New technologies in survey research

July 2019

Ipsos MORI Social Research Institute

Patten Smith and Faith Jones

Contents

1. Executive Summary

1 Executive Summary

Overview

This report aims to explore 'innovative ways of asking survey questions', including alternatives to traditional self-report methods and face to face interviews. We have looked at innovative methods for collecting survey data (rather than just asking questions) to ensure that our discussion covers the full range of new measurement techniques currently being examined by methodologists.

Purpose of research

The overarching purpose of this report is to provide information to help inform thinking about the design and feasibility of innovative data collection methods, in both existing longitudinal studies and future studies. This report discusses the use of new technologies within social research with a specific focus on their suitability for use in high quality in longitudinal surveys.

Description of approach

In order to ensure the most promising innovative methods were included in this report we identified and assessed a wide range of innovative data collection methods. These included innovative ways of administering conventional survey questions as well as data collection using more advanced technological methods. We reviewed a wide range of sources and consulted recognised survey experts to ensure that we identified an extensive range of candidate methods. We used the Total Survey Error (TSE) framework to assess the data quality and overall error associated with each innovative method we discussed.

Summary of findings

We have identified five broadly defined data collection methods that have been recently used or tested for use in market research and social research quantitative surveys. Due to the speed of technological advancement we are unable to predict exactly how technology will impact on data collection in the future, therefore we chose to limit our focus to methods that were in more advanced stages of development. These are:

1. The use of smartphones for online surveys:

The growth in smartphone ownership and use makes it vital to ensure there is a focus on questionnaire adaption for smartphones when designing online surveys to minimise measurement error and non-response error.

2. App based measurement of activity type in diaries:

There are a variety of ways smartphone apps can be used to collect activity data, both as a potential alternative to traditional paper/ web diaries, but also by integrating other smartphone-enabled information like photographs and GPS information. These appear to have a promising future, although it appears that to minimise coverage error and non-response error requires expensive methods such as providing devices and face-to-face recruitment

3. In-the-moment surveys:

In-the-moment surveys enable the capture of data in real-time using questionnaires trigged by either technology (e.g. GPS) or participants themselves. They facilitate data capture and measurement of changes in experience, emotions attitudes and behaviours at specific points of reference. Current limitations are coverage (not everyone has smartphones that enable triggers to be applied) and non-response (e.g. not downloading app, not responding to trigger, trigger failing to work).

4. Health and activity monitoring:

Wearable technology and smartphone sensors can capture data about health and activity. We outline the type of health data that can be collected using smartphones and wearable devices and discuss how this could provide a more objective data collection method and more accurate measurement. Specific limitations (beyond coverage and non-response) are related to the heterogeneity of different devices (making it difficult to use respondents' own devices) and the (current) high cost of providing generic devices.

5. Non-health passive monitoring

Lastly, we looked at the potential use of passive monitoring using smartphone apps and fixed sensors. Passive measurement enables a greater level of detail, accuracy and enables the capture of information which has previously been impossible to measure when using traditional self-report survey methods, such as travel data, energy use, online behaviours. The same limitations apply as for health and activity monitoring – coverage, non-response, heterogeneity and cost.

Conclusions and Recommendations

The majority of innovative methods we discussed were subject to limitations, specifically coverage and non-response, resulting in a potential for fixed and variable errors of representativeness. We believe that, at present, cooperation rates are poor, therefore use of app/ device based technologies in high quality surveys would not be recommended as a matter of course. However, attitudes to technology may change and it is likely that as people become more familiar and engaged with technology their willingness to use apps/ devices will increase. It would be necessary to conduct rigorous preliminary methodological feasibility work, not only to address and minimise errors of representativeness but also to better understand participants' engagement with, and willingness to use technology for research. Until this has been done the use of many innovative data collection methods in high quality surveys will be hindered.

2. Introduction

2 Introduction

The brief for this work states that 'it should assess innovative ways to administer survey questions, including alternatives to face to face interviewing, and the use of technology-based approaches and evaluations of these'. It is intended that it should 'tap into the latest developments and innovative thinking in survey methodology and the practical testing and implementation of methodological innovations.' It is also intended that the work will inform methodological innovations in current and future methodological studies. Particular mention is made of representativeness of results and the need for value for money. The brief was addressed specifically to UK based organisations working in survey data collection and methodology. In this report, we interpret 'innovative ways of asking survey questions' broadly as meaning 'innovative methods for collecting survey data' to ensure that our discussion covers the full range of new survey measurement methods currently being examined by survey methodologists. We focus on fresh data collection and exclude data linkage to 'organic' data (Groves, 2011) such as administrative or social media data.

In recent years technology has significantly altered how the public interacts with businesses and the public sector. The great majority of the UK population (87%) now has a home connection to the internet, 79% of the adult population uses a smartphone, and most people now habitually use advanced technology in many spheres of their everyday lives (Ofcom, 2019). In market research these developments are reflected both in the mushrooming of online survey data collection over the past two decades, and in the emergence of new data collection methods such as activity measurement using wearable accelerometers, the use of GPS to track individuals and trigger place-specific data collection, passive measurement of broadcast media exposure, etc.

Although academic and Government funded survey research has also started adopting new technologies, it has lagged behind the commercial research sector. This lag can be partly justified by the surveys' having more stringent quality requirements, but it remains open as to whether this might also be due to institutional conservatism or lack of familiarity with cutting-edge data collection technologies. This report indirectly examines this assessing new technologies' current readiness for cost-effective high quality survey data collection.

We have identified five broadly defined data collection methods that have been recently used or tested for use in market research and social research quantitative surveys. Because we cannot fully predict how technology will impact on survey data collection over the next decade or so, we have limited our attention to methods in relatively advanced stages of development. We have also restricted our review to methods used to collect primary data and have excluded methods for linking respondent records to administrative or social media data. For each type of method, we describe how it works and what it requires of survey respondents, its advantages and limitations relative to competing methods and assess its future prospects for use in high quality survey research.

In assessing potential advantages and limitations of each method, we are largely guided by the Total Survey Error (TSE) framework (Groves et al, 2009). This framework is universally used by survey methodologists for the assessment of survey quality and involves the methodical identification of all sources of error (defined as the difference between the true population value of a survey parameter and the survey estimate of this value) which can arise at each stage of the survey process. It divides the survey process into two main strands, one concerning the representativeness of the survey sample and one concerning the accuracy of measurements made. Each error source can generate: (i) variable error generated by underlying random processes and (ii) fixed errors (bias) which would remain fixed if identical survey methods were applied repeatedly. We have focused on the three error sources especially relevant to the assessment of innovative data collection methods in the context of the ESRC funded longitudinal surveys:

1. *Non-coverage error* – can the eligible survey sample be approached / accessed using the relevant technology?
2. *Non-response error* – what proportion of approached eligible sample members can be contacted and persuaded to provide data, and do survey responders and non-responders differ in important ways?
3. *Measurement error* – do survey questions collect accurate data?

