may lead to a higher probability to participate, but probably also leads to a more positive smartphone attitude that might be work-related. Future research could shed more light on this relation and the occurrence of nonresponse bias specifically in the smartphone parts.

Unfortunately, we were not able to see directly who turned off GPS tracking. We coded participants with few GPS data points as nonrespondents as we assumed they turned off their GPS tracker. However, it is also possible that these people

ence is that group 1 and 3 now differ significantly on doing sports ($F(3, 3930) = 2.95$, $p = 0.031$, $\eta_p^2 = 0.002$). By excluding incomplete cases, we excluded $n = 148$ for group 1, $n = 17$ for group 2, $n = 29$ for group 3 and $n = 35$ for group 4.

Table 3
Average Time Spent on several Activities by Four Distinctive Groups of (Non)respondents, and the Absolute and Relative Bias induced by Nonresponse..

	Not willing	Willing, no paticipation	Participation in non-smartphone parts	Full Respondents	Sample Mean	Absolute Bias (hours)	Relative Bias (in %)
Work	15.35	15.38	16.27	21.42	16.46	4.96	30.13
Volunteer Work	1.84	1.29	1.66	1.68	1.78	0.10	5.63
Watching TV	21.22	19.21	19.94	16.56	20.26	3.70	18.26
Sports	1.95	2.38	2.47	2.07	2.03	0.04	1.97
Going out	1.20	1.61	1.73	1.12	1.24	0.12	9.68
Music	0.12	0.03	0.18	0.11	0.11	0.00	0.00
Theater	0.11	0.04	0.09	0.05	0.09	0.04	47.78
Creativity	0.58	0.60	0.88	0.51	0.59	0.08	13.53

The absolute and relative bias are calculated by comparing the Full Respondents (respondents) to the Sample Mean (full sample), as described by Groves and Peytcheva (2008).

had their phones switched off during data collection, and only turned them on to complete the diary. Other causes may be technical problems, or empty batteries. Since it is difficult to pinpoint what the reasons are for the missing GPS data, it is difficult to predict how this introduces bias in the estimates. Future research should make this clearer.

A further limitation of our study is that we used respondents who were already participating in the LISS panel. The advantage of using a panel is the large amount of auxiliary variables available for predicting nonresponse and studying nonresponse bias. However, the experienced sample may be generally more willing to participate in surveys, and smartphone surveys in particular. The response rate in our study was comparable to earlier, offline time use surveys. If we take into account though the fact that LISS respondents are used to doing research, it is likely that repeating our smartphone study in an independent cross-sectional sample would result in a lower response rate than found in TUS conducted with paper diaries (Abraham et al., 2006; Van Ingen et al., 2008).

There is a long way to go before we can use smartphones as the sole data collection mode in general population studies. This study proves that we can conduct smartphone-app surveys with some success. Future research should focus on how to increase smartphone familiarity and on ways to convince people to do survey related tasks on smartphones. Consent surveys could shed more light into these issues. Smartphone surveys are promising tools for social research, but if we want to improve response rates and decrease response bias, there is still a lot of work to do on easing respondents into a task that is at least to some degree intrusive.

Acknowledgements

We would like to thank the Netherlands Institute for Social Research (NL: Sociaal en Cultureel Planbureau) and the LISS Panel for the design and collection of the data. In particular we want to thank Henk Fernee, Nathalie Sonck and Ineke Stoop.

References

Abraham, K., Maitland, A., & Bianchi, S. (2006). Nonresponse in the american time use survey: Who is missing from the data and how much does it matter? *Public Opinion Quarterly*, *70*, 676–703. doi:10.1093/poq/nfl037

Belli, R. F., Shay, W., & Stafford, F. (2001). Event history calendars and question list surveys: A direct comparison of interviewing methods. *Public Opinion Quarterly*, *65*, 45–74. doi:10.2307/3078785

Bouchard, C., Tremblay, A., Leblanc, C., Lortie, G., Savard, R., & Thériault, G. (1983). A method to assess energy expenditure in children and adults. *American Journal of Clinical Nutrition*, *37*, 461–467. doi:10.1093/ajcn/37.3.461

Callegaro, M. & Disogra, C. (2008). Computing response metrics for online panels. *Public Opinion Quarterly*, *72*, 1008–1032. doi:10.1093/poq/nfn065

Chatzitheochari, S., Fisher, K., Gilbert, E., Calderwood, L., Huskinson, T., Cleary, A., & Gershuny, J. (2018). Using new technologies for time diary data collection: Instrument design and data quality findings from a mixed-mode pilot survey. *Social Indicators Research*, *137*(1), 379–390. doi:10.1007/s11205-017-1569-5

Cialdini, R. (1993). *Influence: Science and practice* (3rd ed.). New York: HarperCollins Publishers Inc.

