June 2019

Economic and Social Research Council

Methodological Briefing: Mode Effects

AUTHORS
Jonathan Burton
Annette Jäckle
University of Essex

Understanding Society

Push-to web designs that aim to maximise online completion of the survey are popular, given aims to contain costs of data collection. The protocols specifying how the modes are implemented can have large effects on take up of the web survey.

Using web-only is still not possible in the UK, because of bias in the types of people who have access to the internet and in the types of people who will complete a survey online. Therefore, follow-ups of web non respondents in other modes are still key, if the aim is to represent the general population.

Variable costs per interview are much lower with an online questionnaire than with interviewer administered questionnaires. However, using more than one mode of data collection can increase variable costs in the other modes and certainly increases fixed costs of data collection.

The answers respondents give in an online survey seem to be similar to those given to interviewers. There are however four key challenges regarding measurement with push-to-web surveys: sensitive information is more accurately reported in self-completion surveys than with interviewers which limits comparability of answers, tests of cognitive ability produce very different results between modes, consent rates for data linkage are much lower online than with interviewers, and protocols to collect biomeasures without interviewers are still under development.

New methods of data collection can increase the scope and quality of data collected in surveys, providing opportunities for innovative research. The method or technology itself can affect the quality of data collected. In addition, the protocols used to implement that method or technology can have large effects on both selectivity of who participates and quality of measurement.

Acknowledgement:

Understanding Society is an initiative funded by the Economic and Social Research Council and various Government Departments, with scientific leadership by the Institute for Social and Economic Research, University of Essex, and survey delivery by NatCen Social Research and Kantar Public. The research data are distributed by the UK Data Service. This methodological briefing has been funded by the Economic and Social Research Council UK Population Lab Innovation Development Grant (ES/S016651/1).

In this briefing we review existing research on different modes of survey data collection, with a focus on research using *Understanding Society* on the effect of mode on participation, the effect on data quality, and the cost implications of mixing modes. We also provide a brief review of research on the quality of data collected with new methods that go beyond survey questionnaires, focusing on experimental work using the *Understanding Society* Innovation Panel.

The choice of mode is a critical element in the survey design process. It is affected by what we want to ask, and to whom we want to ask it. The mode affects the way in which the survey is implemented, how questionnaires are designed, how the sample are invited to participate, the response rate, the quality of the data, the cost of the survey, and the time it takes for the survey to be conducted.
Different modes may be used at different stages of the survey process: advance notification, invitation, reminders, data collection, and follow-up of non-respondents. Within each stage there may be multiple modes, for example an advance letter and an email may be sent to a sample member with information on how to access an online survey, or a reminder may be sent by letter, email, SMS text message, or a phone call. The focus of this briefing is the mode used in the data collection stage of a survey to administer the questionnaire; this includes the initial mode, and any follow-ups of non-responders to that first mode.

Survey modes differ along several dimensions: (1) the degree of interviewer involvement, (2) the degree of contact with the respondent, (3) the channel of communication (aural/visual), (4) the locus of control over the interview, (5) the degree of privacy, and (6) the degree of technology use. Increasingly, as technology develops, there is scope for a broader range of data collection modes, using the diverse possibilities of mobile devices such as smartphones or wearables. We review an example of the use of mobile devices as a mode of data collection below (pages 9-11). In this briefing we focus on modes which are most commonly used for longitudinal studies in the UK. This means that our focus is on face-to-face, telephone, and online surveys. Stand-alone mail surveys (i.e., excluding a paper self-completion questionnaire as part of a wider interview) are not appropriate for long, complex, general population surveys which use extensive routing and employ data fed forward from a previous interview. However, they may be used for simpler questionnaires which do not employ much routing, such as between-wave supplements. Paper questionnaires have, for example, been used in the birth cohorts on occasion: the BCS70 sweep used a simple mailed questionnaire at age 26 (1996) and the 1958 National Child Development Study (NCDS) used a short-mailed paper questionnaire for their Twins Study in 2008.

We use the Total Survey Error framework (Groves et al, 2011) to review how the quality of a survey estimate may be affected by the survey mode. The framework identifies potential errors at different points in the data collection process. When evaluating a survey statistic, a rule-of-thumb is to think about “selection” and “measurement”. Selection is about who is answering the question, and measurement is about how the question was asked and answered. Both elements can affect the survey estimates.

Each survey mode has its strengths and weaknesses, for example, face-to-face interviewing is expensive but generally has a higher response rate, which potentially reduces selection bias, however it may be prone to a form of measurement bias – for example, social desirability bias (the, often unconscious, desire of the respondent to be seen positively by the interviewer, and so hide any socially ‘undesirable’ behaviour or opinions). A self-completion mode (e.g., online) may encourage respondents to be more honest in their opinions, so reduce measurement bias, but the absence of an interviewer to motivate and help respondents when they are having difficulty may reduce the quality of the data. A self-completion mode may have a lower response which may increase selection bias.