In the next section we examine the use of smartphones to administer self-completion web surveys, the use of diary apps to track type of activity, the use of apps to administer 'in-the-moment' surveys, the use of apps and wearables to collect health and activity data, and the use of apps and wearables to collect other forms of passive data. We then discuss non-coverage and non-response error, which we believe to be the main impediments to the adoption in high quality surveys of many of the technologies discussed in the report. We conclude with a brief assessment of future prospects.

3. The new methodologies

3 The new methodologies

Using smartphones in online surveys

What is meant by smartphone online survey data collection?

As smartphones ownership has increased in recent years, an increasing number of survey respondents are choosing to use them to respond to online administered surveys. This development is inexorable and it is now only possible to implement high quality online surveys if they allow smartphone questionnaire completion. However, conventional online questionnaires require considerable adaptation if they are to be successfully administered on smartphones, mainly to take account of their smaller screen sizes and use of touchscreens. This section deals with smartphone administration of traditional online surveys, rather than with the use of additional smartphone capabilities such as GPS and accelerometer sensors for data collection (these are dealt with later)[1].

Why use smartphones for online survey data collection?

We administer web surveys on smartphones because we must. The Ofcom technology tracker (Ofcom, 2019) found that, of UK internet users, those considering the smartphone to be the most important device for connecting to the internet increased from 33% in 2015 to 52% in 2018. Furthermore, the proportion of online survey respondents using smartphones for survey completion has increased over recent years (e.g. see Gummer et al, 2018; Roberts and Bakker, 2018). This is well illustrated in two recent Ipsos MORI surveys (Ipsos MORI, 2019; personal communication). In the Active Lives push-to-web survey of sport participation among English adults the proportion of online respondents using smartphones for survey completion increased from 16% in 2015/2016 to 25% in 2018/2019. Similarly, in the National Student Survey, a survey of final year (mainly university) students in Britain, the proportion of those completing online questions who used smartphones more than doubled from 23% in 2016 to 49% in 2019.

In the preceding years when smartphones were less common, researchers either ignored them altogether and made no adaptations for smartphone completers, or they encouraged respondents to use desktops/laptops rather than smartphones. Nowadays, neither of these approaches are effective and we must accept that respondents will often try to complete questionnaires on smartphones thus it is up to us to optimise questionnaires accordingly (Couper, 2019; Nicolaas, 2019).

Limitations of smartphone survey administration and how best to address them

If smartphones were used as a single data collection mode it would be important to assess coverage and non-coverage error. However, this is unnecessary because currently we would only consider using smartphones alongside other devices in multiple-device online surveys; for such surveys the only relevant coverage statistics relate to population coverage using *any* data collection device used.

[1] In reality the desktop-smartphone/mobile dichotomy is an oversimplification and for survey purposes devices are better regarded as lying on a continuum desktop->laptop->tablet->phablet->smartphone, based mainly on screen-size. (Couper, 2019) However, much research indicates that tablets behave more like laptops/desktops than smartphones, and we therefore focus our attention on smartphones v. larger screened devices.

Non-response error is of greater concern; however, it is hard to quantify because in a mixed mode survey where many respondents have access to more than response device (e.g. both a laptop and a smartphone), it is impossible to specify the eligible base size for both mode-specific response rates and non-response bias. Despite this, device-choice remains relevant to non-response bias; there is evidence both that the proportion of internet users preferring smartphone access is increasing, but also that for smartphones response rates are lower and breakoff rates higher than they are for desktops/laptops even where mobile users are randomly allocated to device type. A recent review found that response rates for PC-administered surveys were between 3% and 43% higher than they were for smartphone administered ones, and that those who did not respond to a mobile web survey tended to have less advanced phones and used them less did mobile respondents (Couper, Antoun and Mavletova, 2017). In a meta-analysis Mavletova and Couper (2015) found smartphone break-off rates to be around 8 percentage points higher for smartphones than for PCs. Relatively little is known about why these differences arise or how they should be addressed. There is also evidence that proportions responding on mobile devices can be increased by using SMS invitations rather than email ones (De Bruijne and Wijnant, 2014; Mavlevota and Couper. 2014) and by offering higher incentives for smartphone completion (Mavletova and Couper, 2016). However, the implications of these findings for general population mixed device survey practice are unclear.

Measurement error is the other main error type relevant to the administration of online surveys using smartphones. Measurement differences between smartphone and other device completions could arise for several reasons, notably, screen size, input method (finger v. mouse), connectivity, and context of use (location, presence of other people, simultaneous competing activities, etc.). Overall, the evidence is encouraging: researchers have not found pronounced measurements differences by device. Recent reviews (Couper, 2019; Couper, Antoun and Mavetova, 2016) indicate that there is:

- little difference on social desirability bias by device
- little difference between device type in response to and content of open-ended questions
- Mixed evidence on straight-lining, satisficing, response order effects and other indicators of satisficing ('lazy' answering)
- Some evidence that multi-question grids are better unpacked into separate questions when fielded on smartphones; however, formatting grids for smartphone completion is currently an active area of research

Much attention is currently being given to how best to optimise questionnaires for smartphone administration (Couper, 2019; Couper, Antoun and Mavletova, 2016). Recent guidance for high quality 'push-to-web' surveys (Ipsos MORI, 2019) has recommended that for smartphone administration questionnaires should:

- Reduce all non-essential content
- (Generally) avoid progress bars
- Keep question stems short
- Avoid drop-downs
- Use no more than 7 response options per question – or else use expandable headers

- Not use more than 5 points in attitude scale items
- Avoid sliders and small response buttons
- Avoid all horizontal scrolling
- Replace grids with single questions or with mobile-friendly grids
- Minimise use of instructions

More generally a 'mobile first' design approach is recommended for mixed device online surveys in which the questionnaire is firstly designed for small-screen devices and thereafter adapted for larger screen devices (Ipsos MORI, 2019; see Ipsos MORI 2017 for an example).

Future applications.

The numbers of respondents using smartphones to access the internet and wishing to complete online surveys on them will continue to increase. In parallel with this, the use of online data collection in high quality surveys (including ones funded by the ESRC), is also likely to increase, and it will be essential that these questionnaires are fully adapted for smartphone completion. Design principles are quickly being established, although further research is required to examine why response rates remain lower and break-off rates higher for smartphones than for other devices. It is also hoped that eventually it will be possible to extend the kinds of data collected on smartphones largely by means of downloaded apps, as discussed below.

App based measurement of activity type (e.g. time-use, travel, consumption) in diaries

What are they?

Diary data have been collected using paper and standard online questionnaires over many years. However, several recent exploratory studies have investigated how similar data may be collected using smartphone apps. Smartphone diary studies require respondents to download apps to their smartphones and use them to record relevant activities. Usually app-based diaries involve active respondent data input, but efforts have also been made to use them to collect purely passive travel data based on GPS monitoring. We discuss diaries requiring active respondent input here and cover passive data collection in the section on passive measurement below.

Why use them?

App diaries potentially have the advantage over pen and paper and standard online approaches because they encourage more timely data collection and thereby reduce the chance of respondents forgetting activities and other recall errors. Although respondents make diary entries themselves, diary apps can integrate this information with other data collected on smartphones such as GPS coordinates or photographs. Below we give two examples of activity diary app use recently presented at a CLOSER event on using new technologies in longitudinal surveys.