Cloïn, M., Van den Broek, A., Van den Dool, R., De Haan, J., De Hart, J., Van Houwelingen, P., . . . Spit, J. (2013). *Met het oog op de tijd: Een blik op de tijdsbesteding van Nederlanders. [In the interest of time: A look at the Dutch time use.]* The Hague: The Netherlands Institute for Social Research—SCP.

Costa, P. & McCrae, R. (1992). Normal personality assessment in clinical practice: The neo personality inventory. *Psychological Assessment*, *4*, 5–13. doi:10.1037/1040-3590.4.1.5

Cottrill, C. D., Pereira, F., Zhao, F., Dias, I., Lim, H., Ben-Akiva, M., & Zegras, C. (2013). Future mobility survey. *Transportation Research Record: Journal of the Transportation Research Board*, *2354*, 59–67. doi:10.3141/2354-07

De Bruijne, M. & Wijnant, A. (2014). Mobile response in web panels. *Social Science Computer Review*, *32*, 728–742. doi:10.1177/0894439314525918

De Leeuw, E., Hox, J., & Luiten, A. (2018). *International nonresponse trends across countries and years: An analysis of 36 years of Labour Force Survey data.* doi:10.13094/SMIF-2018-00008

De Leeuw, E., Hox, J., Lugtig, P., Vis, C., Göritz, A., Bartsch, S., & Engel, U. (2010). *Does familiarity breed contempt? Measuring and comparing survey attitude among new and repeat respondents cross-culturally.* Paper presented at the 63rd WAPOR Conference, Chicago, IL.

Dufau, S., Duñabeitia, J., Moret-Tatay, C., Mcgonigal, A., Peeters, D., Alario, F., … Grainger, J. (2011). Smart phone, smart science: How the use of smartphones can revolutionize research in cognitive science. *PLoS ONE*, *6*(9). doi:10.1371/journal.pone.0024974

Ermes, M., Parkka, J., Mantyjarvi, J., & Korhonen, I. (2008). Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *IEEE Transactions on Information Technology in Biomedicine*, *12*, 20–26. doi:10.1109/TITB.2007.899496

Eurostat. (2009). *Harmonised european time use surveys (HETUS): Guidelines 2008.* Luxembourg: Office for Official Publications of the European Communities.

Fan, W. & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior*, *26*, 132–139. doi:10.1016/j.chb.2009.10.015

Fernihough, A. (2014). Mfx: Marginal effects, odds ratios and incidence rate ratios for GLMs. R package version 1.1. Retrieved from https://CRAN.R-project.org/package=mfx

Fisher, K. & Gershuny, J. (2013). The 2014–2015 United Kingdom Time Use Survey. *10*, 96–97. doi:10.13085/eIJTUR.10.1.91-111

Galesic, M. (2006). Dropouts on the web: Effects of interest and burden experienced during an online survey. *Journal of Official Statistics*, *22*, 313–328.

Gershuny, J. (2012). Too many zeros: A method for estimating long-term time-use from short diaries. *Annals of Economics and Statistics*, *105/106*, 247–270. doi:10.2307/23646464

Goldberg, L., Johnson, J., Eber, H., Hogan, R., Ashton, M., Cloninger, R., & Gough, H. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, *40*, 84–96. doi:10.1016/j.jrp.2005.08.007

Groves, R. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly. 70*, 646–675. doi:10.2307/4124220

Groves, R., Cialdini, R. B., & Couper, M. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, *56*, 475–495. doi:10.1086/269338

Groves, R. & Heeringa, S. (2006). Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *169*, 439–457.

Groves, R., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: Description and an illustration. *Public Opinion Quarterly*, *64*, 299–308. Retrieved from http://www.jstor.org/stable/3078721.

Haan, M., Lugtig, P., & Toepoel, V. (2019). Can we predict device use? An investigation into mobile device use in surveys. *International Journal of Social Research Methodology*, *22*(5), 517–531. doi:10.1080/13645579.2019.1593340

IBM Corp. (2016). *IBM SPSS statistics for Windows, version 24.0.* Armonk, NY: IBM Corp.

