Willingness of Online Panelists to Perform Additional Tasks

Revilla, Melanie; Couper, Mick P.; Ochoa, Carlos

methods, data, analyses | Vol. 13(2), 2019, pp. 223-252 DOI: 10.12758/mda.2018.01

Willingness of Online Panelists to Perform Additional Tasks

Melanie Revilla[1], *Mick P. Couper*[2] & *Carlos Ochoa*[3]

[1] *RECSM-Universitat Pompeu Fabra, Barcelona, Spain*
[2] *University of Michigan, USA*
[3] *Netquest, Barcelona, Spain*

Abstract

People's willingness to share data with researchers is the fundamental raw material for most social science research. So far, survey researchers have mainly asked respondents to share data in the form of answers to survey questions but there is a growing interest in using alternative sources of data. Less is known about people's willingness to share these other kinds of data. In this study, we aim to: 1) provide information about the willingness of people to share different types of data; 2) explore the reasons for their acceptance or refusal, and 3) try to determine which variables affect the willingness to perform these additional tasks.

We use data from a survey implemented in 2016 in Spain, in which around 1,400 panelists of the Netquest online access panel were asked about their hypothetical willingness to share different types of data: passive measurement on devices they already use; wearing special devices to passively monitor activity; providing them with measurement devices and then having them self-report the results; providing physical specimens or bodily fluids (e.g. saliva); others. Open questions were used to follow up on the reasons for acceptance or refusal in the case of the use of a tracker.

Our results suggest that the acceptance level is quite low in general, but there are large differences across tasks and respondents. The main reasons justifying both acceptance and refusal are related to privacy, security and trust. Our regression models also suggest that we can identify factors associated with such willingness.

Keywords: online panel; respondent willingness; passive data collection; mobile data collection

The widespread adoption of digital technologies, especially those available on mobile devices, is expanding opportunities for survey researchers to enhance and extend survey measurement, whether through active or passive measurement (see Link et al., 2014). Much of the early research on exploiting these technologies for research has focused on small groups of volunteers. The challenge remains of using these features in the context of large-scale survey data collection. This paper extends that work by exploring stated willingness to provide a variety of types of additional information in the context of an opt-in panel in Spain. We explore willingness across different types of requests that vary in the level of effort required and the degree of intrusiveness, to investigate what additional tasks respondents find more or less acceptable. We also explore reasons for willingness or unwillingness to accept one particular task, installing software to passively track browsing behavior. Finally, we examine the correlates of willingness to accept these additional tasks.

Background

The expansion of the Internet and the development of a range of new active and passive measurement tools, particularly on mobile devices, present a number of potentially exciting opportunities for survey researchers. As the AAPOR (2014) task force noted, "there are a wide array of applications and features available on these devices which can augment and in some cases even replace survey data" (see also Link et al., 2014). The AAPOR report addressed five potential uses of technologies to extend or replace surveys: 1) location or geo-positioning, 2) scanning and QR/barcode readers, 3) visual data capture (photos or video), 4) Bluetooth enabled devices and related technologies, and 5) mobile applications or "apps". The report calls for "more assessments of auxiliary data collection capabilities," specifically in terms of "respondent cooperation and compliance, data quality, and potential

Acknowledgments

We thank Gerardo Ortiz for his work on the programming of the survey and Oriol Bosch for his help with coding the open questions. We also thank the Editor, Prof. Edith de Leeuw, and three anonymous reviewers for their very useful comments on a previous draft of this paper.

Direct correspondence to

Melanie Revilla, RECSM-UPF, Edifici Mercè Rodoreda 24.406, Ramón Trías Fargas, 25-27, 08005 Barcelona, Spain

E-mail: melanie.revilla@upf.edu

sources of error" (AAPOR, 2014, p. 9). This paper is focused on the first of these issues.

There are several potential advantages of these new measurement opportunities for supplementing survey data, whether on mobile devices or on PCs. These include 1) reducing respondent burden (i.e., replacing survey questions with passive measurement, or providing easier ways to share information), 2) improving the quality of measurement (i.e., obtaining data that respondents find difficult to report or recall accurately), and 3) measuring new things (i.e., enhancing and extending measurement into new domains).

While a number of papers have argued for the benefits of the enhanced measurement capabilities (see, e.g., Palmer et al., 2013; Raento, Oulasvirta, and Eagle, 2009; Wrzus and Mehl, 2015), much of the research to date has focused on relatively small samples of volunteers. A key challenge for the broader adoption of these new measurement tools in large-scale surveys relates to respondents' willingness to install apps, activate passive tracking, or do the additional tasks researchers ask of them. Further, for those tasks involving ongoing actions beyond the initial consent and installation, continued compliance with the request (or adherence to the protocol) is an additional concern. The mode in which the request is embedded may also be important: Burton (2016) reports a 34 percentage point lower consent rate to administrative record linkages among those surveyed online than those interviewed face-to-face in the *Understanding Society* Innovation Panel.

Several studies have begun to explore these issues and test the feasibility of such additional tasks in the context of ongoing surveys. Most of these studies focus on a single task or technology. For example, some have explored willingness to permit GPS capture. Armoogum and colleagues (2013) asked respondents in the 2007-8 French National Travel Survey (a face-to-face survey) about their willingness to use a GPS device. About one-third (30%) said yes without conditions, while 5% agreed as long as they could turn it off (the rest said no). Biler, Senk, and Winkerova (2013) asked respondents in a face-to-face survey in the Czech Republic about willingness to participate in a travel survey using GPS tracking. Only 8% said they were willing, with 25% uncertain, and 57% not willing. Joh (2017) reports on a pilot study using mail survey recruitment to a travel survey using a GPS-based smartphone app. Of those invited, 5.9% responded to the baseline survey. Of those who reported having a qualifying smartphone, 31.7% downloaded the app and provided some data (representing 1.3% of the original sample).

Turning to online recruitment, Toepoel and Lugtig (2014) asked Dutch panelists for the one-time capture of GPS coordinates: 26% of smartphone participants and 24% of PC participants agreed to such capture. In an online panel study of college students in the U.S., Crawford et al. (2013) reported that 58% said yes to a hypothetical question about GPS capture. In a subsequent wave, between 20% and 33% of survey respondents (depending on the consent condition) provided usable

GPS data. The LISS Mobile Mobility Panel in the Netherlands recruited panelists with smartphones to provide GPS data. Of those who completed the invitation survey (75% of invitees), 37% were willing to participate and 30% (81% of those willing) downloaded the app, activated Wi-Fi and GPS, and provided data for at least one day (Scherpenzeel, 2017). Invitations were restricted to those willing to use their smartphones for research; Antoun, Couper, and Conrad (2017) found that about 41% of LISS panelists were willing to use their smartphones for research.

Other studies have examined the installation of a research app. McGeeney and Weisel (2015) report on a study in the Pew American Trends Panel. Panelists who used an eligible smartphone were randomized to a browser- or app-based version of an experience sampling survey, in which they were asked to complete a short survey twice a day for 7 days. For those in the app group, 76% agreed, and 80% of those installed the app (i.e., 61% of those invited). Completion rates for the 14 surveys were significantly lower for the app group than the browser group. Johnson, Kelley, and Stevens (2012) explored a modular survey design that required installation of an app. Of the panelists who met the eligibility criteria (including use of a smartphone), 43% expressed willingness to do the modular survey and were sent a link to download the app. Of these, 37% (or 16% of qualified panelists) successfully downloaded the app, and 33% (14% of qualified panelists) completed one or more surveys.

