

Response to ESRC: List of research priorities in the area of survey data collection (SDC) methods



16 May 2022

Document was prepared by the **Senior Leadership Group (SLG) of the ESRC-funded UK Survey Data Collection Network (SDC-Net)**: Prof Gabriele Durrant (University of Southampton, UoS, and Director of the National Centre for Research Methods, NCRM), Dr Olga Maslovskaya (University of Southampton, UoS), Gerry Nicolaas (National Centre for Social Research, NatCen), Laura Wilson (Office for National Statistics, ONS) and Prof Lisa Calderwood (Centre for Longitudinal Studies, CLS and the University College London).

Aim of this document:

This document has been prepared for consideration by the *ESRC Research Methods Advisory Group* (Meeting on 27th May 2022). The aim of the document is to provide information to the advisory group of key research priorities in the area of survey data collection methods. This will be used to inform a forthcoming research methods call by the ESRC later on in 2022.

Introduction and context

The Covid-19 pandemic had a significant impact on survey data collection methods, forcing survey agencies to stop collecting information via face-to-face (f2f) interviewing during lockdowns. For some studies this led to a rapid move to other modes of data collection instead, such as to telephone, where phone numbers of participants were available, and to online in some cases, e.g., where preparations to online data collection had already started prior to the Covid-19 pandemic. For other studies, data collection was paused throughout the pandemic and is now re-starting f2f. In addition, the pandemic was a catalyst for development of new, innovative data collection approaches, such as video-assisted personal interviewing (VAPI) and other Covid-secure contact approaches such as 'knock-to-nudge', where an interviewer reminds respondents about the survey and/or collects their phone numbers, but does not conduct an interview. The pandemic also presented challenges to the survey research industry around retention and recruitment of interviewers and their changing role. These changes have had significant impacts on all social surveys including ESRC's longitudinal and cross-sectional studies and major government repeat cross-sectional surveys. Given these significant changes and challenges, it is important to better understand the current survey data collection landscape in social surveys, specifically barriers to online data collection, the mixing of modes and mode effects, issues associated with the return of f2f interviewing, the role of interviewers, and the effects of the changes on data quality and analysis.

In response, the ESRC funded the **Survey Data Collection Network (SDC-Net)** (Dec 2021- Nov 2022) to enable study leadership teams, survey agencies, industry bodies and academics to get together with the aim of collating existing evidence, to share knowledge and good practice and to identify key research areas for the development of research methods in the area of SDC and the improvement of social surveys. (Title of the ESRC-grant: 'The impact of Covid-19 on survey data collection methods in the social sciences'.) The grant is part of the ESRC-funded National Centre for Research Methods (NCRM). The network currently comprises more than 60 members from across 18 institutions, including a wide range of academic and non-

academic partner institutions in the UK, for example, representatives from the University of Southampton, ONS, CLS/UCL, NatCen, Ipsos, Kantar, Understanding Society, Market Research Society (MRS), ESRC and various government departments, local authorities and learned societies (e.g. SRA) and others. The group includes a mixture of senior representatives and leading survey methods experts of these institutions (e.g., directors of methods), as well as those that are leaders on particular topics and operational areas (e.g., those that lead on video interviewing or fieldforce). Further leading experts or directors of institutions are invited to particular meetings of the group, depending on the topic and strategic focus (e.g., the Deputy Director of ONS, the Deputy Director of Office for Statistics Regulation, and the Director of NatCen were invited and presented at the last meeting in May). More specifically, the network identified priority areas for the work of the network and the wider research community in this area at the beginning of the grant, based on input from all network members. The network is led by the **Senior Leadership Group (SLG)** with the members listed above (they are also the investigators of the associated ESRC grant). Further information on the project and the network can be found here: <https://www.ncrm.ac.uk/research/SDC-Net/>

Request by the ESRC

The ESRC (Joanna Lake and Becky Shipman) contacted the SLG to ask them to identify research priority areas in survey data collection methods. They requested that the wider network would be consulted for feedback. The ESRC first announced ESRC's initial thinking and its working with the ONS to all network members during the wider group meeting on 2nd Feb 2022. The consultation, including a list of potential priority areas, was subsequently developed by the network senior leadership group, in response to the initial ESRC request that developed further over time.