Many large social surveys have adapted a design which uses different modes to collect data, in which the weaknesses of one mode may be compensated for by the strengths of another mode. The use of different modes may be consistent across all sample members, for example, *Understanding Society*, like the British Household Panel Survey (BHPS) before it, uses face-to-face interviewing to achieve a high response, combined with a short self-completion questionnaire to collect information which may be perceived as being more sensitive. This design, where all respondents answer one set of questions in one mode, and another set in a second mode, is referred to as ‘multiple modes’. These multiple modes may be adopted within a single wave (as above), or across waves, for example in the US Health and Retirement Study which alternates a face-to-face wave with a telephone wave.

Increasingly, surveys have started to use ‘mixed modes’, where some respondents may answer a set of questions in one mode, and other respondents answer the same questions in a different mode. This may be done as a concurrent design – where respondents are offered a choice of modes – or sequentially, where one mode is offered, and non-respondents in that mode are then allocated to a second mode. Research has shown that the concurrent design is less successful than the sequential design (Medway and Fulton, 2012) and so the focus of this briefing is on this latter design, the use of sequential mixed modes. Mixing modes like this may be done to reduce costs (e.g., by starting with the cheaper mode and only using a more expensive mode for non-response follow-ups) or to increase response (e.g., by starting with the highestresponse mode, and then offering non-respondents an alternative cheaper mode).

We are gaining more information about the experience of using mixed modes on surveys as more longitudinal studies adopt this design. Longitudinal studies are, in some ways, ideal for using a mixed mode design because it is possible to collect the appropriate contact information at the initial wave, to facilitate contact in a second mode at subsequent waves, such as an email address and mobile phone number (Voorpostel et al, 2019). However, there has been some initial caution in the wide adoption of mixed mode designs for many ongoing longitudinal studies because one of the strengths of these studies is the production of data that is comparable over time, and the risk that changing the mode of data collection may affect both the selection (who responds) and the measurement (how they answer) properties of the study.

Understanding Society has recently adopted a sequential mixed-mode design, in which a set of sample members are invited to respond to the survey online. After a number of weeks, those that have not taken part online are then issued to interviewers who attempt to contact and interview them face-to-face. We have adopted this design after extensive testing using the *Understanding Society* Innovation Panel (IP). The IP allows us to experimentally test and evaluate different fieldwork designs before rolling them out on the main sample. (University of Essex, 2019) This gives *Understanding Society* a great advantage in that changes in design are evidence-based, and that this evidence can also be used by other longitudinal studies around the world. This note sets out the lessons we have learned on *Understanding Society* about mode effects on response, data quality, costs and timing.

Mode effects

There may be differences in results obtained in a survey using Mode A and those from a survey using Mode B. This may be due to selection (who is answering the survey) or measurement (how the question affects the answers respondents give) – or a combination of both.

Differences due to selection are because the participants in each mode differ in some way on some characteristic of interest. That is, the responding sample in Mode A is different in a systematic way to the responding sample in Mode B (and both may differ from the target population). For example, an online survey will systematically exclude those who do not have access to, or are not willing to use, the internet to participate in surveys. If those who are excluded are different to those who are included on a characteristic of interest, the survey statistic will be biased.

Differences due to measurement are where the mode in which the question is asked affects the way people respond. That is, the same person may give a difference response in mode A than in mode B, because of the way that the question is asked, the response options presented, the way the answer is recorded, etc. For example, people may be more willing to be honest about drinking, smoking, or other socially ‘undesirable’ activities when answering a survey online than if they are being interviewed in person.

The mode effect on measurement is at the level of the question, there may be some questions in a survey for which there is no mode effect, whilst other questions in the same survey may exhibit a strong mode effect.
In addition to mode effects in the response measures (response rates, response bias), we may see mode effects in the speed or timeliness of the data collection, and the costs of data collection.

Effects of mixed modes on sample selection

If we use a mix of modes in data collection, there is the potential to increase response rates and reduce response bias. Response *rates* would increase if non-respondents to the primary mode are more likely to participate if offered a second mode. Response *bias* would decrease if those who respond in the secondary mode are under-represented in the sample who responded to the primary mode.

On the BHPS, and then on *Understanding Society*, we use a sequential mixed mode design to increase response rates. Towards the end of the fieldwork period, we try to contact by telephone those adults who have not yet been interviewed face-to-face.
More recently, on *Understanding Society,* we wanted to reduce the cost of data collection, however, we wanted to do this without negatively affecting the quality of the study. This requires using a cheaper mode first, and then following up non-respondents using face-to-face interviewers.

We carried out an experiment on the second wave of the Innovation Panel (IP2) which tested two versions of a design which initially attempted to contact and interview sample members by telephone, with nonrespondents then issued to interviewers to contact in-person (two CATI-CAPI groups) against the standard in-person (CAPI) mode. Response rates were significantly lower in the mixed-mode design (Lynn et al, 2010). Furthermore, at the next wave, which was fully face-to-face, one of the CATI-CAPI groups from IP2 still had lower response rates (Lynn, 2011; Lynn, 2013). Analysing the characteristics of those who were less likely to respond in the mixed-mode design indicated that there were significant differences in response for those in older age groups, unemployed, without a mobile phone, and for those who do not use the internet (Lynn, 2011). Thus, the use of this particular mixed mode design had potential to introduce selection biases into the survey. The results of this experiment caused us to rule out the use of telephone first in a mixed-mode design for the main stage of *Understanding Society*.