Examples of app diary studies

Time-use diaries

Smartphone time-use diary apps have been used in several studies of people's time-use (e.g. Fernee and Sonck 2014; Hendriks, Ludwigs and Veenhoven 2016; Vrotsou et al. 2014) including the age 14 wave of the ESRC funded Millennium Cohort Study (MCS) (Gilbert, et al, 2017). In the MCS study, respondents were asked to complete two 24-hour time-use diaries using a smartphone / tablet time-use diary app (supported by both IOS and Android operating systems) on their own devices, or to complete a more conventional web diary on a laptop or desktop. Respondents who had no access to, or who refused to use, a relevant device could complete a paper diary. Because alternative completion modes were on offer it is not possible to calculate a simple app usage response rate. However, 48% of the eligible sample produced time use data and of these 67% used the app, 27% used the web diary and 8% used the paper instrument. 73% of all app diaries returned in the combined pilot and dress rehearsal samples were defined as being of 'good quality' compared to 90% of online diaries and 65% of paper diaries. The app also produced fewer activity episodes than the other modes, but this may have been attributable to structural differences in the ways data were input across modes and / or to sample member self-selection.

Expenditure diaries

Jäckle et al (2017) conducted a study in which the Understanding Society Innovation Panel interview (wave 9) respondents were invited by letter (and email where possible) to download an app and use it to record household expenditure over a month. Both email and postal reminders were sent to non-respondents and conditional gift voucher incentives were offered. They were asked to use the app either (i) to take photographs of shopping receipts or (ii) to enter amount and category of their spending manually or (iii) to record that they had spent nothing on that day. Receipt data were analysed using software created by a market research company to analyse receipts for commercial purposes to extract information on product/service, retailer, price, time, etc. Although the sample approached were all recent participants in a long running longitudinal survey, and despite the additional incentivisation, participation in the study was modest with only 10.4% using the app at least once during each of the study weeks. The app was more likely to be used by women, younger and better educated sample members, and was associated with differences in financial behaviour (such as using online banking).

Limitations of app diary studies

As can be seen from these examples, app-based diary studies are probably too heterogeneous to allow generalisations to be made about measurement error. Errors in measurement of self-reported time-use are likely to differ considerably from errors in the measurement of photographed receipts or of reported travel.

For apps using respondents' own devices coverage error will depend on smartphone ownership rates and on differences in relevant characteristics between smartphone owners and non-owners as discussed in section 4 below. However, if devices are supplied to sample members, coverage error should no longer be an issue. There do not appear to be any additional difficulties (beyond the general ones discussed in section 4) with persuading sample members either to download a diary app or to agree to use a researcher supplied app, *if a face-to-face recruitment method is used* (e.g. Gilbert et al, 2017; Rofique et al, 2011). The low response rates observed for the expenditure study outlined above are likely to have derived from a combination of sensitivity of data collected, length of data collection period and use of a mail-based recruitment method.

Although face to face recruitment appears to be reasonably effective it is very expensive. Therefore, it would be beneficial to explore ways in which diary studies might achieve good response rates whilst using mixed mode recruitment methods.

Future applications

We consider app and supplied-device based diary methods to have a promising future, especially using face-to-face recruitment in established panels, noting the successful collection of time-use data in the most recent wave of the MCS as an example. However, their successful implementation in the absence of prior face-to-face recruitment remains to be proved, and detailed feasibility work is essential before any study of this sort is proposed.

In-the-moment surveys

What are they?

Traditionally, invitations to take part in surveys are launched en masse, approaching all potential respondents at broadly the same time once fieldwork has started. These invitations are initiated by the research team, either via a mass email or postal mailing, or by interviewers as they work through their allocations of sampled addresses or phone numbers. In contrast, in-the-moment surveys are triggered at disparate times relevant to individual respondents. These triggers are intended to allow surveys to capture data on emotions, attitudes and behaviours at specific points of reference (e.g. specific locations or activities). Triggers to participate in a survey can be led by technology (e.g. by using GPS tracking to trigger a survey when a participant has entered or exited a certain geographical area) or led by participants (e.g. every time they take their medicine). Participant-led triggers can be further enhanced through the use of technology such as Near Field Communication (NFC) stickers[2] which can be placed in relevant locations to remind respondents to trigger the survey.

Why use them?

In-the-moment surveys are best suited to research questions that are at risk of recall bias, or where it is difficult and/or expensive for face-to-face field interviewers to cover a geographic area.

In-the-moment surveys are often designed for multiple completions per respondent, reflecting the number of times a participant has had defined experiences. It allows for specific data to be collected as close as possible to the point of interest. This is especially useful where it is desired to receive specific feedback on a regular moment or behaviour, without requiring participants to recall all their relevant experiences (or perhaps their most recent experience) in a one-off questionnaire. This is likely to improve accuracy in relation to behaviours, and further reduces the risk of post-hoc rationalisation, which is particularly important for capturing accurate emotions and attitudes.

Examples of where 'geo-triggering' has been used to trigger a survey based on a location include litter 'hotspots' and major infrastructure roadworks. In both these cases, it was inappropriate and expensive to use face-to-face field interviewers, and, because data were collected from respondents whenever they went through the areas, it allowed the measurement of changes in experience.

[2] NFC stickers contain microchips which push information to a smartphone device. Similar to a QR scanning code, they can direct a participant's smartphone device to a specific internet website address – which in turn could be a survey.

Some examples of in-moment-surveys

Litter hotspots in Manchester

Moss and Ginnis (2014) conducted a study to evaluate the success of Manchester City Council's litter campaign. Participants were recruited from the council's customer contact database, and those who agreed to take part were asked to download a smartphone app which tracked their location. Participants received a notification to take part in a survey each time they entered one of four litter hotspots. The survey asked for real time feedback on the state of the area and views of whether this had improved compared to other experiences. The study received 138 entries from 62 participants and provided useful feedback on the level of cleanliness by time and day. A key challenge of the study was recruitment. Of 669 who completed a screener questionnaire, only 77.4% had a smart phone that was compatible with the technology; and only 37%; downloaded and logged in to the app.

Understanding MS sufferers in Spain

Bailey (2019) conducted a study to monitor adherence to medication and quality of life factors of MS sufferers over the course of one week. The study randomly allocated respondents in to three cells: cell 1 completed a diary survey through a mobile app; cell 2 completed a diary survey prompted by the use of NFC stickers placed in different parts of the home; cell 3 completed a survey at the end of a self-observation period. Each respondent in cell 2 received 4 NFC stickers; common places to put these included by the bedside table, on the medicine cupboard, or on the mirror. The study found that those in cell 2 who used NFC stickers were quicker to engage in the study than those in cell 1 (i.e. they were quicker to complete their first study), and that they completed on average twice as many surveys as cell 1. This enabled respondents to provide a greater volume of data closer to the moment of interest. This exploratory study was subject to coverage problems because the cell 2 app was not useable on the Apple iPhone.

