Resource report

closer
The home of longitudinal research

The use of new technologies to measure socio-economic and environmental concepts in longitudinal studies

Annette Jäckle (University of Essex)
Alessandra Gaia (City, University of London)
Michaela Benzeval (University of Essex)

January 2018

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
MRC Medical Research Council

To cite this report, please use the following reference:
Jäckle A, Gaia A, Benzeval M. The use of new technologies to measure socio-economic and environmental concepts in longitudinal studies. London, UK; 2018.

CLOSER
UCL Institute of Education
20 Bedford Way
London
WC1H 0AL
United Kingdom

Tel: +44 (0)20 7331 5102
Email: closer@ucl.ac.uk
Web: www.closer.ac.uk
Twitter: @CLOSER_UK
YouTube: CLOSER

Contents

1. Introduction

New technologies are increasingly being used in market research and more recently social research to improve the breadth, quality and ease of different kinds of data collection, as well as collecting new kinds of data. New technologies that can be used by researchers to collect data include smartphones, gadgets related to the 'quantified self' movement, or the internet of things. Researchers can also link to external technologies to draw in new kinds of data such as social media, smartmeters, storecards and barcoding. Finally, new technologies mean the way samples and data are processed can reduce the complexity of data collection, for example, using hair samples to obtain measures of cortisol or dried blot spots for metabolomics. New technologies can be used to respond to emerging research needs by measuring concepts that cannot be captured with survey questions, for example air quality, or by measuring key variables, such as household expenditure, more accurately or in a less burdensome way. In the future, the most effective way of collecting data will require a more flexible and heterogeneous approach with different topics requiring different data collection methods and frequencies to achieve the most accurate and consistent data.

These exciting opportunities also present significant challenges for both data collection and analytical methods. For example, the passive measurement of health and other behaviours enabled by new digital technologies offers the possibility of capturing data less susceptible to the biases usually associated with self report, but creates new sources of bias in terms of who might participate and how well they engage. In addition, the intensive measures of behaviours these technologies can provide (frequent sampling over extended periods, ambulatory measurement in the wild, sensors closely coupled to individuals) bring the promise of far richer phenotyping of studies, but very different kinds of data to those traditionally collected in surveys. Advances like these mean that new technologies will change the nature of the data that can be measured. Incorporating them into longitudinal studies creates additional challenges such as ensuring consistent measures over time despite changing technologies.

Drawing on evidence from across the CLOSER longitudinal studies, this report reviews how new technologies are being used to advance survey measurement of socio-economic concepts and features of the environment. The review focuses on practical considerations, implications for data quality and key methodological research needs. The contents of this report are based on a CLOSER workshop held in May 2017 (http://www.closer.ac.uk/event/new-technologies-measure-non-health/).

The workshop and report were funded by a CLOSER Innovation grant awarded to Michaela Benzeval and Annette Jäckle (University of Essex), and Kate Tilling and Dr Andy Skinner (University of Bristol) and is part of a series of three reports (see Jäckle, Gaia and Benzeval 2017; Stone and Skinner 2017).

2. Background

In the quest to provide new data for new science, social science studies are increasingly interested in new methods of data collection. To date information is mainly collected through survey questions, implemented as interviewer administered or self-administered questionnaires. In addition many social science studies collect biological samples, and perform physical or cognitive tests. Such ‘designed’ data collection is often supplemented by ‘organic’ data (Groves 2011) which are generated by the administrative processes of government or other agencies: survey respondents are asked for permission to link their survey responses to data about them held by other organisations. New technologies provide both new ways of implementing designed data collection, and new forms of organic data. Social media, for example, generates data that can create new research opportunities and be linked to survey data. The focus of this report is on the use of new technologies for designed data collection.

New technologies offer potential advantages over questionnaire based data collection in terms of the content and quality of data collected, the burden placed on respondents, and the cost of data collection. Which of these benefits can be realized however depends on the features of the technology. Key distinctions between different technologies are whether the data collected are objective or subjective measures and whether the respondent is actively involved in the data collection or whether data are collected passively.

Subjective measures collected with new technologies can, for example, take the form of questions or diaries administered through a smartphone application. In these cases the respondent is the source of the data, as with traditional survey questionnaires. The advantage over traditional questionnaires is that apps can be used to administer questions about events in real time, reducing recall periods and thereby improving measurement. Objective measures can take many forms. For example longitude, latitude and altitude data collected by GPS services, data about direction and speed of movement collected by accelerometry sensors, environmental data collected by sensor, or data coded from photographs or videos. The advantage of objective measures is that these do not rely on respondent memory and therefore often provide more accurate information. Some technologies, such as GPS, accelerometry or videos, can collect continuous data providing levels of detail that cannot be collected with questionnaires.

