



Social Research Practice

ISSUE 11 SUMMER 2021

The Social Research Association journal for methods in applied social research

Contents

- 03 Editorial
Richard Bartholomew
-

ARTICLES

- 04 Incentivising early responses in a push-to-web survey:
an experiment
Patten Smith, Ipsos MORI; Lucy King, Food Standards Agency; David Candy,
Ipsos MORI; Richard Bridge, Food Standards Agency; and Beth Armstrong,
University of Sheffield
- 13 A disaggregated approach to recruiting
Own Household Fathers in cohort studies
Rebecca Goldman, Fatherhood Institute; Paul Bradshaw, ScotCen Social
Research; Adrienne Burgess, Fatherhood Institute; and Konstantina Vosnaki,
ScotCen Social Research
-

RESEARCH NOTE

- 23 How to replace large-scale, continuous, face-to-face
household surveys during the COVID-19 pandemic:
the Scottish Health Survey example
Shanna Christie, NatCen Social Research
Joanne McLean, ScotCen Social Research

Editorial

Richard Bartholomew
Editor

Welcome to issue 11 of Social Research Practice.

In our first article **Patten Smith et al** describe an ingenious experiment using the Food Standard Agency's 'Food and You 2' survey to explore how incentive payments can be used to increase online responses in push-to-web surveys (a push-to-web survey is one in which an initial postal invitation is sent to sampled addresses to encourage people to complete an online questionnaire). Online data-collection methods can provide higher quality and more complex data. But for research requiring a high-quality random probability sample, the challenge is to be able to successfully combine the use of the postcode address file (PAF) with the advantages of online data collection. Effective (and affordable) conditional incentive payments offer a way of persuading people to access and complete an online questionnaire, and to do so in a timely way. The authors have tested out which structure of incentives works best.

It is a commonplace that today's children grow up within increasingly complex family structures. Around a third of children under 16 live most of the time with one parent, usually the mother. But many birth fathers continue to play a significant role in their child's life even when they do not live in the same household. For the major UK cohort studies, it has always been difficult (and costly) to collect sufficient information from and about these non-resident fathers, yet they can have an important influence on their child's development. In their article on 'Own Household Fathers' **Rebecca Goldman et al** discuss their scoping study, conducted for the Economic and Social Research Council, on adopting a more disaggregated approach to involving this heterogeneous category of fathers in future sweeps of cohort surveys. They examine the merits of focusing primarily on those fathers who are more involved with their children's lives while still maintaining the overall representativeness of samples.

The demand for rapid information on the immediate effects and longer-term consequences of the Covid-19 pandemic has led to many new social studies being commissioned and existing ones being radically adapted. However, existing continuous surveys have also had to adapt their methods to the changed circumstances of national lockdowns and social distancing requirements. Face-to-face interview methods had already become increasingly rare but were particularly suitable for collecting good quality information on more sensitive and personal topics, including physical measurements. The pandemic has made this method virtually impossible, perhaps for the foreseeable future. In our third piece (a research note) **Shanna Christie and Joanne McLean** discuss the adaptations that have had to be made to the ongoing Scottish Health Survey to meet these challenges. It is not just about new methods of data collection but also about assessing how these changes affect time series data – a key issue if you are trying to distinguish the social and health consequences of the pandemic from the effects of a methodological change. The authors have also published an SRA blog to coincide with this research note.

Please get in touch if you think you could offer an article or shorter research note on what is being learned during the pandemic: admin@the-sra.org.uk

We welcome proposals for new articles or shorter research notes. If you are interested in offering a research note or a full article you can find more information on the [SRA website](#).

Incentivising early responses in a push-to-web survey: an experiment

Patten Smith, Ipsos MORI; Lucy King, Food Standards Agency; David Candy, Ipsos MORI; Richard Bridge, Food Standards Agency; and Beth Armstrong, University of Sheffield

Abstract

Push-to-web surveys in the UK typically involve mailing letters to samples of addresses selected randomly from the postcode address file (PAF), asking residents to complete an online questionnaire, then including postal questionnaires with later reminder mailings. We conducted an experiment in wave 1 of the push-to-web 'Food and You 2' survey to encourage online responses by manipulating the incentive regime. Incentives were varied across three experimental groups: control group respondents received a £10 conditional incentive; experimental group E1 received a £15 conditional incentive if they responded before the first reminder despatch or a £10 one if they responded later; experimental group E2 also received a £15 conditional incentive if they responded before the deadline but only a £5 one if they responded later. Offering the larger incentive to those responding before first reminder despatch (£15 instead of £10) increased the number of responses received before that deadline. Compared with the control group, group E1 (£15 or £10) retained a higher questionnaire return rate at the end of fieldwork. In contrast, in group E2 (£15 or £5), the early questionnaire return rate gain was completely lost by the end of fieldwork, and the final questionnaire return rate was below that of the control group. Comparisons of selected demographic and survey variables between the three groups provided no evidence that different incentives affected survey estimates. Costs per achieved respondent were around 5% greater for both the early incentive groups relative to the control.

Introduction

Background

Recent years have seen the emergence of new forms of data-collection methodology for high-quality random probability general population surveys in the UK. Until recently the standard data-collection method for such surveys had been in-home face-to-face interviewing. A number of factors, including increasing survey costs, decreasing response rates and increased population web-literacy, have recently motivated survey practitioners to explore alternative approaches.

One of the most commonly used of these is the so-called 'push-to-web' method (Dillman et al, 2014; Ipsos MORI, 2019; Williams, 2016). As currently practised in the UK, this method typically (although not always) includes the following elements: a random sample of postcode address file (PAF) addresses, postal invitations to participate in an online survey, and two or three reminder mailings, one of which includes a postal questionnaire that can be completed instead of the online one (Ipsos MORI, 2019).

The 'push-to-web' method includes these elements mentioned for a number of reasons. First, online questionnaires are used because (relative to comparably priced paper self-completion questionnaires) they permit the use of more complex question structures and provide higher quality data. Second, because random probability surveys require high coverage sample frames, and because the only readily available one of these in the UK (PAF) includes postal addresses but no electronic contact information, it is necessary to make initial contact using offline methods. Third, research indicates that people who most readily respond online to surveys differ substantially from those who respond using other modes (such as postal questionnaires or interviews), tending to have higher incomes, to be better educated, and to be less likely aged 65+. In order to minimise non-response bias, it is, therefore, necessary to supplement online data collection with another data-collection mode, of which sending postal questionnaires is the most cost effective. The push-to-web methodology is used for a number of major high-quality UK surveys, including the Community Life Survey (Kantar Public, 2020), the Active Lives Survey (Sport England, 2021) and the Food and You 2 survey (Armstrong et al, 2021).

As just described, postal questionnaires have to be administered during reminder phases of push-to-web surveys in order to minimise non-response bias. However, online questionnaires typically have superior measurement properties to postal questionnaires: they allow more questions to be asked, allow more complex question routing to be used, control answer formats, allow real-time checks to be made of respondents' answers, and show considerably lower levels of item non-response. For a given respondent, we would prefer to receive an online response to a postal questionnaire one.

In push-to-web surveys, therefore, it would be advantageous if we could persuade survey participants to use online questionnaires instead of postal ones. The experiment we report on here was designed to investigate the extent to which this might be done through the use of differential incentives.

There is a considerable body of evidence on the impact of incentive payments on response rates (Church, 1993; Nicolaas et al, 2019; Singer and Ye, 2013). Cash, or cash-like (for example monetary vouchers), incentives are generally more effective at increasing response rates than in-kind gifts. Sending unconditional incentives to all selected sample members in advance is more effective than offering equivalent-value incentives which are conditional on taking part in the survey. However, in UK push-to-web surveys, the standard practice has been to offer conditional cash-like (voucher) incentives to respondents. Because response rates in push-to-web surveys have generally been fairly low (Williams, 2016), unconditional incentives have generally been regarded by survey organisations as unaffordable.

Two recent pieces of UK methodological work suggest that an effective way of boosting online response in a push-to-web survey might be to use the initial invitation letter to offer an additional conditional incentive for completing the online questionnaire by a given deadline, and before the postal questionnaire is mailed. In the three-mode Next Steps Age 25 survey, an experimental group was offered a £20 incentive for online completion in the first three weeks of fieldwork and £10 if they responded later than that (Peycheva et al, 2019). The remainder of the sample was offered a £10 conditional incentive regardless of when the questionnaire was completed and served as a control group. The early completion incentive increased the short-term response rates but not the final overall ones. However, by the end of fieldwork the *online questionnaire response rate* remained significantly higher for those who had been offered the early completion incentive. Similarly, in non-experimental methodological work in the Understanding Society longitudinal survey, a £10 bonus conditional on completing an online questionnaire within two weeks appeared to raise the whole household online completion response rate from 19% to 26% (Carpenter and Burton, 2018).