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. doi:10.1007/s11301-014-0111-y

Knulst, W. & Van den Broek, A. (1999). Do time-use surveys succeed in measuring "busyness"? Some observations of the Dutch case. *Loisirs & Société*, *21*, 563–572. doi:10.1080/07053436.1998.10753671

Lai, J., Vanno, L., Link, M., Pearson, J., Makowska, H., Benezra, K., & Green, M. (2010). Life360: Usability of mobile devices for time use surveys. *Survey Practice*, *3*. Retrieved from http://www.surveypractice.org/index.php/SurveyPractice/article/view/120.

Lemay, M. (2010). *Understanding the mechanism of panel attrition.* Unpublished Doctoral thesis, Doctor of Philosophy, University of Maryland, College Park, MD.

Lipps, O. (2009). Attrition of households and individuals in panel surveys. SOEPpapers on Multidisciplinary Panel Data Research, 164. Retrieved from http://hdl.handle.net/10419/150711

Lugtig, P. (2014). Panel attrition: Separating stayers, fast attriters, gradual attriteres, and lurkers. *Sociological Methods & Research*, *43*, 699–723. doi:10.1177/0049124113520305

Mavletova, A. (2013). Data quality in PC and mobile web surveys. *Social Science Computer Review*, *31*, 725–743. doi:10.1177/0894439313485201

Michelson, W. (2005). *Time use: Expanding the explanatory power of the social sciences.* Boulder, Colorado: Paradigm Publishers.

Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*, *7*, 221–237. doi:10.1177/1745691612441215

Minnen, J., Glorieux, I., Van Tienoven, T., Daniels, S., Weenas, D., Deyaert, J., . . . Rymenants, S. (2014). Modular online time use survey (MOTUS)—translating an existing method in the 21st century. *Electronic International Journal of Time Use Research*, *11*, 73–93. doi:10.13085/eIJTUR.11.1.73-93

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. doi:10.1093/esr/jcp006

R Core Team. (2017). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Raento, M., Oulasvirta, A., & Eagle, N. (2009). Smartphones: An emerging tool for social scientists. *Sociological Methods & Research*, *37*, 426–454. doi:10.1177/0049124108330005

Revilla, M., M. & Couper. (2017). *Willingness of online panelists to perform additional tasks.* Paper presented at the General Online Research Conference, Berlin.

Richter, D., Körtner, J., & Saßenroth, D. (2014). Personality has minor effects on panel attrition. *Journal of Research in Personality*, *53*, 31–35. doi:10.1016/j.jrp.2014.08.001

Salthouse, T. (2014). Selectivity of attrition in longitudinal studies of cognitive functioning. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *69*, 567–574. doi:10.1093/geronb/gbt046

Scherpenzeel, A. (2009). Start of the LISS panel: Sample and recruitment of a probability-based internet panel. CentERdata, Tilburg. Retrieved from https://www.lissdata.nl/sites/default/files/bestanden/Sample%5C_and%5C_Recruitment.pdf

Scherpenzeel, A. (2011). Data collection in a probability-based internet panel: How the LISS panel was built and how it can be used. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, *109*, 56–61. doi:10.1177/0759106310387713

Scherpenzeel, A. & Das, J. (2010). ''true'' longitudinal and probability-based internet panels—research portal. In J. Das, P. Ester, & L. Kaczmirek (Eds.), *Social and behavioral research and the internet* (pp. 77–103). Boca Raton: Taylor & Francis.

Sonck, N. & Fernee, H. (2013). *Using smartphones in survey research: A multifunctional tool. Implementation of a time use app; a feasibility study.* The Hague: The Netherlands Institute for Social Research—SCP. doi:10.1007/s13398-014-0173-7.2

Statistics Netherlands (CBS). (2013). Tijdsbstedingsonderzoek 2011/2012 - onderzoeksdocumentatie. [time use survey 2011/2012 - research documentation.

Stocké, V. (2006). Attitudes toward surveys, attitude accessibility and the effect on respondents' susceptibility to nonresponse. *Quality and Quantity*, *4*, 259–288. doi:10.1007/s11135-005-6105-z

Stoop, I. A. L. (2005). *The hunt for the last respondent: Nonresponse in sample surveys.* The Hague: The Netherlands Institute for Social Research—SCP.

Stoop, I. A. L. (2007). No time, too busy. Time strain and survey cooperation. In G. Loosveldt, M. Swijngedouw, & B. Chambré (Eds.), *Measuring meaningful data in social research* (pp. 301–314). Leuven: Acco.