We then used the fifth wave of the Innovation Panel (IP5) to test a web-first (Web-CAPI) design, compared to the single CAPI mode. A random allocation of two-thirds of households were initially invited to take part online, with non-responders followed up some weeks later by an interviewer. The remaining third of households were issued directly to an interviewer (Jäckle et al, 2015).
The web-first design resulted in similar response rates to the single CAPI design. Households issued to the single CAPI mode were significantly more likely to have interviews from *all* eligible adults in the household (Auspurg et al, 2013; Jäckle et al, 2013). Another important indicator in a mixed mode survey, where one of the goals is to reduce the costs of data collection, is the proportion of households that are fully completed online. That is, where there is no requirement to send an interviewer to the household to contact remaining non-respondent household members. Almost one in five (18.5%) of households in the original IP sample who were issued web-first completed all elements of the study online. This proportion was higher in the refreshment sample (for whom this was their second wave of interviews), where just over one-third of households completed online. This was achieved through the use of higher levels of

incentives (23% completed online when there was a £10 incentive, 37% with a £20 incentive, and 43% with a £30 incentive) (Jäckle et al, 2013).

Analysis indicated that there was no group for which the web-first design had a significantly higher response rate than the single mode, although there were groups for whom there was a significantly lower response rate in the mixed-mode sample, suggesting that there was a risk of selection bias (Jäckle et al, 2013).

At subsequent waves of the IP, the mixed mode sample did achieve higher response than the single face-toface mode. However, this was done using higher incentives for those issued to web. At IP6 the response for the CAPI-first sample, who all received £10, was 68%, for the web-first £10 sample it was slightly lower at 66%. However, for the web-first group who were offered £10 plus £20 if everyone in the household participated, the household response was 75%, and for the web-first group where everyone received £30 it was 76%. By IP7, the third mixed-mode wave, the response rate for the mixed-mode sample was significantly higher than the single CAPI mode (Al Baghal (eds), 2015). Again, this was due to the higher incentives for the web-first sample, where the two higher incentive groups achieved a household response rates around 10 percentage points higher than the CAPI group (Hanson et al, 2015). Analysis by Bianchi et al (2017) show that there were no cumulative differences between response rates in the mixed mode and single mode design over three waves, and there were only minimal differences in the sample composition.

With this evidence, we felt confident about introducing mixed modes onto the main sample of *Understanding Society* (University of Essex, 2018). However, the implementation was rolled out slowly to allow us to monitor and evaluate the results. At Wave 7, only adults in households that had not responded at Wave 6 were allocated to web-first, as this was the group who seemed to have been more receptive to an online interview than a CAPI interview in the IP. Using data from the IP, we modelled the propensity of households to respond online. This model was then run using main-stage data to give households a predicted probability of responding online. At Wave 8, aside from 20% of the sample that was ring-fenced to be CAPIfirst, households who were in the top 40% likely to take part online were issued web-first (Lynn, 2017; Knies (ed.), 2018).

At Wave 8, to maximise the proportion of households that completed online we used incentives as a bonus for those who completed their online interview within the initial period, before non-responders were issued to interviewers. This led to a large increase in the proportion of households that completed online from 18% without the bonus to 36% with the bonus (Carpenter and Burton, 2017).
We also experimented with other methods to increase the proportion of households that completed online during Wave 8 and found that increasing the number of reminders – including an additional letter reminder – and increasing the length of the web-only fieldwork also led to higher web-completion rates during the initial web-only period (Carpenter and Burton, 2017). It should be noted, however, that around half of those issued to web-first still required a follow-up in the secondary mode.

The requirement to follow-up in a secondary mode is needed to ensure that the overall sample composition is representative of the population. Analysis of the Innovation Panel has shown that the mixed-mode sample at IP5 was less representative of the population than the CAPI-only sample. However, this difference was smaller at the next couple of waves where the mixed-mode design had higher response rates than it had at IP5 (Bianchi et al, 2017). The CAPI follow-up of web non-respondents was more successful at balancing the sample and making it more representative at IP6 and IP7, than it had at IP5.

A comparison of responses elicited online against those face-to-face indicate that there were significant differences in just over 18% of questions. This may have been either because the same people would have answered differently in different modes (measurement effect) or because the responding sample online were different in some ways to those who responded in-person (selection effect). Controlling for selection effects, to the extent possible, the proportion of questions with significantly different responses online and face-to-face was reduced to just under 3%, suggesting that most of the differences in responses online and face-to-face are due to the different responding samples. An analysis of respondent characteristics shows that at IP5 those less likely to participate in a mixed mode survey than a face-to-face survey were those who lived in rural areas, those aged 21-30, single parents, couples with children, and those who do not have access to the web. The systematic differences between face-to-face and the mixed mode samples reduced at IP6 and IP7 as the response rate for the mixed mode survey increased, and by IP7 it was only those who expressed preference for responding by web that were significantly more likely to participate in a mixed-mode design (Bianchi et al, 2017). However, there were still significant differences between those who responded online and those who responded face-to-face. Thus, a face-to-face follow-up of online nonresponders is still necessary to reduce selection effects (Jäckle 2015).

