



GenPopWeb2: Transitioning from Interviewer-Administered Surveys to Online Data Collection: Experiences, Challenges and Opportunities

Final Report

Olga Maslovskaya (University of Southampton), Lisa Calderwood (CLS, UCL), Gerry Nicolaas (NatCen) and Laura Wilson (ONS)

27 May 2022

Introduction

[GenPopWeb2](#) is a network of UK-based academic and non-academic partners including government departments, survey organisations, academics and major ESRC investments to share knowledge and collaborate in the area of online data collection in social surveys as well as in setting the research agenda in the field. The Principal Investigator is Olga Maslovskaya (University of Southampton). The Co-Investigators are Lisa Calderwood (UCL), Gerry Nicolaas (NatGen) and Laura Wilson (ONS). The network activities were funded by the ESRC via the project “Transitioning from Interviewer-Administered Surveys to Online Data Collection: Experiences, Challenges and Opportunities” (ES/V001051/1). The project had an advisory group which was overseeing the network’s activities (the list of members of the advisory group can be found in Appendix 1).

In 2013 the First GenPopWeb network ran a series of events to bring together international experts to establish the lay of the land at that time regarding online data collection in social surveys. The network also identified new research questions and potential areas for collaboration across sectors. This network was highly effective in focusing the minds of researchers in this field and empowered them to direct research within their organisations (Nicolaas et al., 2014; Wilson and Maslovskaya, 2019). Since 2013, there has been a lot of research and new developments in this field, and the trend towards increasing use of online data collection has continued.

Preceding the GenPopWeb2 network, a one-day international conference on “The Future of Online Data Collection in Social Surveys: Shared Learning on the Challenges, Opportunities and Best Practice” was held at the University of Southampton in June 2019. The event aimed to bring together survey researchers and practitioners from across sectors and key survey organisations, to share ideas and experiences in the area of online data collection, to identify future research priorities in the field, and to discuss gaps in the literature. The conference was jointly organised by the University of Southampton (as part of the ESRC-funded project “Understanding survey response behaviour in a digital age: Mixed-device online surveys and mobile device use” (ES/P010172/1) and the Office for National Statistics (ONS) (Wilson and Maslovskaya, 2019). The event was well received and attended by 78 international researchers from 14 countries from academic, governmental organisations, national statistical institutes, and private survey organisations. The following topics were identified at the time as high priority areas for future research: questionnaire design for online surveys, survey budgets, sampling with special focus on address-based sampling in the UK, data linkage, response rates, digital divide, recruitment of respondents, representativeness and sample compositions, innovations, and use of new forms of data collection such as data obtained via sensors and apps, measurement errors, complex measurements, mode effects and time series (Wilson and Maslovskaya, 2019). This conference also identified a need for further evidence on how to utilise the opportunities that technological change offers and how to ensure that online data collection is effective, efficient, and results in high quality data which can be confidently used for important policy, financial and other decisions. It was also highlighted that activities of the GenPopWeb network should be reinstated.

Main Aims

The GenPopWeb2 network was created in 2020 and the main aim was to address challenges and gaps in knowledge which are of crucial importance for transitioning to online data collection in the UK and to enable sharing knowledge across industries. The network’s main principle is knowledge exchange. To exchange knowledge, various international activities

were organised between February 2020 and September 2021. Three main overarching themes addressed by the network were:

1. **Sampling and Participation** with a particular focus on barriers to transitioning to online data collection for cross-sectional surveys. The main focus was on sampling frames and recruitment approaches for push-to-web surveys, as well as on hard-to-reach and off-line populations.
2. **Measurement Issues** with a particular focus on questionnaire design for mobile devices, including length and modularisation and complex measurements such as bio-measures, cognitive assessments, as well as data linkage consents.
3. **Adjustment Approaches** with a particular focus on accounting and adjusting for measurement differences due to device and mode effects at the analysis stage, and guidance for users, as well as on measurement comparability in the context of time series and longitudinal data.

Literature Reviews

Two **literature reviews** were commissioned. The topics for literature reviews are:

1. Recruitment and participation: Maximising participation in both cross-sectional and longitudinal contexts;
2. Utility of probability-based online surveys: Push-to-web and online panels.

The first literature review was conducted on **strategies to improve response rates in probability-based online surveys**. The review considered the evidence on the effectiveness of different approaches to response maximisation and looked at whether their impact on selection bias had been studied. The main findings were that the use of prenotifications, reminders and incentives have an overall positive impact on participation rates. However, the evidence is limited on the impact on bias of these strategies, and further research is needed in this area.