A few studies have attempted the collection of passive tracking (e.g., browser log) data. For example, de Reuver and Bouwman (2015) tried to recruit participants from a Dutch online access panel. An initial random sample of the panel did not yield sufficient panelists willing to install the tracker. They then targeted panelists who had previously agreed to the collection of log data. Among these, 31% expressed willingness to allow capture of log data, 22% installed the app, and 14% participated for the full four weeks of the study. The primary reason for non-participation provided was related to privacy (16% of those who provided a reason), followed by a variety of situational factors (holidays, illness, etc.; 15%). Reasons for dropping out during the study were primarily related to technical issues such as battery drainage and reduced performance of the phone. Van Duivenvoorde and Dillon (2015) asked eligible respondents in an opt-in panel in the U.S. to participate in a follow-up study which required them to install passive tracking software. Among respondents who completed the baseline survey (32% of those invited), 3.6% expressed willingness and 2.1% installed the software. Kissau and Fischer (2016) similarly invited members of a Swiss panel to install tracking software. Of those invited, 23% of the main and 8% of the boost sample respectively expressed interest in the study, while 10% of the main and 3% of the boost sample installed the tracking app. In a similar study in Spain using the same tracking app (Wakoopa; see https://wakoopa.com/) Revilla, Ochoa, and Loewe (2017) reported that between 30% and 50% of loyal panelists who were invited agreed to install the tracker.

The capture of accelerometry data (often using stand-alone devices) is more common in large face-to-face surveys, with wide variation in agreement and compliance rates. For instance, Lauderdale et al. (2014) report an initial agreement rate of 80.3% for a sleep actigraphy study, with 88.4% of those who consented (69.8% of those invited) providing usable data. Roth and Mindell (2008) reported a similar consent rate of 80.3%, with 47.7% of those consenting (38.3% of the initial sample) providing usable data (for other examples, see Hassani et al., 2014; Menai et al., 2017; Gilbert et al., 2017). Howie and Straker (2016) conducted a review of trials involving accelerometer use among children, and reported compliance rates ranging from 2% to 60%.

We know of only two studies that has attempted accelerometry measurement in an online study. The FLASHE study in the U.S. (see https://cancercontrol.cancer.gov/flashe) recruited dyads of caregivers and their 12-17 year-old children participants from a commercial online access panel. Of those invited, 39% consented and were enrolled in the study. Of those who consented and were randomly assigned to the survey and accelerometer study, 59% completed the study (23% of invitees). In contrast, for those assigned to the survey-only group, 86% completed the study. Scherpenzeel (2017) reports a 57% willingness and 51% adherence rate to an accelerometry study in the Dutch LISS panel.

Even more intrusive biomarker measures are often used in face-to-face surveys (e.g., McFall, Conolly, and Burton, 2014; Sakshaug, Couper, and Ofstedal, 2010), or as a follow up to telephone surveys (e.g., Boyle et al., 2010; Gautier et al. 2016), but few studies have tested biomarker measures in the context of Internet surveys. For instance, while biomonitors are increasingly being used to study alcohol consumption among volunteers (see, e.g., Greenfield, Bond, and Kerr, 2015), we know of no studies that have tested this on general population samples, especially those with online participants. In one exception, Avendano, Scherpenzeel, and Machenbach (2011) undertook a small pilot in the LISS panel. Panelists were recruited for home cholesterol measurement, involving a finger-prick and blood spot measurement using a device designed for self-administration. Of the 200 panelists invited, 38 (19% of invitees) returned a blood sample, 31 of whom (16% of invitees) provided valid data. Another subsample was asked to chew on a cotton swab and return the saliva sample for cortisol measurement. Of the 200 invited, 30 (15% of invitees) completed the task.

Gatny, Couper, and Axinn (2013) tested the collection of saliva in an ongoing Internet diary study of young women. Saliva kits were mailed to 150 respondents who reported the end of a romantic relationship and were eligible to participate in the collection; 65% mailed back a saliva sample. Similarly, Etter and Bullen (2011) recruited 196 users of electronic cigarettes online, and mailed them a saliva kit: 16% mailed back a saliva sample.

Two other recent studies are relevant. Jäckle and colleagues (2017) invited panelists in the *Understanding Society* Innovation Panel to download an app to scan receipts and report their spending over 4 weeks; 16.5% of respondents downloaded the app and completed the registration survey, and 12.8% used it at least once. Similarly, Angrisani, Kapteyn, and Samek (2017) invited panelists in the Understanding America Study (UAS) to sign up to a customized financial aggregator website and provide financial information. Of those invited, 65% consented; 68% of those who consented (32% of those invited) signed up, and 38% of those (12% of those invited) linked one or more financial institutions.

This brief review of selected studies shows a wide range of stated or actual willingness across a variety of tasks and settings. Many of the studies report rates of compliance without looking at reasons behind the decision or examining differences between those who are or are not willing (some exceptions are reviewed below). Further, all these studies examine only a single request for additional data or technology use.

In one of the few studies to both explore reasons for unwillingness and examine socio-demographic correlates, Pinter (2015) asked members of an access panel in Hungary who used smartphones if they were willing to install a research app. In response to the initial request, 42% said they were unwilling, with a further 23% uncertain (the remaining 35% were willing). Those who were uncertain or unwilling to install the research app were asked their reasons for not being willing, in a series of closed questions. Major reasons proffered by this group included (multiple mentions possible): not enough free time (61%); not enough information to decide (53%); concerns about extra costs of using an app (45%); would participate in some research activities but not others (44%); and concerns about battery use (43%). After additional persuasion aimed at these concerns, 57% eventually agreed to install the app. Pinter also found that behavioral variables (frequency of smartphone use, number of apps, use of GPS, etc.) were moderately but significantly correlated with willingness. In addition, significant but weak correlations of willingness with other socio-economic variables (including political orientation, age, labor force status, income, and frequency of socialization) were found.

Armoogum et al. (2013) examined demographic correlates of willingness to use a GPS device. They found that younger persons, males, those in smaller households, with higher income, with a computer in the household, and with more cars were more willing to participate. Biler et al. (2013) reported that those who used a shopping or travel discount card were more willing to agree to GPS tracking, as were those who used navigation features on their smartphone, those who use social networks, younger persons, and those in larger households.

In a study among Netquest panelists in seven countries, Revilla and colleagues (2016) elicited panelists' willingness to do three additional tasks: 1) share GPS location, 2) install an app, and 3) take a photo. Focusing on the data from Spain, 36% of smartphone owners who responded to the survey said they were definitely willing to install an app, while a further 27% said they were probably willing, and only 15% said definitely no. They found consistent significant negative effects of age on tolerance for additional tasks across countries, but not for other variables (gender, education, and household size) in multivariable models.

Keusch et al. (2017) used a vignette approach to vary features of the request to install a tracking app in a study among opt-in panel members in Germany. Overall they found that 64.5% would *not* be willing (0-5 on an 11-point scale), and 34.9% would *definitely not* be willing (0 on the scale) to install a tracking app. Factors that affected willingness included the sponsor of the study, the length of time that the tracker would be used, the size of the incentives, and the ability to turn off the tracker.

Wenz, Jäckle, and Couper (2017) measured willingness to perform a variety of task in the *Understanding Society* Innovation Panel in the U.K. They found that willingness varied by task (e.g., 59.3% accelerometry capture, 36.7% GPS capture, and 25.5% tracking app). They found lower willingness for men, those with lower education, and those with higher security concerns. Jäckle and colleagues (2017) explored demographic and behavioral correlates of participation in a spending app study. They found that frequency of Internet and mobile device use, along with general cooperativeness with research, were predictors of participation in the app study.

This review of the emerging literature illustrates some of the challenges of exploiting the technical capabilities of modern technology to enhance and extend measurement. The studies reveal considerable variation in willingness to use new technologies for research. But because each study looks at only a single technology or data collection activity, it is hard to determine if this variation is due to the type of request or other features of the design (such as the mode in which the request is made, or the sample on which the study is based). Further, there are inconsistent findings with regard to socio-demographic correlates of willingness, which again may vary by task. Relatively little attention has focused on behavioral and attitudinal correlates of willingness.

This paper adds to this literature by examining panel members' stated willingness to perform a variety of additional tasks. This allows us to explore variation both between tasks and between respondents. We expect the key factors affecting the decision to accept a task include privacy concerns and the effort (or burden) required. Because of the concern for privacy, we expect tasks in which respondents have control of the information provided to have higher levels of acceptance than tasks in which the information is provided automatically to the panel company. However, because of the level of effort required, we expect passive measurement to have higher levels of acceptance than active measurement.