The experiment reported here built on this work by using the initial recruitment mailing to offer a larger conditional incentive to those responding before the first reminder letter despatch date than to those responding later. The experiment was conducted on the Food and You 2 survey commissioned by the Food Standards Agency to collect data on people's attitudes, behaviours and knowledge about food

safety and other food issues. For the survey, an issued sample of 21,053 PAF addresses were mailed an initial invitation letter asking up to two address residents aged 16+ to respond to an online survey; a URL and login details for two individuals were provided. Nine days later a first reminder letter was sent to non-responding addresses and to addresses where only one of two or more eligible adults had responded. A similar second reminder was sent 21 days after the first, now accompanied by a postal questionnaire in addition to the online survey login details. A final reminder letter (including the login details but not accompanied by a postal questionnaire) was sent 11 days after the second reminder. Data collection took place between 29 July and 6 October 2020.

The experiment

The issued address sample (21,053 addresses) was ordered by local authority and LSOA (SOA in Northern Ireland) multiple deprivation index, and then systematically allocated (by the method of random start and fixed interval) to three experimental groups. Half the sample was allocated to the control group in which each individual respondent received a flat rate £10 shopping voucher conditional on responding. The two experimental groups each comprised a quarter-sample of addresses (groups E1 and E2). Group E1 respondents received a £15 shopping voucher if they responded by 6 August, eight days after initial despatch, (termed ‘early’ respondents for the remainder of this paper) and a £10 voucher if they responded later. Group E2 early respondents also received a £15 voucher but those responding later received only a £5 voucher. The 6 August date was sufficiently early to ensure that all those receiving the £15 incentive would not be sent any reminders, with a view to partly offsetting the additional expenditure on incentives with reduced expenditure on printing and postage. Because postal questionnaires were sent only in the second reminder mailing, all early respondents were necessarily also online respondents, and the two incentive regimes tested in the experimental conditions were deliberately designed to encourage online responses. The experimental conditions are summarised in Table 1.

Table 1: Experimental conditions

	Control	Experimental E1	Experimental E2
Issued number of addresses	10,527	5,263	5,263
Incentive amount received for response by 6 August	£10	£15	£15
Incentive amount received for response after 6 August	£10	£10	£5

In the following section we discuss the results of the experiment in three subsections. We first show the impact of the experimental treatments on responses received (i) overall (ii) before the higher incentive cut-off date (iii) online and (iv) on the postal questionnaire. Second, we compare the profiles of respondents in each experimental group on demographic variables and important survey variables. Thirdly, we compare the overall administration costs of the three incentive treatments.

Results

Address level response

Table 2 shows by experimental group: (i) the number of addresses from which one or two *early* responses were received and (ii) the number of addresses from which one or two responses were received at any time during fieldwork. The table shows that offering the early completion incentive successfully increased the number of early responses received in both experimental groups (E1: 13.7%, control: 7.9%; $p < 0.001$. E2: 11.8%, control: 7.9%; $p < 0.001$). Furthermore, significantly more addresses responded early in group E1 than in group E2 (13.7% vs. 11.8%; $p < 0.003$) despite the incentive for early responses being the same in the two groups.

Overall, the address level response rate was highest for group E1 (E1: 33.6%, E2: 27.7%, control: 30.3%; E1 vs. control, $p < 0.001$. E1 vs. E2, $p < 0.001$) which is perhaps unsurprising as this group received the highest incentives (£15 for an early response; £10 for a later response). However, more control group addresses than group E2 addresses responded overall (30.3% vs. 27.7%; $p < 0.001$) despite the higher early return response rate for the latter group, suggesting that a £5 incentive for later responses was considerably less effective than the £10 one.

Table 2: Responding addresses by experimental group

	Control		Experimental group E1		Experimental group E2	
	N	%	N	%	N	%
Issued addresses	10,527	100.0%	5,263	100.0%	5,263	100.0%
Addresses with one or two responses before first reminder despatch ^a	833	7.9%	720	13.7%	619	11.8%
Addresses with one or two responses at any time ^b	3,185	30.3%	1,767	33.6%	1,456	27.7%

^a All pairwise inter-group differences significant ($p < 0.01$) using Chi-square test.

^b All pairwise inter-group differences significant ($p < 0.01$) using Chi-square test.

The ultimate aim of offering higher incentives for early completions was to persuade potential respondents to switch from postal to online survey completions. We cannot directly assess whether an experimental treatment had this effect because to do so would require us to know which of our online respondents would have responded by post, and which of our postal respondents would have responded online had they received a different experimental treatment. We can, however, compare the number of online and postal questionnaires received in each condition.

Table 3: Mean number of online and postal returns per issued address

	Control	E1	E2	Sig tests
Issued addresses	10,527	5,263	5,263	
Online returns per issued address	0.269	0.337	0.254	C vs. E1: $t=6.4$, $p<0.001$ C vs. E2: $t=1.4$, n.s. E1 vs. E2: $t=6.7$, $p<0.001$
Postal returns per issued address	0.174	0.156	0.138	C vs. E1: $t=2.2$, $p<0.025$ C vs. E2: $t=4.7$, $p<0.001$ E1 vs. E2: $t=2.2$, $p<0.029$
Total returns per address	0.443	0.493	0.392	C vs. E1: $t=4.0$, $p<0.001$ C vs. E2: $t=4.2$, $p<0.001$ E1 vs. E2: $t=7.2$, $p<0.001$
Ratio of online to postal returns (online: postal)	1.55:1	2.16:1	1.84:1	

Note: t = t -value, n.s. = not statistically significant

Table 3 shows that experimental condition E1 delivered significantly more online responses per issued address than did either of the other conditions, and that E2 and control conditions delivered similar numbers of online responses. The control condition delivered more postal questionnaires than did either experimental condition and, of the latter, E1 delivered more postal returns than E2. In line with the address level response rates shown in Table 2, Table 3 shows that condition E1 delivered the largest number of responses overall, and condition E2 the least.

Condition E1 successfully encouraged online responding: it obtained more online responses, more responses overall, and a higher ratio of online to postal responses than did the control condition (2.16 vs. 1.55). Although condition E2 also increased the *ratio* of online to postal responses from 1.55 (for the control) to 1.84, it did not increase the number of online questionnaires received in absolute terms, and was achieved at the cost of obtaining fewer responses overall (0.392 responses per issued address vs. 0.443 for the control).

Demographic and survey variable comparisons

Response rates are not direct measures of survey error. Rather, they are best interpreted as providing broad indicators of risk of the relevant survey error type: non-response bias. Indeed, there is considerable evidence indicating that response rate predicts non-response bias rather poorly (Brick and Tourangeau, 2017; Groves and Peytcheva, 2008; Sturgis et al, 2017). The response figures discussed above, therefore, do not allow us to draw any conclusions about levels of non-response bias in each condition.

We also note that experimental conditions may vary in levels and types of measurement bias. Different numbers of people complete online and postal questionnaires in each condition and, if measurement bias differs across these two modes, the overall measurement bias would be expected to differ across conditions.

For the great majority of Food and You 2 survey variables, we are unable to estimate the cumulative impact of non-response and measurement bias because, for these variables, no trusted population estimates are available for comparison. This means we are unable to compare levels of bias by condition. We can, however, compare estimates across conditions. If they are found to differ by condition, we can legitimately conclude that levels of bias vary by condition, although we cannot determine which condition provides the least biased estimates.

Making such comparisons is a worthwhile exercise both because (i) if differences by condition are found, this indicates that sources and levels of bias warrant further detailed investigation and (ii) it would indicate that changing the incentive regime could inadvertently disrupt trends in survey estimates.

For these reasons we compared conditions across five demographic variables and 30 survey variables. Estimates were very similar across the three conditions for all variables compared. Table 4 shows a range of illustrative estimates.

Table 4: Selected demographic and survey estimates by condition (unweighted estimates)

Estimate	E1 (%)	E2 (%)	Control (%)
Demographic variables			
No children in household	73	72	73
Female	42	39	41
Aged under 35	21	20	20
Aged 65+	28	28	28
In full-time work	40	40	40
Ethnic group: white	91	91	92
Selected survey variables			
Does most of household food shopping	44	45	45
Ever receives home deliveries from supermarket	42	42	43
Does most household food preparation	48	50	48
Never eats chicken or turkey when meat is pink	92	92	92
Always washes raw chicken	16	17	17
Uses different chopping board for raw meat and other foods	47	45	46
Always checks use-by dates when preparing food	64	67	66
Always washes hands when preparing food	77	77	78
Confident that food buy is safe	93	93	93
Trusts the Food Standards Agency	76	78	77
Food security: low or very low	14	15	14

Note: all differences statistically non-significant.