With active data collection the respondent has to provide the measures of interest. The measures can be subjective responses to questions administered through an app, or objective measures such as uploaded photographs. With active tasks the respondent often has to install the technology, for example by downloading an app or creating an account, and then also has to complete the measurements. In contrast with passive data collection the data are collected without the respondent’s involvement. They do however have to set up the technology, which can include many stages as described in some of the examples in Sections 4 and 6: they have to agree to the data collection, and depending on the technology, install a device, download an app or log on to a service, ensure batteries remain powered, keep the device activated throughout the measurement period, etc. The main advantages of passive measurement are that it provides objective measures and is less burdensome for the respondent than active measurement. New research however suggests that respondents would be more willing to do active tasks for a survey than passive tasks: although active measurement is more burdensome, respondents seem to prefer to have control over what information is captured about them (Wenz, Couper and Jäckle 2017). Similarly, they are more willing to participate in passive data collection if they can switch the data collection off for periods of time (Keusch et al. 2017).

Below we describe the use of new technologies focusing mainly, but not exclusively, on work by CLOSER longitudinal studies and fieldwork agencies in the UK. Section 3 reviews technologies that require active involvement of the respondent. Section 4 focuses on technologies that capture data passively. Section 5 reviews the use of video cameras, and Section 6 the use of GPS based technologies. For each of the studies included in this review, we provide a short description of the technology used, what the aims of introducing this technology were, and how and when it was implemented. Where possible we review key findings in terms of participation rates, biases in participation, measurement and costs. In Section 7 we conclude with a review of the key open issues regarding the use of new technologies for survey data collection.

3. Methods requiring active involvement of the respondent

Several studies have used technologies where the respondent has to actively engage in the measurement: for example, *Understanding Society* implemented an app to measure household expenditure over a month, the Millennium Cohort Study implemented a mixed mode time diary that included an app version, and the Labour Market and Social Security study (PASS) is planning an app study to facilitate in the moment data collection.

The *Understanding Society* spending study was implemented in collaboration with Kantar Worldpanel. Participants were asked to use the app to take pictures of their shopping receipts, to enter the amount and category of spending directly in the app, or to report that they had not spent any money that day (Jäckle et al. 2017). The aim was to test a method of capturing detailed objective data about monthly household spending, which would be coded from receipts, rather than relying on respondent recall. The expectation was that the app would produce more accurate data, and would be less burdensome for respondents, than completing a spending diary or answering retrospective questions about spending in the last month. In a study lacking questionnaire space for a detailed module on household spending, the app was seen as a complementary method to collect data that would augment the value of the questionnaire data. The app was implemented in the Innovation Panel sample in autumn 2016, following the completion of wave 9 fieldwork. All wave 9 respondents were invited to download the app and use it for a month. The invitation letter stated that they would earn the following rewards: £2 or £6 simply for downloading the app (experimental split); £0.50 per day that they used the app including reporting that they had not spent any money that day; £10 if they used the app every day for the month; and, £3 if they completed a short end of project questionnaire. The incentives were sent as Love2Shop vouchers at the end of the project. Participation in the spending app study was low, but stable over the month: 16.5% of respondents completed the registration survey (5 questions asked prior to downloading the app), and 12.8% used the app at least once. Surprisingly there was little drop-off over the month: after 28 days 82% of those who had used the app at least once were still using it. The strongest predictors of participation, controlling for different potential barriers such as access to mobile devices, were willingness to download an app for a survey, using their device every day, giving consent to data linkage and low item non-response rates in the prior interview. Analyses of the representativeness of participants shows biases in socio-demographic characteristics and financial behaviours, but not in correlates of expenditure. There were clear differences between socio-demographic groups in who participated and who did not: women, younger and more educated sample members were more likely to participate. The differences between age groups and education levels were entirely explained by differences in access to and use of mobile devices. There were also clear differences in financial behaviours between respondents who took part in the spending study and those who did not: people who use a computer document or spreadsheet to keep a budget were much more likely to participate in the app study than those who do not. Those who check their bank balance more frequently and those who check their balance online or using an app on a mobile device were also much more likely to participate. Surprisingly however there were no differences between participants and non-participants in correlates of expenditure: personal monthly income, how well the respondent feels they are getting by financially, household spending on groceries and food eaten outside the home, household spending on electricity, gas and oil, or whether the household is struggling or late paying housing costs or bills.