Our focus here is on *stated* willingness, rather than *actual* compliance with the request. We expect that actual compliance rates will be lower than those based on expressed willingness. However, research has shown that stated willingness is a useful measure in its own right, especially if the goal is to examine reasons for and covariates of (un)willingness (see, e.g., Couper and Singer, 2013; Couper et al., 2008, 2010).

Methodology

Data

We use data on the self-reported willingness to complete a variety of different tasks in a web survey. The data was collected from the 15th of September to the 3rd of October 2016 using the Netquest opt-in panel in Spain (www.netquest.com).

Since 2014, Netquest has invited selected panelists to install a tracker (or "meter") on the devices that they are using to go online (PCs, tablets or smartphones), and share (passively) with Netquest the information registered by this tracker (URLs of the web pages visited, time of the visits, ad exposure, and app use in the case of mobile devices) (see Revilla, Ochoa, and Loewe, 2016). In this paper, we only considered panelists who had not yet been invited to install the tracker. In addition, because the survey was also used for other experiments[1], it focused on panelists who have Internet access through both a PC and smartphone. Panel profile information was used to send the invitation to panelists meeting this criterion, and filter questions were used to verify such access. Cross quotas for age and gender were used to ensure that the distribution of these variables in the sample was similar to that observed in the full panel.

The survey contained a maximum of 69 questions. Respondents were able to proceed without providing an answer to the questions; however, they were not able to go back to a previous question. In addition to the questions on willingness to participate in different tasks, the survey included questions on trust and personality traits, as well as socio-demographic questions, and questions about the survey experience/context. The full survey (in Spanish) can be found at the following link: http://ww2.netquest.com/respondent/glinn/mobile2016.

In this paper, we focus on two sets of 10 questions on the willingness to participate in different tasks. These questions were asked in different ways to different respondents, who were randomly assigned to answer through grids or item-by-item questions with vertical or horizontal scales; and on a PC or smartphone. Random-

1 The other experiments compare the answers for PC and smartphone respondents to rank order questions, grids versus item-by-item questions, and agree-disagree versus item-specific formats.

ization was done independently for each experiment, which allows us to test for confounding. A series of Kolmogorov-Smirnov tests for equality of distributions across the different groups showed significant differences in only a very few cases (2 out of 60 showed significant effects for device; 0 out of 40 for grid versus item-by-item; and 1 out of 40 for scale direction). Thus, we see no evidence of confounding, and ignore the other experiments in our subsequent analyses. We describe the 20 questions in more detail below.

From the 5,907 panelists invited to the survey, 3,051 started it (51.7%) but only 1,623 (53.2% of those who started) answered the first main survey question (following the screener questions). The rest were screened out for a variety of reasons (e.g., used a different device, did not have Internet access through both PC and smartphone, quotas full). Another 132 respondents were excluded later because they switched device during the survey or did not pass some basic quality checks (e.g., the answers to the gender and/or age questions differed from the profile information). Finally, 15 participants dropped out after the first four questions. Thus, a total of 1,476 respondents (48.4% of those who started; 90.9% of those who answered the first main survey question) finished the survey using the required device type; these are the focus of our analyses[2].

Data Preparation and Preliminary Analyses

In this section we describe the items used in our analyses, and the preparation of analytic variables.

a) *Proportions of respondents who self-reported that they would be willing to do a series of tasks for a given incentive level, and average willingness score.*

We asked respondents 20 questions about the willingness to collaborate with Netquest beyond answering survey questions. The different activities proposed were classified a priori in different groups:

- Passive measurement on devices they already use, e.g., "Use the accelerometer on your smartphone to measure your physical activity and report it (passively) to Netquest".
- Wearing special devices to passively monitor activity, e.g., "Wear a small device on your wrist that measures your alcohol consumption and directly sends the information to Netquest".
- Providing respondents with measurement devices and then having them self-report the results (i.e., they could see the results, and decide to edit their answers),

2 The final dataset used is available from the first author upon request.

e.g., "Measure your blood cholesterol level using a finger prick we will provide you and self-report the results to Netquest".

- The provision of physical specimens or bodily fluids, e.g., "Measure your saliva cortisol by chewing special gum for 30 seconds, then putting it in a vial and mailing it to Netquest".
- Others, e.g., "Let your children answer surveys that we would send to you for them".

Our goal was to vary the requests on several dimensions including the frequency of measurement (one time versus continuous), the degree of respondent involvement (passive versus active), the sensitivity of the topic (blood alcohol levels versus photos of products), and so on. The full list of items appears in Table 1.

These questions were separated in two sets of 10 questions each, which differed on two additional levels:

- The incentive offered in exchange for collaboration with the request: 30 points in the first set versus 40 points in the second. These points can be exchanged for gifts by the panelists: for instance, with 20 points, a panelist can get an e-book; with 40 points, an online film; with 120 points, a cinema ticket. Across 186 surveys administered to Netquest panelists in Spain in 2016, the number of points received per survey varied from four to 58 with an average of 14 and a median of 12 (Revilla, 2017).
- The number of answer categories: the first set uses partially labeled 5-point scales from "1- Definitely not" to "5- Definitely yes" whereas the second set uses partially labeled 11-point scales with similar labels (on 0 and 10). A "not applicable" (NA) option was also available.

Note that the incentive and response scales are confounded: the first set used a 5-point scale and 30-point incentive, while the second used an 11-point scale and offered 40 points. We address this confounding later. While we did not randomize items across the incentive and response scale conditions, the order of the items within each set of 10 questions was randomized across respondents in order to minimize potential order effects.

For each question, we look at the proportion of missing answers (respondents were not required to answer each item), the proportion of not applicable (NA) answers, the proportion of respondents who would accept the task (i.e., they answered 4 or 5 for the first set of questions and 6 to 10 for the second set of questions) among those who provided an answer different from NA, and the average willingness rating among those providing an answer different from NA (on a 0-10 scale; transforming the score for the first set by subtracting 1 and multiplying by 2.5; following Preston and Colman, 2000, and Dawes, 2008).

b) Self-reported reasons for accepting or not accepting installation of a browser tracker.

For these analyses, we use the answers to two open questions[3]. The first asked about the reasons why respondents said they would accept (or not) the invitation to install an application on their PC which registers the URLs of the websites they visit and report this (passively) to Netquest.

The second, asked only to respondents who said they would not be willing to install the tracking application or who selected the middle answer category, was: "What would Netquest most need to change such that you would accept the invitation to install an application on your PC which registers the URLs of the websites you visit and report this (passively) to Netquest?"

The answers to these questions were coded by a native Spanish speaker. When a respondent provided several reasons, we consider them all.

c) Factor analysis to identify common elements in willingness to participate in different tasks.

Next, we study what affects willingness to participate in different tasks in a more general way. We expected the tasks proposed in these questions to pertain to different categories of activities (see subsection a). In order to empirically examine the grouping of these tasks, we conducted a principal component factor analysis (PCA) based on the 817 respondents who provided a substantive answer (excluding NA) to all items[4]. Three factors with an eigenvalue greater than 1 were identified. Given that we expected the three factors to be correlated, we considered a 3-factor solution with oblique rotation.

The PCA suggested the following classification of tasks. For a full description of the items, we refer to Table 1.

- Factor 1 ("*PhysicalMeasures*") includes six items about sharing different physical measures: *PassiveStress, CholesterolSelfReport, AlcoholSelfReport, PassiveAlcohol, CholesterolVial, CortisolVial.*
- Factor 2 ("*BehaviorTracking*") includes six items about allowing the fieldwork company to track behavior: *PassiveGPS, TrackerPC*, *TrackerMobile, FacebookProfile, Emotion, EyeMovement.*
- Factor 3 ("*RespondentControl*") includes four items where the respondent has control on the reporting: *PhotosProduct*, *ScanBarcodes, TestProduct, PhotosMobile.*

3 To reduce burden, we asked the open questions only about one selected task.

4 Running the PCA on the larger sample with 17 items (excluding the three items about children where the NA proportions are high) yields essentially the same factor structure.

