Designing sampling strategies for social qualitative research: with particular reference to the Office for National Statistics’ Qualitative Respondent Register

Amanda Wilmot
Office for National Statistics, UK

1. Introduction

The design of a sampling strategy for a qualitative study is as important as that for quantitative inquiry. A well-defined sampling strategy that utilises an unbiased and robust frame can provide unbiased and robust results.

There is a tendency, particularly within a quantitative environment, to consider that the sampling strategy for qualitative research is of lesser importance to that where statistical inference is required. Indeed, it is not unknown for those unfamiliar with qualitative research methods to suppose that no more than a convenience sampling strategy is applied. That is to say the researcher makes no attempt, or only a limited attempt, to ensure that the sample is an accurate reflection of the population.

This paper attempts to show why it is as important to develop a robust sampling strategy, from a well-constructed sampling frame, for qualitative research practice. The paper also discusses how the Office for National Statistics (ONS) puts theory into practice using a qualitative ‘Respondent Register’ developed for use for sample frame construction for qualitative social research.

1.1 Understanding qualitative research

It is difficult to discuss the design of qualitative sampling strategies without first discussing the nature and purpose of qualitative research, and how its approach differs from quantitative research.

The terms ‘quantitative’ and ‘qualitative’ research are used as overarching categories covering a wide range of approaches and methods within each. However, the very bases of those approaches differ.

Quantitative research, by definition, implies a measurement or numerical approach. The methodology employed is based on the testing of hypotheses deduced from theory. Using statistical inference the results may be generalised to the population.

Qualitative research aims to provide an in-depth understanding of the world as seen through the eyes of the people being studied. It aims not to impose preordained concepts; hypotheses and theory are generated during the course of conducting the research as the meaning emerges from the data. Statistical inference is not the objective, although within government, results are used to inform policy and therefore some form of generalisation or transferability is implicit.

Qualitative research may stand alone or in conjunction with quantitative research, used before, along side or after. Where the studies are associated then the sampling strategy for both should be considered at the same time.
2. Informing the design of a qualitative sampling strategy

Different qualitative sampling strategies may be used at different stages of the research, or for different research purposes.

Questions which the researcher should ask themselves at the outset, and which will inform the design of the sampling strategy, are the similar for both quantitative and qualitative research. They are:

- What are the research objectives?
- What is the target population?
- Who should be excluded from the sample?
- Who should be included in the sample?
- What is the budget?
- What is the reporting time period?
- How many qualified researchers are available to work on the project?
- What sampling technique(s) should be employed?
- How are the data to be analysed?
- What data collection methods should be employed?
- What are the sample criteria?
- How long will the interview be?
- What size should the sample be?
- What should be used as the sampling frame?
- How should potential respondents/participants recruited?

All of the above are interdependent, however some of the questions require a more detailed discussion with regard to their application in a qualitative research environment.

2.1 Resources

Time, budgetary and other resource constraints may impact on the qualitative sample design but should not be allowed to undermine it. The nature of the data collection method (e.g. cognitive, in-depth, or group interview), the human resources available to the project and their skills base are also important considerations.

2.2 Sampling technique

The sampling technique employed is a crucial element of the overall sampling strategy.

At this point it is important to understand why probability sampling is inappropriate for qualitative research. In probability sampling members of the research population are chosen at random and have a known probability of selection. Groups are represented in the sample in their true proportions; or, where unequal probabilities are used the data are reweighted back to the true proportions. The aim is to produce a statistically representative sample, suitable for hypothesis testing.

Qualitative research uses non-probability sampling as it does not aim to produce a statistically representative sample or draw statistical inference. Indeed, a phenomenon need only appear once in the sample.
Purposive sampling is one technique often employed in qualitative investigation. With a purposive non-random sample the number of people interviewed is less important than the criteria used to select them. The characteristics of individuals are used as the basis of selection, most often chosen to reflect the diversity and breadth of the sample population.

However, there are different approaches to purposive sampling some of which focus on different aspects of the sample members, cases are chosen because they are considered more extreme, for example. One form of purposive sampling is ‘theoretical sampling’, developed from the ‘grounded theory’ approach (Glaser and Strauss, 1967). The term ‘grounded theory’ expresses the idea that theory is generated, through an iterative process, involving the continual sampling, collection and analysis of data to inform the next stage of the sample design, until ‘theoretical saturation’ is achieved; that is, no new ideas or theories emerge. The iterative nature of the theoretical sample design is important. It gives the researcher the opportunity to analyse the data as the sampling progresses and means that the researcher can add to or change the emphasis of the sample design, and in doing so ensure robustness of the theories generated. It is therefore valuable to have considered the analysis technique early on in relation to the qualitative sampling strategy.

