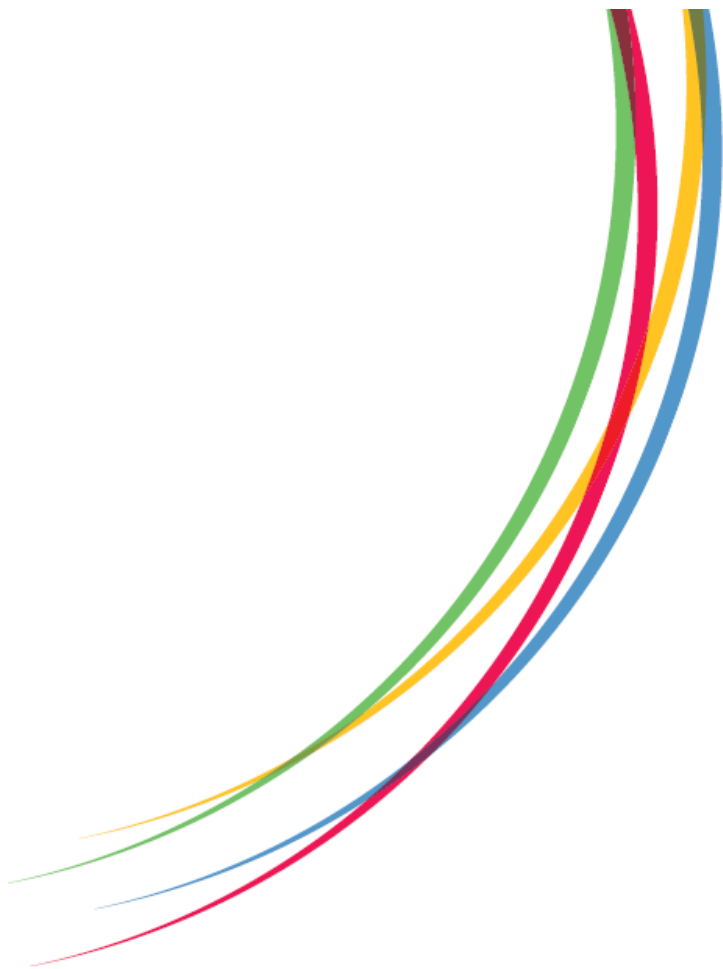




Understanding Society

Working Paper Series

No. 2024 – 08

September 2024

Understanding Society: Minimizing selection biases in data collection using mobile apps

Annette Jäckle (University of Essex)

Jonathan Burton (University of Essex)

Mick P. Couper (University of Essex)



**Economic
and Social
Research Council**

Non-technical summary

We have implemented a series of studies using mobile apps to collect data for the *Understanding Society* survey. In each case, the mobile app enabled us to collect data that we could not collect in the annual interview: detailed information about household expenditure, daily data about relationships, stressors and wellbeing, detailed body measurements, or spatial cognition. However, in each case, only a sub-set of respondents invited to the mobile app study participated and provided data. In this article we review what we have learnt about how to collect data using mobile apps, based on these experimental studies.

When survey respondents are invited to participate in additional data collection using a mobile app, there are several steps they must go through to provide data (from having a compatible device, to finding the app in the app store, to installing it and logging in, etc). At each of these steps, there are potential barriers that lead to some respondents dropping out. We review the different steps and potential reasons for dropping out and discuss the fieldwork protocols that can be used to reduce these barriers.

We then document the experimental studies that we have conducted with mobile apps in the *Understanding Society* Innovation Panel and report the outcomes of our mobile app studies and the effects of different fieldwork protocols aimed at reducing the barriers to participation. We also review results from experimental studies by other researchers.

Our findings reinforce the importance of designing app studies so that they are feasible (e.g. compatible with different OS systems) and easy (e.g. finding and logging into the app) for all. The fieldwork protocols that make a difference are related to how the respondent is invited to the app study. Inviting face-to-face respondents to an app study within an interview significantly increases participation rates compared to sending them an invitation letter after their interview. For web respondents the mode of invitation to the app study however makes no difference. Placing the invitation to the app study early in the interview also increases participation, compared to placing it at the end of the interview.

Understanding Society: Minimizing selection biases in data collection using mobile apps

Annette Jäckle (University of Essex)

Jonathan Burton (University of Essex)

Mick P. Couper (University of Michigan)

Abstract: The UK Household Longitudinal Study: *Understanding Society* has a programme of research and development that underpins innovations in data collection methods. One of our current focuses is on using mobile applications to collect additional data that supplement data collected in annual interviews. To date, we have used mobile apps to collect data on consumer expenditure, well-being, anthropometrics and cognition. In this article we review the potential barriers to data collection using mobile apps and experimental evidence collected with the *Understanding Society* Innovation Panel, on what can be done to reduce these barriers.

Keywords: UKHLS, panel survey, experiments

JEL classification: C81, C83

Acknowledgements: Spending Study 1 and 2 were supported by grants from the UK Economic and Social Research Council (ESRC) and the National Centre for Research Methods (NCRM) (ES/N006534/1) and by ESRC funding for the *Understanding Society* survey (ES/N00812X/1). The Spending Study 1 application was developed by Kantar Worldpanel; the Spending Study 2 application was implemented in QMob by Kantar. The Wellbeing Study was funded by the Economic and Social Research Council (ESRC) grant for the *Understanding Society* survey (ES/N00812X/1). The Wellbeing application was developed by Connect Internet Solutions Ltd. The BodyVolume Study and the Spatial Cognition Study were funded by the Economic and Social Research Council (ESRC) grant for the *Understanding Society* survey (ES/T002611/1). The BodyVolume application was developed by Select Research, the Cognition application was developed by Glitchers. Ethical approval for all studies was obtained from the University of Essex Ethics Committee. The data from all studies are

available from the UK Data Service. We thank Thomas F. Crossley for comments on an earlier version of this article.