Costs

The marginal cost of obtaining a response depends on how much is spent on printing (invitation and reminder letters and postal questionnaires), postage, incentives and processing across the whole issued sample. Increasing the amount offered in incentive is likely to increase the overall amount spent on incentives, but, if increased incentives also increase response propensities, especially early in fieldwork before reminders and postal questionnaires are issued, they will also reduce the amount spent on printing, despatch processes and postage. It is, therefore, of interest to estimate the net cost of achieving a response across conditions. Table 5 shows that the costs per response were around 5% greater for the two experimental groups than for the control group.

Table 5: Relative cost per achieved response by condition

Condition	Cost (control group indexed at 1)
Control	1.000
Experimental group E1	1.049
Experimental group E2	1.054

For group E1, the greater cost can be mainly attributed to the increase in the amount of incentive offered to respondents, although this increase will have, in part, been offset by the increase in response rate. In contrast, the increase in costs relative to the control for group E2 mainly resulted from the greater printing and postage costs associated with its lower response rate and having to issue a greater number of reminders and postal questionnaires.

Discussion

In line with previous research findings, we found that offering a larger (£15 instead of £10) incentive to those responding before a stated deadline increased the number of responses received by the deadline date. The group offered the higher incentive for responding early retained a higher overall questionnaire return rate than the flat rate £10 incentive control group, but only if the incentive did not drop to below that of the control group (£10) after the deadline. In condition E2, where the incentive for later respondents dropped to £5, the early questionnaire return rate gain was completely lost by the end of fieldwork, and the final questionnaire return rate was below that of the control group.

One unexpected result was that the group E1 address level early questionnaire return rate (13.7%) was significantly greater than the group E2 one (11.8%) despite the early completion incentives being the same across the two groups. We speculate that this was because, in both versions of the initial invitation letter, in addition to the £15 early completion amount, the incentive amounts for later responses (£10 in E1 and £5 in E2) were displayed fairly prominently, and that these figures may have caught the attention of those rapidly skimming the letter.

The main purpose of offering an early completion incentive was to persuade those wishing to respond to do so online rather than through a postal questionnaire which was issued later with a second reminder. Although we do not have direct evidence about this, we do have strong circumstantial evidence. Group E1, in which respondents were offered £15 and then £10, delivered more online responses, fewer postal responses and more responses in total per issued address than did the control group in which respondents were offered £10 regardless of when the questionnaire was completed. In group E1 there were 2.16 online responses per postal response; the corresponding figure for the control group was 1.55.

In contrast, relative to the control, group E2 delivered no more online responses, fewer postal responses and fewer responses overall per issued address. It increased the ratio of online to postal responses from 1.55 (for the control) to 1.84, but this came at the cost of a reduced overall questionnaire return rate.

Comparisons of selected demographic and survey variables between the three groups provided no evidence that incentive regime affected survey estimates. This means that the choice between incentive regimes can be based purely upon comparisons in overall response rates, in online response rates and in costs.

The costs per achieved respondent were around 5% greater for both the early incentive groups relative to the control. Taken with the evidence about questionnaire return rates already discussed, this clearly indicates that the group E2 incentive regime (£15 followed by £5) was not cost effective: fewer responses were received overall, the absolute number of online responses was not increased, and the cost per response increased.

The value-for-money judgement for the group E1 incentive regime is more nuanced. Although the cost per response is greater relative to the control, so is the quality of each response: responses are more likely to be online (allowing longer, more complex questionnaires, real-time edit checks, and delivering lower levels of item non-response relative to postal responses) and are associated with a higher response rate. Therefore, the judgement as to whether it is worthwhile to offer an additional incentive for early completion depends upon how much positive value is placed on increasing response rates for the higher quality online questionnaires and response rates overall, relative to the negative value of increasing the survey costs.

References

- Armstrong, B., King, L., Clifford, R. and Jitlal, M. (2021). Food and You 2: wave 1 key findings. https://www.food.gov.uk/sites/default/files/media/document/fy2-wave-1-report-_key-findings_1.pdf
- Brick, J. A. and Tourangeau, R. (2017). 'Responsive survey designs for reducing nonresponse bias'. *Journal of Official Statistics*, 33(3): 735–752.
- Carpenter, H. and Burton, J. (2018). Adaptive push-to-web: experiments in a household panel study. *Understanding Society Working Paper Series*, No. 2018–05.
- Church, A. H. (1993). 'Estimating the effect of incentives on mail survey response rates: a meta-analysis'. *Public Opinion Quarterly*, 57(1): 62–79.
- Dillman, D., Smyth, J. and Christian, L. (2014). *Internet, phone, mail and mixed-mode surveys: the tailored design method*. 4th edition.
- Groves, R. and Peytcheva, E. (2008). 'The impact of nonresponse rates on nonresponse bias: a meta-analysis'. *Public Opinion Quarterly*, 72: 167–189.
- Ipsos MORI. (2019). Push-to-web best practice guide. <https://www.ipsos.com/ipsos-mori/en-uk/push-web-best-practice-guide>
- Kantar Public. (2020). Community Life Survey technical report 2019/20. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/962599/Community_Life_Online_and_Paper_Survey_Technical_Report_-_2019-20_V2.pdf
- Nicolaas, G., Corteen, E. and Davies, B. (2019). The use of incentives to recruit and retain hard-to-get populations in longitudinal studies. <https://esrc.ukri.org/files/funding/funding-opportunities/the-use-of-incentives-to-recruit-and-retain-hard-to-get-populations-in-longitudinal-studies/>
- Peycheva, D., Calderwood, L., and Wong, E.R. (2019). Effects of a time-limited incentive in a mixed-mode longitudinal study of young adults. Unpublished manuscript.
- Singer, E. and Ye, C. (2013). 'The use and effects of incentives in surveys'. *Annals of the American Academy of Political and Social Science*, 645(1): 112–141.
- Sport England. (2021). Active Lives Adult Survey, November 2019/20 report. <https://sportengland-production-files.s3.eu-west-2.amazonaws.com/s3fs-public/2021-04/Active%20Lives%20Adult%20November%202019-20%20Report.pdf?OjWdwCLnl3dNgDwp3X4ukcODJIDVG7Kd>
- Sturgis, P., Williams, J., Brunton-Smith, I. and Moore, J. (2017). 'Fieldwork effort, response rate, and the distribution of survey outcomes: a multilevel meta-analysis'. *Public Opinion Quarterly*, 81(2): 523–542.
- Williams, J. (2016). 'An introduction to address-based online surveying'. *Social Research Practice* 3: 23–35.

A disaggregated approach to recruiting Own Household Fathers in cohort studies

Rebecca Goldman, Fatherhood Institute; Paul Bradshaw, ScotCen Social Research; Adrienne Burgess, Fatherhood Institute; and Konstantina Vosnaki, ScotCen Social Research

Abstract

Own Household Fathers (OHFs) do not live full-time with their child and are often described as ‘non-resident’. As part of a scoping study on recruiting and retaining these fathers in child cohort studies, we investigated the potential of a disaggregated recruitment strategy. This strategy prioritises engaging the more involved of these fathers, with targeted recruitment methods, to give a balanced sample for analysis of their impact on children’s wellbeing and outcomes. It contrasts with previous attempts to recruit them as an undifferentiated category, resulting in response bias towards more involved fathers, and low overall response rates. Following a review of methodological literature and a cognitive pilot of questions to identify involvement sub-categories of OHFs, we conclude that a disaggregated strategy has potential once a cohort study is underway, but a birth cohort study’s first sweep is an opportunity to recruit all birth fathers regardless of their level of involvement.

Funding acknowledgement

This project was funded under the Economic and Social Research Council’s (ESRC) ‘UK Population Lab’ programme of scoping studies and methodological reviews to inform thinking about innovations in longitudinal studies (**Innovation and development in longitudinal studies: outputs from the ‘UK Population Lab’ programme – Economic and Social Research Council (ukri.org)**).