Methodology

For the purpose of this consultation, the SLG identified a list of 10 priority areas, reflecting experiences from the group as leading experts in the field and taking account of the priority areas that the SDC-Net group had already identified at the start of the network. Subsequently, all network members were contacted per email on 21st April 2022 with a request to rank the list of priority areas from 1 to 10 (with 1 representing the top priority) (ranking was anonymously; reminder emails were sent; deadline was 3rd May 2022). In total, 27 responses were received from network members, which indicates a reasonable level of feedback (bearing in mind that a number of organisations have several representatives on the network and we suspect only one person or a small number of representatives per organisation responded). The (initial) results were shared with the ESRC and were briefly discussed in the most recent network meeting on 4th May 2022).

Such an approach also has limitations, and these include:

- Priority areas initially identified by the network were not intended for this purpose and may have been influenced by immediate needs of the survey community rather than mid- to longer-term objectives, which would benefit from ESRC funding;
- The list of priority areas is not mutually exclusive; they are naturally interlinked, and it is often not possible to address one without another;
- Questions that involve ranking items can be cognitively challenging for respondents, which may affect the responses. For example, such questions do not reveal why respondents choose to rank each item in the way they did, and responses do not provide insight into how much weight a respondent gives to each item. We tried to address this during our meeting discussion, but time was limited and responses were not forthcoming in the online meeting environment.

Priority areas: justification, discussion, and conclusion

The following provides a ranked list of interlinked priority areas (**Table 1**), indicating the ranking from 1-10 (with 1 representing the highest priority and 10 the lowest), and further details of the priority area with a **justification why further investment** is needed. The priority areas were also confirmed as important in a very recent ONS Survey Owners Group meeting (at senior (Director and Deputy Director) level).

In summary, the identified interlinked priority areas (PA) are (listed in order of priority as identified by the network; summary score indicated in parenthesis, derived as a simple summary of the response listings received, with the lowest number indicating highest priority):

1. Future of face-to-face survey data collection (75)
2. Investigating survey data quality (80)
3. Innovations in survey data collection (112)
4. Adjustment for mode effects (137)
5. Improved sampling frames for general population surveys (148)
6. Changing role of survey interviewers (150)
7. Complex measurements in online surveys (179)
8. Discontinuity/time series in repeat cross-sectional and longitudinal measurements (184)
9. Development of an inclusive data system across the whole data lifecycle (210)
10. Exploration of innovative methods to achieving this inclusive data system, including respondent centred design (212)

Conclusion

Survey and survey data continue to be of high importance in the social sciences and for policy making and decisions. The COVID-19 pandemic has highlighted just how important good and timely data are for public policy and society. The pandemic has had wide ranging implications and caused major challenges for surveys and survey research, in particular accelerating the **move to online data collection and the implementation of innovations**. It is clear, that we continue to need **high-quality surveys** to be able to inform policy decisions and to achieve strategic goals. It is now important to take stock and to take a whole system approach, with strong partnerships between academics, survey researchers and survey practitioners at the centre, to improve the inclusiveness of UK data and evidence.

The identified priority areas, in the area of survey data collection, are naturally **interlinked**. Specifically, f2f surveys (priority area 1) are closely linked to the role of interviewers (priority area 6) (and could be combined). Data quality (priority area 2) is overarching all topics and is central to all survey developments. Given that the future of surveys will focus most likely on mixed mode designs, with a significant element of online data collection, complex measurements in online surveys (priority area 7), adjustments of mixed mode effects (priority area 4) all need to be tackled (priority areas could be combined). The focus on inclusivity and a much better representation of all groups of society (linked to the government goal on inclusivity) (priority area 9) and a **respondent centred approach** (priority area 10) are fundamentally new thinking frameworks (priority area 10 could be included within priority area 3 'innovations in surveys').

An **integral and comprehensive approach** in survey research, design and data collection, is now needed and any future funding should address this. We advise to fund a coherent body of research in this area, that is **coordinated** (at least loosely), rather than a disparate collection of individual, small research grants. (The idea of a research centre bringing together key organisations and their experiences in the area has been suggested but funding as part of the preparations for this current call may not stretch to this.) One possibility could be that the current survey data collection network, with its senior leadership group, could take on the overarching coordination role, under the umbrella of the National Centre for Research Methods (NCRM).