Sometimes theoretical hypothesis generation is not the primary aim of the research. Where the sample population is clearly defined, such as when testing already operational survey questions, and where resource and time constraints are in place, then a more constrained purposive sampling strategy can be devised that avoids iteration and does not necessarily achieve saturation, on the grounds of diminishing returns.

Whatever approach is used, some advance knowledge of the population under investigation is necessary when carrying out purposive sampling.

2.3 Sample criteria

A decision will be required as to the sample selection criteria. That is, what characteristics will need to be reflected in the sample population to address the research question. The decision on which criterion to use will be informed by the policy advisor and other subject specialists, as well as a review of the current literature. The researcher will need to know whether particular sub-groups need to be included to ensure breadth. The criteria used may be based on demographic characteristics or behaviours or attitudes, and will need to be prioritised if purposive sampling is to be employed. This is partly influenced by the fact that qualitative research is often, but not always, based on a relatively small number of cases so it may not be possible to include all of the sample criteria in the sample design. Some criteria may be considered more important than others in relation to the research objectives.

2.4 Interview length

The intensity and therefore the length of the qualitative interview will also impact on the design of the qualitative sampling strategy and the decision of sample size. Longer interviews may provide more data than shorter interviews. A decision may be taken, depending on the nature of the study, to conduct a larger number of shorter interviews or a smaller number of longer interviews.

2.5 Sample size

A feature of qualitative sampling is this fact that the number of cases sampled is often small. This is because, as mentioned earlier, a phenomenon only need appear once to be of value. There is no need for scale as there is no need for estimates of statistical significance. Furthermore, because
qualitative investigation aims for depth as well as breadth, the analysis of large numbers of in-depth interviews would simply be unmanageable because of a researcher’s ability to effectively analyse large quantities of qualitative data. However, the small-scale approach only works if the researcher has a strong sampling strategy (Ritchie and Lewis 2003).

The issues that should be considered when determining the sample size for qualitative investigation are dependent on the heterogeneous or homogeneous nature of the sample population, or requirements of the data collection methods employed; for example, focus groups tend to be more productive and manageable if participants have some commonality.

The number of selection criterion required and the degree to which criteria are nested (dependent on whether certain characteristics are to be controlled for e.g. age), are important considerations. The intensive nature of the study; whether multiple samples are required, the inclusion of a control sample for instance; and the resources available to conduct the study, are also important for determining sample size.

To provide some idea of the scale of qualitative investigation one might expect to achieve between 20 and 50 interviews for a one-to-one investigation and around 60 to 100 participants at group interview, depending on the research question.

The size of the sample required will of course also feed into the decision about type of sampling frame to use.

### 2.6 Sampling frames

A sampling frame is a list or map that identifies most units within the target population. (Missing units are referred to as undercoverage.)

#### 2.6.1 Frame evaluation

When evaluating the effectiveness and efficiency of any sampling frame for qualitative research, it is important, as with quantitative research, to consider whether the frame is comprehensive. That is, all of the target population are included. The full range of dimensions, and information needed to inform the sample selection, should be covered. This is because sections of society missing from the frame may have different characteristics and indeed different behaviours, opinions and attitudes from those covered by it. This undercoverage may affect the results if associated with the subject of enquiry for example, phenomena may not be raised or survey questions not tested thoroughly.

It is however, also important to consider overcoverage. Sample members may be listed more than once, or the list may contain members considered out of scope for the purposes of the study. Their inclusion in the study could impact on the findings and indeed on resources and the ultimate cost of the project.

Furthermore, the frame should also contain sufficient numbers in each sub-group to provide the sample size required, as not everyone who is eligible will be willing to take part. Three or four people may be contacted who fulfil the sample criteria before one agrees to take part.

A practical consideration is whether the frame can be easily manipulated in order to identify those with the relevant characteristics.
As with frames for quantitative research, geographical clustering is important because if the population are highly dispersed then fieldwork will be more resource intensive. It is also important to know whether the potential respondent contact details are complete and up-to-date. Lastly, whether the time and cost involved in using the frame is justified.

2.6.2 Types of frame

There are basically two types of frame (or list) available for social research practice. These are existing lists that can be used as frames, perhaps after some manipulation, or frames that need to be constructed.

Existing frames

Existing frames usually comprise records which were constructed for administrative purposes, for example, a published list of GP surgeries. As such, they tend not to have been designed with research purposes in mind and may not be very well maintained from a research perspective. Furthermore, some administrative records, for example, benefit records, will be covered by data protection and confidentiality issues, which can make them difficult to access.