The second literature review focused on exploring the **utility of probability-based online surveys**, as well as on comparing the different dimensions in which probability-based online surveys vary from a data quality perspective. The main conclusions from the second literature review are below (the literature review will be available on GenPopWeb2 website shortly):

- 1) Although the internet penetration has been growing during the last years, it is recommended to make special provision for **offliners** than to exclude them. If possible, it is recommended to provide internet connection than to allow the use of offline modes, to avoid mode effects.
- 2) Although allowing to participate with **mobile devices** might introduce nonresponse errors, these seem to be offset by the potential coverage errors of excluding them. Hence, it is recommended that mobile devices are allowed for survey completion. Since optimising the survey design for mobile devices seems to reduce nonresponse and measurement errors, the mobile-first approach to survey design represents the best practice.
- 3) Recruiting off the back of another cross-sectional offline survey or an offline panel might be a feasible alternative comparable or even better than using a fresh sample. Existing research has shown that little extra bias is introduced when recruiting the panellist, mostly carrying away the bias of the offline surveys. Besides, using the rich set of information from the base survey or panel used to obtain the sample can help designing better adjustment strategies than those available for fresh samples.

- 4) Some variables have been linked to nonresponse in most of the literature reviewed, regardless of the one-time or panel nature, or the different recruiting strategies used. Better educated respondents have been found to be more likely to join online panels and/or to participate in all types of surveys. Evidence also shows that younger respondents are more likely to participate, in almost all the literature and contexts reviewed. Substantial evidence also shows that people with higher incomes are more likely to participate, and that men and non-white/non-native English speakers are less likely to participate. Therefore, specific strategies should be explored to tackle these differences in the likelihood of joining online panels and/or participating in online surveys.

Events

Seven **online open speaker events** and one **invited meeting of experts** were organised. All events were well-attended, and international. Slides for all events are available at GenPopWeb2 website. The events were approximately 2 hour long each with exception of the invited expert event which was 4 hour long:

- Invited expert event covered issues associated with **adjustments for mode effects**. A report of this event was also produced. 23 September 2020. 13 participants.
- One event discussed issues associated with **within-household selection methods for probability-based online surveys**. A report related to this topic was also produced and will be shortly available on GenPopWeb2 website. 5 November 2020. 3 speakers, 1 discussant and 144 participants.
- Three events were organised to address issues associated with **complex measurements in online surveys (cognition, data linkage consent and occupational coding)**. 26 February, 5 March, 12 March 2021. 3/4 speakers at each event. 90, 62, 60 participants respectively.
- Two events addressed issues of **designing and delivering online surveys (transitioning long questionnaires and online questionnaire design, development and testing)**. 14 July, 28 September 2021. 3 speakers at each event. 132 and 61 participants respectively.
- The last public event discussed issues associated with **covid-19 pandemic and transitioning to online data collection in social surveys**. 17 September 2021. 4 speakers. 102 participants.

The remainder of this summary report will discuss each event and will also summarise main recommendations for future research and implications for survey practice.

The first event which took place was the closed invited event for survey methodology experts which addressed issues associated with **adjustments for mode effects**. Many social surveys have moved to mixed-mode designs, often motivated by potential cost-savings, but these designs can potentially increase risk of mode effects. In the absence of access to experimental design data, it is very important to address issues of endogeneity and potential mode effects in mixed-mode surveys by isolating mode effects on measurement from selection effects and by making necessary statistical adjustments. Currently there is no agreed best practice guideline and the methods used for isolation and statistical adjustment for mode effects vary. It is very rare for surveys to provide guidance to users on this issue. The choice of an approach is often subject specific or survey specific, or mode effects are often not considered at all by data users during the analysis. This practice increases the risk of reporting unreliable results. Statistical adjustments are not easy to implement as they are currently very technically demanding. Preventing mode effects is, of course, ideal but even with appropriate questionnaire design, prevention is not always possible or is often not fully successful.

Moreover, it was noted that the aim of preventing mode effects may conflict with trying to achieve the ‘best’ measurement in each mode. The experts identified the methods which are available for statistical adjustment and highlighted limitations of each method. It was identified that there is a great need for further methodological research in the area. The complete list of areas for methodological research can be found in the report which was produced after the meeting (the report will be available on [GenPopWeb2](#) website shortly). It was mentioned that currently many data analysts who analyse mixed-mode surveys tend to ignore the problem of mode effects as there is no user-friendly approach to adjustment available for the researchers and data depositors rarely provide guidance on this. It was recommended that this practice should change, and improved guidance should be provided.