Corresponding author: Annette Jäckle, Institute for Social and Economic Research, University of Essex, Wivenhoe Park, Colchester, Essex CO7 9FJ, United Kingdom, aejack@essex.ac.uk.

1. Introduction

Understanding Society takes a very systematic, experiment- and evidence-based approach to innovations in data collection methods. This is enabled by the Innovation Panel, a separate sample of 1,500 households that mirrors the design of the main *Understanding Society* study but is used for experimentation and methods testing (University of Essex. Institute for Social and Economic Research 2023a). We have, for example, used the Innovation Panel to prepare the move from face-to-face interviews to mixed mode data collection including web and telephone (e.g. Jäckle, Lynn and Burton 2015); to test methods of collecting biomeasures from respondents, without nurses or interviewers (Benzeval et al. 2023; Kumari et al. 2023); to investigate why consent rates for data linkage are lower in web than face-to-face surveys (Jäckle et al. 2022a); and, the focus of this paper, to test ways collecting data using mobile applications.

Economists, along with many other empirical social scientists, want more – more granular and longer running – data to fit models in a wide variety of different domains. Some of this demand is being met by rapid advances in and widespread adoption of mobile technologies such as smartphones and wearables. The expansion of sensors in these devices has brought new opportunities for measurement (see Struminskaya and Keusch 2023). *Understanding Society* has, for example, collected data using mobile apps on consumer expenditure, well-being, anthropometrics and cognition (see Section 3). Other examples include travel data, collected using GPS tracking and/or app-based travel diaries (e.g. McCool et al. 2021), time use, *in situ* measurement of subjective states (e.g., ecological momentary assessment or day reconstruction methods; e.g. Kahneman and Krueger 2006), contemporaneous measurement of expenditures and other economic behaviours (e.g. Dubois, Griffith and O'Connell 2022), diet and food consumption, measurement of physical activity (using accelerometers; e.g. Kapteyn et al. 2018).

While many are conducting studies using volunteers, concerns about selection bias and its effect on inference has long been a concern for economists (e.g. Heckman 1979; Meyer, Mok and Sullivan 2015). Traditional population-based surveys are designed to minimize these selection effects, but often conduct infrequent (annual or biennial) measurements, raising concerns about recall biases and timeliness of

measurement. Many longitudinal surveys are therefore exploring the use of new technologies to enhance and extend measurement, with the aims of potentially reducing measurement errors by collecting data in a timely manner and/or collecting objective data that does not rely on respondent recall, and measuring and mitigating selection biases using information collected previously in the panel about non-participants. Our particular focus here is on data collection using mobile apps. Combining both survey-based and app-based measurement has the potential to minimize the drawbacks of each method alone. To that end, the UK Household Longitudinal Study: *Understanding Society* has a program of research to develop methods for data collection using mobile apps that would supplement the annual interviews.

In this article we review what we have learnt about how to collect data using mobile apps, based on a series of experiments in the *Understanding Society* Innovation Panel. Our focus is on the selectivity of data collected with mobile apps. In Section 2 we review the potential barriers that can lead to (selective) non-participation in mobile app studies and the fieldwork protocols that can be used to reduce these barriers; in Section 3 we document the experimental studies that we have conducted with mobile apps in the *Understanding Society* Innovation Panel; in Section 4 we report the outcomes of our mobile app studies and the effects of different fieldwork protocols aimed at reducing the barriers to participation as well as reviewing results from experimental studies by other researchers; in Section 5 we summarize the key findings on effective fieldwork protocols and discuss the outlook and agenda for future methods research on data collection with mobile apps.

2. Barriers to using mobile apps for data collection

When survey respondents are invited to participate in additional data collection using a mobile app, there are several steps they must go through to provide data. At each of these steps, there are potential barriers that lead to some respondents dropping out. The different steps and potential reasons for dropping out are summarized in Table 1, as well as the fieldwork protocols that can be used to reduce these barriers. The asterisks (*) in Table 1 denote fieldwork protocols that we have tested experimentally in the *Understanding Society* Innovation Panel.

Most surveys using mobile apps for data collection rely on respondents installing the app on their own device. This excludes respondents who do not have a compatible device, either because they have no device, or a device with an operating system or version that is not compatible with the app. This barrier can be overcome by loaning devices to respondents who do not have a compatible device themselves (e.g. Scherpenzeel 2017). If the app is compatible with multiple operating systems and devices, for example iOS as well as Android, with tablets as well as smartphones, and with older operating systems, these exclusions can also be minimised.

Respondents who have a compatible device must be willing and able to participate in the app study. In some studies respondents are explicitly asked for consent to participate, in others consent is implied by the respondent's participation in the study. At this stage there are multiple possible reasons for respondents dropping out, for example not knowing how to install and use an app, lack of interest or time, privacy or data security concerns, or feeling that the study is asking too much. Both the nature of the task and the wording of the invitation matter here: how burdensome the task is, in terms of the number of days and time required; the offer of incentives or personal feedback; explanations of the purpose of the additional data collection for research; explanations of how to install and use the app; and assurances for privacy and data security. How and when the respondent is invited to the additional data collection, e.g. whether as part of an interview or with an invitation letter, and where in the interview the invitation is issued, can also affect willingness to participate.

Respondents who are willing to participate must find the app in the app store appropriate for their device. They might fail at this stage, for example, if they mis-type the app name, do not identify the correct app because there are others with similar names, or if a link they use to find the app is broken. Ensuring that the name and logo of the app are unique and not easily confused with others can help at this stage, as well as showing the respondent the app logo and offering them different ways of finding the app, including links and QR codes.

Once the respondent has found the correct app, they need to install it on their device. This might fail, if they have insufficient storage space on their device or are on a limited data plan. In both cases, designing the app with a view of minimising its size will help at this stage.

