



City Research Online

City, University of London Institutional Repository

Citation: Hanson, T. & Matthews, P. (2017). Adapting survey design for smartphones: lessons from usability testing and survey implementation. *Social Research Practice*(3 (Win), pp. 37-44.

This is the published version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/25209/>

Link to published version:

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

Social
Research
Association

Issue 3
Winter 2016/17

Social Research Practice

The SRA journal for methods in applied social research



Social Research Practice

Issue 3 Winter 2016/17

The Social Research Association journal for methods in applied social research

Contents

- 01 Editorial**
Richard Bartholomew
-
- 03 Removing the 'gag': involving people with dementia in research as advisers and participants**
Jenni Brooks, Department of Psychology, Sociology and Politics, Sheffield Hallam University; Nada Savitch, Innovations in Dementia; Kate Gridley, Social Policy Research Unit, University of York
-
- 15 Gypsy and Traveller household formation rates**
Nigel Moore, Opinion Research Services
-
- 23 An introduction to address-based online surveying**
Joel Williams, Kantar Public
-
- 37 Adapting survey design for smartphones: lessons from usability testing and survey implementation**
Tim Hanson and Peter Matthews, Kantar Public
-

Editorial

Richard Bartholomew

Editor

*Welcome to the third issue of **Social Research Practice**, the SRA's journal for methods in applied social research.*

With around one in 14 of those aged 65 and over suffering from dementia, and the number of sufferers of all ages projected to increase to around 1.14 million by 2025¹, the question of how to conduct good, ethically acceptable social research with people with dementia could not be more pertinent. How should researchers respect the right of those with dementia to be involved as active participants in research enquiries, giving them as much opportunity as possible to express their own views and experiences: views which may differ from those of their carers? In their fascinating article 'Removing the 'gag': involving people with dementia in research as advisers and participants', Jenni Brooks and colleagues explore practical and sensitive ways of engaging people with dementia in research and enabling them to communicate their experiences whilst also protecting them from exploitation. This involves probing the practical meanings of informed consent in the context of the requirements of the Mental Capacity Act (2005).

Few local planning issues are guaranteed to provoke as much heat, anxiety and prejudice as the siting of permanent sites (or 'pitches') for the Gypsy and Traveller community. It is an issue which brings out the NIMBY tendencies of even the most liberal and tolerant local residents, but also causes immense frustration and feelings of discrimination amongst Gypsies and Travellers. Local authorities are often caught in the middle as they are required by law to ensure they have identified a sufficient number of Gypsy and Traveller pitches for the next five years to meet local needs. The accuracy of their forecasts of the future level of demand for sites, therefore, becomes critical. In his article on Gypsy and Traveller household formation rates, Nigel Moore discusses how, through a careful analysis of Census of Population data and Opinion Research Services survey data, it is possible to provide local planning authorities with a more realistic estimate of the likely demand for new pitches in their areas. Whilst this will never completely defuse the sensitivity of the issue, it does mean that plans (and the debate about those plans) can now be more grounded. This is a good example of the practical impact that sound research and analysis can have on issues which directly affect the quality of people's lives.

Online panel data-collection has now become the dominant form of survey data-collection in much quantitative market research. Recent events have also led to a sharp focus on the relative accuracy of online panel data in political opinion polling compared with more traditional interview-based surveys using random sampling methods. In the public sector, research commissioners have been much more cautious about relying on online panel survey data because of valid doubts about the coverage of key sub-populations (the elderly, the poor), the use of non-random sampling methods and representativeness. But online methods have the attraction of much lower costs and much greater speed: two features of great interest to public sector organisations. As Joel Williams sets out in his article 'An introduction to ABOS – address-based online surveying', there is, therefore, keen interest in new approaches which can combine random sampling with the convenience and lower costs of online surveying. Joel describes and assesses a possible approach to meeting this need, testing it against seven key criteria including response rates, cost, coverage and non-response bias.

On a related theme, Tim Hanson explores what we know so far about the advantages and possible drawbacks of using smartphones for completing online surveys ('Towards device agnostic survey design: lessons from usability testing, experiments and survey implementation'). With 71% of UK adults now

¹ Dementia UK, 2nd edition. Alzheimer's Society, 2014.

owning a smartphone (rising to nine out of ten 16-34-year-olds) they may well become (or perhaps have already become) the preferred method for completing online surveys. At a minimum, respondents and especially younger respondents will naturally expect to be able to complete all online questionnaires in a smartphone format. But what do we know about the implications of smartphone use for question design and presentation, questionnaire length, the quality of answers provided and drop-out rates? Tim explores these and other questions, and sets out guiding principles for ensuring questionnaire and survey designs are suitable for the age of the smartphone.

We are keen to have a regular flow of new articles, and will be publishing our fourth issue in June 2017. If you are interested in offering an article for a future edition of Social Research Practice, see <http://the-sra.org.uk/journal-social-research-practice>. Read the guidelines for authors and download the template for articles. If you have an idea for an article but are not sure if it is suitable, just drop me a line: rabartholomew@btinternet.com

Removing the 'gag': involving people with dementia in research as advisers and participants

Jenni Brooks, Department of Psychology, Sociology and Politics, Sheffield Hallam University; Nada Savitch, Innovations in Dementia; Kate Gridley, Social Policy Research Unit, University of York

Abstract

People with dementia are often excluded from taking part in research because of perceived difficulties in consent, capacity and communication. We argue that involving people with dementia in research is important, and describe how we involved people with dementia as both advisers and participants in research about the use of life story work.

Researchers worked in partnership with Innovations in Dementia, who supported a network of advisers with dementia. Focus groups were arranged to ensure meaningful contributions by people with dementia. It was difficult to use standardised quality-of-life measures, and we describe the challenges faced with capacity and consent, recruitment and selection, and data collection.

We suggest there is a need for (a) new tools for measuring quality of life of people with dementia which do not require participants to respond in prescribed ways, and (b) ethics and consent processes which are appropriate for non-medical research and which facilitate the involvement of people with dementia.

Funding acknowledgement

This paper presents findings from research funded by the National Institute for Health Research (NIHR) Health Services and Delivery Research (HS&DR) Programme (project number 11/2000/11). The views and opinions expressed in this paper are those of the research team and do not necessarily reflect those of the HS&DR Programme, NIHR, NHS or the Department of Health.

Introduction

In the past, people with dementia have often been excluded from shaping and taking part in research, either because they were considered to be incapable of verbally communicating their thoughts and feelings, or because those thoughts and feelings were considered to be of questionable validity (Hubbard et al, 2003). Instead, carers were asked to rate people's quality of life using proxy measures (for example the Alzheimer Disease Related Quality of Life measure, Rabins et al, 1999), which collected information **about** people with dementia, rather than **from** them.

However, research has demonstrated that there are often discrepancies between carers' views on what is important for people with dementia, and what people with dementia themselves consider to be important (see for example Dröes et al, 2006). Moreover, care-giver ratings of quality of life consistently underestimate people with dementia's own perceptions of their quality of life (Logsdon et al, 2002; Moyle et al, 2012). Indeed, Trigg et al (2011) suggested that self-reported and proxy assessments of quality of life of people with dementia may measure two very different things. This means that proxy measures are not

necessarily an appropriate or accurate way of incorporating the perspectives of people with dementia into research. (For further discussion about quality-of-life measures for people with dementia see Ready and Ott (2003), and for a review of the challenges of using proxy measures see Rand and Caiels (2015)).

Since Kitwood (1997a, 1997b), the view that people with dementia do not have an internal reality worthy of attention has been strongly challenged. However, the assumption that they may not be able to communicate that reality in a meaningful way has lingered. Moreover, the ethics processes around research are complex and can be time consuming, and this may discourage researchers from involving people with dementia. If people with dementia are included in research, their participation is usually limited to those in the early stages of dementia who are able to give informed consent to join the research project themselves. However, the experiences of people with early and later-stage dementia differ greatly.

In this paper, we describe how our research team worked in partnership with Innovations in Dementia¹ to ensure people with dementia were involved in meaningful ways, both as advisers shaping the project, and as participants. The research referred to in this paper was a three-year study, funded by the National Institute for Health Research (NIHR), which explored the use of life story work in dementia care. The full report of the study is available (Gridley et al, 2016).

Life story work

Life story work involves both **recording** aspects of a person's past life, present interests, and future plans and wishes, and **using** that information to improve their care or for their pleasure. It can involve elements of reminiscence (Woods et al, 2005) and life review (Morgan and Woods, 2010), but both reminiscence and life review focus on the past, whereas life story work seeks to include aspects of current preferences and future wishes and plans. Also, unlike reminiscence and life review, the aim of life story work is to create a tangible product, such as a life story book or box which could be shared with others.

Life story work has been used in health and social care for three decades, with children (Ryan and Walker, 1985), people with learning disabilities (Hewitt, 2000) and older people (Clarke et al, 2003). There has been a growing interest in its potential to improve care for people with dementia (Gibson, 1991; Murphy, 1994).

However, at the start of this project, there was limited research evidence on the outcomes or costs of doing life story work with people with dementia (Russell and Timmons, 2009; McKeown et al, 2010).

The study

The purpose of the research was to carry out the development and initial feasibility stages of an evaluation of life story work for people with dementia. The research was carried out between July 2012 and March 2015.

¹ Innovations in Dementia is a Community Interest Company that works collaboratively with people with dementia, and on projects to enhance the lives of people living with dementia and showcase living well with dementia. It also facilitates the Dementia Engagement and Empowerment Project (DEEP).

People with dementia were involved in several stages of the project:

Stage	Involvement of people with dementia	
	As advisers	As participants
Systematic review of the existing literature to produce a narrative synthesis of good practice and theories of change	Gave views and talked about experiences of life story work which informed the search and data extraction	n/a
Qualitative focus groups with people with dementia, family carers and professionals who use life story work	Advised on ethical issues involved in running focus groups with people with dementia Advised on the development of information sheets and consent forms Validated focus group findings	25 people with dementia (and others) took part in focus groups to explore good practice
Feasibility study to test data-gathering processes and instruments	Advised on the development of information sheets and consent forms	39 people with dementia took part by completing quality-of-life measures at several time points 10 people gave qualitative feedback about their life story books

Findings from the research are reported elsewhere (Gridley et al., 2016). This paper considers the involvement of people with dementia both in an advisory capacity, and as participants in the qualitative study and the feasibility study.

Involving people with dementia in research

Involving people in research that is about them has become an important principle in health research. Guidance from INVOLVE, an advisory group funded by NIHR that supports active public involvement in health and social care research, advises that:

'It is a core democratic principle that people who are affected by research have a right to have a say in what and how publicly-funded research is undertaken. Public involvement is an intrinsic part of citizenship, public accountability and transparency. In addition, public involvement in research can lead to empowering people who use health and social care services, providing a route to influencing change and improvement in issues which concern people most'

INVOLVE (2012, p8)

Public involvement in research can both improve the research and also lead to empowerment of those involved. This is particularly important for groups which have traditionally lacked power and voice, such as people with dementia. Assuming that people with dementia cannot contribute to research perpetuates negative stereotypes of incapability (Wilkinson, 2002). There is also a risk that, by not including the voices of people with lived experience, we may lack understanding and, therefore, fail to deliver adequate care (Cohen, 1991).

However, traditional user involvement is done in such a way that it can exclude people with dementia. Service-user involvement in research design and delivery is often in the form of an advisory group. This can exclude people with dementia from taking part, as they may be unwilling or unable to travel; may have decreased confidence; or need support to take part in formal meetings. Because of the nature of dementia as a deteriorating condition, there can be problems with the same people continuing to be involved over a longer-term project.

Ethics procedures designed to ensure that participants are treated well and are not exploited may exacerbate the exclusion of people with dementia from research. For example, the process of gaining consent may cause confusion, and there can be anxiety about 'official' forms (see for example Dewing, 2008).

We were concerned that traditional research processes could effectively 'gag' people with dementia and prevent their voices from being heard. We therefore began our study with the underlying conviction that people at all stages of dementia have a key part to play in research into services or interventions which are designed to help them (Downs, 1997). We had three motives:

- To actively work to include as many people with dementia in our research as possible
- To take a broader, more flexible attitude towards consent, including exploring the concept of assent and the use of consultees
- To use a wide range of ways of engaging with people with dementia throughout the research

We tried to work creatively to ensure that people with dementia could contribute to the research in positive and meaningful ways suited to their abilities; were able to understand, consent and assent to take part in research activities; and were able to withdraw from research activities when they wanted to.

Involving people with dementia as advisers

The project had traditional advisory and steering groups, consisting of representatives of partner organisations (steering group) and other organisations including specialist academics (advisory group). But, as outlined above, we considered that there were more appropriate ways of involving people with dementia in shaping the research (see Corner, 2002).

Innovations in Dementia recruited six advisers with dementia who contributed to the project through individual meetings with members of the Innovations in Dementia team. The research team produced easy-to-read documents at various stages explaining aspects of the research and asking questions. Innovations in Dementia staff then visited each adviser individually, collated their comments and fed back to the research team.

By meeting with people in their own homes, people with different abilities, needs and interests were supported to give advice to the project. Throughout the research, the research team produced feedback telling people how their comments had been used.