Within-Household Selection Methods for Probability Web Surveys event addressed issues associated with challenges for address-based push-to-web online and mixed-mode surveys. In the UK address-based samples are used for cross-sectional general population surveys as individual level sampling frames are not available, and as a result within-household selection of respondents is a key challenge. (Quasi) random methods of respondent selection are not used any longer because research has shown that postal instructions are not necessarily followed and a wrong selection is made in a large proportion of households. Instead, the most common methods are (1) asking all eligible adults to take part and (2) asking up to any two eligible adults to take part. However, these methods have their own challenges. This event, comprising three speakers and a discussant, covered the pros and cons of the current methods of choice in the UK, considered these alongside other methods, and identified areas for further research and experimentation. The discussion concluded that there are various methods of within-household selection available such as one adult, any two adults, all adults and various two-stage approaches (see below) but none are without error. It is still unclear which methods may be better than others and, therefore, further research is required based on the Total Survey Error Framework. Of particular importance is comparing the trade-off between selection effects and non-response error for the different methods. It is also important to take costs into account when different methods are considered in further research. Although the quasi-random next/last birthday selection technique is now rarely used in the UK for online surveys, the leading methodological experts agreed that there is value in revisiting this technique using the so-called “confirmation method” (Olson and Smyth, 2017). The “confirmation method” aims to reduce the risk of error in selection by asking the respondent to actively confirm that they are the target respondent according to the selection criteria. It is also important to further investigate the characteristics of household members choosing to take part when the “any two adults” method has been applied. Further research into various forms of a two-stage approach is needed, i.e. only one household member is initially asked to take part and, after collecting some household information (e.g., number of adults, full household roster), other household members are invited. There is some indication that this may improve address-level response but more evidence on within-household response and bias is required. And, finally, it is important to further explore the ways of how to improve the “hand-off”, i.e. questionnaire placement when the target respondent is someone other than the contact person in the household.

The next three events addressed issues associated with **complex measurements in online surveys**. Each of the events focused on different types of complex measurement: cognition, record linkage consents, and occupational coding. The events showcased current evidence and discussed how the challenges could be overcome in order to optimise the collection of these measures in online surveys.

Measuring Cognition in Online Surveys event addressed issues specifically associated with measuring cognition. ‘Gold-standard’ approaches to measuring cognition in surveys involve

interviewer-administered assessments, and adapting these measures, or identifying new measures, for online administration is a key challenge. The impact of mode on measuring cognition and consequently the data was discussed. Evidence was presented around the existence of mode differences in measurement of cognition, between online and interviewer administered modes, and also device differences comparing touch-screen and keyboards. Higher cognitive scores were observed in online mode in comparison to interviewer-administered modes. The challenges associated with the design of these measures in online context were discussed, and some examples of novel approaches to implementation including development of new measures and the use of gamification. It was also demonstrated that paradata on response times can be used to help understand cognitive performance. Further research is needed into whether calibration across modes is feasible, how to optimise online measurement of cognition, which cognition measures are best suited to online collection and how to develop new measures of cognition for online surveys. The use of paradata for measuring cognition is another potential area of further research.

Collecting Data Linkage Consents in Online Surveys event discussed issues associated with challenges in obtaining consent to data linkage in online surveys. Consent to data linkage, particularly to administrative data sources, is increasingly common in social surveys as linked data sources can provide additional research resources. Evidence was presented showing that consent rates vary by mode and are much lower in web mode in comparison to interviewer-administered modes data collection. It was reported that when respondents complete a survey online, they usually understand the linkage request less well, they are also more concerned about privacy/data security, and they process consent requests less carefully. It was further shown that providing additional information for web respondents and improving understanding of request does not seem to be helpful. However, trust was reported to be a key driver of consent for web respondents. A different presentation suggested that logical arguments for participation, clear explanation of the process, highlighting benefits and reassurances, trustworthy sponsors, relevance of research project and incentives help increase consent rates for data linkage.

The discussion focused on ways to boost consent rates in online surveys, and potential areas for future research related to this. These included the visual design and layout of the consent request, whether telephone call-backs for non-consenters could be implemented appropriately, or alternatively to investigate whether removing consents from online instrument and using a telephone follow up could work better. It was also suggested to explore whether replicating interviewers' function in online survey in a form of "live chats" over messaging service using an avatar or "chat bot" or asking respondents to phone in during the survey to talk them through consent process might help.