Subar, A., Kirkpatrick, S., Mittle, B., Zimmerman, T., Thompson, F., Bingley, C., . . . Potischman, N. (2012). The automated self-administered 24-hour dietary recall (ASA24): A resource for researchers, clinicians, and educators from the National Cancer Institute. *Journal of the Academy of Nutrition and Dietetics*, *112*, 1134–1137.

Thompson, F., Dixit-Joshi, S., Potischman, N., Dodd, K., Kirkpatrick, S., Kushi, L., . . . Subar, A. (2014). Comparison of interviewer-administered and automated self-administered 24-hour dietary recalls in 3 diverse integrated health systems. *American Journal of Epimediology*, *12*, 970–978. doi:10.1093/aje/kwu467

Toepoel, V. (2013). *Informing panel members about study results: Effects of traditional and innovative forms of feedback on participation.* Paper presented at the Workshop Longitudinal Research in Internet Panels, Mannheim.

Van Ingen, E., Stoop, I. A. L., & Breedveld, K. (2008). Nonresponse in the Dutch Time Use Survey: Strategies for response enhancement and bias reduction. *Field Methods*, *21*, 69–90. doi:10.1177/1525822X08323099

Appendix A

Demographical Composition of the LISS panel, Time Use Survey sample and the Dutch Population.

Table A1

Not-imputed, valid percentages of groups

	LISS panel[c]	TUS sample	Dutch population[b]
Age			
Average (in years)	40.6[a]	44.38	41.6[a]
	(21.98)	(17.09)	
65+	15.9	14.0	16.2
Gender			
Male	49.0	46.4	49.5
Female	51.0	53.6	50.5
Ethnicity			
Native Dutch Background	87.2	87.6	79.1
Migration Background	12.8	12.4	20.9
Urbanicity			
Extremely Urban	12.7	13.5	20.5
Very Urban	25.3	25.4	24.0
Moderately Urban	23.9	23.2	18.1
Slightly Urban	22.3	23.0	18.6
Not Urban	15.9	15.0	18.8
Household			
Single HH	28.3	17.4	36.8
Composition per Household			
Couple with children	34.5	44.3	27.3
Couple without children	30.2	30.4	29.2
Single with children	5.6	6.7	6.8
Other	1.4	1.2	0.6

[a] Average age reported as mean instead of percentages. [b] Dutch population statistics correspond to individually based statistics, with January 1st 2012 as the base date. Statistics can be found on http://statline.cbs.nl (Statistics Netherlands, 2012).

[c] LISS panel composition of September 2012.

Appendix B
Factor Scores for the Different Constructs.

Table B1
Factors Sociopsychological Variables.

	Factor				
	1	2	3	4	5
Neuroticism					
Get stressed out easily.	0.70	–0.02	0.05	–0.09	–0.12
Am relaxed most of the time.	–0.61	–0.01	0.01	0.19	0.16
Worry about things.	0.67	0.06	0.09	–0.21	–0.01
Seldom feel blue.	–0.52	–0.06	0.07	0.08	0.06
Am easily disturbed.	0.73	–0.04	0.08	–0.13	–0.09
Have a soft heart.	0.39	0.00	0.26	0.00	0.01
Get upset easily.	0.75	–0.08	0.07	0.05	–0.08
Change my mood a lot.	0.69	0.03	–0.02	0.14	0.07
Have frequent mood swings.	0.68	0.03	–0.02	0.13	0.04
Get irritated easily.	0.54	–0.03	–0.11	–0.05	0.07
Often feel blue.	0.73	0.01	–0.05	0.00	0.04
Extraversion					
Am the life of the party.	0.12	0.64	0.01	0.01	0.15
Don't talk a lot.	–0.07	–0.67	–0.07	0.05	0.10
Feel comfortable around people.	–0.06	0.48	0.23	0.00	0.01
Keep in the background.	0.01	–0.79	0.22	0.06	0.08
Start conversations.	0.00	0.67	0.16	–0.06	–0.02
Have little to say.	0.17	–0.51	–0.13	0.11	–0.01
Talk to a lot of different people at parties.	0.03	0.63	0.22	0.12	–0.05
Don't like to draw attention to myself.	0.00	–0.62	0.34	–0.05	–0.07
Don't mind being the center of attention.	0.01	0.57	–0.11	0.14	0.17
Am quiet around strangers.	0.14	–0.64	–0.08	0.03	0.11
Agreeableness					
Feel little concern for others.	0.02	–0.03	–0.51	0.12	0.06
Insult people.	0.18	0.09	–0.30	0.20	0.22
Sympathize with others' feelings.	0.09	–0.18	0.88	0.12	–0.01
Am not interested in other people's problems.	0.02	–0.02	–0.65	–0.03	0.01
Am not really interested in others.	0.09	–0.06	–0.68	–0.01	0.05
Feel others' emotions.	0.13	–0.04	0.69	0.15	0.14
Make people feel at ease.	0.11	0.31	0.38	–0.06	0.06
Take time out for others.	0.05	0.03	0.66	–0.02	0.04
Am interested in people.	–0.04	0.06	0.71	0.11	0.04