Once installed, the respondent must log into the app. They might fail and drop out at this stage if they mis-type their login credentials or the login is not working for technical reasons. Here, the usability of the login credentials matter: whether the respondent has to enter a user name as well as a password, and the format of the username or password, for example whether it is an alphanumeric code or something more memorable; whether characters that are easily confused (e.g. 0 and O, 1 and l) are excluded.

Depending on the app, respondents who have successfully installed and logged in must set any necessary permissions, for example for the app to access the device's camera, to track geolocation, or to send notifications such as daily reminders. Depending on the app requirements, respondents who are not willing to give these permissions will not produce complete data. At this point, ensuring that the respondents understand why the permissions are required and how they will be used is important.

Once the respondent has logged into the app, they must use it, either to enter data or to launch automated passive data collection. At this point, respondents might look at the app and decide they do not want to participate after all, or indefinitely delay or forget to use it. Here the app design and usability matter, including whether the app is intuitive and pleasing to use and whether the task appears reasonable and easy. Reminders to use the app might also help at this stage.

Finally, the respondent must adhere to the study protocol and complete all tasks, whether this is data entry once a day or at each 'event' (e.g. at each purchase), or whether it is continuous passive data collection over a period of time for which the relevant app permissions must remain on. At this stage, reminders will help, as well as the offer of bonuses if the task is completed according to the protocol, for example, every day.

Table 1: Barriers to using mobile apps for data collection

Steps respondent has to go through to participate	Potential reasons for drop-out	Fieldwork protocols to increase participation
Have a compatible device	• No device • Wrong OS • OS version too old	• Loan devices to participants • Compatibility: OS / OS version / tablet & phone
Willing and able to participate (sometimes: explicit consent)	• Does not know how to • Not interested • Too busy • Privacy concerns / too sensitive • Data security concerns • Survey asking too much	• Incentives* • Feedback* • Length of task (# days / daily time*) • Wording: purpose / privacy / data security • Instructions for installing and using app • Mode of invitation: letter / in interview* • Placement of invitation in interview*
Find app	• Mis-type app name • Do not recognise app • Link broken	• Search / link / QR code • Name and logo of app
Install app	• Lack of storage space • Insufficient internet bandwidth	• App size
Log in to app	• Typo • Login not working	• User name and/or password • Password usability
Set relevant permissions in app/device (e.g. notifications, camera, location tracking)	• Not willing • Misunderstanding why permissions required	• Wording: justification • Incentives
Use app once	• Not willing • Indefinite delay before using • Forgets to use app	• App design & instructions • Reminders
Adhere to study protocol	• Not willing • Forgets • Weak internet signal	• Reminders • Bonus incentives*

Note: * denotes fieldwork protocols that we have tested experimentally in Understanding Society mobile app studies.

3. Data: UKHLS experimental studies using mobile apps

The Innovation Panel is a core feature of *Understanding Society*: The UK Household Longitudinal Study. The survey design and implementation is based on the main *Understanding Society* survey, with all adults (aged 16+) in sample households eligible for annual interviews, and younger household members (aged 10-15) given a self-completion questionnaire. The sample is a probability sample (all households have a known non-zero chance of selection into the sample) of approximately 1,500 households restricted to Great Britain, with refreshment samples added every three years to maintain a viable sample size for experimental studies. The first wave of the Innovation Panel was fielded in 2008 as a face-to-face survey. Since wave 5 (2012), the survey has used a mixed mode design, with adults in a random 2/3 of sample households invited to complete the survey online, with non-respondents followed up by face-to-face interviewers, and 1/3 issued to face-to-face interviewers first.

The Innovation Panel is used for methodological research that underpins the design and implementation of the main *Understanding Society* survey and as general research infrastructure. We hold an annual competition inviting proposals for methods tests, experiments, and new survey content to be implemented on the Innovation Panel (advertised on <https://www.understandingsociety.ac.uk> in February every year). The study leadership proposes experiments as well.

To date we have conducted five studies using mobile devices for data collection. The fifth study, a navigation game that tests cognition, is currently still in the field. Table 2 summarizes the key features of the Innovation Panel studies using mobile apps for data collection:

- Spending Study 1: respondents were asked to use an app for one month to report their daily expenditure, by uploading photos of all shopping receipts. The study included an experiment with respondent incentives (for documentation and data access, see University of Essex. Institute for Social and Economic Research 2021).

- Spending Study 2: respondents were asked to use an app for one month to report all expenditure in a diary. The study was implemented in two samples: the Innovation Panel included experiments with how respondents were invited to the app study; the parallel implementation in the Lightspeed online access panel included an experiment with offering within-app feedback on personal expenditure (see University of Essex. Institute for Social and Economic Research 2022).
- Wellbeing study: respondents were asked to use an app for 14 days to complete a daily questionnaire about relationships, stressors, and wellbeing. The study included experiments with respondent bonus incentives, the length of the daily questionnaire, and the placement of the invitation to the app study (see University of Essex. Institute for Social and Economic Research 2023b).
- BodyVolume study: respondents were asked to use an app once to take two photos of themselves, that were used to calculate measures of body volume. The study included experiments with respondent incentives and the type of feedback promised (see University of Essex. Institute for Social and Economic Research 2023a).
- Cognition study: respondents were asked to use an app to complete a navigation game, consisting of several levels of difficulty, that is used to measure spatial cognition. The study included an experiment with respondent incentives (at the time of writing the data collection still in the field).