Why collecting data from Own Household Fathers in cohort studies matters

Around a third of children under 16 years live for all or most of the time with one birth parent, predominantly their mother. Research has shown that these children’s wellbeing and development are influenced by birth fathers. This is through part-time co-residence, father–child interactions, financial contributions, the father’s co-parental relationship and decisions with the birth mother, genetics, epigenetics, and – where it applies – the child’s experience of a lack of father involvement. The children may have two parental households in which they spend substantial time. A family systems approach is important in understanding pathways to child outcomes and for influencing policies and services to facilitate child wellbeing.

UK cohort studies have followed children from pregnancy or birth into adulthood. They are internationally renowned for providing robust evidence to influence policy and practice. They have successfully collected data from mothers' cohabiting partners (predominantly birth fathers and stepfathers) but have not attempted to recruit the vast majority of Own Household Fathers (OHFs), the term we use for birth fathers who do not live full-time with the cohort member ('cohort child'). They have not tracked or interviewed fathers who become OHFs following parental separation. OHFs are often described as a homogenous category of hard-to-reach 'non-resident' or 'absent' fathers. This weakens the perceived rationale for their inclusion as research participants.

While the cohort studies Fragile Families in the US and the Longitudinal Study of Australian Children (the LSAC – 'Growing Up in Australia') have been successful in recruiting and retaining OHFs, other studies internationally have had low response rates. This is partly due to difficulties in gaining contact details and to OHFs' reluctance to respond. It may also be influenced by a lack of knowledge of and expenditure on effective methods to engage and collect data from these parents.

Instead, UK cohort studies have collected limited data from mothers and older cohort children about OHFs. Such data do not give the father's perspective and are associated with reporting biases and item non-response. It is also a missed opportunity. Involved OHFs can provide rich data on their behaviours, attitudes and activities that involve and have an impact on their child. If the mother drops out, then a longitudinal study may retain the cohort child via their involved OHF. Data from biological fathers, whether resident or not with the child, are needed to explore genetic and epigenetic effects.

The ESRC's Longitudinal Studies Review proposed that a new birth cohort study could collect data from 'non-resident' fathers as well as 'resident' fathers in order 'to adequately address the gene-environment interplay, to better understand the dynamics of separated families, and to enable more research on the intergenerational transmission of inequalities' (Davis-Kean et al, 2017, p. 24). With an ESRC-funded feasibility study for a new early life cohort study launched in April 2021, and longitudinal studies such as ALSPAC and Understanding Society seeking to recruit fathers and separated parents, it is timely to better understand how OHFs may be successfully engaged.

The scoping study

In 2019, the ESRC commissioned ScotGen and the Fatherhood Institute to scope, assess and propose options for childhood sweeps of cohort studies to identify, recruit and retain birth fathers who do not co-reside full-time with their cohort child – OHFs. This would build on and update previous UK methodological reviews (Bradshaw, 2008; Kiernan, 2016).

As part of this scoping study, we investigated the feasibility of a disaggregated approach to maximising response rates and representativeness based on the heterogeneity of OHFs. This paper focuses on that objective, with other study findings reported in a working paper (Goldman et al, 2019).

The heterogeneity of OHFs

The Fatherhood Institute's 'Where's the daddy?' review of fathers in large-scale quantitative datasets¹ had re-conceptualised a substantial proportion of OHFs (typically all described as 'non-resident') as part-time resident with their child/ren, with regular overnight co-residence (Goldman and Burgess, 2018). Co-residence in this context refers to the father and child, in the father's or mother's household or elsewhere.

While household surveys and cohort studies use a binary classification of individuals as primarily resident or not at the sampled address, this obscures the growing phenomenon of second addresses. In population surveys, this can lead to double counting of children living between two households, and classification of involved OHFs as men without dependent children (Toulemon and Penniec, 2010; Waller and Jones, 2014). In family and child development research, it results in an emphasis on 'father absence' rather than exploring the impacts of living arrangements.

¹ Funded by the Nuffield Foundation.

'revolving doors to family life with many parents and children living together only some of the time' (Kiernan, 2006, p. 666)

During preparatory work for our scoping study, we extended our conceptual 'Where's the daddy?' work to describe ten sub-categories of OHFs defined according to co-residence (regular overnight stays), daytime care and 'virtual' interactions between OHF and cohort child. These can be associated with different types and levels of impact on children.

OHFs may have 'minority overnight care' of their child/ren (for example one weekend a fortnight); or less commonly, 'majority overnight care'² (often classified as 'lone fatherhood') or 'equal overnight care' (three or four nights per week). Other OHFs have their own main home but regularly stay in the mother's household with their child/ren. These 'part-time stay' fathers, but not 'minority overnight care' fathers, were termed 'part-time resident' in the Millennium Cohort Study (MCS), referring to part-time residence with the mother. OHFs who do not regularly co-reside with their child may provide substantial daytime care or have frequent 'in-person' time together. Others have regular 'virtual' interactions with their child, via phone, video-call, text, email or social media, sometimes combined with overnight stays and 'in-person' time during holiday periods. Alternatively, OHFs may have infrequent contact or no contact with their child, recently, or since parental separation or birth.

A re-conceptualised disaggregated approach to maximising response rates and representativeness

In cross-sectional and longitudinal studies which treat OHFs as an undifferentiated category, responding fathers are most likely to be those with substantial interactions with their children, and good relationships with the children and/or mother (Bryson et al, 2017; Baxter, 2018). No-contact and infrequent-contact OHFs are especially challenging to recruit. This creates response bias towards involved OHFs. Fieldwork outcomes are judged as unsuccessful, and the achieved sample as unrepresentative of the population of 'non-resident fathers'. This discourages future attempts to include OHFs within family research studies.

An alternative strategy is to use rationed recruitment and fieldwork resources to prioritise engagement of **involved** OHFs. This may be a realistic objective and provide a substantial and balanced sample of these fathers for analysis of children's outcomes in the context of both households in which they are part-time resident or spend substantial time. OHFs could be classified in our involvement sub-categories on the basis of data provided in mother interviews if these precede OHF interviews at each sweep. Response rates and response bias could be measured separately for OHF sub-categories.

The disaggregated approach can extend to questionnaire development, fieldwork and analysis, with the most highly involved sub-categories of OHFs asked questions about father-child relationships and fathering behaviours, similar to data collected from fully resident fathers. These questions may engage these OHFs and increase their retention at future sweeps. Data can also be collected about or from family members in these OHFs' households. For research into the process and impact of parental separation, and for bio-social research seeking to 'address the gene-environment interplay', collection of data from less involved and no-contact OHFs is also important, but achieving a high response rate from these fathers may be overly-ambitious.

Recent research on participant engagement in longitudinal studies has incorporated 'targeted response inducement' strategies (Lynn, 2020) to improve participant engagement, representativeness and cost effectiveness (Jessop et al, 2019). Advance letters, motivational statements, interviewer scripts, telephone calls, incentives, and between-sweep mailings can be targeted at or tailored for participant sub-groups, with promising results (Park et al, 2018). In our context of improving OHF engagement, these strategies could be targeted and tailored according to OHF sub-category.

² These fathers are interviewed as 'primary caregivers' and 'resident fathers' in child cohort studies, so were excluded from the remit of our scoping study.

Research questions and methods

The potential in a disaggregated approach to OHF engagement resulted in our scoping study incorporating these research questions:

- ▶ How prevalent are OHF sub-categories in recent cohort study data? Would sample sizes be sufficient for separate calculation of response rates and separate analysis in a large national cohort study?
- ▶ What questions would we ask in mother interviews to identify OHF sub-categories?
- ▶ Are there precedents for cohort studies setting thresholds of father involvement for OHF recruitment and retention?
- ▶ Are there precedents for cohort studies tailoring recruitment and retention strategies to differentiated subsets of OHFs?

Our study's first phase involved the following elements:

- ▶ Analysis of data from questions in the Growing Up in Scotland (GUS) study which broadly identify our OHF sub-categories when children in the first birth cohort were aged seven to eight years. GUS is the most recent large-scale cohort study in the UK which is representative at a national level. Approximately one-quarter of families participating in this sweep met our inclusion criteria for analysis (Table 1)
- ▶ A scope of fieldwork instruments from UK and international studies to find questions which could identify OHF sub-categories; combined with a review of published critiques of these questions
- ▶ A review of methodological literature (over 120 documents) relating to OHF engagement in large-scale quantitative studies in the UK and abroad
- ▶ An email consultation with directors of longitudinal studies in the UK, US, France and Australia to identify recent technical documentation, unpublished papers and ongoing methodological work

The second phase involved developing a question set for face-to-face mother interviews to identify OHF sub-categories and a small-scale cognitive interviewing pilot to test and refine it.