Table 1: Ranked interlinked priority areas in the area of survey data collection

Final ranked order (based on SDC-Net consultation outcome) 1= highest priority 10= lowest	Priority area	Background and brief justification why further investment is needed
1	Future of face-to-face survey data collection	<p>High-quality survey data still often relies on face-to-face survey data collection, including many UK-based and ESRC funded surveys. The Covid-19 pandemic had a significant impact on survey data collection methods, forcing survey agencies to stop collecting information via face-to-face (f2f) interviewing during lockdowns. For a number of studies, data collection was moved to alternative forms of data collection (e.g., telephone) or paused throughout the pandemic, but they are now re-starting f2f. However, this is mostly with adaptations, such as knock-to-nudge approaches, or using complementary, innovative approaches such as VAPI (Video Assisted Personal Interviewing), all in a changed survey landscape (e.g., high-frequency data collection or much faster data collection environments). Another important impact had the changed role of the interviewer and the availability of the interviewer workforce more generally (with many interviewers having left the profession). To summarise, the f2f data collection landscape has changed significantly and a simple return to pre-pandemic environments is not possible. There is, hence, an important need to investigate the future of f2f data collection, both in terms of its adaptations and their consequences (eg. for survey data quality) and the role of f2f data collection more generally (its future justification, and its mixing with other modes such as online and telephone).</p> <p>This priority area is linked to priority area 6.</p>
2	Investigating survey data quality	<p>The results from survey data are widely used to inform important policy, economic, social and financial decisions in the UK and elsewhere, with wide ranging future (political) implications. There is hence an unchanged need for high-quality survey data. However, the wide-spread changes in data collection and the way surveys are run in (post/)pandemic times have had wide-ranging impacts on the quality of our data. In particular, our understanding of the quality of the resulting data throughout the life-cycle is currently limited. For example, (new) surveys have been implemented very quickly during pandemic times (e.g. the UK Covid-19 Infection Survey), where some surveys did not go through the usual stages of survey preparation, such as testing and piloting due to time constraints, and new innovative methods have been applied (e.g. knock-to-nudge, video interviewing etc), often without much, if any piloting (which was often not possible during the pandemic). This is in addition to shrinking research budgets that may allow data quality investigations. There are also changed user demands on surveys, such as more complex, granular, inclusive and timely data, and the (sometimes ad-hoc) mixing of modes - all in a context of further declining response rates. It is, hence,</p>

		<p>now time to thoroughly investigate the data quality of the resulting data, including, for example, development and trends in response rates, nonresponse bias and representativeness, measurement error of responses etc. Whilst often new methods are being (quickly) implemented, the theoretical underpinnings are often missing. Currently, there are big gaps in the literature on the data quality of new approaches to data collection, which were trialled and/or tested since the onset of the pandemic. This priority area may include experimental designs to be able to compare, for example, face to face and online (including VAPI) data collection, and also detailed assessments of new approaches. In summary, nd the survey landscape needs to work out a new 'normal' (expanding initial ideas from the 2021 Cathie Marsh Memorial Lecture; 'Back to normal or a new normal?').</p>
3	Innovations in survey data collection	<p>The wide ranging changes have also led to the exploration, development and adoptions of innovative, new methods and approaches being applied to survey data collection designs. Examples include, among others, video interviewing, use of electronic questionnaire devices (e.g. in the ESS), knock-to-nudge approaches, greater use of web and mixed-mode within surveys and respondent centred designs (see also priority area 10). Often approaches are applied heuristically without much theoretical grounding or even pilot testing. However, they are hoped to offer breakthroughs or improvements to designs. The consequences of such approaches are not well understood at present, e.g., what are the consequences of these methods on data quality? In-depth investigations of the resulting data and data quality following the application of innovations are urgently needed.</p> <p>Priority area 10 could be included in here.</p>
4	Adjustment for mode effects	<p>Nowadays mixing of modes is the norm in many surveys, which underlines the importance of mixed mode data collection approaches. Cost pressures, for example, lead to the advocacy of a web-first approach with other, alternative modes of data collection following after that. Also, part of a survey may be offered a particular mode for comparison purposes (e.g., f2f). The mixing of modes needs careful design considerations as well as careful considerations on how best to adjust the resulting data, to ensure coherent conclusions across the modes can be drawn. Particular areas that need further research and development in the area of mixed mode (adjustment) include: (The following areas are informed by the outcomes of an expert group meeting on mode effects in 2021. The group had been brought together by the previous GenPopWeb2 network, also funded by the ESRC, Feb 2020 - Sept 2021, https://www.ncrm.ac.uk/research/genpopweb2/)</p> <ol style="list-style-type: none"> 1. development of a coherent definition of mode effects as well as a comprehensive, rounded framework of all (interlinked) aspects of mode effects (e.g., reasons for using different modes, patterns of selection into different modes, measurement of mode effects and adjustment for mode effects), where a better understanding of mode effects and adjustments will in turn inform designs of surveys from the beginning. 2. development of more sophisticated adjustment models and methods, going beyond currently used regression methods, for example: multilevel and survival models, latent class models; improved imputation methods; methods that can be used in the production for large number of surveys and for many variables within the same survey and for aggregate statistics; adjustment for time series; and methods for estimating variances given adjustments, which are currently limited.