Effect of mixed modes on measurement

As noted above, the mode of a survey may affect how a person answers the questions. This may be because of the way the question and/or answer options are presented (verbally, by an interviewer, or visually, onscreen), it may be because there is a (positive or negative) effect of having an interviewer present which may affect the respondent's motivation to answer or influence an answer. For example, item non-response (where a particular question is not answered) is usually higher in self-administered surveys, such as web surveys, than in interviewer-administered surveys (Jäckle et al, 2015).

At IP5 the level of item non-response was significantly higher (although at a low level) for the mixed mode design compared to the single face-to-face design. For a key question about gross pay, the level of item non-response was almost 8 percentage points higher (10.0% for face-to-face, 17.7% for mixed mode) (Jäckle et al, 2013). An experiment was implemented at IP6 to test the effectiveness of different methods of reducing item non-response for a small number of key questions (Al Baghal and Lynn, 2014). Using a motivational statement which appeared if the respondent skipped the question without answering, they were able to bring down the item non-response level on these key questions to that of the face-to-face group.

A concern we had on *Understanding Society* is that the web mode might fail to pick up new entrants to a household, or people who have left the household. On a household panel study, the first task after making contact is to enumerate the household. This involves listing all those who were present in the household at the previous wave and checking whether they are still resident. Where someone has left, interviewers need to collect information about the leaver (when and why they left, where they have moved to). Interviewers also need to check to see if anyone has joined the household and collect information about these new entrants and identify whether the new entrant has been part of the study before and so is a 're-joiner', so that the data from their previous interviews can be correctly matched to the present interview. This can be a complex task and translating this to a self-completed online questionnaire where there was no interviewer available to help the respondent was a challenge. With a complex task such as this, it was possible that a respondent completing it online would just confirm that all the previous residents were still resident, and we may miss new entrants and leavers in the household, resulting in poorer quality data and errors in a key measure of interest (dynamics of household composition). Analysis of the household grid in IP5 and IP6 has indicated that there are no significant differences in the numbers of leavers or new entrants by mode (Jäckle 2015).

We have found that there are significant mode differences when it comes to asking the respondent for their consent to link administrative data held on them by government agencies to their survey responses; with those interviewed online being much less likely to consent (Burton, 2016; Jäckle et al, 2018). This mode difference has also been found on other mixed mode studies (Calderwood, 2016; Thornby et al, 2017). Controlling for selection into mode, the difference between online and in-person consent rates *increases* which suggests that those who complete online would be expected to consent at higher rates, but the act of completing the survey online rather than in-person with an interviewer, reduces their likelihood to consent (Jäckle et al, 2018).
Measurement of cognitive function is another area where results are significantly different for those interviewed online to those interviewed in-person. One challenge is to identify tests which can be carried out online, and so do not rely on interviewer-administration, and where the results are not affected by the type of device used or speed of internet connection. Although there has been relatively little research on the results of cognitive function tests in a mixed mode survey, what has been done has shown that mode has a large effect. Using the IP, Al Baghal (2017) found that even controlling for selection into mode, those who complete cognitive tests online significantly out-performed those who completed in-person. This is echoed in research using the Health and Retirement Study in the US (McClain et al, 2018).

Another measure of survey quality is the timeliness of the data (Biemer and Lyberg, 2003). There is little research on the effect of mixed mode surveys on the time required for a survey. Whilst online-only surveys can produce data relatively quickly, for example using online non-probability access panels, probability sample surveys take longer because of the requirement to make multiple contact attempts (reminders). There is a tension between quality, cost, and timeliness. On a sequential mixed-mode survey that uses the web, the time required to complete the survey is at least the same as that required for the secondary mode. Our experience on *Understanding Society* is that the fieldwork length of a mixed-mode wave is longer than that for a single-mode wave.
To enable the interviewers to have all of their workload at the same time, to allow them to work efficiently, the web period starts earlier than the face-to-face period. Wave 6, which was the last fully face-to-face wave, had a fieldwork period of 23 weeks for each sample month. From Wave 9 onwards, the fieldwork period for each sample month is 28 weeks, which includes the initial 5-week web-only period. In addition to the longer fieldwork period, a mixed mode survey takes longer at either end of the survey process. It takes longer to specify, script, and test a questionnaire which is used in two or more modes. It also takes longer to process the data from two different sources (e.g., interviewer lap-tops and internet servers) and to combine them into one data-set.

Mixed modes and survey costs

Different modes have different cost profiles; a different mix of 'fixed' and 'variable' costs. The fixed costs are those that are required independently of the number of interviews done, whilst variable costs are those that vary as the sample size varies. For example, face-to-face surveys have high proportion of variable costs (interviewer time, travel, payment for interviews) whilst a web survey has a high proportion of fixed costs (programming the questionnaire) after which the cost of each additional respondent is very small (Couper, 2016).