Table 2: Understanding Society Innovation Panel data collection using mobile apps

	Spending Study 1	Spending Study 2	Wellbeing	BodyVolume	Sea Hero Quest
Year	2016	2018	2020	2022	2023
Concepts measured	Monthly expenditure	Monthly expenditure	Relationship quality and wellbeing	Body measurements	Spatial cognition
Task	Report daily spending by photographing shopping receipts	Report direct debits, standing orders and daily spending in app diary	Answer background questions (once) and daily questionnaire	Answer profile questions and take two photos of self (front and side view)	Complete a navigation game with multiple levels of difficulty
Duration	31 days	31 days	14 days	One-off	One-off
Treatments	Incentives	Mode of invitation, feedback	Placement of invitation, incentives,	Incentives, feedback	Incentives
Invitation to app study	Letter after annual interview	Experimental: letter vs. in annual interview	Experimental: early vs. late placement in annual interview	Annual interview	Annual interview
Respondent incentives	Experimental: £2 vs. £6 for installing the app, £1 for every day app was used, £10 bonus if app used on all 31 days	£1 for completing direct debit/standing order section, £0.50 for every day app was used, £10 bonus if diary used	£1 for every day app was used plus experiment: £10 if app used on all 14 days vs. £2.50 on four randomly	Experimental: £5 if app used vs. £5 added to unconditional incentive sent before the annual interview	Experimental: £10 vs. £30 if app used

		every day, £3 if completed debrief questionnaire	selected days if app used that day vs. no bonus		
Feedback on data	No	Experimental: invitation promised feedback in app on cumulative spending vs. no feedback vs. feedback not mentioned but given	No	Experimental: invitation promised feedback on total body fat vs. visceral body fat vs. not mentioned	No
Number of respondents invited to app study	N=2,112	IP N=2,638; Lightspeed online access panel N=2,878	N=2,152	N=2,536	Still in field
% of invited who used app at least once	13%	IP: 17%; Lightspeed: 14%	45%	18%	Still in field
App Developer	Kantar Worldpanel	Understanding Society app programmed by Kantar Public using QMob platform	Understanding Society app programmed by Connect Internet Solutions	Select Research	Glitchers

4. Results: Fieldwork protocols to reduce barriers to participation in mobile app studies

In the following we review empirical findings from our studies testing the effects of different fieldwork protocols for data collection using mobile apps. We review findings according to the steps a respondent has to go through to participate in a mobile app study and also report on results from experimental studies by other researchers. We first review the effects of fieldwork protocols on participation rates; we end this section with a review of the effects on selectiveness of participants.

In the original analyses we have examined different outcomes, for example, whether the respondent completed a registration survey (if there was one), whether they installed the app during the interview at the point of invitation, whether they used the app at least once, whether they used the app every week during the study period, and at what point during the study period they dropped out. In the summary here we focus on the effects of different fieldwork protocols on whether respondents used the app at all. We refer to the original articles for the more detailed analyses of different outcomes.

Have a compatible device

We have not experimented with fieldwork protocols relating to compatible devices. However, increasing penetration of smartphones has worked in our favour. At the time of our first mobile app study in 2016, only 66% of Innovation Panel respondents reported having a smartphone. By 2022 this had increased to 90% of respondents. The apps we have used for data collection have not placed stringent requirements on operating systems. For example, the Wellbeing app implemented in 2020 was compatible with iOS versions 11, 12, 13 and Android versions 5 to 10. As a comparison, the English National Health Service (NHS) COVID-19 Test and Trace app launched in the same year, required iOS 13.5 and Android 6 or higher, leading to exclusions of parts of the population with incompatible devices. In our most recent study, the BodyVolume app fielded in 2022, we estimated from the respondent debrief questions that 6% of respondents did not have a compatible device.

Willing and able to participate

Lack of interest or willingness to do additional tasks have been the main reasons for non-participation in our mobile app studies. We have experimented with several protocols that aimed to influence the respondent's willingness to participate.

Respondent incentives: We offer *Understanding Society* panel members an unconditional incentive for every annual interview and believe that they expect us to offer additional incentives if we ask them to do additional tasks. However, our incentive experiments with mobile app studies have so far had mixed results. In Spending Study 1 we assumed getting respondents to install the app in the first place would be the biggest hurdle. We therefore experimentally offered them £2 or £6, just for installing the app, plus £0.50 for every day on which they used the app and a £10 bonus if they used it for the 31 study days. The incentive experiment, however, had no effect on whether respondents used the app (Jäckle et al. 2022b). In the BodyVolume app we experimentally either offered respondents £5 conditional on using the app, or increased the unconditional incentive that they are sent with the invitation letter for the annual survey by £5. The conditional incentive increased the percentage who used the app by 7 percentage points, compared to the unconditional incentive (Jäckle, Burton and Couper 2023b). Other authors have conducted similar experiments. Keusch and colleagues (2021) implemented an experiment in an app-based study of refugees in Germany. Following an initial web survey, respondents with an Android smartphone were invited to download the app for passive mobile data collection for 3 months. About half were randomly assigned to a condition where they were offered €30 for installing the app and allowing data collection for the full three months. The control group were not offered an incentive. The incentive group had a participation rate of 7%, compared with 3% in the control group (difference not statistically significant). The IAB-SMART study, implemented on a sub-sample of the German PASS study (Panel Study Labour Market and Social Security) included a similar experiment, which did have an effect: Haas et al. (2020) reported a higher participation rate if respondents were offered €20 for installing a survey app (16%), than if they were offered €10 (13%). McCool and colleagues (2021) conducted an incentive experiment in a travel app study in the Netherlands. Sample persons were randomised to three incentive groups, all of whom received an initial €5 incentive: 1) received an extra €5 when they installed the app and another €5 if the app was active for seven

days, 2) received another €10 if the app was active for seven days, and 3) received another €20 if the app was active for seven days. The three incentive conditions achieved device registration rates of 30%, 36%, and 40% respectively. Revilla, Paura and Ochoa (2021) experimented with different methods of recruiting participants in non-probability online panels in three countries to download and use a research app. In a fully-crossed design, they varied the invitation message (emphasising benefits for respondents only, or also benefits for brands and researchers) and the incentive (no incentive or a small point-based incentive) for downloading the app. The token incentive had a positive effect in one country (Spain), but not the other two (USA and Brazil). Emphasising the benefits for respondents only had a slightly but not significantly higher installation rate.