Using this format, our advisers with dementia were able to shape crucial aspects of the research:

- **Outcomes of life story work:** advisers told us what they would expect life story work to achieve, which gave us an idea of what to look for in our literature review
- **Ethical issues involved in doing focus groups:** we listened to advisers' comments about how they would feel about being involved in focus groups, and fed these into our research design and application for ethical approval from the Social Care Research Ethics Committee. For example, our advisers told us that it would be best for people with dementia to meet in a familiar place and with people they knew
- **Information sheets and consent forms:** advisers helped us make our information sheets and consent forms for people with dementia clear and easy to understand
- **Validating focus group findings:** the research team met with advisers to discuss and refine the findings from the first stage of the project

- **Style of film for dissemination:** advisers watched several short films and told us about their preferences, for example for real people rather than animations, and to avoid voiceovers. This informed the brief and will help us to choose a film-maker

The partnership between the research team and Innovations in Dementia to set up a network of advisers in this way ensured that people with dementia were able to contribute in a real and valuable way to this research. This also made for a better experience for those who took part in the research itself as participants, which we describe in the following two sections.

Involving people with dementia in the qualitative research

We conducted focus groups with 25 people with dementia, 21 family carers and 27 professionals (a total of 73 people). The aim of these was to find out:

- The different ways life story work was done
- What people perceived as good practice in life story work
- Challenges faced

In order to make sure that people with dementia could contribute, we conducted the focus groups in an inclusive way (see Bamford and Bruce, 2002 for further discussion of the successes and challenges of involving people with dementia in focus groups). Following advice from our network of advisers and advisory group, we held informal meetings in small groups, at familiar venues where people felt safe and confident, and at times to suit group members. Where appropriate, there was extra support from friends, family or professionals. People were given easily accessible information sheets, and time to read these, and we explained the research in person too.

Researchers took examples of life story books, a memory box, and a tablet with photographs to the focus group sessions. Participants brought in memorabilia, and discussed memories and meanings, and their feelings and wishes about recording and sharing their stories. In general, the focus groups with people with dementia had fewer participants than those with carers and professionals in order to allow and encourage people to take part, especially those with communication impairments.

All these sessions were held in familiar settings such as their day centre, and known and trusted group facilitators were on hand. In the group sessions, we gave people the opportunity to ask questions, and reflected back opinions and statements to make sure people were happy that their views were understood.

These group sessions provided useful information about the way life story work is done and what people consider to be good practice, and also fed directly into the next stage of the research by influencing our decisions on which outcome measures to use (for more details see Gridley, 2016).

Involving people with dementia in the quantitative research

We also involved people with dementia as participants in the feasibility-study phase of the research. Participants for this phase were in care homes or hospitals, and likely to be at a later stage of dementia than the advisers or participants in the qualitative work described above. There were, therefore, different considerations to facilitate their involvement.

This phase of the research had two strands: (i) a study in six care homes, using a stepped wedge² design (comparing quality of life and relationships before and after life story work was introduced); and (ii) a pre-test, post-test design³ in in-patient mental health assessment units across two hospital trusts (comparing outcomes in wards where life story work was done with those where it was not).

² In a stepped wedge design, each setting receives the intervention (life story work training and implementation) at some point during the period of the study (ensuring equity), but roll out is staggered, rendering implementation across multiple sites more straightforward than attempting a simultaneous start (Brown and Lilford, 2006).

³ In a pre-test post-test design, the sites receiving the intervention (in this case the hospital wards already doing life story work) are compared with a control group (a hospital ward not doing life story work).

Participants with dementia were asked to take part at three time points (in hospitals) or four time points (in care homes) over six months. At each point, they were asked to complete three separate questionnaires to measure outcomes, discussed below.

The next sections outline the challenges at the assessing capacity to consent, selection of participants, and data collection stages of the feasibility study.

Capacity and consent

The Mental Capacity Act (2005, Section 2(1)) states:

'For the purposes of this Act, a person lacks capacity in relation to a matter if at the material time he is unable to make a decision for himself in relation to the matter because of an impairment of, or a disturbance in the functioning of, the mind or brain.'

(Department of Health (DH), 2005)

Dementia is listed in the Code of Practice accompanying the Act as an example of 'an impairment or disturbance in the functioning of the mind or brain' that could affect a person's ability to make a decision (Department for Constitutional Affairs, 2007, p44). This may have been seen by some as sufficient justification for excluding all people with dementia from research on the grounds of lack of capacity. However, while such impairment is necessary for an assessment that a person lacks capacity under the Act, it is not sufficient. The act states that:

'a person is unable to make a decision for himself if he is unable:

(a) to understand the information relevant to the decision

(b) to retain that information

(c) to use or weigh that information as part of the process of making the decision or

(d) to communicate his decision (whether by talking, using sign language or any other means)'

(DH, 2005 section 3(1))

Even if a person has dementia, they may still be able to meet these conditions, especially as the Act requires that people be given 'all practical and appropriate support to help them make the decision for themselves' (Department for Constitutional Affairs, 2007, p45). Moreover, assessments of capacity are decision-specific. People with later-stage dementia may not have the capacity to understand all the details of a complex research project in order to give informed consent to become a long-term participant, but they may be capable of deciding to answer some straightforward questions that they know will be used for research purposes, and they may want the opportunity to do so.

Life story work is widely used with people in both the early and later stages of dementia, and therefore, we considered it important to measure outcomes for people at all stages of the disease.

There are already established mechanisms for including people in research who do not have the capacity to give informed consent through the use of consultees (DH, 2008). Consultees can come in two forms.

A 'personal consultee', usually a family member, is:

'someone who knows the person who lacks capacity in a personal capacity who is able to advise the researcher about the person who lacks capacity's wishes and feelings in relation to the project and whether they should join the research'

(DH, 2008, p3)

If such a person is not available, researchers must nominate another person, a 'nominated consultee', who may be a paid professional, providing they have no connection with the research project.

For people without capacity to give informed consent, we took the combined approach of seeking the advice of a consultee about whether to include the person with dementia in the study as a whole, while also seeking the **assent** of the person with dementia to ask them questions for specific outcome measures at

each time point. With advice from the consultee, we were thus able to include people without the capacity to consent in the project, enabling us to collect information from their care records and approach them for their assent to complete the outcome measures.

Assent in research with people with dementia:

'includes an affirmative agreement to participate as expressed verbally (i.e. orally) or a non-verbal indication of willingness to cooperate with study procedures, both at the time of enrolment and over the course of the study'

(Black et al, 2010)

At the start of the research, and at each time point, we asked the person if they would like to answer a few questions for our research project (the outcome measures). They were free to say yes or no, or just answer a few of them. For a wider discussion of consent processes for involving people with dementia in research, including the idea of process consent, see Dewing, 2002, 2007; and McKeown et al, 2010.

Recruitment and selection

Recruitment and selection were approached slightly differently in hospitals and care homes. In hospitals, diagnosis of dementia was recorded in medical records, but in care homes it was not always clear which residents had dementia.

Diagnosis rates for dementia were below 50% in England (DH, 2013) and diagnosis was not routinely recorded in care records in our care homes. We therefore relied instead on the judgement of the care home managers, who were asked to produce an anonymised list of all residents with dementia from which ten in each care home were randomly selected to be approached (directly or through a consultee).

Contacting consultees was not always straightforward, as, for data protection reasons, initial contact had to be made by care home staff to ask if they would mind being contacted by a researcher. This was obviously not a priority for staff, and the research team had to remind staff several times before contact was made in many cases (staff 'gatekeeping': Hellstrom et al, 2007).

This proved to be a critical barrier to recruitment, as it transpired that the majority of people with dementia (43 out of 59) who were selected for the study did not have the capacity to give informed consent, and without agreement from the potential consultee for us to make contact with them we could not progress, effectively excluding the person with dementia from the study. Because of the indirect nature of first contact, we will never know the reasons for refusal or non-response from those who did not agree to a researcher making contact.

In each care home, we randomly selected ten residents to take part – a total of 59 were invited, and 39 eventually recruited. Out of the 16 people with dementia we approached who had capacity to consent for themselves, 14 chose to take part (87.5%) and two refused (12.5%). Of the 43 people without capacity, 25 consultees advised us to include their relatives (58%), seven were excluded on consultee advice (16%) and a further 11 (26%) were effectively excluded because potential consultees did not agree to the researchers contacting them in the first place. Clearly, the people with dementia who had capacity were more likely to give a positive response than consultees. However, further research would be required to ascertain whether this difference reflects true differences in the preferences of people with dementia at different stages of progression, or differences in the perspectives of consultees on the one hand and people with dementia on the other.

Recruitment was slightly different in hospitals – this was not a stable population so people were recruited on admission. Hospital recruitment remained open for six months, and our target was to recruit 60 new patients in total. However, there were only 50 new admissions across all sites during the recruitment period. Of these, 43 were eligible to take part, but none were assessed to have capacity and only 27 consultees were approached by staff. We finally recruited a total of just 12 people with dementia across all hospital sites (44% of those approached, 28% of all those eligible). Here, the combination of staff gatekeeping and consultees advising against inclusion made recruitment doubly difficult.

Data collection

We used three standardised measures to collect data about the quality of life of people with dementia and their relationships with their family carers. To measure quality of life we used two tools designed specifically for use with people with dementia: DEMQOL (Smith et al, 2007) and the Quality of Life in Dementia – Alzheimer's Disease measure (QOL-AD) (Logsdon et al, 2002). We also used The Scale for the Quality of the Current Relationship (Spruyette et al, 2002) to measure the quality of relationships between people with dementia and their carers.

Both QOL-AD and DEMQOL have been shown to be reliable and valid measures of quality of life for people with mild to moderate dementia (Bowling et al, 2015) but there is less evidence of successful use for people with severe dementia. Both consist of a series of questions with a choice of four answers for each question. The Scale for the Quality of Current Relationship consists of 14 statements with answers given on an agree/disagree scale.

In care homes, people with dementia found QOL-AD the easiest of the three sets of questions to answer, and 64% (24/39) of people were able to complete this at baseline (in a face-to-face interview). This was substantially more than were able to give informed consent themselves (36%, 14/39), demonstrating the importance of including people who do not have capacity to give informed consent in research. DEMQOL fared slightly less well, with 31% (12/39) of participants able to complete this at baseline. However, we used a visual aid for QOL-AD (each question printed in large print on coloured paper, with the answer options underneath). It was also the shortest, and often delivered first, which may have influenced the successful response rate. We alternated the order of the measures at the final follow up but this did not appear to influence response rates.

In hospital wards, where participants had more severe dementia, none were able to give consent to take part themselves, and four out of the 12 consultees did not give permission for us to approach their family member directly. Of the eight participants we did approach, at baseline only two were able to complete QOL-AD and one to DEMQOL.

The Scale for the Quality of the Current Relationship was the least successful measure in this study. It consists of a series of statements to be read out by the researcher, and an agree/disagree scale ranging from 'strongly agree' to 'strongly disagree'⁴. Agree/disagree scales are cognitively difficult for anyone as they mix attitudes (whether you agree or disagree with a statement) with strength of feeling (how much you agree or disagree) (Fowler, 1995). The subject matter compounded difficulties, as the questions ask how well the person gets on with their family carer. Asking participants to clarify a response by saying to what degree they 'agreed' or 'disagreed' was inevitably confusing. However, many participants understood the questions and gave an appropriate verbal response, but could not pick a point on the scale, so we recorded verbatim answers (which a member of the research team coded later) to ensure as many people as possible could participate.

We found that some questions in all of the measures could be emotional for some people, and often people wanted to give longer, more detailed answers than was required by the tick-box format. It felt unethical to rush people, particularly bearing in mind we might be their only visitor that day or even that week.

Data collection for this stage of the study was particularly resource-intensive. It takes time to build relationships and trust with people with dementia, and with staff in care homes and hospital wards.

Conducting research in care homes and hospital ward environments is difficult and time consuming. Researchers had to fit data collection around the priorities and activities of the site, and residents, of course, had their own preferences. Data collection had to stop for residents' scheduled meal times, activities or medical appointments, and we arrived for one pre-arranged visit to find many of the residents in the study were out on a day trip.

⁴ The items in this measure are translated for use in English from a validated Dutch measurement tool. The language in the answer scale was particularly difficult for our first set of participants to understand, spanning from '1 = totally not agree' to '5 = totally agree'. After piloting we amended the answer scale to the more conventional '1 = strongly disagree' to '5 = strongly agree'.

We could collect data from several people with dementia in a care home or hospital ward on the same day. But if someone had moved, for example, from a hospital ward to a non-study care home, it could take an entire day to visit and collect information from that one person. Although some measures were designed for self-completion, this was not possible for people with later stages of dementia, so it could take a researcher anything up to two hours to get through the three measures (including the consent process at the start, and general chatting about what the resident wanted to talk about). This is not feasible for off-site researchers in a larger-scale trial, and in such studies, it may instead be more appropriate to have on-site staff administering outcome measures.

Conclusion

Early in this paper, we raised concerns about excluding the voices of people with dementia from research. Processes designed to protect potential research participants from harm can sometimes lead to them being excluded from research altogether. Researchers may be discouraged from including people with dementia as participants, particularly those who cannot themselves consent to take part by lengthy ethics processes largely designed for medical research and sometimes ill-adapted for qualitative work. Involving people with dementia in the design and management of research projects in a meaningful way is also time consuming and can be expensive. Previously, researchers have either neglected to do this or have resorted to tokenistic involvement such as inviting a single individual to join an advisory group.