Time use diaries are another possible use of smartphone application for active data collection (for example, Fernee and Sonck 2014; Hendriks, Ludwigs and Veenhoven 2016; Vrotsou et al. 2014). The Millennium Cohort Study (age 14) implemented both a web-based diary and a smartphone app for respondents to record their main activities, location, whom they were with, and their level of enjoyment. The web diary could be used on netbooks, desktops and laptop computers; the app was optimised for tablets and smartphones and supported both by iOS and Android operating systems. Sample members who owned neither a personal computer nor a mobile device, or who refused to use the web or app versions, were offered a paper diary instead. The app diaries could be completed in real time and without an internet connection; although an internet connection was necessary to upload the data. In the pilot study 75% of participants completed the diary and most chose to do so using the app. For technical reasons there were several differences in the design of the diary between the three modes that could affect measurement:

- First, the online and paper diaries provided cohort members with a ‘time-grid’ approach with a visual display of the whole diary. In contrast the app diary used a question based format which asked about the different diary elements in turn.
- Second, in the online and paper diaries respondents had to choose 10-minute slots as start/end times for each activity, whereas in the app diary the respondent could assign specific start/end times to the exact minute.
- Third, in the online and paper versions respondents could enter multiple contextual elements for one activity, for example if the location changed in the course of an activity. In contrast in the app changes in the contextual features had to be reported as a sequence of multiple activities of the same type, with different contextual elements.
- Fourth, the online diaries included soft and hard edit checks to increase data quality at the data collection stage. The app had fewer checks than the online diary.

Examining the quality of the time use data collected with the three types of diaries (such as number and types of activities reported) in the pilot and dress rehearsal studies suggested that the web diary performed best (producing 71% of high quality diaries and only 7% of unusable diaries), followed by the app (with 65% high quality and 23% of unusable diaries), followed by paper (with 59% of high quality and 32% of unusable diaries). However, the share of diaries submitted with no information was higher in web diaries (21%) compared to app (12%) and paper diaries (only 9%). The design of these pre-tests was however not experimental so differences could be due to the characteristics of respondents who self-selected into the different modes.

Another example of data collection with Smartphone applications is in planning for the Panel Study "Labour Market and Social Security" (PASS): this is a longitudinal study of labour market outcomes and poverty in Germany. PASS sample members will be invited to participate in the associated study MoDeM (Mobile Device Measures), which will use a smartphone application to collect data for labour market research (Bähr et al. 2017). Data will be collected both passively through smartphone sensors and actively through pre-programmed in-app questionnaires. The latter are in-the-moment surveys triggered at certain times or by geographical locations when respondents enter a certain geo-fence. The aims of the study are to improve measurement quality (through passive measurement and asking about events closer in time), and to supplement the questionnaire data by measuring new concepts (for example formal or informal job search methods). Respondents will be invited to answer some survey questions that are pre-programmed in the app that are triggered through specific activities (such as visiting job centres). In addition, participants will be asked for permission to passively collect sensor data to measure position and movements, and to collect data about smartphone app use, phone book entries, signal strength, etc. To maximise participation sample members will be offered a 10 Euro Amazon voucher for installing the app. Each month a certain amount of up 10 Euro can be gained by granting the app permissions (e.g., access to the geolocation, phone book entries, or activity sensors) and participating in surveys is also incentivised with 10 Cents per question. Additional continuous feedback on the respondent’s network quality is provided as a non-monetary incentive. Only sample members who own an Android smartphone will be invited to the study. This is because the Apple operating system does not allow third parties to collect sensor data. This restriction may lead to coverage error, which will be assessed using data already collected at previous surveys waves. Other barriers to participation are the drainage of smartphone battery power and internet allowance for data transfer. Solutions for this are being sought (for example, use of Wi-Fi for data transfer). Finding a suitable app provider has proven a challenge: the choice was limited to German providers because data security regulations require data to be stored in Germany and there are few providers with expertise both in sensor data collection and in programming smartphone app surveys with embedded push messages.