Personalised feedback: If respondents benefit from participating in a study, they might be more willing to participate and motivated to provide complete data. Offering personalised feedback on the data collected could offer such a benefit, however, our experiments with feedback have so far not had any effect on participation. In Spending Study 2 one third of respondents were told that they would see a cumulative report of their spending in the app, one third were not told about the feedback but did have access to it, and a third were not given any feedback (Wenz et al. 2022). These treatments had no effect on participation in the app study or on reporting expenditure in the app. Similarly, Rodenburg et al (2023) tested the effects of different feedback conditions in a mobile app-based version of a Household Budget Survey: participants were either offered feedback immediately within the app, or delayed feedback at the end of the study period. The feedback treatments did not affect participation or drop-out from the study. Since feedback does have large effects on participation in other contexts, where the respondent does not know the information contained in the feedback (e.g. blood analytes, see Benzeval et al. 2023), we conducted a follow-up experiment where we varied what we told respondents about the contents of the feedback. In the BodyVolume study all respondents had access to the same feedback once they completed the app task: total body fat, visceral body fat, and waist to hip ratio. We however varied what we promised respondents when we invited them to the app study: a third of respondents were told they would get feedback on their body fat, a third were told it would be feedback on their visceral body fat, and a third were not

told about the feedback at all. Again, none of the treatments affected participation in the app study (Jäckle, Burton and Couper 2023b).

Length of the study task: In a vignette study varying the characteristics of a hypothetical mobile app study, Keusch et al (2019) found that respondents would be more willing to participate in studies that ran for 1 rather than 6 months. Remmerswaal et al. (2023) experimentally varied the duration of a mobile app travel diary. Surprisingly they found that participation rates in the 7-day group were higher than in the 1-day group, however the experiment was confounded in that the 7-day group had an unrestricted start while the 1-day group were told to use the app on a specific pre-defined day. In the Wellbeing app study we kept the study duration fixed (14 days), but experimentally varied the amount of time required on each of the study days: one half of respondents were asked to complete a daily 10-minute questionnaire, the other half a 2-minute questionnaire. Surprisingly, this had no effect on participation rates (Jäckle et al. 2023).

Mode of invitation to the mobile app study: For the first mobile app study (Spending Study 1) we invited all respondents by sending them a letter after they had completed their annual interview. For Spending Study 2 we experimentally invited half of respondents to the app study within their annual interview, while the control group were invited by letter sent after their interview. The effect of the invitation mode depended on the mode in which respondents completed their annual interview. A random two-thirds of households are invited to complete the annual interview online, with non-respondents followed up by face-to-face interviewers. The remaining third are issued to face-to-face interviewers first, with non-respondents given the option to complete the survey online. Using the randomised mode allocation as an instrumental variable for the actual interview mode, the results indicate that for web respondents the mode of invitation to the mobile app study made no difference (16% in both groups used the app at least once). For face-to-face respondents, however, the participation rate was 9% among those invited by letter and 29% among those invited in the interview (Jäckle et al. 2022b). Rodenburg et al (2023) report similar results from an experiment with a mobile app-based Household Budget Survey. The participation rate for sample members contacted by interviewers was 25%, compared to 12% for those invited by letter.

We tested whether the mode of the annual interview affects app study participation in the BodyVolume study. Using the randomised allocation of sample members to web-first versus face-to-face-first as an instrumental variable for the mode of interview suggests that respondents were 18 percentage points more likely to install the app if they completed the annual interview with a face-to-face interviewer than if they completed it online; there was however no significant effect of the mode on whether respondents used the app (Jäckle, Burton and Couper 2023b).

Lawes et al. (2022) experimented with different invitation methods in an app-based EMA (Ecological Momentary Assessment) study among registered job seekers in Germany. They tested three different conditions: 1) a mailed letter containing a QR code to app stores to download the app, 2) an email with a link to the app stores, or 3) a prenotification letter with a description of the study, followed by an email with a link to download the app. The prenotification + email yielded a significantly higher signup rate (7%) than either the letter (5%) or email-only (5%) conditions.

Placement of the invitation to the mobile app study within the annual interview: Research on respondent consent to data linkage has repeatedly shown that respondents are more likely to consent if asked earlier on in the interview (e.g. Burton, Couper and Jäckle 2023). We implemented a similar experiment with the invitation to the Wellbeing app study. A random half of respondents were invited to the app study early on (after short modules on COVID-19, demographics, and mobile device use), the other half were invited at the end of the interview (on average 45 minutes into the survey). Respondents were more likely to use the app at least once if they were invited early in the annual interview (49%), than if they were invited at the end (40%, see Jäckle et al. 2023).

Find the app in the app store

We have not experimented with protocols aimed at this stage, but have changed our procedures over time. In previous studies we simply told respondents the name of the app, showed them the app logo, and asked them to find it in their app store. In the BodyVolume app study we tailored the instructions to respondents. We captured paradata in real time to detect how the respondent was completing their annual survey. Respondents who were completing the survey online using an iOS or Android smartphone or tablet were shown a link to the app in the corresponding app store.

Respondents who were completing the survey online with a different device or computer, those who said they wanted to install the app on a different device, and those who completed an in-person interview, were shown QR codes for the app stores, as well as the app name and logo. Overall, only 4% of respondents invited to the BodyVolume app study said they did not install the app because they could not find it in the app store (Jäckle, Burton and Couper 2023a).

Install the app and log into the app

These stages in the process also appear to be a lower hurdle. In the BodyVolume study, 4% of respondents said they did not manage to install the app, 3% said they did not manage to log in (Jäckle, Burton and Couper 2023a). We have not yet experimented to test the effects of password usability on participation in app studies.