For a mixed-mode survey, the costs depend on the fixed and variable costs of both modes. So, for *Understanding Society*, we have the fixed costs of developing the questionnaire in both modes and the variable costs of using interviewers to follow-up non-responders to the web survey. In addition to the sum of costs for both modes, there are additional costs – the sample management and data processing of a mixed mode survey incurs additional costs over the cost of each mode. For studies with large sample sizes, such as *Understanding Society*, variable costs make up a larger proportion of total costs. Therefore, one focus for longitudinal studies is the push-to-web, to encourage and motivate sample members to take part online. For a household panel study, like *Understanding Society*, there is the added complication that we aim to interview all adults in the household, and so if there is one adult in a multiple-adult household that does not complete their survey online, we must send an interviewer to the address to complete the interview inperson. Cost savings, therefore, are made if whole households complete online, or even more so if entire interviewer allocations are completed online.

There is not much research in this area, because of the commercial sensitivity in revealing costs, and the difference between costs and price. Bianchi et al (2017) estimate savings in data collection (i.e., variable costs) for IP5-IP7 and indicate that the mixed mode design provides cost savings of around 15% at IP5, 8% at IP6, and 11% at IP7. The slightly lower savings at IP6 and IP7 are due to the increased level of incentive used in these waves.

On Wave 8 of *Understanding Society* we worked with Kantar Public to implement an adaptive design to increase the proportion of households that completed online. With information from Kantar about the costs and the savings, we could see how the additional cost of the early-completion bonus was a little less than the saving made because it led to a higher proportion of whole households completing online. On the adaptive design project, most of the experiments we implemented saved more than they cost (Carpenter and Burton, 2018). However, an increasing web completion rate also affects the costs of face-to-face interviewing; because there are fewer households issued to interviewers, but these households are still spread over the UK, and are likely to be harder-to-reach households. This has meant that the per-interview cost of a face-to-face interview has increased. Following on from the adaptive design for Wave 8, we have worked with Kantar Public to use an ‘Open Book’ system of managing variable costs and the uncertainty about mode uptake. Using the Open Book model to simulate a 100% CAPI-first design and a 30% CAPI-first design whilst keeping the issued sample size and the cost to issue households constant, the model suggests that the 30% CAPI-first design will cost around 84% of the 100% CAPI design. Overall, we have found that cost savings on *Understanding Society* have been so far been relatively modest, given the need to maintain a high-quality study.

Mixed modes: conclusion

To conclude this section, there is an increasing move to web surveys in longitudinal studies. However, in the current context, it is unlikely that we will be able to rely on single-mode web surveys for rigorous academic/government research because of the issues around mode effects due to selection and measurement. The focus is therefore on push-to-web designs to increase the flexibility and convenience for respondents, reduce overall costs and with the potential of increasing response where web non-respondents are followed up in a second mode. There is the potential to reduce costs on large longitudinal studies, although the cost reductions are more modest on designs which attempt to interview all household members but could be more substantial for surveys where only a single participant in a household is required.

- Push-to web designs that aim to maximise online completion of the survey are popular, given aims to contain costs of data collection. The protocols specifying how the modes are implemented can have large effects on take up of the web survey.
- Using web only is still not possible in the UK, because of bias in the types of people who have access to the internet and in the types of people who will complete a survey online. Therefore follow ups of web non-respondents in other modes are still key, if the aim is to represent the general population.
- Variable costs per interview are much lower with an online questionnaire than with interviewer administered questionnaires. However using more than one mode of data collection can increase variable costs in the other modes and certainly increases fixed costs of data collection.
- The answers respondents give in an online survey seem to be similar to those given to interviewers. There are however four key challenges regarding measurement with push-to-web surveys: sensitive

information is more accurately reported in self-completion surveys than with interviewers which limits comparability of answers[1], tests of cognitive ability produce very different results between modes, consent rates for data linkage are much lower online than with interviewers, and protocols to collect biomeasures without interviewers are still under development.

Novel methods of data collection

As technology becomes ever more portable and pervasive, additional opportunities are opening up for data collection. Wearables, such as FitBit and similar kit, allow continuous or near-continuous monitoring of movement, heart rate, temperature, and air pollution. More commonly, mobile phones can be multiplemode survey devices; allowing for CATI and web interviews, as well as enabling users to take photos (e.g., of location, receipts, meals) and to share their geo-location. The expanding use of apps opens up even more opportunities for the collection of information, actively or passively. For more about new technologies and data sources for data collection, see Jäckle et al (2019a).

Compared to survey questionnaires, new methods offer ways of collecting more detailed data (such as continuous monitoring over periods of time), more accurate data (for example by collecting data passively rather than asking respondents to recall information), about concepts that respondents would not be able to report on (such as GPS location or air quality), and in ways that might be less burdensome for respondents. New methods tend to be specific to particular domains. This means that new methods cannot replace questionnaires, but instead complement them. For example, a survey might supplement a questionnaire with asking respondents to use a wearable device to monitor their physical activity and to use an app to record their time use. New methods of data collection can therefore be thought of as new modes – expanding the challenges of mixed mode data collection. As with questionnaire based data collection, how useful the data collected with new methods are for research depends on the selectivity of the data (about whom are data collected?) and measurement (how well are the concepts of interest measured?).