In addition to overnight stays and in-person and 'virtual' father-child interactions, we asked about the relationship (partnership status, amicability, previous cohabitation) and travel time between the birth parents. This information can be used to target and tailor OHF engagement and to plan fieldwork.

Face-to-face interviews were carried out with 11 birth mothers of GUS cohort children who had been in contact with their OHF in the previous year. After asking each section of questions in a structured survey format, ScotCen researchers used qualitative probes to ascertain how mothers had decided on their response and what they thought about the questions.

How prevalent are OHF sub-categories in GUS data?

Table 1 summarises the results of our analysis of GUS data. These findings are specific to mother-reported father involvement for children aged seven to eight years. We would expect higher levels of father-child interactions and co-residence to be reported by OHFs (Waller and Jones, 2014; Baxter, 2018). For older children, greater time since parental separation and greater child independence influence levels and types of involvement. GUS questions did not enable estimation of the size of the 'part-time stay' sub-category.

We have extrapolated the findings to a hypothetical cohort study with an achieved sample size of 14,000 children, similar to that obtained in the MCS when cohort children were aged seven to eight years. Aggregating OHF sub-categories, there would be a sample size of over 2,000 for children staying overnight at least monthly (equal or minority overnight care) or having weekly in-person time with their OHF. There would be over 400 children with a lower level of in-person contact or solely virtual contact, and 1,000 without current contact. These sample sizes are amenable to separate analysis of response rates and response bias, and separate substantive analysis.

Table 1: OHFs by sub-category, using mother-report when cohort child aged 7 to 8 years in 2012-13

<i>Base: cases where child's birth father is alive, is not primarily resident in mother's household, and respondent is not an adoptive parent</i>	Weighted % in GUS sweep 7 data	Extrapolated sample size in cohort sweep with achieved sample of 14,000 cohort children, with 25% having an OHF
Equal/near-equal overnight care (three or four nights per week) – part of this sub-category may be included among 'resident fathers'	4	140
Minority overnight care (at least monthly as no fortnightly response code) – may be additional days with in-person daytime contact	40	1,400
At least weekly in-person contact (no regular overnights)	14	490
Regular in-person contact at least monthly but not as often as weekly (no regular overnights)	4	140
Virtual contact at least monthly (no regular in-person contact)	3	105
Infrequent/occasional contact (less often than monthly)	5	175
No current contact	30	1,050
Unweighted base	673	3,500

Published data shows that fathers who are OHFs in their child's earliest years are much more likely to be highly involved than out-of-contact at that time. For example, amongst the fifth of children in the second GUS birth cohort aged ten months in 2011 who had an OHF, around 70% saw their father at least weekly, with a quarter having no current contact. This high involvement level is consistent with 95% of births being jointly registered by the mother and father (including married couples where only one parent needs to attend registration). In the MCS in 2000-01, 95% of fathers were living with, or otherwise involved romantically or 'just friends' with, the mother at the time of birth.

What questions would we ask in mother interviews to identify OHF sub-categories?

We found in our cognitive interviewing pilot³ that questions routinely used in longitudinal studies to gain data from mothers about OHFs can be difficult to answer. Consistent with methodological literature, accurately capturing frequency (how often), quantity (how much time) and patterns of OHF–child interaction is challenging. Asking about the last fortnight is preferable to asking about a typical week, due to many agreed arrangements being on a fortnightly basis. In-depth interviewer probing showed that OHF–child interactions were often more frequent and extensive than mothers’ initial responses, which tended to exclude ad hoc contact additional to regular arrangements. We could not test our questions about part-time residence of the OHF in the mother’s household as this relatively rare scenario did not apply in our small sample of mothers.

Are there precedents for cohort studies setting thresholds of father involvement for OHF recruitment and retention?

We found examples of studies excluding less involved or no-contact OHFs (according to mother-report) from the scope of recruitment (Table 2). These studies could only approach OHFs for which they had obtained contact details from mothers, so achieved samples are likely to have excluded ‘at risk’ and ‘high conflict’ family situations.

Table 2: Thresholds for OHF recruitment and retention

Longitudinal Study of Australian Children from second sweep	<p>Parents living elsewhere (PLEs) seeing cohort child at least annually</p> <p>Contact details of PLE available from mother/primary parental respondent, who did not explicitly refuse permission for researchers to contact PLE</p>
Early Childhood Longitudinal Study from first sweep	<p>Father seen cohort child at least once in last month, or at least seven days in past three months, or phone or in-person contact with birth mother at least monthly in past three months</p> <p>Main parental respondent is mother, who gave permission for researchers to contact father</p>
Small-scale pilot of OHF survey for Millennium Cohort Study (third sweep)	<p>Any father-cohort child contact at that time</p> <p>Mother gave permission for researchers to send postal questionnaire to father</p>
Millennium Cohort Study from first sweep	<p>Father part-time resident in mother’s household for at least one night weekly</p>

The LSAC’s second sweep postal survey achieved around 40% response from in-contact parents living elsewhere (PLEs) but the study has had successful recruitment and retention from the third sweep onwards using telephone interviews. Approximately two-thirds of in-contact PLEs (half of all PLEs) have been interviewed. Across telephone sweeps, field response rates ranged from 65% to 80%, and mother refusal rates to provide contact details from 3% to 15%.

³ The tested and amended question sets are in our working paper (Goldman et al, 2019).

Success factors may include engaging PLEs as key cohort participants, intensive national study publicity (not specific to PLEs), and particularities of the Australian context. After the third sweep, the study switched from asking mothers for explicit consent to contact PLEs to treating the mother's provision of contact details as implicit informed consent after explaining 'so that we can contact him/her to conduct a short telephone interview and keep in touch by sending him/her newsletters and calendars'.⁴ This was preceded by a motivational script: 'To fully understand [study child]'s development, it is very important to have information from both parents...'.⁴

We adapted this script for our cognitive pilot, and all 11 UK mothers considered a subsequent request for OHF contact details as appropriate, even those who would decline. Some mothers wanted to first gain the father's permission, so interviewers could use a code to indicate a need for follow-up.

LSAC's treatment of PLE recruitment as a priority mirrors the basis for success in father engagement by family services. This depends on the extent to which fathers are perceived as 'core business' by all team members, integrated into briefings, training and support.

'While locating and gaining the involvement of non-resident fathers can be a challenge... it is seen as a priority in LSAC in order to have a more complete picture of the family environments within which children are being raised in Australia' (Baxter et al, 2012, p. 46)

The US Early Childhood Longitudinal Study's (ECLS) first sweep restricted recruitment to involved OHFs (approximately 80% of all OHFs), but achieved a 50% field response rate, even with financial incentives and telephone follow-up. OHF recruitment was discontinued after the second sweep. The MCS had an unsuccessful small-scale pilot survey of in-contact OHFs. A common element in these unsuccessful attempts to recruit in-contact OHFs, including the LSAC's second sweep, is the use of a postal survey of OHFs, or self-completion questionnaires for mothers to give to OHFs. These low-cost recruitment and data-collection modes have consistently given low response rates.

The other cohort study with success in recruiting and retaining OHFs, Fragile Families in the US, did not impose a 'father-child contact' recruitment threshold at any sweep. It set out to interview all birth fathers from the first sweep, with intensive tracking, financial incentives and experienced interviewers. Birth fathers were recruited on postnatal wards and interviewed in hospitals, with telephone recruitment and interviews when needed. Subsequent sweeps used telephone interviews. The response rate for fathers not living with the mother at birth remained at 60% in the fourth sweep, but response from all fathers subsequently declined. The post-birth period is described as a 'magic moment' when, as noted earlier, fathers are most likely to be involved with their new child and the mother, and most likely to engage with the study.

Reflections on use of thresholds

We concluded that the antenatal or infancy sweep of a birth cohort study offers a unique opportunity to use face-to-face or telephone methods (potentially with online methods as fallback) to aim to recruit and interview all birth fathers⁵ in their own right, without setting OHF-child involvement thresholds. Comprehensive father-specific contact details (including stable contact details) can be collected, and consents can be gained for record linkage. Written communications and interviewers can build a relationship between the father and the study. Biological samples could be collected to provide paternal genetic and epigenetic data.

In principle, a UK birth cohort study using birth registration records as a sampling frame could directly recruit OHFs who have jointly registered a birth because their address is compulsory information for registration. OHFs for sole birth registrations would be recruited via mothers. In nearly half of sole birth registrations in the MCS cohort in 2000-1, birth parents were in a relationship or 'just friends'. OHF recruitment would, therefore, be disaggregated according to joint or sole birth registration.