Occupational Coding in Online Surveys event discussed how the collection of occupational coding in online surveys could be designed to improve the quality of information the respondents provide. Occupation is usually coded by specialist office staff, using information collected by interviewers. Online surveys provide the opportunity for respondents to code their own occupation, or if office coding is used, it important that respondents enter sufficient information for office coders. Different approaches to reporting occupation have been tested: respondents were asked to code SOC/SIC using the same method as interviewer (dual coding), respondents were asked to provide free-text which was then coded in the office, respondents also were asked to do text-based search (free text with look-up). The results suggested that in the dual coding approach, about a third of respondents selected a different occupation to the interviewer. It is important to establish here whether the office coding provides the "gold standard" or whether respondents are the experts. Successful use of embedded look-up for coding differed by mode and were worse for face-to-face mode of data

collection when done by interviewers and they also varied by respondent/interviewer characteristics. It took longer for online respondents to complete occupational coding questions and they provided more text when compared to other modes. For cross-country surveys, the free text approach was not feasible as there were too many languages to consider. More work needs to be done in the area and there might be a need for a radical redesign approach to help respondents to provide the right information about occupations in online surveys. If the decision is to use online coding by the respondents, it is important to establish the optimal length of the list, how it is presented, and its functionalities and it should also use the language of respondents. There is a possibility that different instruments might need to be used in different contexts (different length and specificity depending on how the data will be used, sample size and other characteristics). It is important to explore approaches to measuring occupational coding which are specific to online data collection and not just to use the methods which have been developed for interviewer-administered surveys.

The next two events addressed **issues associated with designing and delivering online surveys.**

Transitioning Long Questionnaires to Online event focused on issues associated with long questionnaires. Radical redesign approach to moving long surveys online allowed to improve efficiencies, to design more inclusive surveys, to reduce respondent burden and to improve data quality. Evidence suggested that with a few caveats, it was possible to develop a self-completion version of a relatively long questionnaire (ESS) and that it included most of the content from the face-to-face survey (with limited adaptations) with “acceptable” response rates, sample composition and data quality. It was concluded that keeping surveys short is good for questionnaire design and good for data quality but it is possible to run online surveys which are longer than 20 minutes successfully. It is important to remember that good questionnaire design is crucial for the success. More research is needed in the area of measurement differences and implications of changes on time series. It would be very important to conduct parallel runs to assess these differences fully. More work needs to be done on questionnaire design and length of questionnaire to identify what is optimal for online mode and how this can be achieved successfully.

Online Questionnaire Design, Development and Testing event addressed issues of accessible survey design, online questionnaire testing during covid-19 pandemic and the respondent-centred survey design approach. The key elements of what makes a survey accessible were highlighted. The benefits and limitations of the remote questionnaire testing, which were implemented during the pandemic and are likely to become the default method beyond the pandemic, were presented and discussed. The importance of developing questionnaires based on respondent needs was demonstrated. It was highlighted that the design of the questionnaire content (i.e., question wording and flow) should not be based on any assumptions held by researchers or stakeholders, instead designs should be informed by insights gathered from research. Survey practitioners and researchers need to take on board accessible survey design and its application. It is important to continue to innovate questionnaire testing methods and retain those that worked well but it is also important to consider the limitations further. Researchers and survey practitioners are encouraged to trial and use respondent centred methods in survey design.

The final event **Covid-19 and Transitioning to Online Data Collection in Social Surveys** focused on how existing high quality surveys continued to collect data during the pandemic. The following two questions were addressed during the event: What are the barriers for transitioning to online data collection, especially for cross-sectional surveys? How would the learning from covid-19 experience shape the future of online and web-first data collection?

Despite some expectations that Covid-19 pandemic would speed up the process of transitioning to online data collection for some surveys, the reality was different. Only the surveys which moved to online data collection previously (mostly longitudinal surveys) or were preparing the move prior to the pandemic, were able to successfully move to collecting data online during the pandemic. The majority of cross-sectional surveys either paused data collection or moved to telephone mode where possible. The two main barriers to quickly transitioning to online data collection identified were concerns about inclusivity and adapting long complex questionnaires designed for interviewer-administration into respondent-friendly self-completion instruments. Low response rate in online surveys was another concern. It was concluded at the event that surveys will continue to transition to online data collection with time but only in appropriate contexts and there will still be a need for interviewer-administered surveys. Barriers to transitioning to online data collection that existed before the pandemic have not disappeared and these should be further investigated and understood better. There were different innovations (new combinations of contact and data collection modes) trialled during the pandemic but more evidence about resulted data quality and their effectiveness beyond the pandemic is required. It was also concluded that face-to-face interviewing is here to stay and will continue being used in social surveys although with necessary adaptations and changes which need to be investigated and understood.