Exploring the migrant-local happiness-gap

Hendriks et al (2016) used a range of different data collection methods to explore why internal migrants report lower levels of happiness than locals; this included used of an Experience Sampling Method (ESM) and a Day Reconstruction Method (DRM). Within the ESM approach, respondents were asked to report their present feeling and actions at short notice after receiving each of six signals distributed throughout the day; in contrast, the DRM approach asked the same respondents to complete a dairy each morning of the previous day. The paper demonstrated that the difference in reported happiness was largely explained by a different experience of daily life, and as such the paper demonstrated the potential value of using real-time data and phone applications in research on happiness.

Limitations of in-moment surveys

Currently, the key limitations of in-moment-surveys are coverage and non-response. Not all eligible sample members have smartphones or have smartphones that are sufficiently advanced to enable triggers to be applied. Furthermore, because triggered surveys usually use smartphone apps, non-response error will also arise (as discussed in section 4) because not all sample members will download the apps. Additional non-response error may also arise in surveys using participant-led triggers where participants fail to initiate data-collection, and in surveys using technology-led triggers where the technology fails to trigger all possible invitations to participate (for example where ongoing GPS location of participants fails always to be collected).

Future applications

As with other methods considered in this paper, the technological barriers to participants will reduce over time. We consider in-the-moment surveys to be particularly valuable in capturing accurate data in cases where activities or behaviours of interest are infrequent, thereby rendering both one-off surveys and diary approaches relatively unsuitable. In-the-moment surveys are also extremely useful where areas of geography are of particular interest but hard for participants to define - for example, when collecting data relating to 'your local area', or areas of poor air quality.

Health and Activity Monitoring

What is it?

Smartphones, smart watches and fitness/activity trackers contain advanced inbuilt technologies enabling them to capture several types of data. They commonly contain accelerometers to measure straight line motion, gyroscopes to measure rotational motion, Global Positioning System (GPS) sensors and cameras. These can be used to capture a range of biometric indices including step-count, calories-burned, heart rate (BPM), pulse volume, physical activity and sleep quality. It is also possible to combine machine learning with sensor data to identify patterns in motion sensor data which can be linked to specific behaviours.

The raw data collected by sensors must be converted into readable formats if they are to be used for research. This typically requires the development of bespoke research apps. Once developed, such apps can also be used to administer in-the-moment surveys triggered by the passively collected heath data as discussed above.

Why collect health and activity data?

The main advantage of passive health and activity data collection over traditional self-report methods is that they capture objective data and reduce recall error, and social desirability bias. For example, findings from the Health Survey for England's (HSE) trial of accelerometers for objective measurement of physical activity indicated that when self-reporting physical activity, 39% of men and 29% of women (aged 16 and over) met the Chief Medical Officer's (CMO) recommendation for physical activity among adults[3]. However, accelerometer data demonstrated that contrary to self-report findings only 6% of men and 4% of women were participating in enough physical activity to meet the CMO recommendations (BHFNC, 2010). Thus, a comparison of self-report data and objective accelerometer data showed significant discrepancies between the two measures, with adults significantly over-estimating the amount of activity they do.

An additional benefit to using smartphones, health and activity monitors is that they can be used to trigger health-relevant 'in the moment' surveys (Shiffman et al, 2008; Intille et al, 2016). Lastly, using these devices to measure health and activity data over self-report methods may reduce respondent burden and could eventually result in higher compliance and completion rates (Stone and Skinner, 2017). However, at present this benefit has yet to be realised (see section 4),

[3] At least 30 minutes of moderate intensity physical activity on at least five days of the week

Examples

Wearable activity watches used in the Millennium Cohort Study (MCS) and the 1970 British Cohort Study (BCS70)

Both MCS and BCS70 have used wearable devices to monitor participants' physical activity in sweeps of their longitudinal surveys. MCS participants were given a wrist-worn device which measured light, moderate and vigorous physical activity during a face-to-face interview, and were required to wear the device for two randomly selected 24-hour periods. In contrast, BCS70 used a thigh-worn activity monitor, fitted by a nurse and worn for 7 days, which could distinguish between different types of sedentary activity as well as activity levels. Both surveys had relatively high compliance rates. Of those who were eligible, 80% of MCS and 87% of BCS70 agreed to wear the respective devices and 72% and 92% respectively of those given devices returned them: overall CLS felt that the use of wearable devices had been a success, particularly with the adult cohort (Brown, Gilbert and Calderwood, n.d.).

Smartwatch System for Passive Detection of Cigarette Smoking

Traditional self-report methods of capturing individuals smoking behaviour in free-living conditions are subject to reporting biases and recall errors, therefore passive measurement of cigarette smoking offers significant opportunities (Skinner et al., 2018). Skinner et al. (2018) used the accelerometer and gyroscope in an android smartwatch (Android LG G-Watch) to capture specific hand movements associated with cigarette smoking. Although data from the *StopWatch* system was not as accurate as that from bespoke sensing equipment, it remained significantly more accurate than participant recall. In a 24-hour period in free-living conditions, the *StopWatch* the system achieved precision of 86% compared to the participant recall precision of 71% (Skinner et al., 2018). If the problems discussed in section 4 can be overcome, this technology will have considerable potential for use in high-quality surveys, offering a cheaper alternative to bespoke equipment and a more accurate method of measurement than self-report (Skinner et al., 2018).

Limitations and how to address them

Three limitations pertain to passive health and activity monitoring in high-quality surveys. One concerns an issue that is applicable to innovative data collection methods more generally, coverage and non-response, and is discussed in more depth in section 4. The other two are discussed briefly below.

Heterogeneity

Heterogeneity can be a problem for two reasons. First, in any survey that uses respondents' own devices there will likely be heterogeneity in the hardware, software and algorithmic function of respondent's devices. This means that data collection capability may vary across devices, or certain types of data collection may only be possible for a subset of respondents. This problem can be overcome by equipping respondents with the same device, but of course this can be a costly alternative. A second, less tractable problem, that is particularly relevant to longitudinal surveys derives from the pace of technological innovation. Although a wearable activity tracker may be an acceptable measurement device now, better, more advanced, devices will be developed in the future and this may reduce the comparability of data over time.

Cost and Complexity

As we have just discussed, supplying respondents with devices solves the problem of measurement heterogeneity within a survey implementation; however, it comes at a cost. For example, the devices used in the MCS and BCS70 studies cost £120 and £125 respectively (Brown, Gilbert and Calderwood, n.d.). Not only does the placement of devices by interviewers increase direct fieldwork costs, there are also substantial costs surrounding the development, piloting, distribution, collection of devices, the resolution of technical issues, and the complexity of data management. Substantial allowance should be made in survey budgets if such devices are to be used.

Future applications:

We believe that, if properly implemented and budgeted for, activity monitoring offers considerable potential for use in high-quality survey research. In addition to the approaches discussed above, a number of other applications hold promise for use in the near future. These include:

- The use of smartphones or smartwatches to collect 'in the moment' data about mood and mental state. This could aid a better understanding of mental health/ wellbeing (Stone and Skinner, 2017).