The final four items were not classified because they loaded on two factors equally (*Accelerometer* and *ChildrenStress*) or had low loadings on all factors (*Children-Survey* and *ChildrenWeight*). Based on this classification, we created a willingness score for each of the three factors identified above, using the procedure described below.

First, for the items in the first set, we recoded the answers from 0 to 4 instead of 1 to 5 and multiplied this by 2.5 to get a score from 0 to 10 (as we have for the items in the second set). Then, for each of the three factors identified, we averaged the rescaled items to get an equally-weighted willingness score from 0-10 for each factor. We did not use factor scores because of the varying levels of missing data across items. We excluded those respondents who did not provide substantive answers to at least half of the items in the factor. This means that 7.1% of respondents did not get a summary score for factor 1, while 7.4% did not get one for factor 2, and 6.2% for factor 3.

d) Regression analyses of the willingness to participate on independent variables related to trust, personality and respondent socio-demographics.

The scores on these three factors form the key dependent variables in our analyses. The correlations obtained between the respective factor scores are as follows: 0.62 between factors 1 and 2, 0.63 between factors 1 and 3 and 0.52 between factors 2 and 3. Given the relatively high correlations, we also consider an overall willingness score computed on all 20 items[5]. An examination of the standardized normal probability plots of the summated scales suggests that they approximate normal distributions, justifying the use of OLS regression.

In addition, since the literature does not systematically identify factors affecting willingness (most of the studies are simply descriptive, reporting the rates of willingness or compliance), we selected independent variables which we expected to be associated with these three factors (Appendix A provides details on all the variables and scales):

- Some basic socio-demographic variables: *Men*, *Age*, *Education* and *Income*.
- One question on the frequency of Internet use on a smartphone (*InternetFrequency*). The more frequently respondents use a smartphone to connect to Internet, the more they are likely to use GPS, social media, etc., that already capture this information. Thus we expect they will be more willing to share data of different kinds too.
- Three variables related to the sharing of content (*ShareFB, ShareTwitter, LikeSharingLife*). The more respondents already share content on Facebook and

5 We also created a score based on the 17 items that do not involve children, and obtain equivalent results.

Twitter, and the more they like sharing their personal life, the more we expect that they will be willing to provide Netquest with different kinds of data.

- Three questions about the benefits of market research (*BenefitForMe, Benefit-Consumers* and *BenefitSociety*). The more respondents value market research, the more we expect they will be willing to participate in different tasks.
- Three questions related to trust (*Suspicious*, *SocialTrust*, *TrustAnonymity*). The more trust people have in general and in the anonymity of the information they share, the more we expect them to be willing to share different kinds of data.
- Two questions about the attitude toward safety (*SecureSurroundings* and *Avoid-Risk*). The more respondents are concerned about safety in general, the more we expect they will also be worried about the risks of sharing different data with a panel company.
- Three variables related to the attitude toward answering surveys (*AnswerIn-come*, *LikeAnswering*, *PriorParticipation*). The more positive respondents' attitudes toward answering surveys (i.e., provided an answer to the income question, liked answering the current survey, answered many previous surveys in the panel), the more we expect them to be willing to participate in other tasks too.
- One question about the attitude toward new activities (*LikeNew*). The more respondents are looking for new things to do, the more we expect them to accept new tasks.

Several of the questions described above (*LikeSharingLife, Suspicious, Social-Trust, SecureSurroundings, AvoidRisk, LikeAnswering,* and *LikeNew*) were part of a separate experiment on agree-disagree (AD) versus item-specific (IS) wording. In a series of models (not shown) we tested whether the different formats affected the relationship of these variables with the willingness factor scores. We tested both main effect models (with an indicator for AD/IS) and interactions (of format with items). Our general conclusion was that the format in which these questions were asked did not have an effect on the conclusions drawn from the models, except for two variables (*SocialTrust* and *LikeNew*). Given that these two variables also did not show consistent or significant relationships with willingness, we decided to drop them from the model. For parsimony, for the other questions, we combine the alternative versions and use them as predictors in the models below.

We examined the bivariate associations between all these variables and the scores created for total willingness and for the three factors *PhysicalMeasures*, *BehaviorTracking*, and *RespondentControl*, and fitted a variety of models. We decided to drop two more variables: *Income* because of the high proportions of missing data (24.7% missing or "I prefer not to answer") and *InternetFrequency* because it did not have a significant effect in the models and we had variables more directly related to the sharing of content through social media (*ShareFB* and *ShareTwitter*).

In addition, we also used the same set of variables to estimate three structural equation models (SEM). In each case, our dependent variable is a latent variable, measured by the different items identified as forming one of the three willingness factors. In terms of independent variables, *Men*, *Age* and *Education* are measured with a single indicator each, whereas the others are measured with several indicators: *Share* is measured with three indicators (*ShareFB*, *ShareTwitter* and *LikeSharingLife*), as is *Benefit* (*BenefitForMe*, *BenefitConsumers*, *BenefitSociety*), and *Attitude toward surveys* (*AnswerIncome*, *LikeAnswering*, *PriorParticipation*), whereas *Trust* and *Safety* are measured with two indicators each (respectively, *Suspicious* and *TrustAnonymity*; and *SecureSurroundings* and *AvoidRisk*). The model was estimated in LISREL and tested using global fit measures as well as the JRule software (Van der Veld, Saris and Satorra, 2009). The model is corrected step by step until an acceptable fit is obtained. Appendix B provides an example of the path diagram for the initial model, as well as a list of the extra parameters introduced in each model in order to get an acceptable fit, and the final estimates of the parameters in each model. Only a summary of the main effects of the SEM are presented in the results section.

Main Results

Stated Willingness to Complete Different Tasks in Exchange for Specific Incentives

We first examine the responses to the 20 individual willingness items. Table 1 provides for each item, the percentage not answering that item (% missing), the percentage of NA answers among those who gave an answer, the percentage who say they would accept the task, and the average willingness score (among those who gave an answer different from NA), ranked by percent willing.

Table 1 Stated willingness, ordered by proportions of accepting (highest to lowest)

If you would receive 30 (or 40*) points in exchange, would you accept the invitation to...	% missing	% NA	% would accept	Average (0-10 scale)
... receive a product at home to test and report on it in a survey (*TestProduct*)*	5.3	1.7	73.7	7.4
... take photos of products with your smartphone and send them to Netquest (*PhotosProduct*)	6.2	2.7	56.4	6.2
... scan barcodes of products with your smartphone and share them with Netquest (*ScanBarcodes*)	5.9	3.5	53.7	6.0

Table 1 continued

If you would receive 30 (or 40*) points in exchange, would you accept the invitation to...	% missing	% NA	% would accept	Average (0-10 scale)
... measure the amount of alcohol in your breath using a breathalyzer test kit we would provide you and self-report it to Netquest (*AlcoholSelfReport*)	6.3	5.8	51.0	5.5
... take photos with your smartphone and send them to Netquest (*PhotosMobile*)*	5.7	2.8	49.6	5.3
... wear a small device on your wrist that measures your stress and directly sends the information to Netquest (*PassiveStress*)	5.8	3.2	44.8	5.0
... measure your blood cholesterol level using a finger prick we will provide you and self-report the results to Netquest (*CholesterolSelfReport*)	6.1	3.3	40.5	4.5
... wear a small device on your wrist that measures your alcohol consumption and directly sends the information to Netquest (*PassiveAlcohol*)*	5.5	5.2	37.8	4.1
... use the accelerometer on your smartphone to measure your physical activity and report it (passively) to Netquest (*Accelerometer*)	6.8	4.6	37.4	4.6
... let your children answer surveys that we would send to you for them (*ChildrenSurvey*)	6.2	25.8	30.7	3.7
... measure your blood cholesterol level using a finger prick we will provide you, then putting it in a vial and mailing it to Netquest (*CholesterolVial*)*	6.0	2.9	30.2	3.3
... measure your saliva cortisol by chewing special gum for 30 seconds, then putting it in a vial and mailing it to Netquest (*CortisolVial*)*	5.9	3.2	27.7	3.1
... measure your children‘s weight when we ask you and self-report it to Netquest (*ChildrenWeight*)*	5.8	25.3	27.5	3.2
... share GPS information from your smartphone with Netquest (*PassiveGPS*)	6.0	3.8	20.8	2.7
… let us record your face while you watch a video in your PC in order to measure the movement of your eyes (*EyeMovement*)*	6.1	3.3	19.3	2.3
... give Netquest access to all the information of your profile on Facebook (as if they were one of your friends) (*FacebookProfile*)*	5.6	5.7	19.0	2.4
… let us record your face while you watch a video in your PC in order to measure your emotional response (*Emotion*)*	5.1	3.6	18.0	2.2