There are only few studies that have implemented new methods of data collection and examined the implications for data quality. In the following we provide a brief review of key issues and include findings from our own studies. *Understanding Society* has been testing new methods of data collection as part of a project on “Understanding household finance through better measurement”, funded by the ESRC and NCRM (ES/N006534/1). In the first Spending Study (2016) we invited members of the Innovation Panel to download an app to their mobile device and to use that to report all purchases for one month, by taking photos of their shopping receipts. This method was used as an alternative to either recall questions in the annual interview or paper based spending diaries. Based on learning from this first study, we implemented a second Spending Study in 2018 (see Jäckle et al, 2019b; 2019c).

New methods and non-response/selectivity

One feature that new methods of data collection have in common is that there are more stages at which respondents can drop out of the data collection than there are with a questionnaire. Even with passively collected data, respondents have to comply with several steps in order for the data to be collected and there is dropout at each stage. In the Spending Studies, where participants had to have a compatible mobile device, download the app, and use it to record their purchases for one month, we found that only

[1] Although many of the more sensitive questions on *Understanding Society* have always been self-completion (paper questionnaires Wave 1-2, CASI Waves 3+), there are some areas in the main questionnaire which may be sensitive to some respondents (e.g., reports of discrimination and harassment). See d’Ardenne et al (2017) for more information.

about three quarters of respondents had a mobile device and less than one fifth participated in the app studies. However those who did participate were committed, reported their spending for most of the month (Jäckle et al, 2019d), and reported low levels of burden (Read 2019a). Whether a respondent participated in the
app studies was strongly related to whether they already did similar tasks for their own purposes (in this case, keeping a budget) and whether they already did such types of tasks on their mobile devices (such as using a mobile app to check their bank balance). A key finding from both Spending Studies is that while the participation rates were low, and the types of people who participated were not random sub-sets of the full respondent sample, there were no differences between participants and non-participants in the outcome that the app was designed to measure: expenditure (Jäckle et al, 2019d; 2019e).

Spending Study 2 was designed to test how best to implement the app study, in order to increase participation and reduce selectiveness of participants. In the first study we had already found that offering a higher unconditional bonus simply for downloading the app did not increase participation rates. Other studies have similarly found that offering higher incentives does not seem to help increase participation rates in data collection using new methods (Angrisani et al, 2018; Haas et al, 2018). The requirement to download an app had been an important barrier to participation in the first study. In the second one we found that offering a browser-based alternative increased participation substantially, if it was offered at the point when respondents declined to download the app (rather than being offered as an alternative at a later point). Offering respondents feedback about their spending did not increase participation. However inviting respondents to the app study within the annual Innovation Panel interview nearly doubled the participation rate, compared to sending an invitation letter by post.

New methods and data quality

How well new methods measure concepts of interest is a key challenge. Some methods provide data that map onto concepts of interest directly (for example, sensors measuring air quality provide data about the quantity of certain particles in the air). Other methods provide data that need to be transformed and processed in order to derive the indicators of interest. For example, GPS location data could be used to passively measure travel, however identifying locations visited and modes of transport is challenging.
Similarly, deriving information about the extent and nature of physical exercise from accelerometry data (that simply captures the direction and speed of changes in location) is challenging. In some cases researchers receive the unprocessed data, in others cases the technology includes algorithms to derive indicators and researchers do not receive the raw data (or the algorithms). In these cases, part of the measurement process is in fact not in the control of researchers. Differences in technical specifications and algorithms between devices, and within devices over time as these are updated, can affect the quality of measurement and comparability of data collected with different devices. Similarly, in the first Spending Study we found that when respondents use their own mobile devices for app-based data collection, differences between devices in the technical specifications can impact the quality of data captured (Read 2019b).

The ultimate test of the quality of data collected with new methods is comparing estimates with a gold standard. We have carried out such a comparison with the first Spending Study, comparing estimates of expenditure to estimates derived from the UK Living Cost and Food Survey (LCFS). The comparison suggests that despite different sources of non-response and measurement error in the app study (based on images of shopping receipts) and the paper based diary used in the LCFS, estimates of total expenditure are comparable (Wenz et al 2018).

New methods and survey costs

We are not aware of any documentation of the costs of collecting data with new methods, or how costs compare to questionnaire based data collection. Methods where devices are loaned out to respondents (e.g. wearables) obviously involve higher variable costs than methods that rely on respondents' own devices, such as their smartphones. Most new methods are likely to increase the fixed costs of running a survey. The logistics of implementing data collection with new methods (e.g. developing protocols for data collection; developing information materials for respondents; sending out and retrieving kit; administering additional respondent incentives; processing raw data, etc.) are complicated and time consuming, increasing fixed costs of survey data collection. Programming bespoke apps for data collection is much more expensive and time consuming than one would expect, especially if the app needs to be developed for more than one operating system and needs to be updated over time (based on personal communication with researchers at Statistics Austria and Statistics Netherlands).