- The triangulation of GPS data with other sensor data, allowing researchers to capture more detailed environmental data by linking location to pollution, noise levels, distance from green space, etc. (Stone and Skinner, 2017)

We are aware of two developments evolving in the field of medical research which, whilst not currently ready for use in high quality surveys, hold longer term promise. First, the increasing popularity of in-home sensors (associated with the expansion of 'smart' devices and the 'internet of things') could provide significant opportunity for data capture (for example, measurement of sleep quality). Although the technology is already available within the context of clinical research context it is not necessarily available to consumers. If current trends in smart technology ownership continue, then the opportunity for the use of in-home sensing within survey research should increase, provided the limitations discussed above are adequately addressed.

Second, the development of 'lab-on-skin' wearable health monitoring offers an exciting alternative to commercially established wearable devices. Whilst wearable health monitors provide robust and reliable functionalities they are not optimised for continuous long-term monitoring (e.g. they need to be charged, use bulky devices, etc.). Scientific advances in the use of polymers for skin-integrated electronics could provide more accurate, non-invasive, long-term, and continuous health monitoring. The human skin is regarded as a signal source: both generating and transmitting biological signals that provide important health metrics. These metrics include: biopotential signals emanating from muscle (ECG, EEG), cardiovascular information from the hypodermis and dermis layers (body temperature, heart rate, blood pressure, oxygen level, and pulse wave velocity), biomarkers from the epidermis (pH of sweat which indicates glucose, water, lactic acid, and urea concentrations). These devices are energetically autonomous and capable of continuous remote sensing and communication (Liu, Pharr and Salvatore, 2017).

Non-Health Passive Measurement

What is it?

Passive measurement of health and activity has been discussed above, so in this section we will focus on non-health related passive measurement. Such measurement uses both apps (downloaded to smartphones) and other researcher placed devices (Volkova et al, 2016; Lessof and Sturgis 2018). Alongside the previously discussed smartphone sensors, many smartphones can also collect metrics on app use, app census, Geographic Positioning Systems (GPS), phone memory, 3G/4G usage, Wi-Fi connectivity, decibel level, ambient sound, IP address, operating system and retina display (and this list is not exhaustive). Purpose-built research apps can use one, or a combination of these metrics to monitor activity occurring on and/ or with the device (e.g. travel). It is also possible to carry out passive measurement using devices placed with and/or in participants' homes/ This placement usually occurs during face-to-face interviewer visits, although in some instances the devices may be posted and fixed by respondents. Examples include tablets to monitor household TV viewing behaviours, and fixed sensors to monitor household energy use.

Why do it?

Compared to self-report methods passive measurement is a more objective method of data collection and should generate more accurate data, free from recall error, self-presentational and cognitive biases. Furthermore, some passive measurement technologies, such as GPS and temperature sensors, collect continuous data and provide levels of detail that cannot be collected using questionnaires (Jäckle et al, 2018).

Not only does passive measurement enable greater detail and more accurate measurement, it is also able to capture data that has previously been unmeasurable, particularly when relying on traditional self-report surveys. While there are numerous areas that can be passively measured, we believe the following have special potential for high quality social research:

- Global Positioning System (GPS): travel and visiting specific locations
- Media exposure and usage: TV and radio
- Environmental measures: Air quality, household energy use and temperature
- Online behaviours: browsing behaviour and app use

Below we give a few examples to give a flavour of kinds of non-health passive measurements that might be used.

Examples

GPS Travel diaries

The UK National Travel Survey collects data on people's travel patterns by means of a 7-day travel diary. The diary collects data on trip timings, origin, destination, travel mode, distance, time, and related information. In 2011 a pilot was conducted to ascertain whether by the paper-based exercise could be replaced using accelerometer-enabled GPS monitors placed by interviewers. Supplementary data on important respondent destinations were collected in respondent interviews. Data were analysed using a programme which imputed values for origins, destination, trips and stages using data collected in combination with GIS and survey data. In contrast to the app-based time-use data collection discussed earlier, no active diary input was required from respondents. In comparison with parallel paper-diary data collection the GPS survey obtained a slightly lower response rate (52% vs. 59%) and delivered data showing fewer, but longer, journeys. Discrepancies were also apparent across journey purposes, and it was concluded at the time that data processing techniques were insufficiently advanced to enable the paper instrument to be replaced by passive collection of GPS and accelerometer data. We are aware, that similar development and pilot work is currently being undertaken by Statistics Netherlands (Lugtig, 2019) which may demonstrate that the method can now be used to collect more trustworthy data that it cold at the time of the NTS pilot.

Energy Follow Up Surveys (EFUS) conducted by the Buildings Research Establishment (BRE) with GfK [4]

In the 2007 and 2017 Energy Follow Up Surveys (EFUS), run by Buildings Research Establishment (BRE) with GfK for BEIS, a variety of sensors were installed to increase the quantity and quality of data collected about energy use and temperature in households (Hulme, 2015). The sensors measured energy inputs into the household (through attachments to energy meters) and energy outputs (measured through temperature loggers installed in participant properties). These were used alongside a range of other inputs to model and understand the energy efficiency of the property, the impacts of behaviours and household structure on energy consumption, and to feed into future policy on domestic energy efficiency (e.g. prioritising funding for improving the energy efficiency of housing stock to certain demographic groups, property types or improvements) (see: Report 11: Methodology, 2013).

Survey participants for the EFUS were sampled from participants from the English Housing Survey and were invited to complete a further face to face interview (response rate for the initial interview was 63%). At the end of the interview, some households were asked for permission to install the temperature loggers (which were installed by the face-to-face interviewers), and to have the energy meter monitors installed (by specialist contractors). In total 81% of approached households agreed to have the temperature loggers installed and 62% agreed to have the energy meter monitors installed.

Households had up to 3 temperature loggers installed, depending on their property structure: one in a hallway, a living room and a bedroom. The loggers stored temperature data internally with a memory capacity of 32,000 readings (c. 15 months), taking a reading every 20 minutes. Energy meter monitors collect energy use data every 10 seconds, sending data wirelessly to BRE for analysis. At the end of the monitoring period participants were sent instructions asking them to return the temperature loggers in a reply-paid envelope (77% of all loggers installed were returned).

Cross-media measurement for the BBC with Ipsos MORI

In order to capture media consumption and viewing/listening behaviour the BBC commissioned Ipsos MORI to build and manage a single-source panel for a cross-media measurement (CMM) system called Compass. Compass comprises a combination of audio metering and online/on-mobile metering. Audio metering captures media consumption for radio and TV through passive 'over-the-air' audio-matching technology using MediaCell+, a software developed by Ipsos MORI (Ford and Kulkarni, 2016). Online media consumption and behaviour is captured using online metering software RealityMine, which collects a range of information regarding online behaviour, such as time-spent browsing internet, websites visited and app use (Verspeek and Ford, 2018).

The nationally representative panel is made up of 2,900 weekly compliant respondents, with boosts for specific groups (16-34, BAME and Scotland, Wales and Northern Irish) (Verspeek and Ford, 2018). Recruitment for the panel was made online, with CATI boosts for hard to reach groups. During the 'onboarding' process potential panellists are screened for number of devices owned, demographics, what TV is owned and detail of service provision (e.g. Freeview, Sky, Now TV). Once respondents have completed the screener questions and have given consent they are invited to download the MediaCell+ app via a download link or email weblink (Verspeek and Ford, 2018).