Set relevant permissions in the app

We have not yet examined this hurdle. The IAB-SMART study (Haas et al. 2020) experimented with different incentives for setting permissions in the app. Their app passively captured five types of data from participants' smartphones over a period of 6 months. Participants were asked to set the permissions at the start of the study, but could later revoke them in the app. They were offered €1 per month of permission for each data type. A random half of participants were offered a €5 bonus if they set all five permissions for the month. This treatment however had no significant effect on permission rates.

Use app once

For all our mobile app studies we have sent respondents an email reminder after their annual interview, reminding them to install the app if they have not already, restating their login information, and asking them to use the app. Most of our apps displayed instructions the first time the app was used, to explain the task and how to use the app. We have not yet experimented with features that would encourage respondents to use the app once they have installed it and logged in.

Adherence to study protocol

Reminders: We have not experimented with reminders, but implemented them differently across studies, depending on the capabilities of the app. The Wellbeing app sent participants a daily push notification at 5pm to remind them to complete their questionnaire. Participants who had not completed it by 8pm were sent a second reminder for that day. In contrast, we were not able to send push notifications through the Spending Study 1 app, the only alternative was to send reminders by email. We felt that daily reminders for the 31 study days would be too much, so we sent a weekly email reminder, including a summary of the incentives earned so far. Respondent debrief data collected in the Wellbeing study, however, suggested that not all participants saw the reminders. Unfortunately, there was no way for us to observe whether or not participants have set permissions for the app to send push notifications.

Bonus incentives: Data collected in the Spending Study 1 debrief questionnaire suggested that the £10 bonus for completing all study days motivated respondents. In the Wellbeing app study we therefore experimentally tested the effect of offering a bonus: a third of the sample were offered £10 if they used the app on all 14 study days, a third were offered £2.50 on four randomly pre-selected days if they to used the app that day (to introduce an element of gamification), and a control group were offered no bonus. The experimental treatments, however, had no effect on whether respondents used the app at all or whether they completed all 14 study days (Jäckle et al. 2023).

Selectiveness of app study participants

Socio-demographic characteristics: Across our studies we have found that the younger age groups and those with higher levels of education are over-represented among app study participants compared to the full sample of those invited (Jäckle et al. 2019; 2023; 2022b). We have, however, not found selection on some other socio-demographics related to the study content: personal income (Jäckle et al. 2019), household spending (Jäckle et al. 2019; 2022), mental health (Jäckle et al. 2023), or respondent weight (Jäckle, Burton and Couper 2023b).

Related personal behaviours: We have found that whether the respondent already does a similar task for their own purposes is the most important predictor of participation. For example, in Spending Study 1 respondents who used a mobile app to check their bank balance were over-represented among participants by 20

percentage points, those who kept a budget by 10 percentage points (Jäckle et al. 2019); in the Wellbeing app study respondents who used health apps were over-represented among participants by 16 percentage points (Jäckle et al. 2023).

Cooperation with the survey: Unsurprisingly, participation in app studies correlates with other indicators of cooperativeness. For example, respondents who consented to link their credit rating data to the survey and those with low item non-response rates in the annual interview were over-represented among app participants (Jäckle et al. 2019).

Effects of fieldwork protocols on selection: Our expectation was that inviting respondents to a mobile app study within an annual interview would bring in different types of people than sending them an invitation letter after their interview. However, that was not the case (Jäckle et al. 2022b). We see some differences in the selection biases between experimental treatment groups in the Wellbeing app study and the BodyVolume app study, however, the patterns are not consistent. We have yet to identify treatment protocols that will reduce the overall level of selection.

5. Summary and outlook

We have implemented a series of studies using mobile apps to collect data for the *Understanding Society* survey. In each case, the mobile app enabled us to collect data that we could not collect in the annual interview: detailed information about household expenditure, daily data about relationships, stressors and wellbeing, detailed body measurements, or spatial cognition. However, in each case, only a sub-set of respondents invited to the mobile app study participated and provided data. Our review of the barriers to participation in mobile app studies highlights the many steps that respondents have to go through in order to participate; the results of our experimental testing showcase that we lose respondents at each of these steps. In this concluding section we flag those aspects of fieldwork protocols that were effective at increasing participation and we discuss ideas for future studies to improve data collection using mobile apps.

Effective fieldwork protocols: Although we believe that our respondents expect additional financial incentives if we ask them to do additional tasks for the survey, our experiments to date with different incentive schemes have had mixed effects on

participation rates. This suggests that there are hurdles for which monetary incentives cannot compensate and that we need to design app studies so that they are feasible (e.g. compatible with different OS systems) and easy (e.g. finding and logging into the app) for all. Similarly, although the general belief is that respondents will expect an app to provide feedback, we have not found evidence that offering feedback increases participation in our studies. Respondent burden is a likely predictor of participation, however, varying the length of daily task also did not affect participation. The fieldwork protocols that did make a difference are related to how the respondent is invited to the app study. Inviting face-to-face respondents to an app study within an interview significantly increases participation rates compared to sending them an invitation letter after their interview. For web respondents the mode of invitation to the app study however makes no difference. Placing the invitation to the app study early in the interview also increases participation, compared to placing it at the end of the interview. Other protocols that we think make a difference but have not yet tested experimentally include ensuring that the app is compatible with a range of older operating systems, tailoring the instructions on how to find the app (depending on whether the respondent is using a compatible device to complete their interview), and sending reminders as push notifications within the app in addition to emails.