Table 1 continued

If you would receive 30 (or 40*) points in exchange, would you accept the invitation to...	% missing	% NA	% would accept	Average (0-10 scale)
... install an application on your smartphone which register the URLs of the websites you visit and report this (passively) to Netquest (*TrackerMobile*)	6.4	4.3	17.8	2.4
... install an application on your PC which register the URLs of the websites you visit and report this (passively) to Netquest (*TrackerPC*)	6.2	5.0	16.6	2.3
... let your children wear a small device on their wrist that measures their stress and directly sends the information to Netquest (*ChildrenStress*)*	5.5	24.4	11.8	1.5

Note: Tasks followed by a * correspond to the second set (40 points incentive). N = 1,476 for the % missing; N varies from 1,375 to 1,400 for the % NA and from 1,028 to 1,374 for the % would accept and average scores. The % would accept and average columns are based on those who gave a substantive answer (i.e., excluding both the missing and NA responses).

The levels of item missing values are quite similar (ranging from 5.1% to 6.8%). Concerning the levels of NA, the three items asking about children clearly differ from the others, which is to be expected since some of the panelists do not have children[6].

The proportions of respondents willing to accept the different tasks show large variations. The most accepted task is that of receiving a product at home to test and report on in a survey: 73.7% of respondents who gave an answer said they would be willing to do this. This is followed by taking photos of products with a smartphone (already much lower: 56.4%). At the other extreme, the task with the lowest level of willingness consists of letting one's children wear a small device on their wrist that measures their stress and directly sends the information to Netquest, with only 11.8% expressing willingness.

It is interesting to note that the willingness to use a breathalyzer and self-report the readings (51.0%) and the willingness to measure one's blood cholesterol level and self-report the results (40.5%) are much higher than the willingness to give Netquest access to all the information in one's Facebook profile (19.0%) or to install an application on one's PC to register the URLs of websites visited and report this (passively) to Netquest (16.6%).

It is also interesting that there is a difference of more than 10 percentage points between the stated willingness for doing a cholesterol test depending if the results

6 However, including or excluding the NA answers to compute the willingness has little effect on the rank order and conclusions overall, even for these three items.

are self-reported by the respondents or if the test is directly sent to Netquest via mail. In the second case, where respondents cannot change the results or decide whether or not to share them, stated willingness is lower. Similarly, a difference of 13.2 percentage points is seen in the case of alcohol tests, depending on whether the results are self-reported or directly sent to the panel company. This suggests that respondents make a distinction both on the types of data being measured and on the degree of control they have over what is captured or reported.

Similar results can be seen when considering the average score instead of the proportion of respondents willing to complete the tasks.

In terms of the effect of the differential incentive and response scale, the ordering of items in Table 1 suggests there is not a strong differential effect of the incentive offered. In fact, the average willingness score for the 10 items in the first set (30 points and 5-point scale) is higher than that for the second set (40 points and 10-point scale), likely reflecting differences in the tasks being asked about more than differences in incentives or response scale. This again suggests we can ignore these confounding factors.

Overall, the mean for the total score of willingness is 4.0 (on a 0-10 scale). Considering the three factors, *RespondentControl* has the highest mean (6.2), followed by *PhysicalMeasures* (4.2) and finally *BehaviorTracking* (2.4).

Self-reported Reasons for Being Willing or Not

Next, we focus on one of the tasks proposed in the first set of questions: the willingness to install a passive browser tracking application on one's PC, for which only 16.6% of the respondents who gave an answer expressed willingness to do so. Table 2 reports the main reasons mentioned in a follow-up open question about why they would accept or not the invitation to install a tracking application.

The main reason mentioned for accepting the task was that respondents did not mind or did not feel that this was confidential information (37.4%), followed by interest in getting the incentive (25.1%), altruism (14.0%) and trust (9.9%). On the other side, the main reason for not accepting this task is linked to privacy concerns (72.6%), with a further 7% raising issues of trust.

In order to improve the acceptance of this task, respondents who said that they would not be willing to install the tracking application or chose the middle category were asked what could be done to help them change their decision. While 68.0% of the respondents said that there is nothing that could be done to make them change their mind, 11.7% mentioned improvements in security and 9.7% increased incentives. Even if for a large majority of respondents, it seems unlikely they will be convinced to install a tracking application on their PC, security and incentives are aspects which could improve the overall acceptance of such tasks.

Table 2 Main reasons* why panelists would accept or not accept the invitation to install a tracking application on their PC

Main reasons for accepting	% (based on N= 171 respondents)
I don't mind/not confidential	37.4
Incentive	25.1
Altruism	14.0
Trust	9.9
Main reasons for not accepting	**% (based on N= 829 respondents)**
Privacy	72.6
No trust	7.0
No reason	5.4
I do not own the PC I use	5.8

Note: * We present all reasons that are mentioned by at least 5% of the respondents. When a respondent provided several reasons, we take them all into account.

Predictors of Willingness to Accept Additional Tasks

Table 3 contains a set of four linear regression models, first predicting the total score for all 20 willingness items, and then predicting each of the three scores on the factors identified earlier. In each case a positive coefficient means greater willingness. The overall proportion of variance explained by the set of predictors ranges from 0.22 (for *RespondentControl*) to 0.35 (for the full 20-item scale).

We also run a series of partial F-tests to see if the collective contribution of each group of variables (we have five groups in the final models, besides the socio-demographic variables: attitude toward sharing, benefit of market research, trust, attitude toward safety and toward surveys) was significant, when explaining each of our four dependent variables of interest. All the tests indicated a statistically significant contribution except one: the test for attitude toward safety in the case of the factor *RespondentControl* ($F(2, 1039)=2.73$; $p=.0657$). This suggests that each set of variables has an association with willingness.

Several variables are statistically significant across all four models, with coefficients in a consistent direction. As expected, the more frequently respondents report posting content on Facebook, the more willing they are to accept a variety of additional research tasks. *ShareTwitter* is not statistically significant in any of the models, but this may be because fewer respondents use Twitter relative to Facebook (57.2% versus 90.1%), or that Twitter is a more public social networking service. Similarly, *LikeSharingLife* is positively associated with willingness. Those who

Table 3 Regression analyses

		TotalScore		*Physical Measures*		*Behavior Tracking*		*Respondent Control*	
Explanatory variables		Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
Demo-graphics	Men	.342	**.019**	.475	**.018**	.463	**.003**	-.073	.692
	Age	-.004	.521	.006	.474	-.001	.909	-.019	**.019**
	Education	-.165	**.035**	-.302	**.005**	-.123	.141	-.010	.920
Share	ShareFB	.131	**.000**	.153	**.002**	.119	**.002**	.141	**.002**
	ShareTwitter	-.018	.618	-.056	.259	.064	.100	-.028	.543
	LikeSharingLife	.400	**.000**	.400	**.000**	.464	**.000**	.285	**.002**
Benefit	BenefitForMe	.136	.139	.103	.408	.206	**.034**	-.004	.973
	BenefitConsumers	.248	**.013**	.425	**.002**	.110	.298	.332	**.008**
	BenefitSociety	.101	.296	-.062	.642	.205	**.048**	.141	.251
Trust	Suspicious	.054	.451	.012	.904	.102	.185	.018	.843
	TrustAnonymity	.607	**.000**	.645	**.000**	.584	**.000**	.402	**.006**
Safety	SecureSurroundings	.058	.402	.113	.232	-.078	.291	.202	**.022**
	AvoidRisk	-.139	**.008**	-.236	**.001**	-.088	.113	-.087	.190
Attitude Toward Surveys	AnswerIncome	.537	**.002**	.509	**.030**	.582	**.001**	.506	**.018**
	LikeAnswering	1.232	**.000**	1.242	**.000**	1.108	**.000**	1.334	**.000**
	PriorParticipation	.277	**.000**	.250	**.008**	.306	**.000**	.208	**.016**
	Constant	-4.333	**.000**	-3.561	**.002**	-5.891	**.000**	-2.475	**.020**
	No. observations	1,044		1,052		1,049		1,056	
	R-squared	.345		.237		.330		.216	
	Adj. R-squared	.335		.225		.320		.204	