Geo-triggered surveys, where invitations to a survey are triggered when sample members are in predefined geographic locations, have been used in social research, for example to study the respondents’ access to job centres (Bähr et al. 2017), and in commercial research, for example to evaluate consumers’ exposure to advertisement campaigns, participation in recreational events, or access to health services (see Clemens and Ginnis 2017). Ipsos-Mori have implemented geo-triggered surveys in commercial social research, for example in a study commissioned by Manchester City Council evaluating a campaign to reduce litter in the streets, a study commissioned by Highways England to evaluate user experience of driving through roadworks, and a study commissioned by a large petrol station brand to analyse the use of petrol stations by drivers. Evidence from these case studies shows that barriers to participation in geo-triggered surveys are high. There are many stages required of participants, each offering opportunities to drop out of the study: sample members have to open the survey invitation email, complete the initial survey, and provide a valid telephone number. They then have to have a device with sufficient storage space to download the app and with an operating system that is compatible with the app. They must download the app, log in to the survey, keep their GPS sensor activated, ensure the battery does not run out, and carry their device with them. The geo-trigger has to function, which can be problematic if respondents move through the geo-fenced area within the interval at which GPS position is measured by the app, if the geo-fence is located between high rise buildings, or indoors. Finally, the respondent has to complete the survey once it is triggered. Overall, in the three case studies considered, only a fraction of the samples ranging between 3% and 0.3% took part in the geo-triggered survey. For example, in the study on petrol stations, of the 11,000 sample members invited to take part, 8% (930) agreed to participate, 6% (690) downloaded the smartphone application, 4% (490) kept the application live for the entire duration of the study, 3% (376) entered a geo-fenced area and had the survey triggered, and 2% (241) participated in the geo-triggered survey (Clemens and Ginnis 2017). Ongoing research is tackling the technical problems with geo-fencing and, with time, technological change is likely to overcome current limitations. Aside from technical issues, the relationship with participants and whether the incentives are relevant are likely to be key drivers of participation. Looking forward, geo-triggered surveys are most promising where it is impractical to manage geographical sampling in any other way, for example to catch commuters travelling from one destination to another, or to survey many small sample points such as parks and open spaces.

4. Passive data collection

Within social surveys, new technologies have been used to collect data passively on different topics ranging from personal finances to air quality, from electricity use to surrounding noise, from human interactions to transportation use. In this section we describe methods that collect data passively using sensors or online platforms. In Section 5 we discuss passive methods of data collection that use videos to record human interactions, and in Section 6 we discuss GPS based technologies. Some of these kinds of technologies have been used to monitor health status (e.g. Berenguer et al. 2008; Kaye et al. 2011). Such use is discussed in the CLOSER report by Stone and Skinner (2017).

The Avon Longitudinal Study of Parents and Children — Generation 2 (ALSPAC-G2) trialled passive data collection of air quality (Zahra 2017). The study used a wearable air monitor that collects data on volatile organic compounds, carbon monoxide, temperature and humidity, and reports an overall air quality score. The aim was to supplement the survey with data on air quality, which cannot be obtained from survey questions, to enable analysis of child exposure to pollutants. A subset of 30 sample members (parents and their children aged 2 weeks to 6 years) were asked to wear the monitor for five days. Parents were asked to fill in an activity diary to report their own and their children's locations and activities. At the end of the study participants were followed-up with a telephone interview and asked to reflect on their experience with the device. Overall, the air monitor was accepted by respondents and the participation rate was high: 80% of invited parents participated. Most parents reported that the task was not particularly burdensome. Given the small size of the device and its light weight, it was easy to carry. However, some participants at times forgot to charge the device, take it with them, or give it to their children. Also, participants reported that they would have preferred to use a smartphone application instead of filling in a diary. There were technical glitches, for example a couple of devices broke, and data could not be downloaded without a smartphone application, and, thus, had to be emailed to the study centre or sent by Bluetooth, which led to data protection concerns. These types of problems are to be expected when working with cutting edge technologies and will hopefully improve as the technology matures.

The HomeSense Project in the UK has trialled data collection with a combination of fixed and personal sensors (Jiang et al. 2017). The fixed sensors measured temperature, humidity, particulate density, light, noise, and electricity use, as well as objects and bodies coming in and out of range. The personal sensors measured physical activity. These sensors send high frequency data over the internet to a central server for visualizations and analyses. The data collection was trialled on a small number of UK households occupied by couples. Data were collected for four days, and one household member was invited to fill in a time-use diary and data from the two sources were compared. Overall, the activities inferred from the data collected by sensors matched closely with those recorded in the time use diary. The only exception is dining, which could not be disentangled from entertainment in the sensor data, as these two activities often occur at the same time.

The Understanding America Study, a probability based online panel, has been trialling the use of a financial aggregator service to collect financial data passively (Angrisani, Kapteyn and Samek 2017). Financial aggregators are online platforms designed to help people manage their finances: the user provides the login details of all their bank accounts; the aggregator scrapes data from the user's financial institutions and provides categorized summaries of the inflows and outflows across all accounts. Understanding America sample members were asked to sign-up to the financial aggregator Yodlee. The aim of using a financial aggregator was to supplement the survey with detailed real-time information about income and spending, using a method that would be less burdensome for respondents and less affected by recall error and item non-response than detailed survey questions. Sample members were experimentally offered different conditional incentives for creating an account with Yodlee (either \$10, \$25 or \$50), for every financial institution they added to the account (\$2, \$3, \$5, \$10, or \$15), and then every month (either \$1 or \$2) for keeping their account updated. In terms of participation, 46% of sample members consented to taking part in the study, 32% created an account with Yodlee, and 12% entered details of at least one of their bank accounts. The experimental manipulations of the value of incentives did not affect participation. Younger and more educated sample members were more likely to participate, as were those who use online banking. Future research includes exploring strategies to mitigate barriers to participation, including phone calls to promote participation and visual aides to address security and privacy concerns and provide guidance on how to sign-up and link institutions. In terms of measurement, the aggregator data scraped from financial institutions contain objective measures of the dates and value of financial transactions which are unlikely to be affected by measurement errors.