		3. development of methods on how best to evaluate adjustment models and investigations on what impacts adjustment models have on statistical outputs (including on descriptive statistics and impacts on particular variables; including development of quantitative measures, indicating what type of variables and to what extent different variables are susceptible to mode effects; this work would be valuable, for example, for future questionnaire designs). Also, (formal) statistical simulations are needed to test new mode adjustment approaches and to investigate nuances and subtleties of existing approaches, such as various regression methods.
5	Improved sampling frames for general population surveys to enable more effective use of remote and new data collection methods	One of the major, longstanding limitations and barriers to making greater use of remote methods for online and mixed-mode surveys is the lack of an individual level sampling frame for general population surveys in the UK. Some studies and organisations have carried out work on enhancement to the PAF sampling frame, and the use of administrative data for sampling is also a strategic area of interest for ESRC and links to several of their previous and ongoing initiatives. Hence, there is a strategic need for cross-sector collaboration to address this barrier (e.g., PAF enhanced with individual information such as phone numbers, work on administrative data sampling frames), and to bring together previous work and learnings.
6	Changing role of survey interviewers	Interviewers play a crucial role in many surveys and survey designs. This is of course particularly the case for f2f and telephone surveys, but also applies to new or adapted (online) designs, for example when using new approaches such as VAPI (Video Assisted Personal Interviews), and knock-to-nudge. Making contact and persuading survey members to take part has usually meant securing an interview to complete face-to-face with the participant. The role that the interviewer plays nowadays can be much wider and the role is changing. In their new role, they may be asking respondents to complete web surveys or placing devices for them to complete themselves. Moreover, interviewers themselves may now be carrying out phone and video-interviews as well as face-to-face. In addition, the interviewer fieldforce has been subjected to wide ranging changes, such as interviewers resigning at scale to find other type of work during lockdowns, with many interviewers now often not returning to interviewing tasks. Also, interviewers are finding alternative, more attractive employment in a generally changed working environment, allowing for more flexibility and higher pay elsewhere. Hence, survey agencies have and continue to lose a significant number of experienced members of staff. The industry is hence facing significant changes and the changing role of the interviewer needs further investigation in survey data collection.
7	Complex measurements in online surveys	The topic of complex measurements in online surveys continues to be an important developing field. Measurement approaches originally developed for, say, f2f surveys may not work well in an online context and often need to be adapted or even changed completely. The previous ESRC-funded GenPopWeb2 network identified this as a key topic area and organised three events in this area in 2021. This research priority area includes topics on, for example: development, assessment and optimisation of measures of cognition in online surveys (e.g., which cognition measures are best suited to online collection and how best to develop new measures of cognition in online surveys); use of paradata (e.g. for measuring cognition); collection of data linkage consent (e.g., how to boost consent rates; how to optimise visual designs and layouts); mixing of modes to improve measurement outcomes (e.g. would removing consent from an online instrument and using a telephone follow-up instead work better than an online only request?; Is calibration across modes feasible for measuring cognition?); complex coding (e.g. occupational coding; optimal length