Continues on next page

Table B2
Continued from previous page

	Factor				
	1	2	3	4	5
Conscientiousness					
Am always prepared.	0.03	0.03	−0.09	−0.51	0.09
Leave my belongings around.	0.00	0.02	0.14	0.71	0.11
Make a mess of things.	0.30	−0.06	0.01	0.61	0.10
Get chores done right away.	0.07	0.10	−0.01	−0.55	−0.05
Like order.	0.26	0.00	−0.10	−0.76	0.05
Often forget to put things back in their proper place.	0.08	0.05	0.04	0.61	0.02
Shirk my duties.	0.26	0.06	−0.09	0.48	−0.07
Follow a schedule.	0.25	−0.03	0.03	−0.52	0.13
Spend time reflecting on things.	0.09	−0.12	0.10	−0.31	0.29
Am exacting in my work.	0.16	0.00	−0.03	−0.36	0.29
Openness					
Have difficulty understanding abstract ideas.	0.28	0.02	0.02	0.05	−0.43
Pay attention to details.	0.12	−0.03	0.13	−0.29	0.37
Have a vivid imagination.	0.11	0.15	−0.08	0.17	0.40
Am not interested in abstract ideas.	0.17	0.00	−0.05	0.02	−0.36
Have a rich vocabulary.	−0.10	−0.02	0.09	−0.01	0.48
Have excellent ideas.	−0.08	0.04	−0.02	−0.11	0.56
Do not have a good imagination.	0.25	0.00	−0.07	0.13	−0.28
Am quick to understand things.	−0.17	−0.13	0.07	−0.11	0.57
Use difficult words.	0.00	−0.07	−0.09	0.18	0.55
Am full of ideas.	0.05	0.19	0.02	−0.08	0.54

Standardized factor loadings. We used an EFA with Promax Rotation in IBM SPSS 24. The Cumulative Explained Variance is 38.53%. Factor correlations range between -0.285 and 0.370.

Table B3
Factors Survey Attitude.

	Factor		
	1	2	3
Survey Value			
Surveys are important for society.	0.80	−0.01	0.19
A lot can be learned from information collected through surveys.	0.79	−0.01	0.13
Survey Burden			
Completing surveys is a waste of time.	−0.40	0.51	0.06
I receive far too many requests to participate in surveys.	0.06	0.55	−0.05
Opinion polls are an invasion of privacy.	−0.13	0.55	0.11
It is exhaustive to answer so many questions in a survey.	0.15	0.53	−0.24
Survey Enjoyment			
I really enjoy responding to questionnaires through the mail or Internet.	0.12	−0.05	0.80
I really enjoy being interviewed for a survey.	0.07	−0.02	0.56
Surveys are interesting in themselves.	0.48	−0.04	0.50

Standardized factor loadings. We used an EFA with Promax Rotation in IBM SPSS 24. The Cumulative Explained Variance is 54.26%. Factor correlations range between -0.327 and 0.424.

Table B4
Factors Privacy.

	Factor	
	1	2
Trust		
How much do you trust each of the following to keep the information they collect from you confidential: public opinion research companies	1.01	0.05
How much do you trust each of the following to keep the information they collect from you confidential: market research companies	0.63	−0.03
How much do you trust each of the following to keep the information they collect from you confidential: government agencies, like Statistics Netherlands	0.49	−0.06
Worries		
In general, how worried are you about your personal privacy?	0.01	0.76
Different private and public organizations have personal information about us. How concerned are you about whether or not they keep this information confidential?	−0.05	0.81

Standardized factor loadings. We used an EFA with Promax Rotation in IBM SPSS 24. The Cumulative Explained Variance is 58.65%. Factor correlation is -0.230.