Note: coefficients in bold when statistically significant (p-value<.050)

have greater trust in the anonymity of their data are more willing to accept additional tasks. Those who answered the income question (indicating a degree of trust or willingness to disclose) also show higher levels of willingness on all four measures. Finally, two indicators of survey engagement are positively associated with willingness: those who liked answering the survey and those who have responded to more prior Netquest surveys have higher levels of willingness.

Several other variables are statistically significant in some but not all of the models. The effect of gender is statistically significant (men more willing) for three of the four models. The coefficient for education is negative (those with higher education less willing) for all four models but only reaches statistical significance for the *TotalScore* and *PhysicalMeasures* models. Those who perceive greater benefit of research for consumers have significantly higher willingness for three of the four measures (*TotalScore*, *PhysicalMeasures*, and *BehaviorTracking*), but the direction of the effect is consistent across all four models. Those who are inclined to avoid risk have significantly lower levels of willingness on *TotalScore* and *PhysicalMeasures,* but not on the other two factors (although, again, the effect is in a consistent direction).

Finally a few variables reach statistical significance in only one of the models. Age has a significant negative effect (older people less willing) only for the *RespondentControl* factor. Both those who see a personal benefit and those who see a societal benefit of market research are more willing to agree to *BehaviorTracking*. Finally, those who rate *SecureSurroundings* as more important are *more* willing to agree to tasks that permit respondent control.

Overall, we see largely consistent effects of predictors across the different types of activities, although there is enough variation among the models to suggest that different types of people react differently to the different types of additional tasks being asked about.

Considering the SEM analyses, quite similar results are obtained, even if there are few differences. The latent variables *Share*, *Benefit* and *AttitudeTowardSurveys* have statistically significant positive effects on the three willingness factors. In addition, men have higher willingness on *PhysicalMeasures* and *BehaviorTracking*. Finally, *Safety* has a statistically significant negative effect on *BehaviorTracking* (the more one cares about safety, the less willing). Details of the SEM are presented in Appendix B.

Discussion

In this paper we investigated the willingness to perform additional tasks among panelists of an opt-in online panel in Spain. We found that the willingness to perform additional tasks is not a unitary phenomenon. Respondents distinguish between different types of tasks, and are more willing to do some but not others. In general, willingness is higher for tasks where respondents have control over the reporting of the results (e.g., taking pictures, measuring one's blood cholesterol level and reporting the results) than for passive tracking behaviors (e.g., installing a tracking app on one's PC or smartphone), even if this means that respondents have to do more work than with passive measurements where they only need to give their permission once. This is probably due to high privacy concerns, which is also what the answers to the open questions suggest: most respondents mentioned reasons related to the issue of trust/security/privacy both for accepting or not accepting the installation of a tracking app.

Our factor analysis revealed three distinct but related types of tasks: *PhysicalMeasures*, *BehaviorTracking*, and *RespondentControl*. Our models also suggest that there are variables that reliably predict willingness, as measured by these factors. This implies that restricting a sample to only those willing to accept a specific task is likely to result in both demographic and attitudinal biases.

The study has several limitations. The results are based on an opt-in panel (already generally cooperative, self-selected, already have a relationship with the panel), and further restricted to those who have Internet access through both a PC and a smartphone. We are studying stated willingness, not actual willingness. The results are restricted to a single panel (Netquest) in a single country (Spain). Thus we should be cautious about generalizing the results to different panels and countries. There are also some limitations in the analyses performed: some questions were asked in different formats (AD versus IS) for random subsets of respondents; some answered the survey on a PC, others on a smartphone (again, randomly assigned). Also we could not really take the difference in incentives into account (i.e., we did not randomly assign respondents to different incentive conditions); however, our primary focus was not on the incentives but the tasks.

In addition, we identified a variety of different tasks, but did not systematically try to vary the features or elements of these tasks, such as the degree of intrusiveness, the potential burden, the degree of respondent control, etc. More work is needed to explore the various dimensions that affect willingness to perform some tasks but not others. Our research has started looking at the "what" (i.e., what people are willing to do and not do) but not as much at the "why" (why people are willing or not, although our open question started to address this issue). More research is needed to explore the reasons behind differential willingness of panelists to accept different tasks, and to understand how stated willingness translates to actual compliance.

Researchers are increasingly exploiting the measurement capabilities of modern technologies. Understanding how consumers react to these requests, and understanding the differences between those who are willing and those who are not, are important steps in evaluating the utility of these additional tasks or measures. Most of the prior studies have focused only on a single task (e.g., GPS capture, or installing a browser tracker). Our research finds that treating all tasks as the same, and making inference from one type of request to all other requests, is risky. The stated willingness to use new technologies to provide additional data to researchers varies according to the nature of the task. A first step to overcoming the barriers to accepting new technologies is understanding within- and between-respondent differences in willingness.

References

Angrisani, M., Kapteyn, A., & Samek, S. (2017). Real time measurement of household electronic financial transactions in a population representative panel. Paper presented at the ESRA conference, Lisbon, July.

Antoun, C., Couper, M.P., & Conrad, F.G. (2017). Effects of mobile versus PC web on survey response quality: A crossover experiment in a probability web panel. Public Opinion Quarterly, 81(5), 280-306.

Armoogum, J., Roux, S., & Pham, T.H.T. (2013). Total nonresponse of a GPS-based travel surveys. Paper presented at the conference on New Techniques and Technologies for Statistics, Brussels, March.

Avendano, M., Scherpenzeel, A., & Mackenbach, J.P. (2011). Can biomarkers be collected in an Internet survey? A pilot study in the LISS panel. In M. Das, P. Ester, & L. Kaczmirek (Eds.), Social Research and the Internet (pp. 371-412). New York: Taylor and Francis.

Biler, S., Šenk, P., & Winklerová, L. (2013). Willingness of individuals to participate in a travel behavior survey using GPS devices. Paper presented at the conference on New Techniques and Technologies for Statistics, Brussels, March.

Boase, J. (2016). Augmenting survey and experimental designs with digital trace data. Communication Methods and Measures, 10(2-3), 165-166.

Boase, J, & Ling, R., (2013). Measuring mobile phone use: Self-report versus log data. Journal of Computer-Mediated Communication, 18(4), 508-519.

Boyle, J., Kilpatrick, D., Acinerno, R., Ruggiero, K., Resnick, H., Galea, S., Koenan, K., & Galernter, J. (2010). Biological specimen collection in an RDD telephone survey: 2004 Florida hurricanes gene and environment study. In L.A. Aday & M. Cynamon (Eds.), Ninth Conference on Health Survey Research Methods (pp. 176-184). Hyattsville, MD: National Center for Health Statistics.

Burton, J. (2016). Results for Web/face-to-Face linkage consent questions in the Innovation Panel. Paper presented at the Mixing Modes and Measurement Methods in Longitudinal Studies Workshop. London: CLOSER.

Couper, M.P., & Singer, E. (2013). Informed consent for web paradata use. Survey Research Methods, 7(1), 57-67.