The classification of transactions, i.e. whether money was spent on utilities, groceries, or transport, is however done by the aggregator service and subject to errors. The aggregator data also miss cash transactions that are not reflected in bank statements.

5. Video cameras

Video cameras can be used on their own as a measurement instrument or in combination with other technologies such as GPS (see Section 8). One example is the use of wearable cameras to measure human interactions, including early life interactions. Traditionally, human interactions have been collected through parents' self-reports or through observation by a third person – either in laboratory settings or in participants' homes (Lewcock 2017). Using head cameras instead offers several advantages: videos provide objective measures of behaviours that are not affected by misreporting (e.g. under-reporting of undesirable behaviours or due to lack of awareness of behaviours); using head cameras might reduce the probability that respondents modify their behaviour, as they tend to do in the presence of an observer and videos enable observing interactions over long time periods in participants' homes.

Internationally, head cameras were first used in a small scale Canadian study of infants' reaction to exposure to faces (see Sugden, Mohamed-Ali and Moulson 2014). This technology has also been trialled by the University of Bristol, to compare data quality on parent-child interactions using third person cameras and head cameras (Lee et al. 2017). A purposive sample of 15 mothers and infants (aged between 3 and 12 months) were asked to wear a broadband camera, and the scene was simultaneously recorded by an observer using a camera. Analyses of the data quality showed overall concordance between the two but, because of a wider field of vision, the third person camera picked up more body gestures, while the head camera was more effective in recording data on subtle vocalisation and gestures. Also, head cameras were more accurate than third party cameras at identifying less sensitive maternal behaviour (e.g. being distracted from the child). It is not clear whether this was due to the absence of the researcher, or to longer recording times by the head cameras.

This pilot study informed the implementation of head cameras in the Avon Longitudinal Study of Parents and Children – Generation 2. After the clinical interview, participants (mothers and children aged 6 months) were invited to use the camera at home for one week. Participants were given suggested activities to record (e.g. breakfast, lunch, play sessions, etc.) and asked to fill in a diary to document which activities they had recorded and who was present. The aim was to obtain five recorded sessions over two days (Lewcock 2017). In terms of respondent's cooperativeness, among the 59 study members invited to use the head cameras 51% agreed to take part, and 34% provided at least one recording. Among those that participated, on average 6 sessions were recorded. In terms of measurement, head cameras were considered a reliable method in recording mother-infant interactions. The technology also promises future improvements: although the quality of audio and video of head cameras is still sub-optimal, technological innovations are evolving rapidly, and manufactures are producing smaller and lighter devices, with wider fields of view, higher definition video, longer battery life and more capacity to store videos to other devices and networks (Lee et al. 2017). In terms of practical implementation, as the device was not very user friendly, providing clear instructions for participants on how to use the camera was both important and challenging. Another challenge was the transfer of data from the cameras to the study laboratory: fieldworkers collected cameras from participants' homes to avoid the need to send sensitive data by post or over the internet.

6. Geographic Positioning Systems (GPS) based technologies

Geographic Positioning Systems (GPS) technologies have been used for data collection in two ways: as well as the GPS location triggering an invitation to a survey (see Section 3 above) they can provide data that is useful in itself for social research, for example to determine activities or the mode of travelling.

An example of the latter is the Studying Physical Activity in Children's Environments across Scotland (SPACE) study (McCrorie 2017) where GPS technology is used to address research questions such as: Does traffic volume determine transportation mode to school? How long do children spend indoors and outdoors? Do children have access to greenspace, and if they do, for which kind of activities are greenspaces utilised? Sample members were loaned a GPS accelerometer and invited to wear it for 8 days, in 2015. The device recorded longitude, latitude, altitude, and timestamps. Two pilot studies were conducted to assess the acceptability of different devices, identifying which were most user friendly, and to trial the logistics of sending and retrieving the devices. The trials brought to light practical challenges, for example, the first device charger exceeded the size of a standard letterbox. The trials also uncovered that that storage space and battery life were insufficient to collect the amount of data initially planned. Consequently the study requirements were changed to collect GPS data every 10 seconds, instead of every 5, and to record waking hours only, instead of 24 hours. In addition respondents were asked to switch the device off and charge it overnight. Among the study population of 2,402 sample members, 2162 were contacted, 1096 agreed to take part; 78% (859) of those returned data but only 36% of all contacted respondents provided the requested five days' worth of data. Participants were more likely to be from high income and highly educated families, and belong to the least deprived areas according to the Scottish Index of Multiple Deprivation. No differences in response rates were found between rural and urban area, and for different levels of respondents' Body Mass Indexes. Large part of the costs of implementing this new technology was for purchasing the devices that were loaned to respondents.