After downloading, panellists must accept relevant permissions (access to location, ability to receive push notifications and use of microphone) and then set up the VPN to allow for the harvesting of websites and app usage data from device.

[4] GfK shared fieldwork with NatCen, however the data was not shared between organisations so the response and return rates are based on GfK fieldwork only.

Respondents need to register their device via the app and download additional metering software on other household devices used (e.g. desktop PC, laptop or tablet), with an average of 3 devices used (Verspeek and Ford, 2018). Once downloads are complete panellists are required to charge their mobile devices once per day and have it with them as much as possible to ensure complete data capture.

The combination of online and audio metering enables the tracking of minute by minute patterns of online, radio and TV audiences, the combined reach and overlaps in services (e.g. media consumption via a combination of devices). This provides an invaluable insight into media consumption at a granular level (Verspeek and Ford, 2018).

Limitations and how to address them:

The same three limitations pertain to the collection of passive non-health data in high-quality surveys as did to that of health and activity monitoring, namely coverage and non-response, heterogeneity and cost and complexity. The first of these is discussed in section 4, and the other two in the preceding section on collection of health relevant data.

The future of passive measurement: *expansion of geo-tracking capabilities*

The passive tracking of GPS is one of the main methods for passive data capture, allowing researchers to collect detailed location, travel and mobility information from people who are outdoors. The most promising advances in non-health passive measurement we found in this study involved the extension of such tracking to indoor environments. The commercial and market-research sectors are working in this by means of geo-fencing, Bluetooth, geo-locking (tracking exits and entrances), barometers and infrastructure mapping.

4. The common issue of representativeness: coverage and nonresponse

4 The common issue of representativeness: coverage and nonresponse

To minimise errors of representativeness, survey samples must exhibit both high coverage (eligible sample members must in-principle be approachable for data collection) and high response rates (approached sample members agree to provide requested information). Under-coverage and non-response lead to both variable and fixed errors. The former, usually measured in terms of sample variance, vary inversely with the size of the sample size providing data. Although in cross-sectional surveys increases in sample variance caused by under-coverage or non-response are easily addressed by increasing the size of the start-sample, this is not possible for longitudinal surveys after the first data collection wave because the start sample is limited to those previously approached for data collection. Fixed errors of representativeness vary across survey variables and arise whenever there are differences in characteristics between those from whom data are and are not collected (whether through under-coverage or non-response). The magnitude of such bias for a survey variable increases with (i) the correlation between that variable and the likelihood of participating and (ii) the non-participation rate. Such bias is hard to measure directly (because we do not usually have data for non-participating cases) and is therefore hard to address. Currently, the main approaches to bias minimisation are (i) prevention through maximising coverage and response rates and (ii) correction through weighting adjustment. More recently researchers have explored using targeted, adaptive or responsive approaches to bias prevention; these involve targeting fieldwork resource on the cases whose participation will most effectively reduce bias (Schouten et al, 2017; Tourangeau, et al 2017). In the remainder of this section we focus primarily on participation rates, both because these directly impact on variable error and because they are readily measurable indicators of the risk of fixed representativeness errors.

Generally, the innovative methods discussed in this report require respondents either to download an app for use on their own device (usually a smartphone) or to use a device provided by the study. Studies involving app downloads for respondents' use on their own devices are subject to non-coverage of the non-device owning population. Latest UK figures show that 79% of the adult population, and well over 90% of 16 to 54 year olds, use a smartphone. In contrast, only 55% of those aged 55+, six in ten of those earning less than £15.6K, and 66% of those in social group DE use one (Ofcom, 2019). There are therefore natural limits to the representativeness that is achievable for the general population, and especially for older and less well-off sample groups. It is also worth adding that coverage in surveys using respondents' own smartphones will probably be slightly lower than these figures suggest because survey apps used may not function on all smartphone operating systems or on older versions of common operating systems. However, this coverage reduction will not be substantial for apps developed for use with recent versions of IOS and Android operating systems.

Moving on to non-response error, we note that research has recently begun to examine participation rates and reasons for non-participation in studies using apps and researcher supplied devices (see Couper, Antoun and Mavletova, 2017 for a recent review). Levels of willingness to use apps / devices in representative surveys have been moderate at best, as the following recent examples indicate. Toepoel and Lugtig (2014) found that 26% of smartphone participants in a Dutch probability panel agreed to one-time capture of GPS coordinates. Rofique, Humphrey and Killpack (2011) obtained a higher, 52%, household level response rate (with all household members cooperating) using GPS monitoring devices in a UK random probability travel survey, but this involved face-to-face placement. Revilla et al (2019) found that a little over half an online volunteer sample in Spain (itself comprising the 24% of panel members answering the invitation questionnaire) said they would, in exchange for a moderate incentive, be willing to take photographs of products and scan barcodes with their smartphones, but fewer would be willing to share smartphone accelerometer data (37%), GPS data (21%) or to install an app to track URLs visited on smartphones tracked (18%). The study indicated that the main reasons for declining to download browser tracking apps related to privacy concerns. The Dutch LISS Panel in a series of studies asked panellists who were willing to use their smartphones for research if they would be willing to use smartphones or supplied devices for research. They found that of those completing the invitation questionnaire (around three quarters of those asked) a little over a third agreed to participation in app-based GPS and time-use data collection and a little over half agreed to being weighed and to wearing a separate accelerometer. Actual participation was lower: 25% for time use, 30% for GPS, 42% for being weighed and 51% for the accelerometer (Scherpenzeel, 2017). Jäckle and colleagues (2019) found that 13% of invited Understanding Society Innovation Panel respondents downloaded the receipt scanning software and used it at least once. They found that frequency of Internet and mobile device use, along with general cooperativeness with research and willingness to share personal information, were predictors of participation in the app study. McGeeney and Weisel (2015) found that of those who used an eligible smartphone in a random online panel, 61% of those invited installed an experience sampling app. Keusch and colleagues (2019) used vignettes on a German non-probability panel to manipulate six characteristics of a request for passive measurement. Averaging across vignettes 35% said they would probably agree to take part, but in contrast 39% said they would not download any app under any of the circumstances presented. The main reasons cited for non-participation concerned privacy and data security. Hypothetical consent rates were greater if the proposed study had a shorter duration, offered incentives, involved university sponsorship and allowed the respondent to switch the app off and on. Wenz, Jäckle, and Couper (2019) measured willingness to perform a variety of tasks using smartphones among smartphone users in the Understanding Society Innovation Panel in the U.K. They found that willingness varied by task (e.g., 65% to take pictures or scan barcodes, 61% for accelerometery capture, 39% for GPS capture, and 28% for a smartphone use tracking app). They found willingness was greater for those who used smartphones with greater intensity and those with lower expressed security concerns, and for tasks involving active respondent participation, tasks collecting less private data and tasks not requiring app downloads.