⁴ LSAC wave 7 'parent 1' interview schedule p14, and also used in waves 4, 5 and 6. Downloaded from [Study questionnaires | Growing Up in Australia](#).

⁵ Specific OHFs could be excluded from scope if data collected from the mother indicated a high-risk situation, but we would expect this to apply to a very small minority of cases.

Initial recruitment thresholds aiming to recruit only OHFs in contact with the child (according to mother or child report) may be relevant for cohorts beginning in later childhood, as with an accelerated cohort design, or when OHF recruitment begins at a later childhood sweep, as in the LSAC. By this time, a greater proportion of OHFs (still a minority) will not have been in contact with the child or mother for some time.

Where fathers (including OHFs) have been recruited in the first sweep of a birth cohort study, decisions on using an OHF–child involvement threshold at a particular sweep can fit resources and study objectives. Where father–child relationships are the key focus, there may be best value in using face-to-face or telephone interviews and resource-heavy recruitment strategies (such as monetary incentives, doorstep visits and phone calls by interviewers) in a targeted approach towards more involved OHF sub-categories. Where family separation is a focus, resources can be allocated to data collection from OHFs who have recently separated from the child’s mother. Any ‘core’ or ‘main’ approach could be supplemented by less costly additional exercises such as online surveys of less involved and no-contact OHFs aiming to maintain communication with the study and to collect basic data, especially from those who were previously study respondents living with the cohort child. We did not find evidence about the use of online data collection or ‘virtual’ keeping-in-touch technologies with OHFs.

Whatever decisions are made about each sweep’s recruitment and data-collection strategy, keeping in touch with all fathers enrolled in the study, including OHFs, is crucial to update contact details and to maintain study commitment. OHFs may move home frequently in the period following relationship separation, so more regular keeping in touch may be needed at this time. Mothers can be asked afresh for OHF contact details at each sweep if contact has not been maintained directly with the father, including if she has previously not given contact details, and if the father has not participated previously. The frequency and types of OHF involvement and residency status, and friendliness of his relationship with the mother, may change through childhood.

‘biological fathers who have minimal contact with their children at one point in time may reappear later and become important presences in the children’s lives... Such fathers could be added to the study’ (National Center for Education Statistics, 2001, p. 32)

Are there precedents for cohort studies tailoring recruitment and retention strategies for differentiated subsets of OHFs?

We found that cohort studies internationally have generally used a single strategy for OHF recruitment and retention. Findings of methodological research on participant engagement in cohort and panel studies are rarely differentiated by respondent characteristics. Yet bespoke and more expensive methods may be affordable for relatively small subsets of respondents (Bryson et al, 2017).

The disaggregated recruitment strategies we found were based on whether the father was expected to visit the mother’s household during the fieldwork period. The preferred recruitment option in the ECLS’s first sweep, but unsuccessful, was to leave a self-completion OHF questionnaire with the mother if the father was likely to visit in the next seven days, along with an unconditional incentive. In contrast, Life Study had planned an option for direct OHF recruitment at the mother’s address (for an interview) if he would be there during fieldwork. Survey information packs would be left with mothers to pass on only as a last resort. In the MCS postal survey of OHFs, when mothers insisted that they preferred to pass the questionnaire to fathers, interviewers attempted to collect father contact details so he could directly receive reminders.

We found no examples of written scripts for participants or interviewers being tailored according to OHF involvement with the cohort child. Life Study had planned to use alternative scripts to ask mothers about study recruitment of the father according to whether the OHF was the mother’s partner.

Reflections on tailoring of recruitment

Tailoring of recruitment according to the level and type of OHF–child involvement is appropriate if OHF interviews are timed for closely after mother interviews, because OHF involvement may change in the intervening period.

Methodological literature on recruiting OHFs and other participants in longitudinal studies and our consultation with international cohort studies underline the importance of interviewer skills, experience and training. Persistence and a ‘can do’ attitude are important, in combination with sensitivity to delicate relationships in some separated families. In service contexts, success in engaging fathers and gaining mothers’ support as gatekeepers is affected by practitioners’ language and behaviours, influenced by their beliefs (Burgess, 2009; Symonds, 2019). Therefore, to build appreciation of the value of data from OHFs, recruiter and interviewer briefings could draw attention to the high prevalence of part-time co-residence and regular involvement among so-called ‘non-resident’ fathers.

Discussion

Our scoping study has uncovered potential in a disaggregated approach to recruitment and retention of OHFs in cohort studies. It can also inform OHF engagement in cross-sectional surveys and household panel studies. Sample sizes are likely to be sufficient in a large national cohort study to calculate separate response rates for more involved OHF sub-categories and to analyse them separately. There are precedents among cohort studies in setting recruitment thresholds based on father–child involvement, but few examples of targeting or tailoring recruitment strategies and participant communications. We concluded that, if possible, first sweeps of birth cohort studies should endeavour to recruit all birth fathers, including OHFs, but revisit inclusion at later sweeps according to resources and research questions.

Our work was designed as a first-stage study to give high-level conclusions and sketch a way forward. It does not provide evidence on whether or not a disaggregated approach is effective in improving response rates and representativeness. The methodological literature did not suggest that there is any single consistently effective OHF recruitment and retention method across studies. As the next stage, a larger scale feasibility study for a specific cohort study can develop implementable OHF recruitment and fieldwork strategies grounded in the practical context and aims of that study, including a disaggregated approach if relevant. Effectiveness would be assessed, including analysis of response and response bias for more and less involved OHF sub-categories. If the budget allowed, specific methodological pilots and experiments would be valuable.

We propose that development of strategies to recruit and retain OHFs would be preceded by qualitative research with these fathers. We did not find any published qualitative research with OHFs about their participation in cohort studies. We found equivalent qualitative work with fully resident fathers about their own participation, and with ‘lone mothers’ about OHF participation.

In the context of cohort and panel studies in the UK, statisticians have discussed to what extent samples for causal analysis in longitudinal studies (compared with nationally representative prevalence estimates) need be random samples (Smith and Dawber, 2019; Benzeval et al, 2019). A disaggregated recruitment strategy could extend to quota sampling or responsive sampling strategies incorporating OHF sub-categories, for example for specific sweeps.

If OHFs can be successfully recruited and retained in high-quality research studies such as national birth cohorts, this will enable social science and health researchers analysing these datasets to strengthen evidence about fathers’ social, behavioural, genetic and epigenetic influences on children, and to influence policies and services to support children, fathers and mothers.

References

- Baxter, J. (2018). Parents living elsewhere data. Longitudinal Study of Australian Children Technical Paper No. 22. Melbourne: Australian Institute of Family Studies.
- Baxter, J., Edwards, B. and Maguire, B. (2012). New father figures and fathers who live elsewhere. Canberra: Australian Government Department of Families, Housing, Community Services and Indigenous Affairs.
- Benzeval, M., Burton, J., Bollinger, J. and Crossley, T. (2019). The representativeness of Understanding Society. Understanding Society Working Paper Series, No. 2020-08. Colchester: University of Essex.
- Bradshaw, P., Reid, S. and Cleghorn, N. (2008). Growing Up in Scotland study: research with non-resident fathers: options review. Edinburgh: Scottish Centre for Social Research.
- Bryson, C., Purdon, S. and Skipp, A. (2017). Understanding the lives of separating and separated parents in the UK: what evidence do we need? London: Nuffield Foundation.
- Burgess, A. (2009). Fathers and parenting interventions: what works? Abergavenny: Fatherhood Institute.
- Davis-Kean P., Chambers R., Davidson, L., Kleinert, C., Ren, Q. and Tang, S. (2017). Longitudinal studies strategic review. Swindon: Economic and Social Research Council.
- Goldman, R. and Burgess, A. (2018). Where's the daddy? Fathers and father figures in UK datasets. Marlborough: Fatherhood Institute.
- Goldman, R., Bradshaw, P., Burgess, A. and Vosnaki, K. (2019). Working paper – recruitment and retention of birth fathers in split-off households: methodological considerations for cohort and longitudinal studies. Swindon: UK Research and Innovation.
- Jessop, C., Hussey, D., Lubian, K. and Wood, M. (2019). Using a targeted design and R-indicators to enhance sample quality in a web/CATI longitudinal study. Swindon: UK Research and Innovation.
- Kiernan, K. (2006). 'Non-residential fatherhood and child involvement: evidence from the Millennium Cohort Study'. *Journal of Social Policy* 35(4): 651-669.
- Kiernan, K. (2016). Fathers and partners in life study: groundwork, outcomes and lessons. London: Nuffield Foundation.
- Lynn, P. (2020). Methodological briefing – methods for recruitment and retention. Understanding Society Working Paper Series No. 2020-07. Colchester: University of Essex.
- National Center for Education Statistics. (2001). Measuring father involvement in young children's lives: recommendations for a fatherhood module for the ECLS-B. Washington, DC: US Department of Education.
- Park, A., Calderwood, L. and Wong, E. (2018). 'Participant engagement in longitudinal studies: current practice, opportunities and challenges'. *Social Research Practice* 7 (Winter 2019).
- Smith, P. and Dawber, J. (2019). Random probability vs quota sampling. Southampton: University of Southampton.
- Symonds, J. (2019). 'Making fathers relevant: how practitioners include both parents in talk about parenting programmes'. *Child & Family Social Work* 25(1): 144-153.
- Toulemon, L. and Penniec, S. (2010). 'Multi-residence in France and Australia: why count them? What is at stake? Double counting and actual family situations'. *Demographic Research* 23(1): 1-40.
- Waller, M. and Jones, M. (2014). 'Who is the residential parent? Understanding discrepancies in unmarried parents' reports'. *Journal of Marriage and Family* 76(1): 73-93.