The quality of GPS data has been compared experimentally with data from traditional travel diaries by NatCen Social Research in the context of the National Travel Survey. Sample members were asked to complete a paper travel diary for all journeys over seven days, except short walks (under a mile) which are reported only on day seven. In 2011, a random group of sample members (aged 12+) were instead invited to carry a GPS device with an embedded accelerometer (Department for Transport 2011; Feng, Moiseeva and Timmermans 2011; Rofique, Humphrey and Killpack 2011). There were several logistical challenges. The software of the devices was not compatible with interviewers' laptops, and thus it was not possible for interviewers to check whether data were being recorded correctly. The devices were not set correctly to collect all the required information such as speed. Finally, due to the short time scale of the pilot the number of devices the supplier had in stock was limited, such that devices had to be used by multiple respondents: once respondents had returned them for data download they had to be re-issued to the field within 48 hours. Respondent participation was slightly lower (52%) than participation in the paper diaries (59%). In addition, the number of daily journeys declined over the week in the GPS based measures, while it was constant in the paper diaries. This suggests that respondents were more likely to forget to use or charge the device as the week progressed. Measurement also presented some problems. The GPS devices did not collect some of the information that is collected in the paper diaries, such as the cost of travel or the type of ticket purchased. The GPS data sometimes had gaps, and was not very accurate in the presence of high buildings in urban areas (so-called urban canyons). Comparing key summary measures: in the GPS data the average number of journeys per person year was smaller (645 *vs*. 934), average journey distances were longer (24 miles *vs*. 6 miles) and average journey time was longer (51 min *vs*. 21 min) than in the paper diaries. Discrepancies were also evident in the mode of travelling and travel purposes, and in the GPS data the travel purpose could not be coded in 14% of journeys, a problem that did not occur in the paper diaries. The overall conclusion was that the quality of data collected with the GPS devices seemed inferior to the quality of data from the paper diaries. This new technology was therefore not taken forward.

While the studies presented so far in this section discuss the use of GPS and accelerometers as stand-alone technologies, these can also be implemented in combinations with other innovative data collection technique. For example, Kantar Public have adopted a mix of GPS technology and videos to analyse driving behaviour in England (Angle 2017). The study collected data on when drivers brake before bends, allowing the evaluation of a government campaign for safe driving on country roads. Passive data collection was considered crucial in this context as braking before bends seems to be an unconscious behaviour, which in most cases is not salient to sample members and therefore not reported accurately in interviews. The first study took place in 2014, when 30 young males aged 18-35 were invited to install a telematic device in their cars for one week before and one week after the campaign. The device collected data on driving behaviour, such as speed and breaking, and included a camera that recorded the road to collect contextual information on how driving is influenced by external factors, such as other vehicles and weather conditions. The device imposed little burden on the respondent, it turned on and off automatically when the respondent turned the car on and off, and did not require any particular maintenance. In a second study which investigated driving behaviour in 2015, a smartphone application was used instead of the telematics device. The key advantage of this was that it was cheaper and would therefore allow data collection on a larger scale: instead of purchasing devices and loaning them to participants, participants were asked to use their own devices. However data quality seemed lower compared to the telematics devices, with fewer short journeys being recorded. This might reflect the additional burden on the respondent with smartphone apps: they had to remember to take it with them on all journeys, and keep both the app and the GPS sensor activated, which would have significantly drained their battery. Consequently, a third study conducted in 2016 again used telematic devices. In the meantime, technology had improved and experience in analysing data in previous years had allowed the process to be more streamlined, resulting in lower costs. Compared to the smartphone application recruitment was easier and attrition was lower.

7. Key research needs

Given the interest associated with innovations, it is important to keep the ultimate goal in mind. Is the aim to collect new content that cannot be reliably collected with survey questionnaires? Or is it to collect more accurate or more detailed data? Is the full granularity that can be collected with some technologies actually needed to address the research questions? What concepts for what populations can be effectively measured with new technologies? These are determined both by the measurement properties of the technology – and by how sample members use the technology to provide information about themselves.