Table B5
Factor Smartphone Usage

	Factor 1
Please indicate whether you ever use a mobile phone for…	
Watching television	0.42
Watching films online	0.73
Listening to the radio	0.46
Listening to your own music	0.59
Reading news sites and daily newspaper	0.66
Reading magazines	0.46
Playing online games	0.46
Playing offline games	0.48
Emailing	0.71
Reading other people's twitter messages	0.47
Sending your own twitter messages	0.41
Visiting social media network sites	0.68
Visiting online forums or discussion groups	0.34
Sending short text messages via the Internet	0.69
Telephoning via the Internet or making video calls	0.39
Downloading music or video files	0.49
Uploading videos, photos or music	0.56
Online banking	0.57
Shopping or ordering goods via the Internet	0.48
For navigation services	0.65
To search specific information on the Internet	0.75
Just to surf around on the Internet	0.64

Standardized factor loadings. We used an EFA with Promax Rotation in IBM SPSS 24. The Cumulative Explained Variance is 31.56%.

Appendix C
Models without participation history or prior decision.

Table C1
Average Marginal Effects for Predicting Willingness to Participate.

	AME %	Std. err.
Sociodemographics		
Gender	-1.82^{***}	1.57
Age	-0.54^{***}	0.05
Educational level	3.83	0.50
Number of children	0.98	0.68
Income	-0.02	0.02
Personality		
Neuroticism	0.04	0.76
Extraversion	-2.14^{**}	0.76
Agreeableness	-0.97	0.84
Conscientiousness	-4.29^{***}	0.76
Openness	3.34^{***}	0.76
Survey Attitude		
Survey value	3.81^{***}	1.14
Survey enjoyment	11.05^{***}	1.21
Survey burden	-2.65^{*}	1.17
Privacy		
Trust	1.47	0.95
Worries	-3.69^{***}	0.85
Smartphone Use		
Smartphone Ownership	21.33^{***}	1.87
Nagelkerke R^2	0.190	

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

Table C2

Average Marginal Effects for Participants' Willingness to Participate in the Different Stages Without Prior Decision.

	Pre Questionnaire		Diary		Pop-up		GPS		Post Questionnaire	
	AME %	Std. Err.	AME %	Std. Err.	AME %	Std. Err.	AME %	Std. Err.	AME %	Std. Err.
Sociodemographics										
Gender	−1.38	2.02	−3.40	2.21	−4.21	2.47	0.55	2.58	0.31	2.51
Age	0.05	0.07	−0.28***	0.08	−0.49***	0.09	−0.21	0.09	0.01	0.09
Educational level	0.59	0.68	1.94**	0.74	1.71*	0.82	1.20	0.87	1.71*	0.84
Number of children	−0.05	0.77	0.61	0.86	0.56	0.96	−0.70	0.99	−0.58	0.96
Income	0.13	0.11	0.16	0.12	0.31*	0.14	0.26	0.14	0.14	0.13
Personality										
Neuroticism	−0.02	0.94	0.26	1.03	−0.06	1.14	−1.06	1.20	0.46	1.16
Extraversion	−2.09*	0.91	−2.17*	1.01	−1.90	1.11	−4.57***	1.17	−1.69	1.13
Agreeableness	0.69	1.00	1.70	1.09	1.17	1.21	0.53	1.28	1.69	1.24
Conscientiousness	3.04***	0.90	3.20**	0.99	2.79*	1.11	4.27***	1.17	2.11	1.13
Openness	−2.03*	0.92	−2.60**	1.01	−2.59*	1.12	−1.43	1.17	−1.64	1.14
Survey Attitude										
Survey value	0.04	1.43	1.87	1.53	2.83	1.73	2.00	1.84	0.79	1.77
Survey enjoyment 0.78	1.46	−0.87	1.58	−3.13	1.77	−3.30	1.87	0.55	1.81	
Survey burden	−0.65	1.51	−1.97	1.61	−1.84	1.81	−2.44	1.92	−1.29	1.87
Privacy										
Trust	0.91	1.15	1.80	1.24	1.22	1.38	2.34	1.46	1.55	1.42
Worries	−0.78	1.02	1.25	1.08	1.28	1.20	1.03	1.28	0.80	1.24
Smartphone Use										
Smartphone use	−0.51	1.08	0.26	1.20	1.37	1.34	3.33*	1.40	−2.18	1.35
Nagelkerke R2	0.035		0.041		0.055		0.044		0.027	

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