The last event links closely to the new ESRC project “The impact of Covid-19 on survey data collection methods in the social sciences” which funds activities of Survey Data Collection Network (SDC-Net). This network addresses issues in survey data collection some of which existed prior to covid-19 pandemic and/or were accelerated by it such as changing role of interviewers and interview capacity, and the remit of this project goes beyond transitioning to online data collection and investigates the entire survey data collection landscape.

Recommendations for Further Research

Within-Household Selection

- Cost comparisons;
- The trade-off between selection effects and non-response error for the different methods of within-household selection;
- Revisit the next/last birthday selection technique using the “confirmation method” (Olson and Smyth, 2017);
- Examine the characteristics of household members choosing to take part when the “any two adults” method has been applied;
- Further research into two-stage approaches;
- Explore ways of improving the hand-off when the target respondent is someone other than the contact person in the household.

Complex Measurements in Online Surveys

- Assess calibration across modes and its feasibility for measuring cognition;
- Investigate optimisation of online measurement of cognition;
- Develop new measures of cognition in online surveys;
- Use of paradata for measuring cognition;
- Explore strategies on how to boost consent rates in online surveys further;
- Explore visual designs and layouts of the consent request;

- Explore the use of telephone follow up or “live chat” functions or “chat bots” for obtaining consent in online surveys;
- Apply radical redesign approach to help respondents to provide the right information about occupations in online surveys;
- Explore the optimal length of the list for occupations if respondents use online coding;
- Explore approaches to measuring occupational coding which are specific to online data collection and not just use the borrowed ones from interviewer-administered surveys.

Designing and Delivering Online Surveys

- Investigate measurement differences and implications of changes on time series;
- Conduct parallel runs to assess these differences fully;
- Explore optimal questionnaire design and length of questionnaire for online mode further;
- Continue to innovate questionnaire testing methods and retain those that worked well but also consider the limitations of those which did not work well;
- Trial and use respondent centred methods in survey design.

Impact of Covid-19 Pandemic on Online Surveys

- Investigate barriers to transitioning to online data collection that existed before the pandemic and have not disappeared;
- Gather evidence about data quality and effectiveness of different innovations which were trialled during the pandemic;
- Investigate necessary adaptations and changes for face-to-face mode of data collection.

Adjustments for Mode Effects (more details could be found in the report which will be available on GenPopWeb2 website shortly)

- Develop a definition of mode effects as well as a comprehensive framework of all aspects of mode effects;
- Develop new adjustment methods, going beyond currently used regression methods, for example, for multilevel and survival models, latent class models;
- Develop methods for estimating variances given adjustments, which are currently limited;
- Investigate what impact adjustments have on statistical outputs including on descriptive statistics and on particular variables;
- Develop quantitative measures, indicating what type of variables and to what extent different variables are susceptible to mode effects;
- Conduct simulations to test new mode effect adjustment approaches and to investigate nuances and subtleties of existing adjustment approaches;
- Produce user-friendly methods for adjustment for mode effects and user necessary user guides.

Recruitment and Participation in Online Surveys

- Gather evidence on the impact of use of prenotifications, reminders and incentives on bias.

References

- Nicolaas, G., Calderwood, L., Lynn, P., Roberts, C.** (2014) Web surveys for the general population: How, why and when? ESRC National Centre for Research Methods. Final Report. <http://eprints.ncrm.ac.uk/3309/3/GenPopWeb.pdf>
- Olson, K., & Smyth, J. D.** (2017). Within-Household Selection in Mail Surveys Explicit Questions Are Better Than Cover Letter Instructions. *Public Opinion Quarterly*, 81(3), 688-713.
- Wilson, L., Maslovskaya, O.** (2019) *A summary of the agreed challenges, opportunities and best practice for online data collection in the UK in 2019 and the future*. Final Conference Report. University of Southampton and ONS. <http://eprints.ncrm.ac.uk/4283/>.

Appendix 1

GenPopWeb2 Advisory Group

Colin Beavan-Seymour (ONS)
Gianmaria Bottoni (ESS)
Peter Cornick (NatCen)
Mick Couper (University of Michigan)
Gabriele Durrant (University of Southampton)
Rory Fitzgerald (ESS)
Annette Jackle (Essex University)
Curtis Jessop (NatCen)
Peter Lynn (Essex University)
Andrew Phelps (ONS)
Patten Smith (Ipsos MORI)
Luke Taylor (Kantar Public)
Joel Williams (Kantar Public)