New methods: conclusion

New methods of data collection can increase the scope and quality of data collected in surveys, providing opportunities for innovative research. It is clear that the method or technology itself can affect the quality of data collected. In addition, the protocols used to implement that method or technology can have large effects on both selectivity of who participates and quality of measurement. The issues are essentially an extension of the questions around the effects of mixed mode data collection on data quality. To date there is however very little empirically based guidance on how best to implement data collection using new methods to ensure the quality of data collected with new methods. This is an area where there is a great deal of research interest and where there is the potential for innovative experimental research.

References

Al Baghal, Tarek, and Peter Lynn (2014) “Using motivational statements in web instrument design to reduce item missing rates in a mixed-mode context”, *Understanding Society* Working Paper, 2014-02.

Al Baghal (eds) (2015) *Understanding Society* Innovation Panel Wave 7: Results from Methodological Experiments, *Understanding Society* Working Paper, 2015-03.

Al Baghal, Tarek, (2017) “The effect of online and mixed-mode measurement of cognitive ability.” *Social Science Computer Review*,
1-15. (Published online)

Angrisani, Marco, Arie Kapteyn, Swaroop Samek (2018)“Real Time Measurement of Household Electronic Financial Transactions in a Population Representative Panel”, Paper prepared for the 35th IARIW General Conference, Copenhagen, Denmark.

Bianchi, Annamaria, Silvia Biffignandi, Peter Lynn (2017) “Web-Face-to-Face Mixed-Mode Design in a Longitudinal Survey: Effects on Participation Rates, Sample Composition and Costs”, in *Journal of Official Statistics,* Vol. 33, No. 2, 2017, pp. 385–408, http://dx.doi.org/10.1515/JOS-2017-0019.

Biemer, Paul P. and Lars E. Lyberg (2003) *Introduction to Survey Quality*, Wiley.

Burton, Jonathan (eds.) (2013) “*Understanding Society* Innovation Panel Wave 5: Results from Methodological Experiments”, *Understanding Society* Working Paper, 2013-06.

Burton, Jonathan (2016) "Results for Web/Face-to-Face Linkage Consent Questions in the Innovation Panel.", presented at the *Mixing Modes and Measurement Methods in Longitudinal Studies Workshop*. London: CLOSER.

Calderwood, Lisa (2016) "Asking for Consent in a Mixed Mode Study – Issues and Outcomes." in *Presented at the Mixing Modes and Measurement Methods in Longitudinal Studies Workshop*. London: CLOSER.

Carpenter, Hannah, and Jonathan Burton (2018) “Adaptive push-to-web: experiments in a household panel study”, *Understanding Society* Working Paper, 2018-05.

Couper, Mick P. (2016) “Costs of Mixed Mode Surveys: Some Considerations”, presented at the *Mixing Modes and Measurement Methods in Longitudinal Studies Workshop*. London: CLOSER.

D’Ardenne, Jo, Debbie Collins, Michelle Gray, Curtis Jessop, and Sophia Pilley (2017) “Assessing the risk of mode effects: review of proposed survey questions for waves 7-10 of Understanding Society”, *Understanding Society* Working Paper, 2017-04.

Groves, Robert M., Floyd J. Fowler Jr., Mick P. Couper, James M. Lepkowski, Eleanor Singer, Roger Tourangeau (2011) Survey Methodology, 2nd edition, Wiley, 2011.

Haas, Georg-Christoph, Frauke Kreuter, Florian Keusch, Mark Trappmann, Sebastian Bähr (2018) *What do researchers have to invest to collect smartphone data?*. Paper presented at the BigSurv18 conference, Barcelona.

Hanson, Tim, Peter Matthews, Alice McGee (2015) “*Understanding Society* Innovation Panel Wave 7: Technical Report”. https://www.understandingsociety.ac.uk/sites/default/files/downloads/documentation/innovation-panel/technicalreports/IP7_TechReport_v4.pdf

Jäckle, Annette (2015) *Impact of sequential Web/FTF mixed modes on survey error.* Internal Memo. ISER, University of Essex.

Jäckle, Annette (2016) *Identifying and Predicting the Effects of Data Collection Mode on Measurement*. Paper presented at the CLOSER Mixing Modes and Measurement Methods in Longitudinal Studies Workshop, London.

Jäckle, Annette, Peter Lynn, Jonathan Burton (2013) “Going Online with a Face-to-Face Household Panel: Initial Results from an Experiment on the *Understanding Society* Innovation Panel”, *Understanding Society* Working Paper, 2013-03.

Jäckle, Annette, Peter Lynn, Jonathan Burton (2015) “Going Online with a Face-to-Face Household Panel: Effects of a Mixed Mode Design on Item and Unit Non-Response”, in *Survey Research Methods,* 9(1), pp. 57-70.
Jäckle, Annette, Alessandra Gaia, Michaela Benzeval (2017) “Mixing modes and measurement methods in longitudinal studies”, CLOSER Resource Report. http://www.closer.ac.uk/wp-content/uploads/Mixing-modes-and-measurement-methods-in-longitudinalstudies.pdf

Jäckle, Annette, Kelsey Beninger, Jonathan Burton, Mick P. Couper (2018) “Understanding data linkage consent in longitudinal surveys”, *Understanding Society* Working Paper, 2018-07.