In this paper, we have described how we were able to involve people with dementia as advisers and participants in this research project through a partnership between the research team and Innovations in Dementia. By planning this partnership in advance, we ensured Innovations in Dementia had enough resources, and were able to spend adequate time involving people in a meaningful way. In addition, through the use of consultees, we were able to include people with more advanced dementia as participants, particularly in the feasibility-study stage of the research. As this is a group which is often excluded from research, we felt it was extremely important that we were able to collect information about quality of life and relationships from this group directly, rather than relying on proxy measures. Our completion rates showed that more people were able to complete the outcome measures than were capable of giving consent to take part, demonstrating that people with more advanced dementia can give useful information even when not able to complete a complex consent process. This is particularly important in light of research demonstrating that there are often differences between how people with dementia and carers rate the quality of life of people with dementia, with carers relying on functional performance, and people with dementia on current mood (Trigg et al, 2011).

However, there were still aspects of involving people with dementia which did not work as well as we had hoped. The process of using consultees, while it allowed some people with dementia to be involved when they otherwise may not have been, was lengthy and complex. We relied heavily on busy care home and hospital staff to contact potential consultees initially, and this was often unsuccessful, meaning several people were not given the opportunity to participate at all.

Outcome measurement tools, while widely recognised standards for measuring quality of life, did not necessarily work for all participants. People with the later stages of dementia in particular could find prescribed 'tick box' answers restrictive and unsuitable, even when they were able to understand the questions being asked. If outcome measures are being used, we would recommend a more flexible approach of recording verbatim answers and, if necessary, coding these later.

We encourage the wider research community to:

- Include as many people with dementia in research as possible
- Take a broader, more flexible attitude towards consent, including exploring the concept of assent and the use of consultees
- Use a wide range of ways of engaging people with dementia throughout the research process

In order for this to happen, there is an urgent need to develop: (a) new tools for measuring quality of life for people with dementia which do not require participants to respond in prescribed ways, but which are still acceptable for use in cost-effectiveness studies; and (b) ethics and consent processes which are more appropriate for non-medical research to facilitate involvement of people with dementia.

Acknowledgments

The research team would like to thank all who took part in research, particularly people with dementia and those who helped to facilitate their involvement.

References

- Bamford, C. and Bruce, E. (2002). 'Successes and challenges in using focus groups with older people with dementia'. In Wilkinson, H. (ed) *The perspectives of people with dementia: research methods and motivations*. London: Jessica Kingsley Publishers.
- Black, B. S., Rabins, P. V., Sugarman, J., and Karlawish, J. H. (2010). 'Seeking assent and respecting dissent in dementia research'. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, 18(1), 77–85.
- Bowling, A., Rowe, G., Adams, S., Sands, P., Samsi, K., Crane, M., Joly, L., Manthorpe, J. (2015) 'Quality of life in dementia: a systematically conducted narrative review of dementia-specific measurement scales'. *Aging and Mental Health* 19(1): 13-31.
- Brown, C.A. and Lilford, R. J. (2006). 'The stepped wedge trial design: a systematic review'. *BMC Med Res Methodol*; 6:54. <http://dx.doi.org/10.1186/1471-2288-6-54>.
- Clarke, A., Hanson, E., Ross, H. (2003). 'Seeing the person behind the patient: enhancing the care of older people using a biographical approach'. *Journal of Clinical Nursing* 12: 697-706.
- Cohen, D. (1991). 'The subjective experience of Alzheimer disease: the anatomy of an illness as perceived by patients and their families.' *The American Journal of Alzheimer's Care and Related Disorders and Research*. May/June: 6-11.
- Corner, L. (2002). 'Including people with dementia: advisory networks and user panels'. In Wilkinson, H. (ed) *The perspectives of people with dementia: research methods and motivations*. London: Jessica Kingsley Publishers.
- Department for Constitutional Affairs (2007). *Mental Capacity Act 2005 Code of Practice*. London: TSO.
- Department of Health. (2005). *Mental Capacity Act*. London: HMSO.
- Department of Health. (2008). *Guidance on nominating a consultee for research involving adults who lack capacity to consent*. http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_083131 [Accessed 9 January 2017]
- Department of Health. (2013). *Dementia: a state of the nation report on dementia care and support in England*. Available online at: www.gov.uk/government/publications/dementia-care-and-support [Accessed 22 August 2016]
- Dewing, J. (2002). 'From ritual to relationship: a person-centred approach to consent in qualitative research with older people who have a dementia.' *Dementia* 1(2): 157-171.
- Dewing, J. (2007). 'Participatory research: a method for process consent with people who have dementia'. *Dementia* 6(1): 11-35.
- Dewing, J. (2008). 'Process consent and research with older persons living with dementia'. *Research Ethics Review* 4: 59-64.
- Downs, M. (1997). 'The emergence of the person in dementia research'. *Ageing and Society* 17: 597-607.

- Dröes, R. M., Boelens-Van Der Knoop, E. C., Bos, J., Meihuizen, L., Ettema, T. P., Gerritsen, D. L., Hoogeveen, F., De Lange, J. and SchöLzel-Dorenbos, C. J. (2006). 'Quality of life in dementia in perspective: an explorative study of variations in opinions among people with dementia and their professional caregivers, and in literature'. *Dementia* 5(4): 533-558.
- Fowler, F. J. (1995). 'Improving survey questions: design and evaluation'. *Applied Social Research Methods Series* 38: 56-57.
- Gibson, F. (1991). *The lost ones: recovering the past to help their present*. Stirling: Dementia Services Development Centre.
- Gridley, K., Brooks, J., Birks, Y., Baxter, K., Parker, G. (2016). 'Improving care for people with dementia: development and initial feasibility study for evaluation of life story work in dementia care'. *Health Services and Delivery Research* 4(23).
- Gridley, K. (2016). 'Understanding the outcomes of life story work' in Kaiser, P. and Eley, R. (eds) *Life Story Work with People with Dementia: Ordinary Lives, Extraordinary People*. London: Jessica Kingsley Publishers.
- Hellstrom, I., Nolan, M., Nordenfelt, L. and Lundh, U. (2007). 'Ethical and methodological issues in interviewing persons with dementia'. *Nursing Ethics* 14: 608-619.
- Hewitt, H. (2000). 'A life story approach for people with profound learning disabilities'. *British Journal of Nursing* 9(2): 90-95.
- Hubbard, G., Downs, M. G. and Tester, S. (2003). 'Including older people with dementia in research: challenges and strategies'. *Aging & Mental Health* 7(5): 351-362.
- INVOLVE (2012). *Briefing notes for researchers: involving the public in NHS, public health and social care research*. INVOLVE, Eastleigh.
- Kitwood, T. (1997a). 'The experience of dementia'. *Aging & Mental Health* 1(1): 13-22.
- Kitwood, T. (1997b). *Dementia reconsidered*. Buckingham: Open University Press.
- Logsdon, R. G., Gibbons, L. E., McCurry, S. M., Logsdon, R. G., Gibbons, L. E., McCurry, S. M. et al (2002). 'Assessing quality of life in older adults with cognitive impairment'. *Psychosomatic Medicine* 64 :510-519.
- McKeown, J., Clarke, A., Ingleton, C., Ryan, T., Repper, J. (2010). 'The use of life story work with people with dementia to enhance person-centred care'. *International Journal of Older People Nursing* 5: 148-158.
- Morgan, S. and Woods, R. T. (2010). 'Life review with people with dementia in care homes: a preliminary randomized controlled trial'. *Non-pharmacological Therapies in Dementia*, 1(1): 43-60.
- Moyle, W., Murfield, J. E., Griffiths, S. G. and Venturato, I. (2012). 'Assessing quality of life of older people with dementia: a comparison of quantitative self-report and proxy accounts'. *Journal of Advanced Nursing* 68(10).
- Murphy, C. (1994). *'It started with a sea-shell': life story work and people with dementia*. Stirling: University of Stirling Dementia Services Development Centre.
- Rand, S., and Caiels, J. (2015). 'Using proxies to assess quality of life: a review of the issues and challenges'. Discussion Paper 2899. QORU: Personal Social Services Research Unit, University of Kent, Canterbury.
- Ready, R. E., and Ott, B. R. (2003). 'Quality of life measures for dementia'. *Health and Quality of Life Outcomes* 1(11).
- Russell, C. and Timmons, S. (2009). 'Life story work and nursing home residents with dementia'. *Nursing Older People* 21: 28-32.
- Ryan, T. and Walker, R. (1985). *Making life story books*. London: British Agencies for Adoption and Fostering.

Smith, S. C., Lamping, D. L., Banerjee, S., Harwood, R. H., Foley, B., Smith, P., Cook, J. C., Murray, J., Prince, M., Levin, E., Mann, A., Knapp, A. (2007). 'The development of a new measure of health related quality of life for people with dementia: DEMQOL'. *Psychological Medicine* 37: 737–746.

Spruytte, N., Van Audenhove, C., Lammertyn, F. and Storms, G. (2002). 'The quality of the caregiving relationship in informal care for older adults with dementia and chronic psychiatric patients' *Psychology and Psychotherapy: Theory Research and Practice* 75(3): 295-311.

Trigg, R., Watts, S., Jones, R. and Tod, A. (2011). 'Predictors of quality of life ratings from persons with dementia: the role of insight'. *International Journal of Geriatric Psychiatry* 26(1): 83-91.

Wilkinson, H. (2002). 'Including people with dementia in research: methods and motivations'. In: Wilkinson, H. (ed) *The perspectives of people with dementia: research methods and motivations*. London: Jessica Kingsley Publishers.

Woods, B., Spector, A. E., Jones, C. A., Orrell, M., Davies, S. P. (2005). 'Reminiscence therapy for dementia'. *Cochrane Database of Systematic Reviews* 2. DOI: 10.1002/14651858.CD001120.pub2.

Gypsy and Traveller household formation rates

Nigel Moore, Opinion Research Services

Abstract

The Housing Act 2004 introduced a requirement for planning authorities to assess the need for additional Gypsy and Traveller pitches. These studies became known as Gypsy and Traveller Accommodation Assessments (GTAAAs), and for many years they used an extremely high net household formation rate of 3% a year to project future household growth rates. This was based on no evidence, and the result was to inflate local requirements for additional pitches unrealistically.

ORS has undertaken GTAAAs for half of all planning authorities across England and Wales. Our findings are that the net growth rate for both Gypsy and Traveller population and households is more plausibly in the range of 1.25% to 1.5%. Applying this rate of projected growth on behalf of planning authorities has significantly reduced their identified needs for new pitches, but has also encouraged many to fulfil their planning requirements because the figures are more plausible.

Introduction

Paragraph 225 of the Housing Act 2004 introduced a requirement for local planning authorities to, 'Carry out an assessment of the accommodation needs of Gypsies and Travellers residing in or resorting to their district.' These studies became known as Gypsy and Traveller Accommodation Assessments (GTAAAs). Subsequent planning policies defined Gypsies and Travellers as anyone who was either culturally a Gypsy and Traveller because of their family background, or who travelled to earn their livelihood. Policies also required the planning authority to have an identified five-year land supply of Gypsy and Traveller pitches to meet their needs, with a pitch being sufficient space for one household to occupy. Authorities which do not have a five-year land supply of Gypsy and Traveller pitches are very vulnerable to losing planning appeals for new pitches, and also are at risk of having their Local Plan found not to be sound.

The rate of household growth is a key element in all housing assessments, including GTAAAs. Compared with the general population, the relative youthfulness of many Gypsy and Traveller populations means that their birth rates are likely to generate higher-than-average population growth, and proportionately higher gross household formation rates. However, while their gross rate of household growth might be high, Gypsy and Traveller communities' future accommodation needs are, in practice, affected by any reduction in the number of households because of dissolution. Therefore, the net rate of household growth is the gross rate of formation minus any reductions in households because of such factors. It is the net rate that is important in determining future accommodation needs for Gypsies and Travellers.

Following the introduction of the Housing Act 2004, there was no evidence on how quickly Gypsy and Traveller households would grow in the future. There was also significant confusion between the gross and net household formation rate. For example, in 2007 the Department of Communities and Local Government issued guidance for conducting GTAAAs (DCLG, 2007) which used a 3% formation rate within a worked example. However, this was ambiguous as to whether planning authorities should treat the 3% as a gross or net formation rate, and was also based on no demographic evidence.

The 3% figure was quickly adopted as a norm for net household formation rates for most GTAAAs despite having no evidence to support it. Adopting a 3% net household formation rate implies a very high growth

rate in future household numbers, which in turn, implies a very high need for Gypsy and Traveller pitches. For example, a 3% net formation rate implies that household numbers will double over a 23.5-year period, and over a typical 20-year local plan period, a planning authority which has 100 Gypsy and Traveller pitches at the outset would need an additional 81 pitches. These figures represent extraordinary rates of growth in households in the context that household projections for England as a whole show growth rates of less than 1% a year. The high rate of projected growth also made it very difficult for planning authorities to comply with the need to have a five-year land supply of pitches because the needs are so high, and therefore, their planning policies were vulnerable to the planning appeal system. This, in turn, has seen many planning authorities losing planning appeals for new Gypsy and Traveller sites in areas such as green-belt locations where housing sites would typically not be granted permission. Planning authorities, therefore, require robust evidence bases to help plan for the future needs of Gypsies and Travellers.

Because of a past lack of evidence about the demographic growth rate for Gypsy and Traveller households, GTAAs adopted an extremely high assumption about future household growth. This paper highlights why the 3% growth rate assumption is very implausible, and also presents alternative and more realistic estimates for the future population and household growth rates for Gypsies and Travellers.

Modelling population growth for Gypsies and Travellers

The basic equation for calculating the national domestic rate of Gypsy and Traveller population growth is simple: start with the base population and then calculate the average increase/decrease by allowing for births and deaths. Given that we are initially seeking to understand the national picture for the existing population, we can assume that net migration to England as a whole is zero and has no impact on population growth. Nevertheless, deriving satisfactory estimates is difficult because of a lack of data for the Gypsy and Traveller population for both birth rates and life expectancy.