Couper, M.P., Singer, E., Conrad, F.G., & Groves, R.M. (2008). Risk of disclosure, perceptions of risk, and concerns about privacy and confidentiality as factors in survey participation. Journal of Official Statistics, 24(2), 255-275.

Couper, M.P., Singer, E., Conrad, F.G., & Groves, R.M. (2010). An experimental study of disclosure risk, disclosure harm, incentives, and survey participation. Journal of Official Statistics, 26(2), 287-300.

Crawford, S.D., McClain, C., Young, R.H., & Nelson, T.F. (2013). Understanding mobility: Consent and capture of geolocation data in web surveys. Paper presented at the annual meeting of the American Association for Public Opinion Research, Boston, May.

Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. International Journal of Market Research, 50 (1), 61-77.

De Reuver, M., & Bouwman, H. (2015). Dealing with self-report bias in mobile Internet acceptance and usage studies. Information & Management, 52(3), 287-294.

Etter, J.-F., & Bullen, C. (2011). Saliva cotitine levels in users of electronic cigarettes. European Respiratory Journal, 38(5), 1219-1220.

Gatny, H.H., Couper, M.P., & Axinn, W.G. (2013). New strategies for biosample collection in population-based social research. Social Science Research, 42, 1402-1409.

Gautier, A., Rahib, D., Brouard, C., Saboni, L., Blineau, V., El Malti, F., David, C., Chevaliez, S., Barin, F., Larsen, C., Lot, F., & Lydién N. (2016). Proposition d'un volet biologique à l'issue d'une enquête téléphonique: retour d'expérience du BaroTest. 9ème colloque francophone sur les sondages, Gatineau, octobre 2016. Available at: http://sondages2016.sfds.asso.fr/programme/programme-avec-presentations/

Gilbert, E., Conolly, A., Tietz, S., Calderwood, L., & Rose, N. (2017). Measuring young people's physical activity using accelerometers in the UK Millennium Cohort Study. London: Centre for Longitudinal Studies, CLS working paper 2017/15.

Greenfield, T.K., Bond, J., & Kerr, W.C. (2014). Biomonitoring for improving alcohol consumption surveys: The new gold standard? Alcohol Research: Current Reviews, 36(1), 39-45.

Hassani, M., Kivimaki, M., Elbaz, A., Shipley, M., Singh-Manoux, A., et al. (2014). Non-consent to a wrist-worn accelerometer in older adults: the role of socio-demographic, behavioural and health factors. PLoS ONE, 9 (10), e110816.

Howie, E.K., & Straker, L.M. (2016). Rates of attrition, non-compliance and missingness in randomized controlled trials of child physical activity interventions using accelerometers: a brief methodological review. Journal of Science and Medicine in Sport, 19, 830–836.

Jäckle, A., Burton, J., Couper, M.P., & Lessof, C. (2017). Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: response rates and response biases. Institute for Social and Economic Research, University of Essex: Understanding Society Working Paper Series No. 2017-09.

Joh, K. (2017). 2017-2018 regional household travel survey. Presentation to the National Capital Region Transportation Planning Board, Travel Forecasting Subcommittee, May 19th.

Johnson, A., Kelly, F., & Stevens, S. (2012). Modular survey design for mobile devices. Paper presented at the CASRO Online Research Conference, Las Vegas, February.

Keusch, F., Antoun, C., Couper, M.P., Kreuter, F., & Struminskaya, B. (2017). Willingness to participate in passive mobile data collection. Paper presented at the annual meeting of the American Association for Public Opinion Research, New Orleans, LA, May.

Kissau, K., & Fischer, D. (2016). Pitfalls and opportunities of research using passive metering software. Paper presented at the General Online Research conference, Dresden, March.

Lauderdale, D.S., Schumm, P.L., Kurina, L.M., McClintock, M., Thisted, R.A., Chen, J.H., & Waite, L. (2014). Assessment of sleep in the National Social Life, Health, and Aging Project. Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 69(8), S125–S133.

Link, M.W., Murphy, J., Schober, M.F., Buskirk, T.D., Childs, J.H., & Tesfaye, C. (2014). Mobile technologies for conducting, augmenting and potentially replacing surveys: Report of the AAPOR task force on emerging technologies in public opinion research. Public Opinion Quarterly, 78(4), 779-787.

McFall, S.L., Conolly, A., & Burton, J. (2014). Collecting biomarkers using trained interviewers. Lessons learned from a pilot study. Survey Research Methods, 8(1), 57-66.

McGeeney, K., & Weisel, R. (2015). App vs. web for surveys of smartphone users experimenting with mobile apps for signal-contingent experience sampling method surveys. Washington, DC: Pew Research Center report, http://www.pewresearch.org/2015/04/01/app-vs-web-for-surveys-of-smartphone-users/

Menai, M., van Hees, V.T., Elbaz, A., Kivimaki, M., Singh-Manoux, A., & Sabiaa, S. (2017). Accelerometer assessed moderate-to-vigorous physical activity and successful ageing: results from the Whitehall II study. Science Reports, 7, 45772; doi: 10.1038/srep45772.

Miller, G. (2012). The smartphone psychology manifesto. Perspectives on Psychological Science, 7(3), 221-237.

Palmer, J.R.B., Espenshade, T.J., Bartumeus, F., & Chung, C.Y. (2013). New approaches to human mobility: Using mobile phones for demographic research. Demography, 50(3), 1105-1128.

Pinter, R. (2015), Willingness of online access panel members to participate in smartphone application-based research. In D. Toninelli, R. Pinter, & P. de Pedraza (Eds.), Mobile Research Methods: Opportunities and Challenges of Mobile Research Methodologies (pp. 141-156). London: Ubiquity Press.

Preston, C.C., & Coleman, A. (2000). Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. Acta Psychologica, 104, 1-15.

Raento, M., Oulasvirta, A., & Eagle, N. (2009). Smartphones: An emerging tool for social scientists. Sociological Methods & Research, 37, 426-54.

Revilla, M. (2017). Analyzing survey characteristics, participation, and evaluation across 186 surveys in an online opt-in panel in Spain. Methods, Data, and Analysis, 11(2), 135-162.

Revilla, M., Toninelli, D., Ochoa, C., & Loewe, G. (2016). Do online access panels really need to allow and adapt surveys to mobile devices? Internet Research, 26(5), 1209-1227.

Revilla, M., Ochoa, C., & Loewe, G. (2017). Using passive data from a meter to complement survey data in order to study online behavior. Social Science Computer Review, 35(4), 521-536.

Roth, A., & Mindell, J.S. (2013). Who provides accelerometry data? Correlates of adherence to wearing an accelerometry motion sensor: the 2008 Health Study for England. Journal of Physical Activity and Health, 10, 70-78.

Sakshaug, J.W., Couper, M.P., & Ofstedal, M.B. (2010). Characteristics of physical measurement consent in a population-based survey of older adults. Medical Care, 48(1), 64-71.

Scherpenzeel, A. (2017). Mixing online panel data collection with innovative methods. In S. Eifler & F. Faulbaum (Eds.), Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung (pp. 27-49). Wiesbaden: Springer.

Toepoel, V., & Lugtig, P. (2014). What happens if you offer a mobile option to your web panel? Evidence from a probability-based panel of Internet users. Social Science Computer Review, 32(4), 544-560.

Van der Veld, W.M., Saris, W.E., & Satorra, A. (2009). Jrule 2.0: User Manual. Radboud University Nijmegen, The Netherlands.

Van Duivenvoorde, S., & Dillon, A. (2015). The best of both worlds? Combining passive data with survey data, its opportunities, challenges and upside. Paper presented at the CASRO Digital Research Conference, February 11-12, Nashville, TN.

Wenz, A., Jäckle, A., & Couper, M.P. (2017). Willingness to use mobile technologies for data collection in a probability household panel. Institute for Social and Economic Research, University of Essex: Understanding Society Working Paper Series No. 2017-10.

Wrzus, C., & Mehl, M.R. (2015). Lab and/or field? Measuring personality processes and their social consequences. European Journal of Personality, 29, 250–271.