How to replace large-scale, continuous, face-to-face household surveys during the COVID-19 pandemic: the Scottish Health Survey example

Shanna Christie, NatCen Social Research
Joanne McLean, ScotCen Social Research

Acknowledgements

Thanks to the survey teams at ScotCen/NatCen and the Scottish Government for their contribution to the article as well to developing and implementing the SHeS 2020 and 2021 telephone surveys.

Abstract

The Scottish Health Survey (SHeS) is a large-scale household survey of the health of people living in Scotland. It is conducted annually. Data is collected over a 12-month period by field interviewers who gather the information in participants' homes. The survey was first conducted in 1995, then again in 1998 and 2003, and conducted annually and continuously for 12 years since 2008. However, when the pandemic resulted in halting face-to-face (F2F) fieldwork in March 2020, it was not possible to continue the survey using the same methods. At the same time, the response to the pandemic led to an increase in demand for high-quality, robust data particularly about health. Consequently, the Scottish Government committed to continuing the SHeS. This research note describes the research design we used for the 2020 survey and the learning from this which informed the design of the 2021 survey. It also discusses next steps and considers the challenges of analysing and interpreting the data in the context of the pandemic, using alternative methods, and analysing data given changing modes.

Introduction

The **Scottish Health Survey** (SHeS) is an annual household survey of the health and health behaviours of the population living in private households in Scotland.¹ It aims to provide health data not available from other sources, and is used to monitor national health indicators.

¹ SHeS is funded by the Scottish Government. ScotCen has been awarded the 2018–2021 contract in collaboration with:

- The MRC/CSO Social and Public Health Sciences Unit, University of Glasgow
- The Centre for Population Health Sciences, University of Edinburgh
- Global Academy of Agriculture and Food Security, University of Edinburgh
- NIHR BRC Diet, Anthropometry and Physical Activity Group, University of Cambridge (Intake24)
- Office for National Statistics (F2F fieldwork)

Like many large-scale population surveys, SHeS is complex, involving multiple data-collection methods. Key to its success is the work of face-to-face (F2F) interviewers who are skilled in motivating people to participate and supporting them through the lengthy interview process.

It involves:

- ▶ Conducting interviews with all adults in the household and up to two children
- ▶ Collecting sensitive information about physical health conditions, mental health and wellbeing, alcohol and drug consumption, diet, loneliness and social capital
- ▶ Collecting objective physical measurements (for example height, weight and blood pressure) and biological samples (for example saliva)
- ▶ Data linkage with NHS data on visits to hospital; diagnosis, treatments and hospital stays for various conditions; registration with a general practitioner; date and cause of death

A random stratified sample² is drawn from the postcode address file (PAF). Each year a response rate of around 60% is achieved, and over 5,000 adult and 2,000 child interviews conducted. The probability sampling approach used, and the achieved sample size is crucial in providing Scottish population health estimates as well as examining health data for sub-groups (for example particular age groups or those living in areas of deprivation). It is the 'gold-standard' in survey research, and is widely used by public health policymakers, academics and practitioners.

The pandemic and associated physical distancing restrictions generated immense challenges in continuing to conduct large-scale surveys such as the SHeS. However, the demand for high-quality, robust data, particularly in the health field, has arguably never been greater. This research note details the challenges the research team working on SHeS has faced so far during the pandemic and how we overcame these to ensure that the SHeS remains the main source of high-quality public health data in Scotland for now and in future.

SHeS 2020

With F2F fieldwork suspended on 17 March 2020, SHeS, in its 16th year of the series, halted. We rapidly developed a telephone survey and conducted fieldwork between 5 August and 23 September. The telephone survey aimed to quickly capture national data on the health of the population and how lifestyle factors and health-related behaviour were associated with this during the pandemic. It focused on national health indicators. While retaining many SHeS features, it differed in approach and questionnaire design:

Approach

- ▶ The development timetable was tight, two months compared to around nine months for F2F
- ▶ Change of mode from F2F to telephone
- ▶ Although the same sampling frame was used (PAF), the sample design differed in that it was un-clustered³ and was based on participants opting-in on receipt of a letter which was assumed to achieve a response rate of around 15% rather than the usual 60% response rate with the F2F opt-out approach
- ▶ Fieldwork data collection was conducted as a six-week snapshot with an adult interview target of 1,000 compared to the usual 13.5 months with an interview target of 5,112 for F2F
- ▶ Children were not included in the telephone survey but have been included in the F2F survey
- ▶ The interview was shorter to encourage participation (30 minutes compared to 45 minutes for F2F) with questions prioritised for inclusion by the Scottish Government
- ▶ Interviewers read out response categories as show cards could not be used (SHeS relies on response cards for numerous questions)

² It is also clustered in the F2F approach to make the fieldwork more efficient.

³ F2F samples are typically clustered in order to make fieldwork more efficient whereas this approach is not necessary for a telephone survey as interviews are conducted from the interviewer's home.

Questionnaire design

- ▶ We reduced some modules (for example smoking, alcohol) but we ensured that we kept the questions needed to measure key indicators
- ▶ We substituted the physical activity module with a shorter, validated measure
- ▶ We made adaptations to account for Covid-19 (for example we changed answer options to employment questions, and we added questions on Covid-19 diagnosis, shielding and changes to health-related behaviours since lockdown)
- ▶ Questions considered high priority but usually included in the self-complete in F2F (for example mental health) were asked by the interviewer
- ▶ Heights and weights were self-reported.

Impact of the changes

Adapting the data-collection method meant that the 2020 telephone survey estimates were likely to differ to some extent from previous F2F estimates. It was difficult to separate the impact of the methodological changes on the 2020 survey estimates and actual changes in population health and behaviours as a result of the pandemic. Therefore, the findings from the 2020 telephone report were classed as experimental statistics.

Although many of the findings were close to usual estimates and others differed in a way that might be expected in the context of the pandemic, some more than others appeared to be affected by the change of mode. The most notable example was the estimate for current smoking prevalence among adults. This was 9% in the 2020 telephone survey whereas prevalence in recent F2F SHeS has been between 17% and 19% since 2017, albeit with a general downward trend. The finding from the 2020 telephone survey is probably an underestimate that could be largely attributed to methodological differences. For instance, in the 2020 telephone survey, younger adults were asked by interview about smoking rather than by self-complete; self-selection bias may also have contributed to lower reporting; the most deprived areas, where smoking prevalence tends to be much higher, were under-represented in the achieved sample, and weighting the data did not seem to fully adjust for this suggesting that we did not obtain a representative sample of those in the most deprived areas.