There are several particular challenges in developing innovations for data collection methods in longitudinal surveys. Longitudinal surveys are concerned with collecting data about a set of concepts in a comparable way over long periods of time. The pace of technological change is however fast, and technologies that might be implemented today are quickly outdated. This has implications not only for the cost of investing in and maintaining equipment, but also for the comparability of measures over time. Technical glitches are unavoidable when working with emerging technologies. Planning and budgeting for studies to test new technologies is difficult: to date there is only modest information in the literature about participation rates at different stages of the data collection on which to base cost assumptions. Some technologies produce large volumes of data that require new methods and skills for handling, storage and analysis. Finally, many studies to date have been on small samples, leaving open questions as to how to scale data collection up to large sample sizes.

Adopting new technologies adds a further dimension to survey design, akin to the move to mixed mode data collection. Understanding the implications of different measurement technologies for who can be measured (representativeness of the sample), and how well concepts are measured (match between the measurement and the concept of interest), are key to enabling effective survey design. What are the trade-offs between coverage and non-response error on the one side and reduction of measurement error on the other? Understanding such trade-offs requires an examination of the different stages of data collection, to understand what can go wrong, and what the implications are for variance and bias of estimates. For traditional questionnaire-based surveys the Total Survey Error framework (Groves et al. 2009) provides guiding principles along which to examine potential sources of error and their trade-offs. For new technologies this framework needs to be adapted, to understand potential sources of error and implications for how best to implement new technologies.

Understanding how respondents engage with and use new technologies is key. This includes considering the barriers to using new technologies; what motivates participation; how engagement with a task can be sustained over periods of time; what preferences respondents have that are relevant to the data collection task; and understanding biases in who participates and who does not. As with any survey, full participation by all sample members is not achievable. Indeed participation rates in studies using new technologies are still low. However if we understand the barriers to participation we can design features of the data collection in such a way that barriers are reduced. This includes features affecting the ergonomics of a task, as well as incentive schemes. In addition we can use survey questionnaires administered to the full sample to collect information about the drivers of participation. These data can be used to derive weights to adjust for differences between participants and non-participants.

New ethical considerations arise with the use of new technologies. For example how to ensure that respondents give informed consent to the passive collection of data and linkage to their survey responses, and about the storage and sharing of data, including incidental data, particularly when the volume and type of data may increase the potential for disclosure.

The final question is whether it is worth investing in data collection using new technologies. What is the value of such data in producing new research compared to other data collection approaches? What are the variable and fixed costs of collecting data using new technologies? How to extract and analyse video or photo data in an automated way? What are the cost-quality trade-offs between purchasing devices to loan to participants, versus using the respondents' own devices? If devices are purchased, the measurement device is standardised across all participants. If everyone uses their own device survey researchers have no control over the specification of the device. What is the value of using existing devices that have been developed for personal use, and therefore might give participants feedback that could influence their behaviours, or not have easy ways of extracting data, versus developing purpose built devices for research?

Developing best practice in the use of new technologies for data collection requires understanding the general principles of what respondents are willing to do, which are transferrable as technologies evolve. Such understanding is enhanced by empirical evidence derived from incorporating experimental comparisons of different design features into data collection protocols. This report highlights some of the challenges from early implementation of a range of technologies in different studies. Further research is now building on these to identify ways of overcoming the barriers identified. Much work remains to be done before new technologies can be implemented on large scale surveys effectively; sharing lessons across studies to building on each other’s experiments to develop this field is key. We are at the very early stages of technological developments. We are often testing the first generation of new technologies, but measurement and reliability are continually improving. Adoption, use, comfort, acceptance, etc., of these technologies is increasing in the general population. Our research knowledge is expanding. It’s possible that many of these barriers could be overcome as we move forward to develop, test and evaluate new tools and technologies.

References

Angle, Helen. 2017. "When self-report is not enough Measuring unconscious behaviour change in response to a campaign designed to change where people brake at bends on country roads." Presented at the CLOSER Workshop: New technologies to measure non-health topics in longitudinal studies, London.

Angrisani, Marco, Arie Kapteyn, and Swaroop Samek. 2017. "Alternatives to Periodic Data Collection." Presented at the Workshop on Developing a Methodological Research Program for Longitudinal Studies, Washington DC.

Bähr, Sebastian, Georg Haas, Florian Keusch, Frauke Kreuter, and Mark Trappmann. 2017. "Enriching an Ongoing Panel Survey with Mobile Phone Measures: The MoDeM Study." Presented at the CLOSER Workshop on New Technologies for Measuring Non-health Topics in Longitudinal Studies, London.