It appears that whilst moderately high response rates for some tasks may be achievable, at present a substantial number of people are unwilling to allow data-collection via downloaded apps or placed devices (although there is considerable variation across studies and data collection type). Without high participation levels, no new technology will prove very useful in high quality longitudinal surveys (unless good quality alternative collection methods can be used for app/device non-respondents) because of their need to maximise response rate to minimise variable non-response error. Furthermore, there is also evidence that participation is correlated with smartphone familiarity and privacy / security concerns indicating that app / device based data are likely to be subject to non-response bias (Wenz et al, 2019; Keusch et al, 2019, Jäckle et al, 2019; Revilla et al, 2010).

At present the best approach to maximising participation rates is likely to involve face-to-face placement, reassuring respondents about privacy / data security, offering respondents the ability to switch the app off and providing monetary incentivisation. However, we believe that the current cooperation rates are insufficiently high to enable app-based / device-based technologies to be used as a matter of course in high quality surveys. There will be exceptions, as the recent collection of time-use and accelerometer data in MCS demonstrates, but to identify these rigorous preliminary feasibility work will be essential.

This may well change over the next few years. An optimistic view would that as people's familiarity and engagement with technology evolves, their willingness to use apps/devices will increase. However, there may well be countervailing trends, such as increased concerns with privacy and data security.

Recent qualitative research indicates that longitudinal survey participants are currently sceptical of passive measurement, particularly when they perceive data collection to resemble surveillance with GPS tracking eliciting the strongest negative response (Ipsos MORI, 2019b). Although the notion of downloading an app was considered acceptable, there were more general concerns regarding the use of their personal phone as a data collection tool (Ipsos MORI, 2019b). One way to address this might be to provide them with a device to make a clear separation between their personal device and their personal data (Ipsos MORI, 2019b). Regardless, in any future app or device-based study it will clearly be essential to provide comprehensive information to support informed consent, and convincing assurances around what information is being collected and why.

in the light of the foregoing we believe it rash to make predictions about people's future willingness to participate in app-based and device-based data collection exercises.

5. Conclusions

5 Conclusions

In this report we have examined a range of innovative data collection methods that may be applicable to high-quality surveys. These have included both the administration of traditional survey questions on mobile devices and various forms of device- or app-based data collection.

Online data collection has now become an important form of data collection in high-quality surveys. For example, the great bulk of Understanding data are now collected using web-first data collection methods. If this development is to continue, then in the light of recent trends in device usage, extending the range of devices in which data can be collected will be essential. Fortunately, recent methodological research on smartphone questionnaire administration indicates that these developments should be relatively unproblematic in the near term.

Unfortunately, the same cannot be said of the various forms of app-based and bespoke-device-based data collection discussed here. Undoubtedly these forms of data collection can significantly reduce measurement error, but often this will come at the cost of representativeness error in both its fixed and variable forms. Those who are able and willing to allow the use of newer app-based and placed-device-based data collection methods differ in a range of ways from non-participants meaning that any data collected by these means will inevitable be subject to potentially significant non-coverage and non-response biases. Furthermore, because participation rates are relatively low, variable non-coverage and non-response errors are also likely to be significant in longitudinal surveys because the size of their start samples is fixed in advance. These difficulties can sometimes be surmounted, notably in recent CLS data collection using time-use diaries and accelerometers, and in the NTS GPS pilot, but we note that this has come at the cost of face-to-face device / app placement. The effective use of apps and placed devices in high quality surveys without interviewer placement has yet to be proven.

Of course, both technology itself and public engagement with technology are changing very rapidly and what is difficult today may prove easier tomorrow and we therefore believe it important for researchers to continue to actively investigate the use of innovative data collection methods in high quality surveys. However, for the reasons just given, in this research urgent attention should be paid to errors of representativeness and to finding effective methods for minimising these. Until such methods can be found our ability to use many innovative data collection methods in high quality surveys will remain severely hampered.

References

References

Bailey (2019). Leveraging tech to access in-the-moment and drive validity in insights. https://cls.ucl.ac.uk/wp-content/uploads/2017/02/Bailey-Pippa-IPSOS-MORI-Session-3.pdf

BHFNC (2010). *Health Survey for England 2008: Focus on physical activity and fitness*. BHFNC summary of the physical activity results. [online] British Heart Foundation National Centre. Available at: https://www.be-activeltd.co.uk/assets/BHFNC-Summary-Feb-2010.pdf [Accessed 29 Jul. 2019].

Brown, M., Gilbert, E. and Calderwood, L. (n.d.). *Using wearable technology to measure physical activity in BCS70 and MCS*. https://www.closer.ac.uk/wp-content/uploads/Using-wearable-technology-to-measure-physical-activity-in-BCS70-and-MCS-Matt-Brown.pdf. [Accessed 28 Jul. 2019].

Couper, M. P.(2019). *Designing and implementing mobile web surveys*. NCRM course. Southampton University. June

Couper, M.P., Antoun, C., & Mavletova, A. (2016), "Mobile Web Surveys: A Total Survey Error Perspective." In P. Biemer, S. Eckman, B. Edwards, E. de Leeuw, F. Kreuter, L. Lyberg, C. Tucker, and B. West (eds.), *Total Survey Error in Practice*. New York: Wiley, pp. 133-154.

De Bruijne, M., & Wijnant, A. (2014b), Improving Response Rates and Questionnaire Design for Mobile Web Surveys. *Public Opinion Quarterly*, 78 (4): 951-962.

Fernee, H., and N. Sonck. 2014. Measuring Smarter: Time Use Data Collected by Smartphones. *Electronic* International Journal of Time Use Research 11(1):94–96.

Ford, J. and Kulkarni, P. (2016). *MEDIACELL: On screen, online & on air*. Available at: https://www.ipsos.com/ipsos-mori/en-uk/mediacell

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

Groves, R. M. (2011). Three Eras of Survey Research. *Public Opinion Quarterly*, 75 (5), 861–871, https://doi.org/10.1093/poq/nfr057

Groves, R. M., Fowler F. J. Couper, M. P., Lepkowski, J. L., Singer, E. and Tourangeau, R. (2009). *Survey Methodology* (2nd ed.). Wiley

Gummer, T., & Roßmann, J. (2015), Explaining Interview Duration in Web Surveys: A Multilevel Approach. *Social Science Computer Review*, 33 (2): 217-234.

Hendriks, M., Kai L., and Veenhoven. R. (2016). Why are locals happier than internal migrants? The role of daily life. *Social Indicators Research* 125(2):481-508.

Henriksen, A., Haugen Mikalsen, M., Woldaregay, A., Muzny, M., Hartvigsen, G., Hopstock, L. and Grimsgaard, S. (2018). Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn

Wearables. *Journal of Medical Internet Research*, 20(3) pp.110.

Hulme, J. (2015). *The Energy Follow Up Survey (EFUS) Summary of key findings from the survey*.

Intille, S., Haynes, C., Maniar, D., Ponnada, A., Manjourides, J. (2016). µEMA: Microinteraction-based ecological momentary assessment (EMA) using a smartwatch. *Ubicomp '16: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, DOI http://dx.doi.org/10.1145/2971648.2971717

Ipsos MORI (2017). *Active lives survey: year 1 technical report.*

Ipsos MORI, (2019a). *Push to web best practice guide*. https://www.ipsos.com/ipsos-mori/en-uk/push-web-best-practice-guide

Ipsos MORI, (2019b). *CLS Cohorts Qualitative Research*. Unpublished.

