

# ESRC Longitudinal Studies Review 2017

## Further analysis of responses to the consultation

Paper 8:

New forms of data collection

April 2017

Report author:

Ruth Townsley, Independent Researcher

The views represented in this report are from those who responded to the consultation and do not represent the views of ESRC

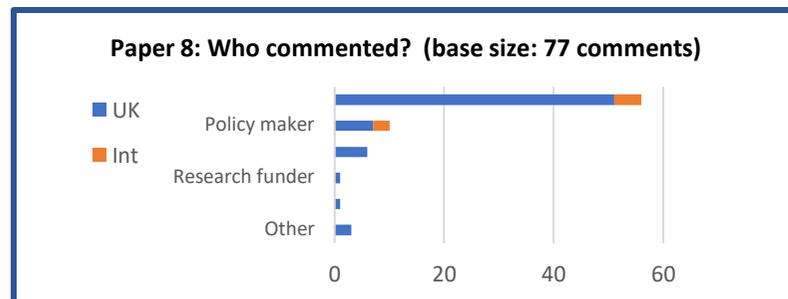


## Introduction

The ESRC Longitudinal Studies Review 2017 is exploring the current and future scientific and policy-relevant need for longitudinal research resources. The review started in October 2016 and will report to ESRC Council early in 2018. An open online consultation in autumn 2016 sought input broadly, resulting in 637 completed responses from UK (83.4%) and international (16.6%) respondents. Respondents were predominantly from the academic sector (80%) as well as government, civil society and business sectors (20%). The main findings of the consultation were published in December 2016 in an [initial report](#). This report is supplemented by short briefing papers that examine key themes from the consultation data in more detail.

## Paper 8: New forms of data collection

Question 9 of the consultation survey asked respondents to suggest up to three methodological or technological priorities for the future development of longitudinal resources. A total of 661 open-ended answers were received, of which 77 related specifically to new forms of data collection. Within their answers, respondents highlighted a range of priority issues, including online and digital data capture; collection of biomedical data; and other innovative methods for data collection.



"An effective social media presence for contacting, interacting with, and disseminating findings to, participants has to be at the centre of a 21st century strategy for reducing losses from attrition. People live much of their lives online these days and the credibility of a longitudinal study is dependent on a strong online presence." (ID 424)

## Online and digital data capture

Over two-thirds (69%) of the 77 comments about new forms of data collection focussed on online and digital data capture. Many respondents expressed awareness of the considerable technological advances that have taken place since the 2006 Longitudinal Studies review, particularly in terms of smartphone ownership and the use of social media. Some longitudinal resources have developed web-based data collection and online survey tools, alongside smartphone apps and wearables to collect real time and continuous data on social, emotional and health measures. Respondents stated the cultural imperative of using online and digital means to enhance data collection, whilst urging for careful consideration of the challenges. Their comments reflected the following underlying issues:

- > Given widespread advances in online and digital technology, alongside changes in public use and attitudes, there is a methodological and cultural imperative for longitudinal resources to embrace innovative forms of data capture alongside traditional survey methods – people expect to be offered choice as to when and how they participate
- > There was an awareness that some longitudinal resources and investments have already embraced these changes – respondents mentioned current review work by CLOSER investigating the use of new technologies for measuring health and non-health behaviours – others referred to the use of online/smartphone apps and accelerometers by MCS and BCS70
- > There was a sense that a shift to online and digital data collection is inevitable, but that this should enhance, not completely replace face-to-face survey work – respondents acknowledged the infancy of these methodologies and stressed a need for investment, testing and evaluation before they can be rolled out for longer term use
- > Longitudinal studies need to adapt to the online and digital world and establish an effective social media presence for contacting, interacting with and disseminating findings to participants – however awareness of the demographics of low social media use is needed

In their comments, respondents mentioned a number of potential advantages of online and digital data capture:

- > Costs are generally cheaper than face-to-face data collection
- > It can be less intrusive and less burdensome for participants and it may be helpful in collecting data from some hard-to-reach groups - this may improve attrition, but awareness of technology use amongst different demographic groups is important
- > Connecting to participants via social media may help to improve credibility and attrition, particularly for younger people and/or those who are geographically mobile – one person suggested that digital identities (e.g. Facebook and Twitter ids) have more longevity than postal addresses for some demographic groups
- > There is also the potential to collect additional data between and within waves, and for more regular data collection as costs are lower
- > Some data collected digitally (especially continuous and real time data) may be more complete, more objective and more accurate
- > There is the potential to collect new forms of data and to add new dimensions to data collected via traditional survey methods – this is particularly the case for smartphone apps which can deliver supplementary measures or collect data on real time experiences
- > Some respondents gave examples of the new data that digital and online methods could capture: measures of information exposure; passive monitoring of device activity and online/digital behaviours; passive monitoring and objectives measurement of biomarkers (e.g. pulse rates) and other health measures (e.g. physical activity, sleep, sitting time); monitoring time use; real world data on mood, emotions, regulation, coping and life events, tracking people soon after life events such as a fall, or hospital episode; delivering ‘in the moment’ surveys.