Appendix A

List of all independent variables considered: exact formulation and scales

- *Men* (1="Male", 0="Female")
- *Age* (in years)
- *Education* (in six categories, from no education to university degree)
- *Income* (in six categories, from lowest to highest)
- *Internet Frequency*: "On average, how frequently do you connect to the Internet using a smartphone" ("1=Once a month or less" to "6=Daily")
- Sharing of content:
  - *Share FB*: "In general, how frequently do you share content on your personal Facebook account?" (8 response options ranging from "I don't have a personal account" to "I share content every day")
  - *Share Twitter*: "In general, how frequently do you share content on your personal Twitter account?" (same 8 response options)
  - *Like sharing life*: either asked in an agree-disagree format ("I like sharing my personal life", 1="Completely disagree" to 5="Completely agree") or an item-specific format ("How much do you like sharing your private life?" 1="Don't like at all" to 5="Like extremely").
- Benefits of market research for:
  - the respondent him/herself (*Benefit for me*)
  - consumers (*Benefit consumers*)
  - the society/citizens (*Benefit society*)

  (1="Does not benefit at all" to 5="Benefits a great deal").
- Trust:
  - *Suspicious*: "I get suspicious easily" (1="Completely disagree" to 5="Completely agree") or "How easily do you get suspicious?" (1="Not at all easily" to 5=" Extremely easily")
  - *Social Trust*: "I don't trust people in general" (1="Completely agree" to 7="Completely disagree") or "How much do you trust people in general?" (1="Do not trust at all" to 7="Trust completely")
  - *Trust anonymity*: "To what extent do you trust that this survey guarantees anonymity?" (1="Do not trust at all" to 4="Trust completely")

- Attitude toward safety:
  - *Secure surroundings*: “It is important for me to live in secure surroundings” (1=“Completely disagree” to 7=“Completely agree”) or “How important is it for you to live in secure surroundings?” (1=“Not important at all” to 7=“ Extremely important”)
  - *Avoid risk*: “I always avoid anything that can endanger my safety” (1=“Completely disagree to 7=“Completely agree) or “How often do you avoid anything that can endanger your safety?” (1=“Never” to 7=“Always”)

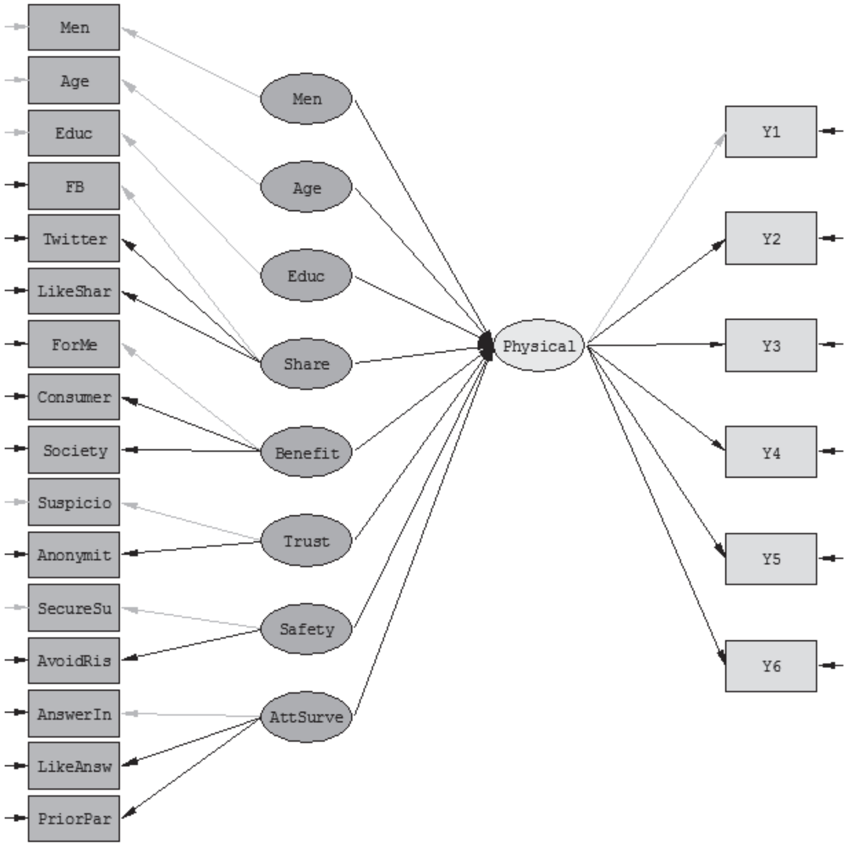
- Attitude toward answering surveys:
  - *Answer Income:* dummy variable coded 1 if the respondent provided a substantive answer to the income question, and 0 otherwise (no answer at all, or “I prefer not to answer” option).
  - *Like Answering*: “How much did you like or not to fill in this questionnaire? (1=“Did not like it at all” to 4=“Liked it very much”)
  - *Prior participation*: number of Netquest surveys completed before this one, recoded from lowest to highest into quartiles.

- Attitude toward new activities:
  - *Like New:* “I am never looking for new things to do” (1=“Completely agree” to 5=“Completely disagree”) or “How often are you looking for new things to do?” (1=“Never” to 7=“Always”)

Appendix B

More information about the SEM analyses

a) Initial model for factor 1 (LISREL path diagram); models are similar for factors 2&3



b) Extra parameters introduced in each model in order to get an acceptable fit

Model Factor 1: correlated error terms for Y5 and Y6; Y2 and Y5; Age and Prior Participation; and cross-loading Att. Survey and Anonymity.

Model Factor 2: correlated error terms for Y5 and Y6; Y2 and Y3; Age and Prior Participation; and cross-loading Att. Survey and Anonymity.

Model Factor 3: correlated error terms for Age and Prior Participation; and cross-loading Att. Survey and Anonymity.

c) Estimates of the parameters in each model (completely standardized solution).

		Physical Measures	Tracking Behavior	Respondent Control
Measurement model	F by Y1	.79 NA	.80 NA	.88 NA
	F by Y2	.77*	.76*	.81*
	F by Y3	.79*	.72*	.64*
	F by Y4	.88*	.71*	.85*
	F by Y5	.75*	.66*	Not present
	F by Y6	.73*	.66*	Not present
	Share by FB	.76NA	.71 NA	.75 NA
	Share by Twitter	.49*	.52*	.50*
	Share by Like Sharing	.39*	.45*	.39*
	Benefit by For Me	.84 NA	.85 NA	.84 NA
	Benefit by Consumer	.86*	.86*	.86*
	Benefit by Society	.82*	.82*	.82*
	Trust by Suspicious	1.00 NA	1.00 NA	1.00 NA
	Trust by Anonymity	-.14*	-.14*	-.14*
	Safety by Secure Surroundings	1.00 NA	1.00 NA	1.00 NA
	Safety by Avoid Risk	.42*	.43*	.43*
	Att. Survey by Anonymity	.49*	.51*	.50*
	Att. Survey by Answer Income	.26 NA	.26 NA	.26 NA
	Att. Survey by Like Answering	.64*	.60*	.67*
	Att. Survey by Prior Participation	.22*	.24*	.20*
Structural model	Men on F	.07*	.11*	-.03
	Age on F	.02	.01	-.05
	Education on F	-.09*	-.05	-.02
	Share on F	.15*	.30*	.16*
	Benefit on F	.17*	.19*	.17*
	Trust on F	.00	.02	-.01
	Safety on F	.02	-.06*	.04
	Att. Survey on F	.51*	.60*	.49*
Fit	Chi-Square	$\chi^2(202)=742.94$	$\chi^2(202)=728.30$	$\chi^2(165)=672.81$
	RMSEA	.053	.052	.056

Note: F refers to the factor of interest (*PhysicalMeasures* or *TrackingBehavior* or *RespondentControl*). Y1 to Y6 refer to the items used to measure this F factor (thus there are different in each model). * Indicates a coefficient statistically significantly different from 0 (t-ratio >1.96). NA indicates that no values are available for the t-ratio because the corresponding loading was fixed to 1 (unstandardized) for identification purposes.