Planning for evaluation: In order to shed light on the barriers to participation in a particular mobile app study, and to identify ways of improving the fieldwork protocols to reduce those barriers, it is essential to plan for evaluation. Collecting background information about everyone who is invited to an app study enables detailed analyses of the selection biases among app study participants. *Socio-demographic characteristics*, such as age, gender, education, and income, can be used to test for selection biases of app study participants (i.e. representativeness of the participant sample). *Substantive variables* that correlate with the information collected by the app, can be used to test the representativeness of the data collected by the app. *Behavioural data* about activities or tasks respondents already do for their own purposes, can help identify barriers to participation, as can questions about the *characteristics of devices* owned by the respondent. *Debrief questions* are invaluable in identifying problems with fieldwork protocols that are not otherwise visible to the researchers. Debrief questions can be asked immediately after the invitation to the mobile app study, for example to identify reasons for refusal or problems with finding

and installing the app. Open debrief questions asked at the end of the mobile app study, asking respondents about their experiences and what would have made it easier for them to participate, are particularly useful in generating insights about how to improve fieldwork protocols. *Paradata* collected passively by the app are another useful source of information, for example, to study how participants navigate and use the app, how much time it takes them to complete tasks in the app, or the timing of when they complete the tasks.

Bespoke apps versus use of third-party apps: A key decision is whether to create a custom-designed app or use an existing third-party app (see Sezgin 2021). Using an existing app reduces development and maintenance costs, by benefiting from a platform that is maintained and updated to remain compatible with any operating system updates that might occur during the study period. Ideally, the existing app can be customised to some extent, for example to use the study's branding, relevant FAQs, customised debriefing questions, as well as dropping features of the app that are not relevant. In contrast, using a custom-designed app provides the greatest flexibility in terms of design and functionality of the app, and greater transparency, as the collection of paradata about the respondents' interactions with the app can be built into the app design. However, disadvantages of a custom-designed app include the additional development and programming costs, and the costs of maintaining the app over time. A third alternative is to use a customizable mobile app questionnaire platform (such as QMob) to create the study app. The disadvantage is that the design is constrained by the capabilities and limitations of the app platform; the advantage is that the programming is quicker and cheaper than a custom-designed app, and the platform deals with maintaining and updating the app.

Agenda for future research: The focus of *Understanding Society* research on using mobile applications for data collection has so far been on designing fieldwork protocols to reduce barriers to participation. Going forward, we are aiming to focus more specifically on how to include sub-groups who are under-represented. In the debrief information we and colleagues working on similar studies collect, we routinely see evidence of sample members who think the mobile app study 'is not for them'. For example, respondents who are not the main shopper in the household thinking there is no point in them participating in a spending study. For the design of future studies,

we are shifting our attention to the problem of how to design fieldwork protocols to make a study inclusive and attractive for everyone.

We have, in some cases, also examined the quality of data collected with apps. Read (2019) examined how the technical specification of the devices used by our participants in Spending Study 1 impacted on the quality of the photos of shopping receipts. He found that Android operating systems (as opposed to iOS), tablets (as opposed to smartphones), and devices with lower RAM were more likely to produce images that were partially or wholly unreadable. Wenz et al (2023) compared the expenditure data collected in Spending Study 1 with benchmark estimates from the UK Living Cost and Food survey. Total expenditure collected with the app mapped the benchmark estimates well, if the expenditures that participants reported directly in the app was added to the data coded from the shopping receipts. Serodio et al. (2023) compared waist and hip measurements derived by the app from photos of participants with those collected by asking respondents to measure themselves with a tape measure. They found that the app produced more outliers and less discrimination between men and women. Assessing the quality of data collected with mobile apps requires more attention going forward.

Understanding Society continually strives to innovate in data collection methods, to provide accurate, reliable and useful data for our users and enable cutting edge research. The Innovation Panel, which mirrors the main *Understanding Society* study but is used for methods testing and experimentation, is an invaluable tool. It enables us to carefully evaluate innovations and make evidence-based decisions, which underpin the quality of *Understanding Society* data.

6. References

- Benzeval, M., A. Andrayas, J. Mazza, T. Al Baghal, J. Burton, T.F. Crossley, and M. Kumari (2023) Does the feedback of blood results in observational studies influence response and consent? A randomised study of the Understanding Society Innovation Panel. *BMC Medical Research Methodology*, 23(1):134.
- Burton, J., M.P. Couper, and A. Jäckle (2023) The effects of position, context, and order on consent to data linkage in a survey. Understanding Society Working Paper 2023-07. Colchester: UK: University of Essex.
- Dubois, P., R. Griffith, and M. O'Connell (2022) The use of scanner data for economics research. *Annual Review of Economics*, 14:723-45.
- Haas, G., F. Kreuter, F. Keusch, M. Trappmann, and S. Bähr (2020) Effects of incentives in smartphone data collection. in *Big Data Meets Survey Science: A Collection of Innovative Methods*, edited by C.A. Hill, P.P. Biemer, T.D. Buskirk, L. Japiec, A. Kirchner, S. Kolenikov, and L.E. Lyberg. Hoboken, NJ: Wiley.
- Heckman, J.J. (1979) Sample selection bias as a specification error. *Econometrica: Journal of the econometric society*:153-61.
- Jäckle, A., J. Burton, and M.P. Couper (2023a) Protocols to invite survey respondents to data collection using mobile apps: Effects of survey mode, incentives and feedback on participation rates and bias. European Survey Research Association.
- Jäckle, A., J. Burton, and M.P. Couper (2023b) Protocols to invite survey respondents to data collection using mobile apps: Effects of survey mode, incentives and feedback on participation rates and bias. Section 4.7 in Vine, J. (ed.) Understanding Society Innovation Panel Wave 15: Results from Methodological Experiments. Understanding Society Working Paper 2023-10. Colchester: University of Essex.
- Jäckle, A., J. Burton, M.P. Couper, T.F. Crossley, and S. Walzenbach (2022a) How and why does the mode of data collection affect consent to data linkage? *Survey Research Methods*, 16(3):387-408.
- Jäckle, A., J. Burton, M.P. Couper, and C. Lessof (2019) Participation in a mobile app survey to collect expenditure data as part of a large-scale probability