The main source for the rate of Gypsy and Traveller population growth is the UK Census of Population 2011. However, the 2011 Census data can be supplemented by household survey data collected by Opinion Research Services (ORS) which is derived from more than 2,000 face-to-face interviews with gypsy and traveller households, and represents a comprehensive source of demographic information. The survey data provides support for the data found in the 2011 Census by being consistent with its age profile and household types. The ethnicity question in the 2011 Census included for the first time 'Gypsy and Irish Traveller' as a specific category. While non-response bias probably means that the size of the population was underestimated, the age profile the Census provides is not necessarily distorted and matches the profile derived from the extensive household surveys. No significant new data relating to the demographic structure of the Gypsy and Traveller population of England is expected until the 2021 Census, so the information contained within this paper is unlikely to be supplanted in the foreseeable future.

Birth and fertility rates

The UK Census of Population provides a way of understanding the rate of population growth through births. For England as a whole, the UK Census of Population 2011 shows that surviving children aged 0-4 years comprise 10.4% of the Gypsy and Traveller population – which means that, on average, 2.1% of the total population was born each year (over the last five years). The same estimate is confirmed if we consider that those aged 0-14 comprise 29.8% of the Gypsy and Traveller population – which also means that almost exactly 2% of the population was born each year.

A key tool used in demographic projections is the total fertility rate (TFR) which is a measure of the average number of children a woman can expect to have during her lifetime. For the whole UK population, the TFR is just below two – which means that, on average, each woman can be expected to have just less than two children who reach adulthood. An analysis of the UK Census of Population 2011 data by researchers at the University of Manchester indicated an overall fertility rate of 2.75 children for each female for Gypsies and Travellers (Jivraj and Simpson, 2015). Meanwhile, the household survey data shows that, on average, Gypsy and Traveller women aged 32 years have 2.5 children. However, because the children of women above this

age point tend to leave home progressively, a full TFR could not be calculated. Despite this, based on these figures, it is reasonable to assume an average of three children for each woman during her lifetime, which is consistent with the 2.75 TFR obtained by the University of Manchester. A TFR of three children for each woman can be considered to be a reasonable upper end of the estimate for calculating population growth. Therefore, this paper proceeds with an estimate for Gypsy and Traveller population growth rate based on women having an average of three children each.

Death rates

Although the above data imply an annual growth rate through births of about 2%, the death rate has also to be taken into account, which means that net population growth cannot conceivably achieve 2% a year. In England and Wales, there are nearly half-a-million deaths each year – about 0.85% of the total population of 56.1 million in 2011. If this death rate is applied to the Gypsy and Traveller community, then the resulting projected growth rate is in the region of 1.25% a year if the birth rate provides for 2.1% of the population being born each year.

However, the Gypsy and Traveller population is significantly younger than average and may be expected to have a lower percentage death rate overall (even though a smaller than average proportion of the population lives beyond 70 years). Whereas the average life expectancy across the whole population of the UK is currently just over 80 years, a University of Sheffield University report found that Gypsy and Traveller life expectancy is about 10-12 years less than average (Parry et al, 2004). So, in our population growth modelling we have used a conservative estimate of average life expectancy as 72 years, which is entirely consistent with the lower-than-average number of Gypsies and Travellers aged over 70 years in the 2011 Census and also the very low numbers of people aged over 70 years in the household survey data. The death rate as a percentage of the population varies from year to year, but is around 0.75% on average.

Modelling outputs

Combining the above information, it is possible to model the growth of the national Gypsy and Traveller population based on the most likely birth and death rates by using PopGroup (the leading software for population and household forecasting). PopGroup takes the existing population age structure from the 2011 Census and applies the TFR and projected life expectancy to the population to project how large it will be in the future.

Based upon recognised academic studies and the very detailed household survey data, if we assume a birth rate which consistently generates a TFR of three and an average life expectancy of 72 years for Gypsies and Travellers, then the modelling projects the population to increase by 66% over the next 40 years. This implies an average population compound growth rate of 1.25% a year, but this varies from year to year. However, if we assume that Gypsy and Traveller life expectancy increases to 77 years by 2050, then the projected population growth rate rises to nearly 1.5% a year while again varying from year to year.

It is possible to compare the projected Gypsy and Traveller population growth rate with the wider population of England. For example, in the ONS's 2012-based sub-national population projections, the projected population growth rate for England to 2037 is 0.6% a year, of which 60% is caused by natural change and 40% by migration. Therefore, the natural population growth rate for England is almost exactly 0.35% a year, meaning that our estimate of the Gypsy and Traveller population growth rate is four times greater than that of the general population of England.

Converting population to household growth

In addition to population growth influencing the number of households, the size of households also affects the number. If we assume that there is to be no change in household sizes for the foreseeable future, the growth rate for population will also be the growth rate for households. The evidence for Gypsy and Traveller households is that this represents a reasonable assumption, but it is necessary to consider factors which may mean households grow more quickly than population.

Population and household growth rates do not necessarily match directly, mainly because of the tendency for people to live in smaller (childless or single person) households (including, of course, older people (following divorce or as surviving partners)). Based on such factors, the CLG 2012-based projections convert current population data to a projected household growth rate of 0.85% a year (compared with a population growth rate of 0.6% a year).

Given that the Gypsy and Traveller population is relatively young and has many single parent households, a 1.5% annual population growth could yield higher-than-average household growth rates, particularly if average household sizes fall or if younger-than-average households form. However, while there is evidence that Gypsy and Traveller households already form at an earlier age than in the general population, the scope for a more rapid rate of growth through even earlier household formation is limited.

Based on the 2011 Census, table 1 compares the age of household representatives in English households with those in Gypsy and Traveller households, showing that the latter have many more household representatives aged under-25 years. In the general English population, 3.6% of household representatives are aged 16-24, compared with 8.7% in the Gypsy and Traveller population. The detailed household survey data shows that about 10% of Gypsy and Traveller households have household representatives aged under-25 years, in line with the UK Census of Population 2011.

Table 1: Age of head of household

Age of household representative	All households in England		Gypsy and Traveller households in England	
	Number of households	Percentage of households	Number of households	Percentage of households
Age 24 and under	790,974	3.6%	1,698	8.7%
Age 25 to 34	3,158,258	14.3%	4,232	21.7%
Age 35 to 49	6,563,651	29.7%	6,899	35.5%
Age 50 to 64	5,828,761	26.4%	4,310	22.2%
Age 65 to 74	2,764,474	12.5%	1,473	7.6%
Age 75 to 84	2,097,807	9.5%	682	3.5%
Age 85 and over	859,443	3.9%	164	0.8%
Total	22,063,368	100%	19,458	100%

Source: UK Census of Population 2011

Table 2 shows that the proportion of single person Gypsy and Traveller households is not dissimilar to the wider population of England but that there are more lone parents, fewer couples without children, and fewer households with non-dependent children amongst Gypsies and Travellers. This suggests that Gypsy and Traveller households form at an earlier age than the general population.

Table 2: Household type

Household Type	All households in England		Gypsy and Traveller households in England	
	Number of households	Percentage of households	Number of households	Percentage of households
Single person	6,666,493	30.3%	5,741	29.5%
Couple with no children	5,681,847	25.7%	2,345	12.1%
Couple with dependent children	4,266,670	19.3%	3,683	18.9%
Couple with non-dependent children	1,342,841	6.1%	822	4.2%
Lone parent: Dependent children	1,573,255	7.1%	3,949	20.3%
Lone parent: All children non-dependent	766,569	3.5%	795	4.1%
Other households	1,765,693	8.0%	2,123	10.9%
Total	22,063,368	100%	19,458	100%

Source: UK Census of Population 2011

The household survey data is broadly compatible with the data above. We have found that around 50% of pitches have dependent children compared with 45% in the Census; there is a high proportion of lone parents; and about a fifth of Gypsy and Traveller households appear to be single person households.

The main point, though, is that since 20% of Gypsy and Traveller households are lone parents with no adult children at home, and up to 30% are single persons, there is limited potential for further reductions in average household size to increase current household formation rates significantly. There is no reason to think that earlier household formation or increasing divorce rates will, in the medium term, affect household formation rates. While there are differences with the general population, a 1.5% a year Gypsy and Traveller population growth rate is also likely to lead to a household growth rate of around 1.5% a year. This is more than the 0.85% for the English population as a whole, but much less than the previously used 3% rate for Gypsies and Travellers.

Key outcome for local authorities

The work undertaken using the UK Census of Population 2011 and household survey data clearly suggests lower needs for future Gypsy and Traveller pitches than if the 3% net household formation rate had continued to be used. However, this may not have reduced the provision of new Gypsy and Traveller pitches because planning authorities are more likely to be motivated to meet plausible levels of need. For example, a 2005 study for Cambridgeshire identified a need for over 1,000 new pitches in the county over a 15-year period, but this was not implemented. However, if need figures are lower, then planning authorities have shown more willingness to implement them.

The following two case studies illustrate the impact of the work contained in this paper.

Case study A

The local planning authority had a total of 259 Gypsy and Traveller households which were either living on sites at the time of the GTAA or were known to want to live on sites in the area. Many of the Gypsies and Travellers were living on sites without full planning permission, and the planning authority had lost a number of planning appeals for new Gypsy and Traveller sites because of its lack of provision.

The planning authority's local plan was running from 2013 to 2033 when ORS undertook a GTAA for it in 2014. If a net household formation rate of 3% had been applied to households in the area, then a further 188 pitches would have been required for future households. However, the GTAA applied a net household formation rate of 2% because the ORS household survey showed that the area had a relatively youthful population when compared to the national position for Gypsies and Travellers. The application of a 2% net formation rate led to the council having an identified need for 114 additional pitches.

Therefore, the GTAA reduced the need for new pitches by 74 across the local plan period because of more robust evidence on household formation rates. This reduction in need presented the council with a more plausible and achievable need figure. So, it endeavoured to meet the needs of households which did not have planning permission on their existing dwelling and provided new sites without the need for planning appeals. The more robust GTAA, when combined with a proactive planning authority, has helped to meet the needs of Gypsies and Travellers, but has also reduced the risk and cost of the planning authority being subject to planning appeals.

Case study B

The local planning authority had a total of 132 Gypsy and Traveller households which were either living on site at the time of the GTAA or were known to want to live on sites in the area. The planning authority had experienced repeated defeats at planning appeals in the years before the GTAA.

The planning authority wanted to develop a local plan which ran from 2014 to 2039. The GTAA identified a need for 46 new pitches using a 2% formation rate, because the population was slightly younger than the national average. This was 28 pitches fewer than would have been the case if using 3% as a formation rate.

The council worked to meet the needs of households in its area and to approve planning permission for several sites. Following a long period of repeated defeats at planning appeals, the planning authority has recently been successful in appeals and has been able to convince a planning inspector that it is meeting the identified need. The planning authority has met the needs of Gypsies and Travellers, and been able to successfully defend its planning policies.

Conclusion

Future Gypsy and Traveller accommodation needs have typically been over-estimated because population and household growth rates have been projected on the basis of assumed 3% a year net growth rates.

The growth of the national Gypsy and Traveller population may be as low as 1.25% a year, which is still four times greater than in the settled community. The often-assumed 3% net household growth rate a year for Gypsies and Travellers is unrealistic.

The best available evidence suggests that the net annual Gypsy and Traveller household growth rate is around 1.5% a year. The consequences of using a rate around 1.5% compared to 3% can be significant over the period of a local plan. Reducing the 'need figures' for a planning authority does not necessarily mean that fewer Gypsy and Traveller pitches are provided. A more realistic need figure in a GTAA means that local authorities are more likely to try to meet the need. An unrealistic figure means that local authorities have little chance of meeting needs, and this acts as a disincentive.

More realistic needs figures in GTAAs may encourage authorities to provide more pitches to meet real needs, while at the same time reducing the risks of adverse consequences for planning authorities at local plan hearings and planning appeals.

References

Department of Communities and Local Government (2007). Gypsy and Traveller accommodation needs assessments: guidance. London: Department for Communities and Local Government.

Jivraj, S. and Simpson, L. (Editors) (2015). Ethnic identity and inequalities in Britain: the dynamics of diversity. Bristol: Policy Press.

Parry, G., van Cleemput, P., Peters, J., Moore, J., Walters, S., Thomas, K. and Cooper, C. (2004). The health status of Gypsies and Travellers: report of Department of Health Inequalities in Health Research Initiative. Sheffield: University of Sheffield.

An introduction to address-based online surveying

Joel Williams, Kantar Public

Abstract

There is a strong demand in the UK for an affordable method of surveying the general population which still uses random sampling techniques. In this paper, we describe an alternative with antecedents in the US but new to the UK: address-based online surveying (ABOS). This paper covers all aspects of the method, seeking to answer questions about its coverage levels, its data quality, likely response rates and the potential for bias in substantive estimates.

Motivation for a new general population survey model

UK business and government are great consumers of information about their populations of interest. Traditionally, much of this information has been derived from questionnaire surveys and, although other data forms are now becoming influential, this is still the dominant method of information gathering in the UK.

Since the late 1940s, the gold standard for sample quality has been to use random sampling methods with in-person data collection. However, although response rates remain fairly high (45-75%), in-person data collection is expensive and requires long timeframes. Consequently, there has always been a demand for more affordable and more agile alternatives but which still use random sampling techniques.