Jäckle, A., Burton, J., Couper, M. P., Lessof, L. (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 Vol. 13, No. 1, pp. 23-44

Jäckle, A., Gaia, A. and Benzeval, M. (2018). *The use of new technologies to measure socioeconomic and environmental concepts in longitudinal studies*. London: CLOSER

Jäckle, A., Gaia, A., Lessof, C. and Couper, M. (2019). *A review of new technologies and data sources for measuring household finances: Implications for total survey error. Understanding Society*. [online] E.S.R.C. Available at: https://www.understandingsociety.ac.uk/sites/default/files/downloads/working-papers/2019-02.pdf [Accessed 28 Jul. 2019].

Josi Rofique, Alun Humphrey and Caroline Killpack (2011). *National Travel Survey 2011 GPS Pilot Field Report*. NatCen.

Keusch, F., Struminskaya, B., Antoun, C., Couper, M.P. and Kreuter, F. (2017), Willingness to Participate in Passive Mobile Data Collection. *Public Opinion Quarterly*, Volume 83, Issue Supplement_1, 2019, Pages 210–235

Lessof, C., and P. Sturgis (2018) New Kinds of Survey Measurement. in *The Palgrave Handbook of Survey Research*, edited by D. Vannette and J. Krosnick: Palgrave Macmillan.

Liu, Y., Pharr, M. and Salvatore, G. (2017). Lab-on-Skin: A Review of Flexible and Stretchable Electronics for Wearable Health Monitoring. ACS Nano, 11(10), pp.9614-9635.

Lugtig, P. (2019). *The TABI travel app: feasibility of data collection via a smartphone app*. The future of online surveys. Southampton. June.

McGeeney, K., & Weisel, R. (2015), App vs. Web for Surveys of Smartphone Users Experimenting with Mobile Apps for Signal-*contingent Experience Sampling Method Surveys*. Washington, DC: Pew Research Center report, http://www.pewresearch.org/2015/04/01/app-vs-web-for-surveys-of-smartphone-users/

Mavletova, A., & Couper, M.P. (2014), Mobile Web Survey Design: Scrolling versus Paging, SMS versus E-mail Invitations. *Journal of Survey Statistics and Methodology*,2 (4): 498-518.

Mavletova, A., & Couper, M.P. (2016), Device Use in Web Surveys: The Effect of Differential Incentives. *International Journal of Market Research*, 58 (4): 523-544

Moss and Ginnis (2014). Moss, N. and Ginnis, S. (2014). *The use of geo-trigger technology in research using smartphones*. http://the-sra.org.uk/wp-content/uploads/moss-ginnis.pdf

Nicolaas, G. (2019). *Online data collection in social surveys: back to basics.* The future of online surveys. Southampton. June.

Ofcom (2019). https://www.ofcom.org.uk/__data/assets/pdf_file/0026/143981/technology-tracker-2019-uk-data-tables.pdf

Report 11: Methodology. (2013). Energy Follow Up Survey 2011. [online] London: Department of Energy and Climate Change. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/274780/11_Methodology.pdf [Accessed 23 Jul. 2019].

Revilla, M., Couper, M.P., & Ochoa, C. (2019), Willingness of Online Panelists to Perform Additional Tasks. *Methods, Data, Analyses* Vol. 13(2), 2019, pp. 223-252

Scherpenzeel, A. (2017), Mixing Online Panel Data Collection with Innovative Methods. In S. Eifler & F. Faulbaum (Eds.), *Methodische Probleme von Mixed-Mode-Ansätzen in der Umfrageforschung*. Wiesbaden: Springer, pp. 27-49.

Schouten, B., Peytchev, A., Wager, J. (2017). *Adaptive survey design*. CRC Press

Shiffman S, Stone AA, Hufford M.R. (2008) Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4, pp.1-32

Skinner, A., Stone, C., Doughty, H. and Munafò, M. (2018). StopWatch: The Preliminary Evaluation of a Smartwatch-Based System for Passive Detection of Cigarette Smoking. *Nicotine & Tobacco Research*, 21(2), pp.257-261.

Stone, C.J., Skinner, A.L. (2017) New technology and novel methods for capturing health-related data in longitudinal and cohort studies. London: CLOSER.

Strain, T., Milton, K., Dall, P., Standage, M. and Mutrie, N. (2019). How are we measuring physical activity and sedentary behaviour in the four home nations of the UK? A narrative review of current surveillance measures and future directions. *British Journal of Sports Medicine*, DOI: 10.1136/bjsports-2018-100355

Strain, T., Wijndaele, K. and Brage, S. (2019). Physical Activity Surveillance Through Smartphone Apps and Wearable Trackers: Examining the UK Potential for Nationally Representative Sampling. *JMIR mHealth and uHealth*, 7(1), DOI: 10.2196/11898

Toepoel, V., & Lugtig, P. (2014), What Happens If You Offer a Mobile Option to Your Web Panel? Evidence from a Probability-Based Panel of Internet Users. *Social Science Computer Review*, 32 (4): 544-560.

Tourangeau, R., Brick, M., Lohr, S., and Li, J. (2017). Adaptive and responsive survey designs: a review and assessment. *Journal of the Royal Statistical Society* 180 (1) 203-223.

Verspeek, J, and Ford, J. (2018) *CMM (Cross-media measurement) - first fruits of single-source measurement for the BBC*. [video]. Athens: ASI International Television and Video conference.

Volkova, E., N. Li, E. Dunford, H. Eyles, M. Crino, J. Michie, and C.N. Mhurchu (2016) “Smart” Rcts: Development of a Smartphone App for Fully Automated Nutrition-Labelling Intervention Trials. *JMIR mHealth and uHealth*, 4(1)

Vrotsou, K., M. Bergqvist, M. Cooper, and K. Ellegard. (2014). PODD: A portable diary data collection system. pp. 381–82 in *2014 Working Conference on Advanced Visual Interfaces*

Wells, T., Bailey, J., & Link, M.W. (2014), Comparison of Smartphone and Online Computer Survey Administration. *Social Science Computer Review*, 32(2): 238-255

Wenz, A., Jäckle, A. and Couper, M. P. (2019). Willingness to use mobile technologies for data collection in a probability household panel. *Survey Research Methods*. Vol. 13, No. 1, pp. 1-22

Dr Patten Smith
Director of Research Methods
Patten.Smith@ipsos.com

Faith Jones
Graduate Research Executive
Faith.Jones@ipsos.com

For more information

3 Thomas More Square
London
E1W 1YW

t: +44 (0)20 3059 5000

www.ipsos-mori.com
http://twitter.com/IpsosMORI

About Ipsos MORI's Social Research Institute

The Social Research Institute works closely with national governments, local public services and the not-for-profit sector. Its c.200 research staff focus on public service and policy issues. Each has expertise in a particular part of the public sector, ensuring we have a detailed understanding of specific sectors and policy challenges. This, combined with our methods and communications expertise, helps ensure that our research makes a difference for decision makers and communities.