Jäckle, Annette, Alessandra Gaia, Carli Lessof, Mick P. Couper (2019a) “A review of new technologies and data sources for measuring household finances: Implications for total survey error”, *Understanding Society* Working Paper, 2019-02.

Jäckle, Annette, Jonathan Burton, Alexander Wenz, Brendan Read, Tim Hanson, David Xu (2019b) *Understanding Society*: The UK Household Longitudinal Study Spending Study 2 User Guide. Colchester: University of Essex.

Jäckle, Annette, Jonathan Burton, Alexander Wenz, Brendan Read (2019c) *Understanding Society*: The UK Household Longitudinal Study Spending Study 2. User Guide – Appendix A: App and Online Diary Screenshots. Colchester: University of Essex.

Jäckle, Annette, Mick P. Couper, Jonathan Burton, Carli Lessof (2019d) “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*, Vol. 13(1), pp. 23-44.

Jäckle, Annette, Alexander Wenz, Jonathan Burton, Mick P. Couper (2019e) “Increasing participation in a mobile app study: the effects of a sequential mixed-mode design and in-interview invitation”, *Understanding Society* Working Paper, 2019-04.

Knies, Gundi (ed.) (2018) *Understanding Society*: Waves 1-8, 2009-2017 and Harmonised BHPS: Waves 1-18, 1991-2009, User Guide, Colchester: University of Essex. https://www.understandingsociety.ac.uk/documentation/mainstage/user-guide

Lynn, Peter, S.C. Noah Uhrig, Jonathan Burton (2010) “Lessons from a randomised experiment with mixed-mode designs for a household panel survey”, *Understanding Society* Working Paper, 2010-03.

Lynn, Peter (2011) “The effect of a mixed mode wave on subsequent attrition in a panel survey: evidence from the *Understanding Society* Innovation Panel”, *Understanding Society* Working Paper, 2011-06.

Lynn, Peter (2012) “Mode-switch protocols: how a seemingly small design difference can affect attrition rates and attrition bias”, *Understanding Society* Working Paper, 2012-07.

Lynn, Peter (2017) “Pushing Household Panel Survey Participants from CAPI to Web”, paper presented at the 28th International Workshop on Household Survey Nonresponse, Utrecht.

McClain, Colleen A., Mary Beth Ofstedal and Mick P. Couper (2018) “Measuring Cognition in a Multi-mode Context”, Survey Research Center, Institute for Social Research, University of Michigan. https://hrs.isr.umich.edu/publications/biblio/9606

Medway, Rebecca L., and Jenna Fulton (2012) When More Gets You Less: A Meta-Analysis of the Effect of Concurrent Web Options on Mail Survey Response Rates”, in *Public Opinion Quarterly*, Vol. 76(4), pp. 733-746.

Read, Brendan (2019a) “Respondent burden in a mobile app: evidence from a shopping receipt scanning study”, *Survey Research Methods*, Vol. 13(1), pp.45-71.

Read, Brendan (2019b) “The influence of device characteristics on data collection using a mobile app”, *Understanding Society* Working Paper, 2019-01.

Thornby, Marie, Lisa Calderwood, Mehul Kotecha, Kelsey Beninger, Alessandra Gaia (2016) "Collecting Multiple Data Linkage Consents in a Mixed Mode Survey: Evidence and Lessons Learnt from Next Steps." in *Unpublished manuscript*. London: Institute of Education.

University of Essex, 2018. Institute for Social and Economic Research, NatCen Social Research, Kantar Public. *Understanding Society: Waves 1-8, 2009-2017 and Harmonised BHPS: Waves 1-18, 1991-2009*. [data collection]. *11th Edition.* UK Data Service. SN: 6614, http://doi.org/10.5255/UKDA-SN-6614-12

University of Essex, 2019. Institute for Social and Economic Research. *Understanding Society: Innovation Panel, Waves 1-11, 20082018.* [data collection]. 10th Edition. UK Data Service. SN: 6849 http://doi.org/10.5255/UKDA-SN-6849-11
Voorpostel, Marieke, Oliver Lipps, Caroline Roberts (2019) "Mixing modes in household panel surveys: recent developments and new findings", chapter in Peter Lynn (ed.) *Advances in the Methodology of Longitudinal Studies*, Wiley.

Wenz, Alexander, Annette Jäckle, Thomas F. Crossley, Mick P. Couper, Jonathan Burton, Joachim Winter (2018) "Quality of expenditure data collected with a receipt scanning app in a probability household panel", paper presented at the Bureau of Labor Statistics, Washington DC.

Understanding Society
Institute for Social and Economic Research (ISER)
University of Essex
Wivenhoe Park
Colchester
CO4 3SQ
Tel: +44 (0) 1206 872957

www.understandingsociety.ac.uk

@usociety

Understanding Society –
UK Household Longitudinal Study

University of Essex

Understanding Society has been commissioned by the Economic and Social Research Council (ESRC). The Scientific Research Team is led by the Institute for Social and Economic Research (ISER) at the University of Essex.