From the early 1970s until relatively recently, RDD ('random digit dial') telephone surveying was the primary alternative. However, the modern requirement to sample mobile numbers as well as landline numbers has made the method more expensive than it was. Contact and cooperation rates have also dropped in recent years, adding to the method's cost and taking away from its quality.

Since the early 2000s, online panel data collection has been used for 'volume' research in the commercial sector. This method has much lower costs and greater timeliness compared to other data collection methods. However, these panels usually exclude everyone who is offline, and random sampling methods are rarely used in their construction. This makes it hard to be *generally* confident about the accuracy and precision of their estimates, despite the occasional success.

An alternative that combines the best features of random sampling (sample unbiasedness for *all* survey variables, not just those for which population parameters are available) with the convenience and (relatively) low cost of online panels would be a very popular option.

A random sample online panel has potential – several exist abroad plus one modest-sized one in the UK¹ – but the substantial set-up costs have discouraged most research agencies from building their own. At Kantar Public (formerly TNS BMRB in the UK), we have, instead, developed a non-panel method – address-based online surveying (or 'ABOS') – as a way to meet demand for low-cost random sample surveying in the UK. It is not strictly new, having its antecedents in 'push-to-web' US studies², but its specific implementation in the UK *is* new.

¹ This has been developed by NatCen Social Research.

² See Messer (2012) for a summary.

Introduction to address-based online surveying (ABOS)

The core ABOS design is a simple one: a stratified random sample of addresses is drawn from the Royal Mail's postcode address file and an invitation letter is sent to 'the residents' of each one, containing username(s) and password(s) plus the url of the survey website. Respondents can log on using this information and complete the survey as they might any other online survey. Once the questionnaire is complete, the specific username and password cannot be used again, ensuring data confidentiality from others with access to this information.

However, this core design must be augmented with several other features to make it workable. Over the course of the last three to four years, we have carried out experiments and gathered other evidence to help us understand which features work best.

This paper is intended as an introduction to the method and as a summary of what we know and what we do not know at the time of writing (autumn 2016). Over the course of this paper, we intend to answer a series of (hypothetical!) questions about the ABOS method, presenting evidence where we have it and giving a current viewpoint about what constitutes best practice. We will also note any plans for the future, including features likely be tested through already commissioned studies. The questions are:

Q1: If the sample is of addresses, how do you convert this into a sample of individuals?

Q2: How do you verify that the data is from the sampled individual(s)?

Q3: How do you cover 'offline' individuals?

Q4: What response rate does the ABOS method get, and what is the impact of the design features you have tested?

Q5: How does response rate vary between subpopulations, and what (if anything) can you do about it?

Q6: What evidence do you have for non-response bias?

Q7: How much does it cost?

Q1: If the sample is of addresses, how do you convert this into a sample of individuals?

The postcode address file (PAF) is thought to provide a highly comprehensive link to the general population of adult individuals living in residential households but the ABOS method has no interviewer to facilitate that link. Instead, we must rely on one or more residents at the address to do this job for us.

The first thing to note is the fact that a small fraction (probably 2-3% in England) of addresses contains more than one household, and there is no way to 'sample' one in a controlled manner. Whoever picks up the letter effectively self-selects their household into the sample. While a weakness, this departure from random sampling is, in our view, small enough to be accommodated in most cases.

Accepting this uncontrolled conversion from a sample of addresses into a sample of households, the question is how to get from here to a sample of individuals while respecting random sampling principles. As part of a test of European Social Survey (ESS) methods, Park and Humphrey (2014) used a variant of the ABOS method in which the first adult to read the letter was asked to log on to the survey website and complete a short questionnaire on household composition. At the end of this, the survey software randomly selected one resident adult and requested that the initial respondent facilitate a transfer to

this selected person (if different). Subsequent analysis suggested that, in many households, this selection stage was ignored and that the goal of a random within-household sample was not obtained.

In our early tests of an ABOS version of the Cabinet Office Community Life Survey³, we instead tested the quasi-random 'birthday' selection method in which the adult resident with the last – or next – birthday⁴ is asked to complete the questionnaire. This is not a true random sampling method but, if implemented accurately, should provide functional equivalence. Its theoretical advantage over the ESS selection method is its simplicity: it does not take much thought to work out who has the last/next birthday and there is no two-stage responding process. However, simplicity does not guarantee compliance. Do households – or more accurately, the individuals picking up the letters – bother with this part?

To test compliance, we included a question on month-of-birth of each adult resident in the household. We hypothesised that those ignoring the 'birthday' sampling instruction would still provide this data (if known). We could then use the date of questionnaire completion to work out which individual *should* have been selected, or at least identify the majority of 'wrong' respondents. If we had sent no sampling instruction – or alternatively, every household ignored the sampling instruction – we would expect the 'right' respondent every n^{th} time where n is the number of eligible individuals in the household. This is the baseline against which to measure the success or otherwise of the sampling instruction. We concluded from our test that the success rate was not a great deal above baseline, and that this success rate was lower for larger households than for smaller ones. Overall, c25% of respondents were identified as 'wrong' respondents.

There are a number of possible ways forward from this. Option 1 is to use the birthday selection method but accept a significant level of non-compliance. Option 2 is to use the birthday selection method but to identify and exclude 'wrong' respondents from the analysis base, effectively converting the problem from one of sampling error to one of non-response error. Option 3 is to ask *all* eligible individuals in the household to participate in the survey, eliminating the flawed within-household sampling process altogether but introducing other challenges in its place.

Option 1 has been used for one of our most recent ABOS studies, motivated by the fact that the Community Life data did not suggest any systematic difference between 'right' and 'wrong' respondents for demographic profile or the substantive variables in that survey. However, not all research commissioners will think this sufficient evidence, and some will find the non-compliance with the sampling instructions fundamentally problematic. Option 2 – the exclusion of 'wrong' respondents – might be more palatable but would lead to a reduction of c25% in the analytical base, and a corresponding increase in costs to maintain the intended sample size. Furthermore, it would produce a responding sample biased towards one-person-households in which the probability of a 'wrong' respondent is zero.

In the end, we recommended to the Cabinet Office a test of option 3 (Williams, 2014), in which *all* adult residents are asked to complete the questionnaire. This is achieved by supplying four sets of login details (with more available on request) which can be used in any sequence. However, although this solves the 'wrong respondent' problem, it introduces others.

Although there are several *general* risks when surveying multiple individuals in the same household, our main concern with option 3 was the risk that one individual would complete the questionnaire multiple times, especially if each completion was incentivised. Of course, the incentive could be dropped or limited to one for each household – removing (most of) the motivation for proxy completion – but doing so would replace the risk of proxy completion with a risk of lower response motivation. Ideally, we would *identify and exclude* proxy completions rather than remove the incentive to respond. The question is how? This forms part of a larger question about how to verify ABOS data so that it is of sufficient quality for users (see Q2).

³ The relevant part of the Cabinet Office – The Office for Civil Society – is now situated within the Department for Culture, Media and Sport. This includes the responsibility for the Community Life Survey.

⁴ Best practice is to allocate a random half of the addresses to 'last' birthday and the other half to 'next' birthday, in an effort to minimise 'season-of-birth' sample effects that affect some survey variables.

One final option – but one untested by us – is to use a hybrid selection mechanism. In Sport England’s new Active Lives Survey⁵, which uses a variant of the ABOS method, *any two* residents are invited to take part, a compromise between random sampling principles and the desire to limit the maximum incentive available for each household (and thus the motivation for proxy completion). This method only departs from true random sampling in households with three or more adults (c18% of households in England) which might be a minor enough exception for research commissioners to accept. However, households with three or more adults are distinctive in numerous ways. With this method, there is a small risk that real differences between these types of household and other types of household are confounded with differences due to the sampling mechanism.

In sum, there is no perfect way of converting a sample of addresses to a sample of individuals, only a set of imperfect ways. For most of our newer ABOS studies, we have used option 3 (all adults can take part) combined with algorithmic weeding out of probable proxy completions. However, we consider this a live topic for research as the empirical evidence for each method is still rather thin.

Q2: How do you verify that the data is from the sampled individual(s)?

With interview-based surveys we have confidence that almost all the data is collected in a controlled manner and from the right individual. Interviewers ensure that the survey protocol is followed, and they themselves are monitored by survey operations staff to minimise the risk of departures from protocol and to catch the occasional cheat.

With ABOS and most other self-completion survey methods, there is no interviewer to do this work so it must be accomplished by other methods. For a start, respondents should be made aware that we expect them to supply data in good faith. This can be partly achieved through (for example) asking the respondent to confirm the conditions of questionnaire completion (non-proxy, in some privacy), asking them to ‘sign’ it as their own work, and by asking for additional contact details to facilitate post-fieldwork verification checks. All these methods make it clear to the respondent that we take data quality seriously, and this in itself may deter some proxy or careless completions of the questionnaire.

However, these design features ought to be combined with a programme of post-fieldwork verification. This can take two forms. The first is to re-contact respondents by telephone to check that the named person completed the questionnaire and (if so) to confirm a few characteristics that ought to be known only to the individual. The second form of verification is to use an algorithm to identify poor data *afterwards*. The implicit assumption underpinning the use of this algorithm is that proxy data will usually be of poor enough quality to be detectable – and discarded if desired.

As it stands, the first form of verification has been implemented only once for one ABOS study we know of (Community Life). No problems were found on that occasion but the low re-contact agreement rate – typical of self-completion surveys – is a major limitation to this form of verification. Furthermore, for cost reasons, this kind of verification can only be applied to a sample of cases so it is a far from *sufficient* method of verification.

Consequently, we are largely reliant on the second form of verification – the bad-data detection algorithm – and must do so without strong evidence of its efficacy for ABOS studies. Instead, the algorithm has been built based upon a more generic understanding of measurement error in a self-completion context.

Our algorithm varies slightly between different ABOS studies but, at its core, it uses a variety of classic indicators of proxy/careless completion, and if a small number of these indicators light up, the case is removed from the data file. This seems to us a proportionate approach to data verification given that no one indicator is *certainly* a sign of invalid data. For the record, this approach led us to remove less than 5% of cases from the 2015-16 Community Life Survey, a rate that may prove to be indicative for ABOS studies in general, and a rate that seems low enough for us to be largely confident of the data’s veracity.

⁵ See www.sportengland.org/research/about-our-research/active-lives-survey for more details.

Although each ABOS questionnaire is different, there are a number of indicators which we use across all studies. These include (i) inconsistencies in household data when multiple completed questionnaires have been received from the same household, (ii) use of the same email address by multiple respondents when providing the necessary details to receive the e-incentive, (iii) suspiciously short completion times, (iv) only a few minutes between one questionnaire being completed and another being started within the same household, and (v) excessive missing data rates.

We also pay special attention to households which have completed the maximum number of questionnaires (four in Community Life). From the development work, we know these questionnaires tend to have been completed more quickly than average (a median of 28 minutes rather than 38 minutes in one Community Life test), and that respondents also tend to select fewer than average items from multiple response lists. However, the missing data rates are average, as is the length of open-ended text, and there is no additional primacy effect that we can detect. For the most part, these completed questionnaires do not look particularly different from others, so we take the view that four completions from a single household do not *necessarily* mean proxy/careless completion in order to obtain a larger incentive. Nevertheless, to be on the safe side, we tend to discard these cases based on fewer lit indicators than are required to discard other cases.

It is an open question whether this combination of 'nudging' respondents to complete the questionnaire truthfully (and with care) together with an algorithmic method of post-fieldwork case-removal is *sufficient*, even if it is *proportionate* and this is certainly an area for further development. Nevertheless, the ABOS method is intended as a low-cost way of obtaining a random sample of the general population; some level of proportionality – some level of compromise – is necessary to ensure that the cost of data verification does not transform ABOS from a low-cost to a high-cost survey model.

Q3: How do you cover 'offline' individuals?

According to weighted Crime Survey of England & Wales data from 2012-15⁶, 17% of the adult population in those countries have either never used the internet or use it so infrequently that they are effectively not covered by an online survey method. This group is shrinking slowly over time, but more because of its demographic decline than a change in behaviour among the group. This group is particularly distinctive in birth cohort and educational level, tending to be older and, controlling for age, disproportionately likely to have no academic qualifications.

Although the size of the offline subpopulation is shrinking, excluding a highly distinctive 17% of the adult population is not acceptable for surveys that aspire to the status of official statistics (unless the survey topic is exclusively concerned with online behaviour). Consequently, ABOS studies need to cover offline subpopulations using an offline data-collection mode. We have experimented with offering paper questionnaires and telephone interviews on request, and have also used paper questionnaires more directly, including one or more in some reminder packs (see Q4).

A different approach – to be tested at scale in early 2017 – is to use a dual sample design in which a standard ABOS study is combined with a smaller interview study in which sampled households are screened for individuals who are either (i) aged 70+ or (ii) have not used the internet in the last year⁷. Our analysis suggests internet-using people in their 70s and 80s are not particularly well covered by the ABOS method, hence their inclusion in both samples. The next age group down – the internet-using 60-69s – is covered as well as any other age group.

Given the additional costs of a separate interview study, we recommend that researchers consider under-sampling the target population, and then applying larger-than-average design weights to the data, rather than seeking absolute proportionality. Either PAF-based in-person interviewing or dual-frame

⁶ It is more usual to use Ofcom survey data to estimate online and offline population sizes but the Ofcom survey is not large enough for a reliable estimation of the age/educational level distribution of the offline population.

⁷ An alternative is to use in-person interviewers to contact sampled households which have not responded to the ABOS study. However, this would enforce pre-clustering of the ABOS sample and (probably) some sub-sampling of non-respondents to control costs.