Berenguer, M, M Giordani, F Giraud-By, and N Noury. 2008. "Automatic Detection of Activities of Daily Living from Detecting and Classifying Electrical Events on the Residential Power Line." Pp. 29-32 in *10th International Conference on E-health Networking, Applications and Services, 2008. HealthCom 2008.* : IEEE.

Clemens, Sam, and Steven Ginnis. 2017. "Mobile-based geo-triggered surveys: experiences from the field." Presented at the CLOSER Workshop: New technologies to measure non-health topics in longitudinal studies, London.

Department for Transport. 2011. "National Travel Survey 2011 GPS pilot: summary analysis". London: Department for Transport.

Feng, Tao, Anastasia Moiseeva, and Harry Timmermans. 2011. "Processing of National Travel Survey GPS Pilot Data, a Technical Report Prepared on behalf of the Department for Transport". Eindhoven: Eindhoven University of Technology.

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.

Groves, Robert M, Floyd J Fowler Jr, Mick P Couper, James M Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. *Survey Methodology*: John Wiley & Sons.

Groves, Robert M. 2011. "Three eras of survey research." *Public Opinion Quarterly* 75(5):861–71.

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

Jäckle, Annette, Jonathan Burton, Mick P. Couper, and Carli Lessof. 2017. "Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household pa4nel: response rates and response biases". *Understanding Society Working Paper 2017-09*. Colchester: University of Essex.

Jäckle, Annette, Alessandra Gaia, and Michaela Benzeval. 2017. "Mixing Modes and Measurement Methods in Longitudinal Studies". *CLOSER Resource Report*. London: CLOSER.

Jiang, J, R Pozza, K Gunnarsdóttir, N Gilbert, and K Moessner. 2017. "Recognising Activities at Home: Digital and Human Sensors." Presented at the ICFNDS ’17, Cambridge, UK.

Kaye, Jeffrey A., Shoshana A. Maxwell, Nora Mattek, Tamara L. Hayes, Hiroko Dodge, Misha Pavel, Holly B. Jimison, Katherine Wild, Linda Boise, and Tracy A. Zitzelberger. 2011. "Intelligent Systems for Assessing Aging Changes: Home-Based, Unobtrusive, and Continuous Assessment of Aging." *The Journals of Gerontology: Series B* 66B(suppl_1):i180-i90.

Keusch, Florian, Christopher Antoun, Mick P. Couper, Frauke Kreuter, and Bella Struminskaya. 2017. "Willingness to Participate in Passive Mobile Data Collection." Presented at the 72nd conference of the American Association for Public Opinion Research, New Orleans, LA.

Lee, R., A. Skinner, M. H. Bornstein, A. N. Radford, A. Campbell, K. Graham, and R. M. Pearson. 2017. "Through Babies’ Eyes: Practical and Theoretical Considerations of Using Wearable Technology to Measure Parent–infant Behaviour from the Mothers’ and Infants’ View Points." *Infant Behavior & Development* 47:62-71.

Lewcock, Melanie. 2017. "Through babies' eyes: practical and theoretical considerations of using wearable technology to measure parent-infant behaviour from the mothers' and infants' view points." Presented at the CLOSER Workshop: New technologies to measure non-health topics in longitudinal studies, London.

McCrorie, Paul. 2017. "Integrating GPS technology into large scale, population level, data collections: practical utility for science, and concerns and considerations regarding its application in 10-11 year old children." Presented at the CLOSER Workshop: New technologies to measure non-health topics in longitudinal studies, London.

Rofique, Josi, Alun Humphrey, and Caroline Killpack. 2011. "National Travel Survey 2011 GPS Pilot, a Technical Report on the Pilot Survey Management and Data Collection". London: NatCen.

Stone, Chris, and Andy Skinner. 2017. "New Technology and Novel Methods for Capturing Health-Related Data in Longitudinal and Cohort Studies". *CLOSER Resource Report*. London: CLOSER.

Sugden, Nicole A., Marwan I. Mohamed-Ali, and Margaret C. Moulson. 2014. "I Spy with my Little Eye: Typical, Daily Exposure to Faces Documented from a First-Person Infant Perspective." *Developmental Psychobiology* 56(2):249-61.

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*.

Wenz, Alexander, Mick P. Couper, and Annette Jäckle. 2017. "Willingness to Use Mobile Technologies for Data Collection in a Probability Household Panel". *Understanding Society Working Paper 2017-10*. Colchester: University of Essex.

Zahra, Jez. 2017. "Incorporating Wearable Air Monitors Into a Research Study Of Young Children: Reflections From Parents And Researchers." Presented at the CLOSER Workshop: New technologies to measure non-health topics in longitudinal studies, London.