- household panel: Coverage and participation rates and biases. *Survey Research Methods*, 13(1):23-44.
- Jäckle, A., J. Burton, M.P. Couper, and B. Perelli-Harris (2023) Participation of household panel members in daily burst measurement using a mobile app: Effects of position of the invitation, bonus incentives, and number of daily questions. Understanding Society Working Paper 2023-01. Colchester, UK: University of Essex.
- Jäckle, A., P. Lynn, and J. Burton (2015) Going Online with a Face-to-Face Household Panel: Effects of a Mixed Mode Design on Item and Unit Non-Response. *Survey Research Methods*, 9(1):57-70.
- Jäckle, A., A. Wenz, J. Burton, and M.P. Couper (2022b) Increasing participation in a mobile app study: The effects of a sequential mixed-mobile design and in-interview invitation *Journal of Survey Statistics and Methodology*, 10(4):898–922.
- Kahneman, D., and A.B. Krueger (2006) Developments in the measurement of subjective well-being. *Journal of Economic perspectives*, 20(1):3-24.
- Kapteyn, A., J. Banks, M. Hamer, J.P. Smith, A. Steptoe, A. Van Soest, A. Koster, and S.H. Wah (2018) What they say and what they do: comparing physical activity across the USA, England and the Netherlands. *J Epidemiol Community Health*, 72(6):471-76.
- Keusch, F., M.M. Leonard, C. Sajons, and S. Steiner (2021) Using smartphone technology for research on refugees: Evidence from Germany. *Sociological Methods & Research*, 50(4):1863-94.
- Keusch, F., B. Struminskaya, C. Antoun, M.P. Couper, and F. Kreuter (2019) Willingness to participate in passive mobile data collection. *Public Opinion Quarterly*, 83(S1):210-35.
- Kumari, M., A. Andrayas, T. Al Baghal, J. Burton, T.F. Crossley, K.S. Jones, D.A. Parkington, A. Koulman, and M. Benzeval (2023) Nurse- and self-collected dried blood spots for the assessment of cardiovascular risk factors: A randomised study of the Understanding Society Innovation Panel. PREPRINT (Version 1) available at Research Square [<https://doi.org/10.21203/rs.3.rs-2596340/v1>].

- Lawes, M., C. Hetschko, J.W. Sakshaug, and S. Griebemer (2022) Contact modes and participation in app-based smartphone surveys: Evidence from a large-scale experiment. *Social Science Computer Review*, 40(5):1076-92.
- McCool, D., P. Lugtig, O. Mussmann, and B. Schouten (2021) An app-assisted travel survey in official statistics. Possibilities and challenges. *Journal of Official Statistics*, 37(1):149-70.
- Meyer, B.D., W.K. Mok, and J.X. Sullivan (2015) Household surveys in crisis. *Journal of Economic Perspectives*, 29(4):199-226.
- Read, B. (2019) The influence of device characteristics on data collection using a mobile app. Understanding Society Working Paper 2019-01. Colchester, UK: Univeristy of Essex.
- Remmerswaal, D., B. Schouten, P. Lugtig, and B. Struminskaya (2023) Effects of study duration on response rates and representativity in smartphone travel app of Statistics Netherlands. MASS Workshop.
- Revilla, M., E. Paura, and C. Ochoa (2021) Use of a research app in an online opt-in panel: The Netquest case. *Methodological Innovations*, 14(1):2059799120985373.
- Rodenburg , E., B. Struminskaya, B. Schouten, and T. Oerlemans (2023) The effects of interviewer administered recruitment and personalized feedback on participation and data quality in an app based Household Budget Survey. presented at the European Survey Research Association conference.
- Scherpenzeel, A. (2017) Mixing online panel data collection with innovative methods. Pp. 27-49 in *Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung*, edited by S. Eifler and F. Faulbaum. Wiesbaden: Springer.
- Serodio, P., T. Al Baghal, A. Jäckle, J. Burton, and M. Kumari (2023) Can we rely on self-reported biomarker data? Benchmarking the accuracy and feasibility of self-reported hip and waist measurements using analogue and digital methods. Section 4.8 in Vine , J. (ed.) *Understanding Society Innovation Panel Wave 15: Results from methodological experiments*. Understanding Society Working Paper 2023-10. Colchester: University of Essex.
- Sezgin, E. (2021) Can we use commercial mobile apps instead of research mobile apps in healthcare research? *Frontiers in Public Health*, 9:685439.

- Struminskaya, B., and F. Keusch (2023) Mobile devices and the collection of social research data. Pp. 101-13 in *Research Handbook on Digital Sociology*, edited by J. Skopek. Cheltenham, UK: Edward Elgar Publishing.
- University of Essex. Institute for Social and Economic Research (2021) Understanding Society: Spending Study 1, 2016-2017. [data collection]. UK Data Service. SN: 8749, DOI: <http://doi.org/10.5255/UKDA-SN-8749-1>
- (2022) Understanding Society: Spending Study 2, 2018-2019. [data collection]. UK Data Service. SN: 8909, DOI: <http://doi.org/10.5255/UKDA-SN-8909-1>
- (2023a) Understanding Society: Innovation Panel waves 1-15, 2008-2022. [data collection]. 12th Edition. UK Data Service. SN: 6849, DOI: <http://doi.org/10.5255/UKDA-SN-6849-15>.
- (2023b) Understanding Society: Innovation Panel Wellbeing App Study, 2020. [data collection]. UK Data Service. SN: 9065, DOI: <http://doi.org/10.5255/UKDA-SN-9065-1>
- Wenz, A., A. Jäckle, J. Burton, and M.P. Couper (2022) The effects of personalized feedback on participation and reporting in mobile app data collection. *Social Science Computer Review*, 40(1):165–78.
- Wenz, A., A. Jäckle, J. Burton, M.P. Couper, and B. Read (2023) Quality of expenditure data collected with a mobile receipt scanning app in a probability household panel. Understanding Society Working Paper 2023-02. Colchester, UK: University of Essex