RDD telephone interviewing (if suitable) might be used for the 'offline' sample. The 'offline' and ABOS samples can then be combined for analysis purposes with weights to deal with the slight overlap in target populations (the c7% of the population who are aged 70+ *and* use the internet).

The choice between (i) using paper questionnaires as an alternative data collection mode within the ABOS sample and (ii) supplementing the ABOS sample with an 'offline' interview survey is largely determined by the complexity of the measurement objectives. Given a straightforward questionnaire, option (i) is a good choice. Firstly, it does not demand the complexity of a dual sample design; secondly, paper data collection is less costly than interview data collection; thirdly, there is plenty of evidence that paper and online questionnaires yield data with similar measurement characteristics (despite inevitable layout differences and the lack of control over the order in which a respondent completes a paper questionnaire).

However, paper cannot readily accommodate complex filtering, loop structures or any responsive pre-population of question and response texts. Simpler versions of the questionnaire might be produced to get around this problem but, in doing so, researchers accept offline non-coverage for the parts of the questionnaire not reproduced on paper. If such structural complexity is necessary, a separate interview survey is the only alternative⁸ despite the additional costs, additional design complexity and the occasional risk to inference of combining interview data with online data.

Q4: What response rate does the ABOS method get, and what is the impact of the design features you have tested?

The calculation of an ABOS response rate is only approximate but we can estimate it by assuming that c8% of sampled addresses will not contain a household, and that an average of c1.9 adults will be resident in each household. These estimates are robust, derived from the Census and from contemporary random sample interview surveys. For the current version of the ABOS Community Life Survey, which asks all resident adults to complete a questionnaire, it is a simple task to divide the number of validated completed questionnaires by this estimated denominator. For variants that seek just one adult respondent for each sampled household, the denominator is simply the estimated number of households. For both variants, we exclude 'rejected' completed questionnaires and partially completed questionnaires from the numerator.

The specific combination of ABOS design features, plus the identity of the sponsor and/or topic of the survey, appears to make a significant difference to the response rate. In 2015-16, the response rate for the Cabinet Office Community Life Survey was 24% but in a contemporaneous survey for a different sponsor (a 'third sector' organisation which must remain anonymous for now), the response rate was only 9%, averaged across experimental conditions. Given this observed variation in response rates, for each new ABOS study we strongly recommend a pilot or a 'soft launch' phase to establish the likely response rate so that the cost for each completed questionnaire can be estimated precisely.

Over the years, we have tested many different design features in an effort either to boost the response rate or to reduce costs. From this we know that:

- (i) Conditional incentives increase the response rate, albeit not in a linear fashion and with some accompanying increase in costs
- (ii) Sending a reminder can almost double the response rate without increasing the cost for each completed questionnaire;
- (iii) Sending a second reminder has half the effect of the first reminder, and thus increases the cost for each completed questionnaire, but if this reminder includes one or more paper questionnaires the impact can be greater, and it can also alter the responding sample profile (and not just through including the offline population). These qualities make it a useful tool for manipulating sample composition as well as for increasing the response rate.

⁸ An alternative is to make the interview method available 'on request' to the ABOS sample. However, in our experience, very few people will contact the research agency to arrange such an interview. Consequently, coverage of the offline population is no more than nominal if this approach is taken.

Beyond these general findings, we have some evidence from specific ABOS studies that may prove to be generalisable to other ABOS studies. One is that sending a vivid 'survey promotion' postcard (without login details) just before one of the letters can cost-effectively prompt people to take part when the detailed letter arrives. This reflects findings from name-based postal surveys in the UK and elsewhere (see Dillman et al (2014) for a thorough review). The second finding is that the identity of the sponsor can have an impact *even with an otherwise identical survey offer*. Combining this evidence with the observed variation in response rate between different ABOS studies, we conclude that sponsors with little name recognition should (if possible) link up with a partner organisation that can lend to the study greater name recognition or reputation.

Q5: How does the response rate vary between subpopulations, and what (if anything) can you do about it?

The postcode address file is itself a 'bare' sample frame but neighbourhood-level data can be attached via the postcode, allowing response rates to be estimated for different strata. Beyond this, we can compare the gender, age and regional profile of ABOS responding samples against the relevant ONS mid-year population totals, allowing us to estimate response rates for post-strata defined by these characteristics. Furthermore, we can also compare an ABOS-responding sample against contemporary high response rate random samples to gauge relative bias on a wider range of characteristics⁹.

Although we have accumulated response information of this type across multiple ABOS studies, we are wary of over-generalising findings given the present small number of studies.

Stratum level response rates

One reasonably consistent feature is that the *online* response rate is inversely correlated with the local Index of Multiple Deprivation (IMD), available at Lower Layer Super Output Area (LSOA) level. No other variable that can be attached to the postcode address file appears to be as strong a predictor of response (although there is some evidence that two other variables have additional predictive value: (i) the census proportion living in flats and (ii) the Census proportion self-classifying into one of the black ethnic groups¹⁰).

This variation in stratum response rates can be reduced by selectively applying design features known to influence the response rate. For example, the incentive level can be varied between strata or, more subtly, the proportion of addresses which receive an incentive, receive a second reminder or receive a set of paper questionnaires can be varied between strata. At the time of writing, we have only employed this kind of responsive design for the Community Life Survey because we have more evidence about the impact of design features for this study than for any other ABOS study. Although preliminary evidence from other studies suggests some consistency in the additive effect of each of these design features, the evidence is not strong enough for us to determine a clear set of rules that would apply to *all* ABOS studies.

Socio-demographic profiles

Although we may manipulate the response rate using the design features described above, the impact of each design feature on the sample's demographic profile is much harder to detect. Furthermore, while in most ABOS studies we have included experiments to test the impact of one, two or even three design features, some features remain under-evaluated. For example, only the very first ABOS study (for the Cabinet Office) tested the impact of *not* offering an incentive. All of our ABOS studies since then have offered at least £5 in return for completing the questionnaire and most have offered £10.

As it stands, the only design feature that we *know* will change the sample profile is the inclusion or otherwise of paper questionnaires in the second reminder. However, even here, all we can say for certain is that paper questionnaires bring in more people aged 60+ and especially those aged 75+ (who will

⁹ Both of these 'indirect' methods of estimating bias are reliant on an assumption of measurement equivalence between the ABOS study and the benchmark study. This assumption should hold sufficiently well for us to safely assess sample balance for demographics but not for opinions or (in some cases) behaviours.

¹⁰ Both are *negatively* correlated with ABOS response probabilities.

otherwise take part in very low numbers). In addition, there are indications that paper questionnaires help bring in people aged under 60 who have long-term illnesses or disabilities and/or live in social rented accommodation.

Given our lack of robust evidence about the impact of specific design features, it is perhaps most instructive to simply compare the demographic profiles of ABOS studies with different sponsors, topics and design features to see if we can identify any consistent outcomes. For illustration, we show data from three 2016 ABOS studies, two of which must remain anonymous for now. Table 1 provides the demographic profiles for each of these studies, although slightly different demographic data was collected for each study, hence some cells are empty.

The two surveys with a substantial number of paper returns (the Community Life Survey and the anonymous survey 2, also for a government department) have very similar demographic profiles, despite the Community Life Survey having a much higher response rate than survey 2 (24% compared to 15%). Across the four common dimensions (gender, age group, working status and ethnic group), the mean absolute marginal error for each category is 3.9 percentage points (%pts) for the Community Life Survey and 2.9%pts for survey 2. However, survey 2 has no data for highest educational qualification or for housing tenure, variables where we can expect higher error scores. If these variables are included, the Community Life Survey's mean absolute marginal error for each category increases from 3.9%pts to 4.6%pts. Given the similarity of the two surveys for the common dimensions, it is reasonable to expect that the survey 2 sample is biased to a similar degree in these other (unmeasured) respects.

Survey 3 (for a 'third sector' organisation) has a distinctively different profile, largely because no paper questionnaires were included in the second reminder package. Consequently, paper completions make up only 2% of the total responding sample. For survey 3, the mean absolute marginal error for each category is a much higher 6.0%pts across all dimensions. In particular, the sample is younger than it should be and too highly educated. It shares these traits with the online-responding subsets of the Community Life Survey and of survey 2. However, the most notable bias is in the gender profile: only 37% of respondents were male. This gender bias is not found in the online-responding subsets of the Community Life Survey (47% male) or of survey 2 (49% male) so it has nothing to do with the almost online-only nature of survey 3.

As a point of comparison, the 2015-16 Community Life in-person interview survey has a mean absolute marginal error for each category of just 1.4%pts. Clearly, the ABOS method produces a less accurate demographic profile than the face-to-face interview method but that is to be expected given its lower response rate. In fact, the accuracy of the ABOS profile is similar (albeit with a different error distribution) to that of a mid-fieldwork in-person interview survey after two or three visits to each sampled address. The accuracy of the profile is also similar to that of contemporary dual-frame RDD surveys, for which 2-5 percentage points of error for each category is typical.

Table 1: Demographic profiles for three 2016 ABOS studies plus a contemporary benchmark

Variable	Survey 1: Community Life, 2015-16 (ABOS version)	ABOS Survey 2	ABOS Survey 3	Post-stratified benchmark survey estimate (Community Life, 2015-16, in-person interview version)
% Responding online	74%	77%	98%	
Responding sample size	3,016	1,170	968	3,027
Response rate	24%	15%	9%	61%
Gender				
* Male	46.1%	48.4%	37.2%	48.8%
* Female	53.9%	51.6%	62.8%	51.2%
Age group				
* 16-24	8.5%	8.4%	15.2%	14.4%
* 25-34	12.9%	15.8%	16.4%	16.8%
* 35-44	14.0%	15.2%	13.8%	16.8%
* 45-54	18.1%	15.8%	17.8%	17.2%
* 55-64	18.2%	17.5%	18.4%	14.0%
* 65-74	18.3%	18.0%	13.5%	11.2%
* 75+	9.9%	9.4%	4.9%	9.6%
Working status				
* Working	54.2%	56.5%	52.4%	57.8%
* Not working	45.8%	43.5%	47.6%	42.2%
Ethnic group				
* White British	86.2%	86.7%*	n/a	79.6%
* Other	13.8%	13.3%	n/a	20.4%
Highest qualification (if aged <70)				
* Degree or higher	33.2%	n/a	41.4%	28.3%
* Other qualification	59.0%	n/a	46.7%	57.9%
* No qualifications	7.8%	n/a	11.9%	13.8%
Housing tenure				
* Outright ownership	37.7%	n/a	32.4%	25.5%
* Mortgaged	27.5%	n/a	27.4%	39.3%
* Renting/other	34.8%	n/a	40.3%	35.2%

For all of these variables, benchmark values are available so marginal calibration methods can reduce category-level errors to approximately 0.0%pts so long as a sufficient number of responding cases are present for each category. At the time of writing both the Community Life Survey and survey 3 responding samples have been calibrated in this way, although not to exactly the same set of marginal totals¹¹. In both cases, the design effect due to calibration is modest: 1.30 for the Community Life Survey and 1.46 for survey 3. However, cell level errors may persist even after marginal calibration, as will non-response errors that are uncorrelated with the variables used in the calibration procedure. Inevitably, the question remains: what level of non-response bias can we expect *after* calibration?

Q6: What evidence do you have for non-response bias?

Non-response bias can be identified with some confidence for a sample's demographic profile but with much less confidence for the substantive data, since benchmarks are usually unavailable. Our evidence for this is rather limited, but in one of our development phases for the Community Life Survey, we were able to shed some light on this, at least for that particular study.

For several years (2012-16), at least one ABOS variant of the Community Life Survey ran alongside the standard in-person interview survey used to produce official statistics. The two designs produced significantly different results even when the samples were weighted to the same population parameters. The question that arose was this: was the difference in results due primarily to (i) *measurement effects* related to the two different modes of data collection (online and paper self-completion questionnaires rather than in-person interviews) or (ii) residual *selection effects*, despite weighting the two samples to the same population parameters? To answer this question, it was vital to disentangle selection and measurement effects in order to determine which had the strongest influence on the results.

Williams (2015) describes the investigation in detail elsewhere but, in summary, the evidence suggested that the difference in data collection mode (that is measurement effects) was responsible for the bulk of the mismatch observed between the results. Selection effects appeared to be small in comparison.

Naturally, the study has some limiting assumptions, and there are questions it could not answer. For example, sample-size constraints limited analysis to total population estimates only. Findings might be different if sub-groups were assessed separately so it is possible that selection effects are meaningful for some parts of the population even if not in aggregate for the total population. Absence of evidence for selection effects does not imply that none exist.

Chart 1 plots the estimated measurement effects against what we might call 'system effects': the difference in results between the ABOS version of the Community Life Survey and the contemporary in-person interview version of the survey. The correlation between the estimated measurement effects and these system effects was very strong ($R = .86$), leaving only a small amount of residual variance that might be explained by selection effects. Furthermore, as table 2 shows, the distribution of estimated measurement effects, in magnitude, almost exactly matched the distribution of the observed system effects.

¹¹ *Community Life*: gender/age, education/age, housing tenure, region, household size and ethnic group; Survey 3: gender, age, age/working status, education/age and housing tenure.