Despite the apparent advantages, many respondents expressed a need for further exploration and evaluation of the outcomes and impact of online, digital and mixed mode data collection, versus postal surveys and/or face-to-face structured interviews. A number of significant and current challenges for online and digital methodologies were mentioned, all of which respondents felt still need further exploration:

- > Mode effects and response bias – the impact of mixed modes on who responds and how
- > Data quality and comparability - completeness and veracity in the absence of an interviewer
- > Consent for data linkage – there were concerns that this is lower when collected online rather than face-to-face
- > The expense and technical challenge of developing online and digital applications for use by social scientists – the market is dominated by consumer devices which do not always translate well into data collection tools.

“Making greater use of technology to engage young people. We need to move with the times and make the experience of taking part in research easy and engaging.” (ID 564)

“Portable health devices or wearables, as well as Smartphones, can be used to track physical activity and to measure other bio-markers such as pulse rate. However, they are not without their challenges including cost, replicability and consistency of measurement, and may be prone to relatively low take-up rates.” (ID 688)

“Using the web in longitudinal surveys has the potential to reduce costs and improve the accuracy of survey measurement. However, it is not without its challenges: it may have mode effects on measurement, and the longer-term effects on retention remain poorly understood. The challenge for the social science and survey research community is how to make most effective use of the web in longitudinal surveys, whilst maintaining data quality, longitudinal and cross-sectional consistency, and longer-term participation in studies.” (ID 394)

“High cost pressure to move to online data collection. However, very little is known about potential mode effects or approaches to deal with low response rates. Better understanding of how online methods could enhance other data collection modes rather than only substituting for them are required.” (ID 329)

## Collection of biomedical data

Prioritising the collection of biomedical data for current and future longitudinal studies was highlighted by 18% of the 77 comments relating to new forms of data collection. This includes direct measures of physical and mental health and collection of biomarker and genetic data. The key issues mentioned by respondents included:

- > The collection of biomedical data and tracking of biomarkers over time is central to biosocial research and can supplement and enhance interview data. In their answers to question 8 of the survey, respondents had already highlighted a need for increased longitudinal biosocial data about social and environmental influences on health and for improved understanding of interactions between biology and social/environmental experiences
- > The collection of biomedical data can be invasive and non-invasive and includes collection of venous blood samples; saliva samples; measures of grip strength; pulse rate; blood pressure; laterality; hair samples; measures of physical activity, sleep and movement via body-worn sensors. There is also a need to explore new forms of biomarkers
- > Regular, repeat biomedical and biosocial measures over time are needed to measure changes and development
- > How can biomedical data be collected more cost effectively? - some respondents suggested self-collection by participants, or collection by trained interviewers – others advocated for maintaining/extending the use of nurses and clinics
- > Is there scope for genotyping BCS70 and MCS? A few respondents made a case for linking of whole genome data with social science longitudinal data in order to help understand genetic bases of disease and how genetic effects vary over the life-course
- > How can omics platforms (metabolomics and proteomics) be better integrated to optimise the use of biomedical data?

“Capacity for interviewers (rather than nurses) to collect biometric measures creates scope for cohort studies to collect new medical data that could be extremely useful in looking at how/why particular illnesses/health issues develop over the life course.” (ID 292)

“More use to be made of biomarker and objective health (for both physical and mental health) data - perhaps needs more prior planning to weave it more effectively together with interview data.” (ID 48)

“Continual advances in technology for the collection and analysis of biomedical samples and data - and the important insight it offers - mean this is an important area that could be further exploited in future years. Longitudinal surveys are important vehicles for collection of these data, including venous blood samples (amongst many other samples). Further work to use the latest technology but also explore new bio-markers and methods (such as clinics) are important methodological issues.” (ID 17)

## Other new forms of data collection

A few comments related to the need to prioritise and explore other new forms of data collection, some of which were covered in Briefing Paper 7 on data linkage:

- > Facial coding to stimuli
- > Automated collection of financial information
- > Methods for collecting dietary data
- > Geographical information systems and remote sensing  
Intergenerational methodologies to link participants’ data with the exposures of their ancestors
- > Ensuring that data collection can evolve and adapt to the needs of ageing cohorts who might have different levels of engagement, health management and access as they age.

“It would be helpful to obtain as much identifying information as possible for members of cohorts concerning their ancestors so that linkage with censuses, births, marriages and deaths can be used to trace the influences of various exposures (psychological or physical) of their ancestors on subsequent generations.” (ID 591)