Unravelling the impact of the methodological adaptations is complex, not least in the context of the ever-evolving and multi-factorial impacts of Covid-19 on different population groups at different times. As we progress with the 2021 survey, we will analyse further to understand and account for this on our survey data. For now, we know that, in the 2020 telephone survey, amongst others, the key impacts on estimate differences were:

- ▶ Low response rates due to the telephone-only approach (F2F typically much higher)
- ▶ Less representative sample as a result of the recruitment approach (F2F achieves a more representative sample)
- ▶ Switch in mode from self-complete to interviewer-administered meaning less privacy to answer sensitive questions (such as about mental wellbeing)

The short development phase did not allow for any substantial adaptations to be made to overcome mode-effects. However, we undertook a full mode-effect analysis following the launch of the 2020 telephone survey allowing us to consider predicted mode-effects alongside findings. This involved assessing questions that were likely to be affected by the change of mode, from F2F and self-complete to telephone and interviewer-administered. We identified a range of possible mode-effects including:

- ▮ Less opportunity for interviewer/participant rapport can lead to less inclination to provide sensitive information which can lead to under-reporting of mental health problems for example, this would be compounded by the absence of a self-complete mode
- ▮ Social desirability bias (participants giving answers that they think will be perceived as more acceptable, in this context healthier) more likely than in F2F
- ▮ Potential for over-reporting of good general health
- ▮ Self-reported heights and weights can lead to under-reporting of overweight and obesity
- ▮ Higher risk of participants satisficing (giving the minimum information required to move through the interview more quickly) than in F2F, more likely on complex questions with many response options or those that involve a calculation

Further factors that could have affected estimates were the time of year and the context of Covid-19. For example, these may both have affected perceptions of general health (summer and lifted restrictions (before the second wave) = good, not having Covid-19 = good) and may also have provided unrepresentative data on the health-related behaviours that tend to have a seasonal pattern (for example physical activity).

Given the impact of other factors such as sample design (see below) and the possible impact of Covid-19 on measures such as mental health, the extent to which differences in estimates could be attributed to mode-effect cannot be quantified (a parallel run of telephone and F2F methods would be the only way to fully assess this). This highlights the importance of establishing an alternative method for SHeS, when F2F is not possible, that mitigates as fully as possible for alternative mode-effects.

Self-selection bias within the 2020 telephone survey presents probably the greatest challenge to comparability. It is influenced by the demographic profile of the respondents (for example a higher proportion of females participated in the telephone survey than the 2019 F2F, those aged 55 and over were over-represented compared to the 2019 F2F, at least one person was interviewed in 10% of eligible households in the most deprived Scottish Index of Multiple Deprivation (SIMD) quintile compared to 20% in the least deprived quintile). Demographic differences can, to a large extent, be controlled for through weighting. However, opt-in bias to a health study is likely to have been influenced by health-related factors (those who perceive that they are healthy and have healthy lifestyles may be more willing to participate than those who do not) and the effect of this cannot be determined in the data or weighted for. So, achieving a representative sample is more challenging for telephone survey than for a F2F survey. In 2021, achieving a more representative sample within Covid-19 restrictions is a top priority.

The 2020 telephone study, therefore, provided a valuable insight into the health of the population in Scotland (including those who have not used health services) during the pandemic but it stands separately from the SHeS time series. It may have value as a test-phase in the SHeS time series. Learning from the 2020 experience, SHeS 2021 aims, under 2021 Covid-19 restrictions, to provide population health data that is as close as possible to the F2F survey, such that the data can be more reliably compared with the SHeS F2F time series and/or to provide data that can be compared to future SHeS telephone surveys should conditions dictate this.

Developing SHeS 2021

We have adapted the 2021 survey based on what we learned from 2020 and some of the innovative methodologies being introduced and tested.

Targeted incentives

We have introduced a targeted incentive strategy (with a differential incentive value) to increase the level of participation among those living in more deprived areas. This is to mitigate against the sample bias by deprivation identified as a source of bias in the 2020 survey and one which was not possible to fully correct in the weighting. We considered the ethical issues associated with the targeted incentive strategy as part of our application, and argued that it was necessary to boost response in more deprived areas in order to produce more representative data used to improve services, often for harder-to-reach groups. Based on our experience from the 2020 survey, it should be possible to adjust for other biases in the sample (for example age and gender) by weighting the data, as we also do when the survey is conducted F2F.

Reinstatement of the full questionnaire

The 2021 telephone survey uses the full questionnaire. This means that the interview will take longer to complete than the 2020 telephone survey (by around 15 minutes). We weighed up the drawbacks of the longer interview, in participant burden and cost, against the need for the survey measures to be the same as those in the F2F survey. This will enable us to derive the same variables in the 2021 data as previously used in the F2F survey time series (for example on physical activity). Also, including the full questionnaire, means we can quickly switch to or offer a F2F interview should restrictions allow and if there is public acceptability for F2F interviewing.

Response cards

We are reintroducing response cards for 2021. We will send these to participants by post after they have opted into the survey. (We will trial an online approach on a larger scale as findings from the pilot suggest that this is a viable option). The response cards will enable participants to select their answer(s) from a pre-defined list, often with numerous or complex response options, as they do in the F2F SHeS. This will negate the requirement to amend the F2F questionnaire (as was the case in 2020) to ensure that the measures are consistent to enable some comparison of trends.

Including children

Children are included in the 2021 telephone survey because of the demand for data on children's health and lifestyles both during and after the pandemic. To collect a sufficient child sample, a child boost sample is being used for the telephone survey (as is the case when the survey is conducted face-to-face).⁴ Around half of the children taking part will be from the child boost survey, and the other half will respond as part of the core sample along with adult members of the household. Up to two children per household can participate with random selection determining which two if there are more than two children in the household. Parents/guardians will answer on behalf of children aged 0 to 12; children aged 13 to 15 can answer the questions themselves (with a parent/guardian's consent).

⁴ The child boost sample is included in the F2F sampling strategy and a similar approach was taken for the child boost sampling in the 2021 telephone survey. At addresses selected for the child boost sample, children only are included in the survey – other members are not included as they are in the main sample.

Concurrent interviewing

As the survey is a household interview with all adults and up to two children eligible to take part, it has always been possible in the F2F approach for multiple people to be interviewed concurrently.⁵ This was not straightforward to implement for the telephone approach. Proxy interviews are not permitted on SHeS. Therefore, it was not possible/desirable for interviewers to ask one member of the household the question on the phone and for that participant to then ask and collect the information from other household members. Feedback from the pilot indicated that allowing some form of concurrent interviewing would make the interview process more efficient. We have introduced a modified approach to concurrent interviewing for the 2021 telephone survey: up to three people can be interviewed concurrently if the household can use speakerphone capabilities.

Online self-completion questionnaire

In response to the mode-effects analysis of the 2020 telephone survey (where questions previously asked in self-complete were asked by the telephone interviewer), we have developed an online self-completion questionnaire for the 2021 survey enabling the more sensitive questions to be moved back to self-complete with no significant mode-effect expected.

This will be introduced at the end of the telephone questionnaire. Participants who agree to complete this element will be asked for a mobile number and email address so that they can be sent a unique link to their self-complete questionnaire. They will then be expected to complete it independently after the telephone interview. Two reminders will be sent to non-completers. A paper version is also available.

SHeS 2021 data and beyond

We will monitor the impact of these adaptations for practicality of administering them and the quality of the achieved sample and data. The adaptations will also enable us to enhance the telephone element of the survey which may be needed alongside F2F if members of the public are reluctant to take part in a F2F interview. We anticipate that some of the enhancements, such as online self-completion, will be included in the F2F survey if a suitable response rate is achieved for this element.

By implementing these adaptations, it should be possible to draw closer comparisons between the 2021 and previous years of SHeS than in the 2020 telephone survey. This should enable us to start examining the impact of the pandemic on population health and present evidence on the interplay between health behaviours and lifestyle risk factors and key health measures during the pandemic. This is critical data which is in high demand from a variety of stakeholders. In the longer-term, the true value of the SHeS time series is that the return to F2F data collection, or as close to it as possible, will be one of the best sources of information on whether Covid-19 has a lasting impact on population health and, if so, the nature of this.

Many of the issues and challenges discussed here are faced by researchers and statisticians working on other large-scale surveys with trends going back decades.

The following questions are being, and will continue to be, debated for years to come:

- For large-scale F2F surveys, where the methodological approach had to change during the pandemic, what has the impact of changes in mode been (that is from F2F to online/telephone) and how does this compare between surveys?
- What has been the impact of changes in sampling and recruitment (for example from opt-out to opt-in)? Will survey sample profiles achieved during the pandemic continue to differ from those achieved in F2F surveys and how can we continue to moderate for this?
- What are the short, medium and long-term effects of the pandemic and the social restrictions on population health, lifestyle factors associated with health and other areas of life (transport, public attitudes, work and income and so on)?

⁵ Up to four people from the household can take part in an interviewing session with questions asked concurrently.

While working through these research questions is a formidable task, there are positives:

- ▶ There is an unprecedented demand for robust data particularly in the health field
- ▶ The impetus to transform, adapt and innovate in survey methodology has probably never been greater
- ▶ The amount of data, spanning so many topic areas, and going back decades, provides researchers with a baseline with which to measure the long-term impact of the pandemic, one of the biggest challenges faced by a generation, once it is feasible to return to a closer model of F2F fieldwork