Chart 1: Measurement effects (called 'mode effects' here) plotted against system effects ('web - F2F') in the Community Life Survey

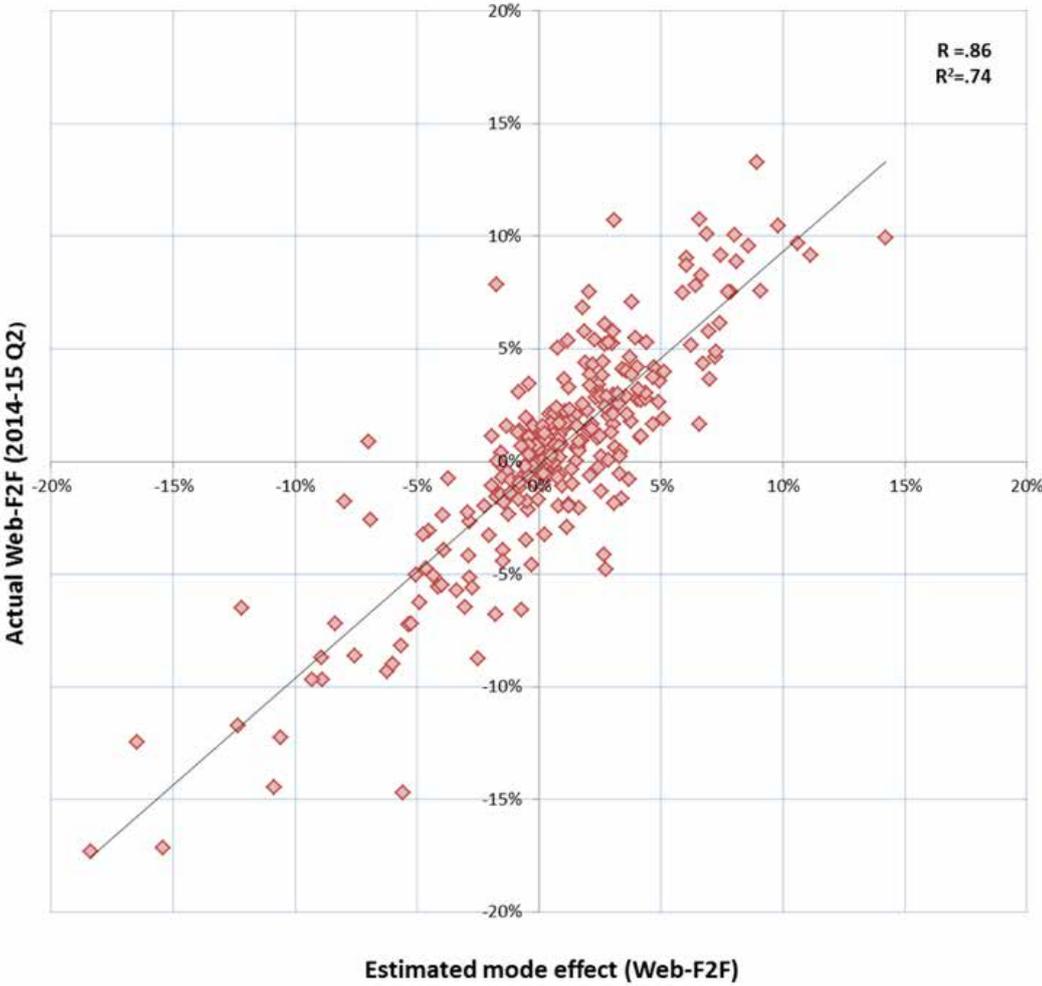


Table 2: Aggregated analysis of estimated measurement effects (online/paper self-completion vs in-person interview) against observed differences between ABOS and in-person interview survey results (July-September 2014)

	Observed difference between data collection systems (July-Sept 2014)	Estimated measurement effect
Mean absolute difference	3.4%pts	3.0%pts
Median absolute difference	2.1%pts	2.1%pts
% of differences <1%pts	26%	32%
% of differences <2%pts	48%	49%
% of differences <3%pts	60%	63%
% of differences <4%pts	69%	74%
% of differences <5%pts	75%	82%
% of differences 5%pts+	25%	18%
% of differences 10%pts+	5%	4%
% of differences that are statistically significant (null expectation = 5%)	40%	38%

The same kind of method was also used to disentangle selection and measurement effects causing differences between the RDD telephone interview and ABOS versions of Sport England's Active People Survey. That study also found that measurement effects were stronger than selection effects but concluded that modest selection effects were probably still present in the data. However, in this case, the benchmark – an RDD landline-only telephone-interview survey – was not of gold-standard quality, so the presence of selection effects was not the cause for concern it would have been had the benchmark been an in-person interview survey.

Two studies are by no means enough to make general conclusions about the nature of ABOS samples. Although it seems fair to say that selection effects are minimal within the ABOS Community Life Survey (once it has been calibrated to population parameters), it does not follow that they will be minimal for *all* ABOS studies. Nevertheless, it seems fair to say that the relatively low response rates obtained from ABOS studies are not necessarily indicative of strong selection effects. This conclusion aligns with studies of in-person interview surveys which have demonstrated high levels of convergence between estimates based on early data (when the response rate was low) and estimates based on final data. See for example Williams, Sturgis, Brunton-Smith and Moore (2016).

Q7: How much does it cost?

ABOS is primarily intended as an alternative to RDD telephone interviewing. So far, we have two examples of 'parallel runs' and have found the cost for each completed ABOS questionnaire to be roughly 60-80% of the cost of a same-survey dual-frame RDD telephone interview. Naturally, the specific combination of design features adopted will influence this cost ratio.

Conclusion

Although the ABOS method has its antecedents in the US, it is a relatively new method for UK survey research, and the details will undoubtedly be refined over the next few years. It appears to obtain reasonably balanced samples at response rates that are similar to those achieved with RDD telephone surveying. Selection effects seem modest – where we have been able to estimate them – but we do not yet have enough evidence to make a general statement about the relative robustness of the method compared to the gold standard of in-person interview surveys. Nevertheless, there is enough positive news to continue developing this as a genuine option for survey research studies.

References

Dillman, D. A., Smyth, J. D. and Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method*. Hoboken, NJ: John Wiley & Sons.

Messer, B. (2012). *Pushing households to the web: experiments of a 'web+mail' methodology for conducting general public surveys*. PhD diss. Washington State University. Available at: <https://research.libraries.wsu.edu/xmlui/handle/2376/4653> [Accessed 19/12/16]

Park, A. and Humphrey, A. (2014). *Mixed-mode surveys of the general population: results from the European Social Survey mixed-mode experiment*. NatCen Social Research. Available at: www.natcen.ac.uk/media/541183/ess-mixed-mode-natcen-report.pdf [Accessed 19/12/16]

Williams, J. (2014). *Community Life Survey: investigating the feasibility of sampling all adults in the household*. Cabinet Office. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/466925/The_Community_Life_Survey_Investigating_the_Feasibility_of_Sampling_All_Adults_in_the_Household_FINAL.pdf [Accessed 19/12/16]

Williams, J. (2015). *Community Life Survey: disentangling sample and mode effects*. Cabinet Office. Available at: www.gov.uk/government/uploads/system/uploads/attachment_data/file/466922/The_Community_Life_Survey_Disentangling_Sample_and_Mode_Effects_FINAL.pdf [Accessed 19/12/16]

Williams, J., Sturgis, P., Brunton-Smith, I. and Moore, J. (2016). *Fieldwork effort, response rate, and the distribution of survey outcomes: a multi-level meta-analysis*. NCRM working paper. Available at: eprints.ncrm.ac.uk/3771/ [Accessed 19/12/16]

Adapting survey design for smartphones: lessons from usability testing and survey implementation

Tim Hanson and Peter Matthews, Kantar Public

Abstract

A growing proportion of people are choosing to complete online social surveys using smartphones. As ownership and use of smartphones continues to grow, it is crucial that we adapt survey design to enable people to complete questionnaires on their device of choice, without having a negative effect on the respondent's experience.

This paper draws on data from a range of UK social surveys to show how device use affects people's experience and behaviour when taking part in surveys. Results from usability testing on Understanding Society, in which respondents completed the survey on smartphones, allow us to illustrate the questionnaire design challenges associated with completing social surveys on mobile devices. Based on this testing, and other evidence from social research, we highlight core principles to improve the experience of completing surveys on smartphones; draw conclusions based on our work to date; and consider possible next steps.

Introduction and background

There has been a steep rise in ownership and use of smartphones in recent years. According to Ofcom's 2016 Communications Market Report, 71% of UK adults owned a smartphone in 2016, almost double the proportion in 2012 (39%). Levels of smartphone ownership were particularly high in younger age groups; nine in ten for 16 to 24-year-olds (90%) and 25 to 34-year-olds (91%).

The importance of smartphones also continues to grow. In 2016, 36% cited their smartphone as their most important device for accessing the internet, compared with 29% selecting a laptop, 20% a tablet and 12% a desktop PC. This represents a substantial change over the last two years: in 2014, 40% of UK adults selected a laptop as their most important internet enabled device, almost twice the proportion selecting a smartphone (23%). Smartphones are seen as especially important among 16 to 24 and 25 to 34-year-olds, with more than half of those in each age group citing these as their most important device for accessing the internet (in both cases, 56%).

The increase in ownership and stated importance of smartphones is matched by high levels of use. Based on Ofcom's 'Digital Day' study in 2016, UK adults spent an average of 93 minutes a day using their smartphones. For 16 to 24-year-olds, the figure was 287 minutes – almost five hours a day spent using a smartphone.

All of this data points to an increasing reliance on smartphones for a wide range of activities. This is particularly the case for young people, who are far more likely to use a smartphone for online activity than any other device.

This clearly has implications for survey researchers, especially in an environment where many of our most important surveys have already moved online, or may do so over the next few years. This includes the UK's largest social research study, Understanding Society, which has included a large online component since the start of 2016. The Community Life Survey, an Official Statistic commissioned by the Cabinet Office, has recently adopted a solely online and postal method. In addition, the Office for National Statistics is currently investigating the feasibility of adopting online approaches for some of its surveys as part of its Data Collection Transformation Programme. And a predominantly online method is planned for the 2021 UK Census.

Survey respondents increasingly expect to be able to complete surveys on any internet-enabled device they choose, including their smartphones. Furthermore, the general importance of smartphones is heightened among young people, who are less likely to participate in social research studies. To encourage the participation of young people – or at least stop their level of participation from dropping further – we need to make surveys accessible on smartphones, reflecting how young people use the internet now.

This article highlights some issues to consider when designing new surveys, or adapting existing surveys, which may be completed on smartphones. It looks at current activity in social research studies and reports on a programme of usability testing using smartphones.

Our own analysis and the wider literature suggest that there are few distinct issues associated with tablet devices, and that they perform similarly to desktop PCs and laptops (for example Lugtig and Toepoel, 2015). As a result, while this article refers in places to mobile devices – which can also include tablets – it focuses on smartphones.

Device choice on social surveys: current practice

While large-scale social surveys are increasingly being moved online, the focus has generally been on making the questionnaire accessible for PCs or laptops rather than smartphones. As a result, researchers have often been more comfortable with surveys being completed on PCs or laptops, and attempts have been made to limit the inclusion of mobile devices. This is typically done by encouraging respondents in survey invitations not to use smartphones, and in some cases smartphones have been blocked from accessing surveys.

There are broadly two areas of concern about smartphones. First, it is generally not feasible to test a survey across the whole range of mobile devices which could be used ahead of fieldwork. There are, therefore, concerns that some respondents might experience technical problems, for example questions not appearing properly, which could make data unobtainable, unusable or unreliable. Second, there are concerns that respondents might answer differently by smartphone, and that data obtained might be of lower quality than that from other devices (for example Struminskaya et al, 2015).

The first concern is alleviated to some extent by modern survey software which generally supports mobile optimisation. At a basic level, this means that the appearance of questions will be automatically resized to fit the different screen sizes, reducing the need for zooming or scrolling. However, appearance will still vary from one mobile device to another. In addition, any non-standard question designs are unlikely to be fully supported and require more extensive testing.

Various studies have addressed the second concern about quality of data obtained, with the results somewhat mixed. Broad conclusions are:

- Surveys typically take longer to complete on a smartphone than on other devices (for example Andreadis, 2015)
- Respondents using a smartphone are significantly more likely to drop out of a survey before completing it (for example Matletova and Couper, 2015)
- Respondents using a smartphone tend to give less detailed responses to open-ended questions (for example Struminskaya et al., 2015)

Other data quality indicators such as straight-lining¹ the number of items selected at multi-code questions, and rates of non-substantive answers² show mixed results, with further analysis needed.

Researchers' concerns about smartphones, however, need to be balanced against the likely expectations of respondents that anything available online should be accessible by smartphone. As smartphones become the 'most important' device for internet access, it will become increasingly untenable to block or even discourage them.

Table 1 shows the proportion of respondents using smartphones in various online random probability social surveys conducted by Kantar Public. The level of smartphone use varied substantially, although in all but one of these surveys, respondents were advised to use PCs or laptops rather than smartphones.

The levels shown in Table 1 suggest that respondents are now completing long and complex surveys on smartphones. For example, so far 14% of respondents who have completed the Understanding Society Wave 8 survey online have used a smartphone. This survey includes a 40-minute individual interview; an additional household element; and covers a varied set of often quite detailed questions, so it is a fairly demanding task.

The final two studies in the table are of particular interest because they were targeted at young people and both saw a much higher rate of smartphone completion. The Wellcome Trust Science Education Tracker covered 14-to-18-year-olds, with 25% of respondents completing by smartphone. A similar proportion (27%) of 16-to-17-year-olds at Wave 4 of the 2nd cohort of the Longitudinal Study of Young People in England used a smartphone. This was despite being advised to use a larger-screen device in their survey invitation. In conjunction with the Ofcom results above, it seems reasonable to conclude that smartphone accessibility is especially important for online studies involving young people.