		<p>of list of coding options); how best to replicate interviewer functions in an online survey environment (for example, using a form of 'live chat', messaging services based on an 'avatar' or 'chat bot', or asking respondents to phone in during or after filling in the survey online); and the collection of bio-measures that are generally difficult to obtain in online survey environments.</p> <p>This priority area is linked to priority area 4, mode effects/use of different modes.</p>
8	<p>Discontinuity/time series in repeat cross-sectional and longitudinal measurement</p>	<p>All surveys have seen a break in their data collection method, either because data collection, such as f2f, being stopped all together or because data collection was switched to alternative modes during the lockdowns. Now in the (post/) pandemic phase, data collection may continue with further adaptations and changes and under revised designs (e.g. adaptations to f2f survey collection, moving from a sole online approach during the pandemic to a mixed-mode or web-first approach). All of the changes have led and continue to lead to discontinuities in time series in repeat cross-sectional and longitudinal measurements. Research is needed to identify the impact of the changes and breaks, as well as to investigate adjustment methods to overcome estimation issues following breaks in designs.</p>
9	<p>Development of an inclusive data system across the whole data lifecycle from collection through to outputs</p>	<p>Inclusivity is high on the agenda in research and government and to be able to make any significant progress on this goal, this must be reflected in our data and methodology systems. Inclusive data helps us to understand how events (e.g. the Covid-19 pandemic, or the current energy crisis) impact differentially on individuals, groups and communities. In turn, this will enable those responsible in government, local authorities and the wider society, individuals and communities, to address the disparities and inequalities, which exist in the UK. It is, therefore, important to research how inclusivity can be improved across the research process from study design, data mapping, data collection, data analysis through to the presentation of findings.</p> <p>In 2020 the UK Statistics Authority launched its new 5-year strategy, Statistics for the Public Good, which included inclusivity as one of four goals. Subsequently, the ONS launched a taskforce on inclusivity in 2020 with the goal of improving the UK's inclusive data holdings in a broad range of areas and to ensure that data and evidence across the UK is reflective and inclusive of all. The taskforce recently published its recommendations on inclusivity. The Market Research Society (MRS) has also been encouraging greater consideration of inclusivity to ensure minority groups are represented in their samples (see industry group launch, inclusion and diversity pages and the Diversity and Inclusion best practice guide from May 2022 to help ensure diverse sampling and inclusive methodologies).</p> <p>However, more needs to be done in the survey space to ensure that data are representative of society and that everyone can take part in surveys. Inclusivity begins at the start of the data lifecycle, at collection, and goes right the way through to outputs - the end product being an inclusive data system. More research and investment are needed into exploring how to create more inclusive surveys, in particular, with focus on how to achieve this in the first phases of the data lifecycle (i.e., at the design and data collection phases). Part of inclusivity is 'accessibility', and further importance should be given to this via new research on how to design and develop accessible surveys so that all parts of society can provide their data without being excluded through design. This includes strategies for contacting and engaging with groups that are hard-to-reach or hard-to-engage-with (for example due to cultural issues/ethnicity, language barriers, age, gender, disability, living in remote areas (e.g. in Scottish highlands and islands). In summary, a more inclusive data system will lead to higher quality data and more effective policy interventions.</p>

<p>10</p>	<p>Exploration of innovative methods that help in achieving this inclusive data system, including respondent centred design</p>	<p>The same or similar challenges and issues in surveys and survey development, such as falling response rates and issues around data quality, continue to persist, and have done so for many years. Although some progress has been made in addressing them, further work and innovations are needed in order to elevate that progress to the next level. Typically, a lot of effort is invested in assessing and improving data quality at the post-data collection phase. However, to be able to address persisting issues, more research is needed focussing on significant improvements and innovative thinking at the design and data collection phases. This is where thinking outside the box and innovation in how we tackle these problems is needed, as only marginal gains are being made using traditional solutions, such as the use of incentives.</p> <p>One example is 'respondent centred design' (Wilson and Dickinson, 2021), which has the respondent at the centre of the question writing and survey implementation process. The approach aims to improve the respondent experience and hence also increases inclusivity (see priority area 9) and presents a radical innovative approach in survey data collection designs (see priority area 3 innovations in surveys). For far too long survey development has prioritised the data user need at the expense of the respondent experience, which has left us with unwieldy surveys with a poor respondent experience. Prioritising good respondent experience in survey designs and data collection phases and investing in a better understanding of the impacts of respondent experiences on data quality (see priority area 2) and data outcomes should be pursued. Some researchers and organisations are already seeing initial results of focusing on the respondent experience and needs. However, much more research is needed to develop these ideas and initial projects further. To stress the importance of this area, Wilson and Dickinson were recently asked by the UN Global Network and Inter Secretariate Working Group for Household Surveys to do a masterclass webinar on Respondent Centred Design, to showcase these new methods. They have subsequently decided to set up, jointly with ONS, a yearlong UN Taskforce on Respondent Centred Design. Survey methodologists and practitioners need ring-fenced time and funding to generate new ideas in this space and to implement these approaches into practice and to base them on solid theoretical and practical grounds, with the aim of finding new ways to tackle persisting problems.</p>
-----------	---	---