¹ Straight-lining refers to cases when respondents provide the same response option to all items in a battery of questions. For example, they may select 'Strongly agree' to a series of attitudes statements, even where responses may be contradictory. Straight-lining can be seen as a sign of low cognitive effort.

² That is 'Don't know', 'Refused', or skipping items (if the survey allows this).

Table 1: Smartphone use for online social surveys

Survey	Survey type	Year	% using smartphone	Base	Contact method	Approx. interview length	Device guidance
Community Life	Cross-sectional, address based sampling	2014-15	3%	2,325	Letter	30 mins	Advised to use PC, laptop or tablet
Understanding Society Innovation Panel W8 ³	Longitudinal	2015	5%	776	Email/letter	45 mins	Advised to use PC, laptop or tablet
Understanding Society Innovation Panel W9	Longitudinal	2016	9%	1,103	Email/letter	45 mins	Advised to use PC, laptop or tablet
Understanding Society W8 ⁴	Longitudinal	2016	14%	5,699	Email/letter	45 mins	Advised not to use mobile device
Longitudinal Study of Young People in England 2 (Wave 4)	Longitudinal; 16-to-17-year-olds	2016	27%	2,851	Email/letter	25 mins	Advised to use PC, laptop or tablet
Wellcome Trust Science Education Tracker	Cross sectional; named sample; 14-to-18-year-olds	2016	25%	4,081	Letter	20 mins	Told they can use any device

³ For the Understanding Society Innovation Panel, smartphones were blocked until Wave 8. In this way, respondents may have been conditioned against using smartphones, which might explain the lower smartphone use compared to the main Understanding Survey.

⁴ Based on interim data.

Understanding the experience of completing surveys by smartphone

In this section, we report on work on understanding the respondent experience of completing social surveys by smartphone. We use the example of Understanding Society as a large study with a lengthy questionnaire, which was originally designed for interviewer administration, but which now offers online completion, alongside face-to-face and telephone.

When the Understanding Society Innovation Panel first opened up to mobile devices, the questionnaire asked respondents about their experience of completing the survey. Respondents were asked to provide feedback on various aspects of the survey including choice of device, ease of completion and any technical problems they experienced (Table 2).

While this was the first time mobile devices were permitted in the study, nothing was done to make the survey more accessible for smartphones beyond the basic automatic resizing of the screen for different devices. The intention here was to see how user-friendly the standard script was when completed on a smartphone.

Although the number of respondents using a smartphone was small, the findings were very positive. Almost all of those using a smartphone said they found the survey 'easy' or 'somewhat easy' to complete, and that they would use the same device if completing the survey again. These findings indicated that there were no major difficulties or technical obstacles for completing by smartphone.

Despite these being 'mobile' devices, most respondents using a smartphone completed the survey at home.

Table 2: Survey experience questions from Understanding Society Innovation Panel 8

Survey experience question	Responses of those answering by smartphone
How easy or difficult was it to complete the survey on this device?	37 out of 40 answered 'easy' or 'somewhat easy'
If you were to complete this survey online again, would you use the same type of device or would you choose to use a different type of device?	37 out of 40 said they would use the same type of device
Where were you while completing this survey? Please select all that apply.	36 out of 40 answered 'at home'

In addition to these 'survey experience' questions, we have conducted two stages of usability testing to assess the experience of completing the Understanding Society Survey on smartphones.

Usability testing is similar to cognitive testing in that it is a questionnaire pre-testing method of assessing and addressing survey design issues. However, rather than focusing on participants' interpretation of question wording, it focuses on the survey instrument and how 'user-friendly' it is. Respondents were asked to complete a slightly abridged version of the Understanding Society questionnaire on their smartphones. A researcher was present with respondents and encouraged them to discuss their experience as they completed the survey. Respondents were asked to comment on any aspects of the survey they found confusing, unclear or frustrating and anything they thought could be enhanced in the layout or interface of the questionnaire. Researchers also asked about points of interest, for example, where non-standard question formats were used.

For the first stage of testing (July 2015) eight respondents were interviewed. For the second stage of testing (October and November 2016) a total of 25 respondents were interviewed over three rounds of testing, with a small number of changes made to the questionnaire at each round to take account of respondents' feedback⁵. For both rounds of testing, we recruited respondents who regularly used their

⁵ The findings from the second stage of usability testing will be written up in a forthcoming publication as part of the Understanding Society Working Paper Series.

smartphones for online activity as we wanted to gain feedback from those who may be likely to choose to complete surveys on mobile devices.

In most cases, we found that respondents completed the survey with few problems, often commenting that the process was easy or straightforward. While some issues did arise, these tended to be minor and did not present substantial obstacles to respondents. However, there were clearly parts of the survey design which were not optimal for smartphones, and there is scope to improve the survey experience for those completing on these devices. In the next section, we discuss some of the issues which arose from our testing. We also draw together some key principles for adapting survey design for smartphones.

Key principles to improve the experience of completing surveys on mobile devices

Evidence from current surveys and our usability testing shows that respondents can complete fairly lengthy surveys on smartphones without major problems. However, the following principles should help improve the experience:

1. Researchers should minimise the length of online surveys to make the respondent's experience easier. The literature shows that, while results vary from survey to survey, respondents completing on smartphones are more likely to break-off midway through the survey (Mavletova and Couper, 2015). The risk of breaking off grows as survey length increases. Break-off rates can be reduced by designing surveys with smartphones in mind (including using these principles)
2. Researchers need to consider the length of questions and response lists when designing questions. In our usability testing, respondents commented negatively about the length of questions. There are often issues when surveys which were administered by interviewers are moved over to a partly or wholly online approach. Some questions include text which may not be necessary, such as preambles, qualifiers, lengthy definitions, repetition between question text and response codes; this may over-burden self-completion respondents. This does not only apply to smartphones – shorter questions are advisable across all modes and devices – but problems are likely to be heightened on small screens
3. The size of text is important. In our testing, some respondents commented that the text was too small and hard to read. Researchers should achieve a balance between a font size large enough for respondents to read easily while not making the font so large that extensive vertical scrolling is needed to read the question
4. Response buttons should be spaced sufficiently to reduce the risk of respondents selecting the wrong response code by mistake. Some respondents in our testing commented that response buttons could be too close together, leading to frustration and risk of miscoding
5. Researchers should consider the value of open questions, since the level of detail obtained from respondents using smartphones may be minimal. In our testing, we found that respondents were reluctant to provide much detail at open questions. This is consistent with other studies showing that those completing on smartphones tend to provide less detailed open responses (for example Struminskaya et al., 2015). We recommend including larger text boxes to encourage respondents to provide more detail and to make it easier for them to enter and review their responses. However, this will only partly get around the issues, and it would be sensible to minimise the amount of typing needed
6. The inclusion of 'non-standard' question formats should be treated with caution. We found problems with slider scales and look-up functions, which could appear differently across different devices. We recommend sticking to basic and standard question formats. There should be considerable testing across devices when using any non-standard formats
7. It is important to clearly distinguish between questions. When blocks of questions looked very similar (for example, a series of grid questions) respondents in our testing sometimes did not realise they had moved to a new question

8. Grid-format questions can be particularly challenging for those completing on smartphones. In our testing, grids appeared cluttered on a narrow smartphone screen, and frustrated respondents as they struggled to select responses. This is consistent with other work showing that 'traditional' grid formats do not work very well on smartphones (for example McClain and Crawford, 2013). Researchers should consider alternative formats
9. Researchers should think about the level of detail asked of respondents. Some respondents in our testing said they would struggle to provide some of the information we asked for, for example about their finances, if they were completing away from home. As noted, most people who complete on smartphones do this at home. However, if they are using a mobile device they may be completing on the move, so researchers need to think about what is reasonable to ask.
10. Researchers should make it clear to respondents that their responses will be saved, and that they do not need to complete surveys in one go. While our testing found some demand for progress bars, they are often not appropriate for social surveys that frequently include extensive filtering. Most surveys do auto-save responses, so if people stop, they can return later and start from the point they left. But researchers may need to make this more explicit to respondents, as some may not want to complete lengthy surveys on their phones in one go.

There are other points to consider, and issues will vary from survey to survey. These will be based on various factors, for example, whether the survey is longitudinal or cross-sectional, the survey topic, the population of interest and questionnaire length. However, the above principles are a checklist of common issues associated with designing and adapting surveys for smartphones.

Conclusions and next steps

There has been a huge recent increase in ownership and use of smartphones for online activities, and survey design needs to reflect this. We draw five key conclusions based on our work to date, before considering next steps.

1. Increasing numbers of respondents are choosing to complete lengthy and complex social surveys using smartphones, and it is likely this proportion will continue to rise. It therefore seems clear that all online surveys should now be designed with smartphones in mind. By blocking smartphones, or failing to optimise surveys for mobile devices, there is a clear risk that respondents will opt not to take part, or drop out midway through.
2. The work conducted to date is encouraging in showing that those respondents who choose to use smartphones are able to complete lengthy and complex surveys on these devices. While we have outlined a number of principles to improve the experience of completing surveys on smartphones, the issues to address do not appear to be insurmountable.
3. Usability testing plays an important role in identifying potential issues with survey design for smartphones, and is recommended when developing or adapting any online survey. Researchers should also test scripts across the most common types of smartphone prior to launching online surveys, with particular emphasis on any 'non-standard' question formats.
4. There are different levels of mobile optimisation. Most scripting software now 'optimises' to mobile devices but this may simply mean that questions resize to the width of screens, so that no horizontal scrolling is needed. Beyond this, there is a need to optimise question presentation, for example, adapting traditional grids to formats more suitable for smartphone screen. Questionnaire content may also need to be optimised, for example, reducing the length of questions and response lists. When researchers talk about mobile optimisation they need to be specific about what this means and the limitations.
5. Many of these issues are not just applicable to smartphones. The increased use of these devices for social surveys has encouraged researchers to focus more on best practice for designing online surveys, which is certainly a good thing. However, making questions clear, concise and clearly presented is something that should be considered across devices and survey modes.

This remains a new area, with further work required to understand how using different devices impacts on data quality, and what more should be done to optimise questionnaires for smartphones. There has been relatively little analysis of this issue in relation to UK social surveys to date, partly because levels of completions by smartphone have, until recently, been very low. However, now that more surveys, including Understanding Society, include online components, and more respondents are choosing to complete using smartphones, there are far greater opportunities for analysis.

Further experimentation is also required, to understand the impact of different forms of question presentation on response. We have recently conducted two new experiments, to assess the impact of alternative presentations of grid format questions, and to compare shorter and longer versions of questions. Other areas to investigate may include:

- Assessing the impact of device guidance in advance communication on device choice and on overall response rates;
- Testing the impact of splitting surveys into ‘chunks’, to be completed in stages, as compared with a single survey.

Research commissioners and practitioners should be encouraged to build experiments into the design of online surveys to provide robust measurement around these issues and to contribute to the evidence.

The increased use of smartphones, alongside the movement of social surveys to an online approach, means that the issues discussed above will become increasingly significant. This presents many opportunities for collaboration and shared learning across the social research community, as we collectively seek to effectively design and adapt our surveys for smartphones.

Acknowledgements

We would like to thank Alexander Wenz at the Institute for Social and Economic Research (ISER), based at the University of Essex, for his analysis of survey response and data quality across devices, which we have drawn on in this article.

We also thank the Understanding Society survey team, at the University of Essex, for allowing us to share early survey data and usability testing findings.

References

- Andreadis, I. (2015). ‘Comparison of response times between desktop and smartphone users’. In Toninelli, D., Pinter, R. and de Pedraza, P. (eds.) *Mobile research methods: opportunities and challenges of mobile research methodologies*. London: Ubiquity Press.
- Lugtig, P., and Toepoel, V. (2015). ‘The use of pcs, smartphones and tablets in a probability-based panel survey: effects on survey measurement error’. *Social Science Computer Review*. Available at: http://www.academia.edu/20076628/The_use_of_Pcs_smartphones_and_tablets_in_a_probability_based_panel_survey_Effects_on_survey_measurement_error. [Accessed 19/12/16]
- Mavletova, A., and Couper, M.P. (2015). ‘A meta-analysis of breakoff rates in mobile web surveys’. In Toninelli, D., Pinter, R., and de Pedraza, P. (eds.) *Mobile research methods: opportunities and challenges of mobile research methodologies*. London: Ubiquity Press.
- McClain, C., and Crawford, S.D. (2013). Grid formats, data quality and mobile device use: towards a questionnaire design approach. AAPOR conference, Boston, MA. Available at: http://www.aapor.org/AAPOR_Main/media/AnnualMeetingProceedings/2013/Session_F-4-4-McClain.pdf. [Accessed 19/12/16]
- Struminskaya, B., Weyandt, K., and Bosnjak, M. (2015). ‘The effects of questionnaire completion using mobile devices on data quality: evidence from a probability-based general population panel’. *Methods, Data Analyses* 9(2): 261-292.
- Ofcom (2016). The communications market report. Available at: <https://www.ofcom.org.uk/research-and-data/cmr/cmr16> [Accessed 19/12/16]

The Social Research Association (SRA)
24-32 Stephenson Way
London NW1 2HX

0207 998 0304
admin@the-sra.org.uk
www.the-sra.org.uk

 [@TheSRAOrg](https://twitter.com/TheSRAOrg)

the-sra.org.uk/journal-social-research-practice