Existing survey samples can provide a frame, but they may also not have been designed with the current research interest in mind. (Although if the qualitative research is a follow-up to a quantitative survey and has been planned as such, then the qualitative sample criteria can be built into the quantitative survey in advance.) Existing survey samples may be affected by undercoverage and response bias, not least because they are dependent on the survey response rate.

For ethical reasons, the survey respondent will have been required to give permission for a survey organisation to contact them again for future research. Not everyone taking part in the original survey will have given their consent, particularly if the first survey interview was long or not interesting. There are therefore additional attrition issues to be considered. Ethical considerations might also include the burden on the respondent taking part in more than one interview.

Constructed frames

Where an existing frame or list is not available then the researcher may have to create their own. Sampling may then take place directly from the constructed frame or indirect methods of obtaining a sample may be used to either construct or to supplement the frame. In qualitative, as with some quantitative, studies the actual sampling and recruitment may take place in the field, as potential respondents fulfil the sample criteria and the overall sampling strategy.

The use of household or organisation screening methods is not uncommon. However, both are subject to the ‘gatekeeper syndrome’ where the person initially contacted may not have available, or may be unwilling to provide, information about others. Particularly where an organisation is concerned, there may be a tendency for the organisation’s representative to select potential respondents who are in more regular contact with the organisation or who may present the organisation in a good light.

Another method of frame construction that can be useful, particularly where the population is clustered geographically, is focussed enumeration. Here interviewers call at a certain number of addresses living either side of respondents who fulfil specified criteria. They will administer a screening questionnaire and attempt to gain an interview with people who may also fulfil the specified criteria. This method is more effective where certain characteristics are known to cluster such as high income earners or certain ethnic groups.
Snowballing, where respondents tell the researcher about others they know with the same or specified characteristics, is a method also used in qualitative research. As potential respondents often know each other, the disadvantage with snowballing is that a sampling frame created using this method alone may be prone to bias. This is because those who know each other may have similar behaviours and attitudes or may influence each other in relation to the research. Those that are missed may have quite different characteristics. As a result the sample may not be particularly diverse and important phenomena may be missed. It is therefore usually used as a supplement to other methods.

Advertising for potential respondents is not advised unless other methods of frame construction are unavailable. If used, a rigorous screening process should be in place but even then the sample members are, to a certain extent, somewhat self-selecting.

2.7 Respondent recruitment

The method of respondent recruitment and its effectiveness is also an important part of the sampling strategy.

2.7.1 Recruitment specification

For each recruitment project, a comprehensive set of specifications should be drawn up, particularly if an external recruitment agency is to be used.

It is therefore of paramount importance to provide a detailed specification of the project to those doing the recruiting. This detailed specification is also useful for the research commissioners in order to alleviate fears about the quality of the research.

The recruitment specification will detail the recruitment strategy and explain the purpose of the research and how the research will be conducted and for whom. It will also provide a schedule of events, as well as the detail of any screening questionnaire. Following this, a quota is drawn up according to the sampling strategy and, as people are contacted and meet the criteria and agree to take part, a note is made of which quota they fill. This in turn informs the recruitment of the next person on the list.

At this point it is worth briefly discussing the use of recruitment agencies. Many research organisations use professional recruitment agencies to conduct respondent recruitment for qualitative research. This approach can certainly save time and be cost effective since such agencies specialise in recruitment and have the in-house expertise to manage it. However, it is important periodically to take the time to quality assure the work of these agencies, either directly or through discussions with respondents about their experience of the recruitment process.

It is also worth discussing the use of incentive payments in relation to recruitment for qualitative research. Due to the intensive (at least cognitively) and sometimes intrusive nature of the interviews, incentive payments are commonplace in qualitative research. At the ONS this is kept to a minimum and is really no more than a token of our appreciation for the time people have spent talking with us, and to cover any travel costs involved on the part of the interviewee. It is important not to mention the incentive to potential respondents at the outset as there is a chance that someone may sift themselves into the study on that basis alone. Certainly in the United States for example, attendance at focus groups is becoming a regular way of earning money. Such professional focus group attendees are to be avoided as their participation will most likely bias the results.
2.7.2 Documenting outcomes

It is important to document the outcome of the recruitment process in the same way as it is important to document outcomes in quantitative survey research. This is because it is essential to know whether there is any attrition or deficiencies and biases in the sample and the sampling frame used with respect to the research objectives. Furthermore, the information is valuable for informing future research using the same frame.

It is important to know the characteristics of those who refuse to take part, particularly if their refusal is related to the research topic. Furthermore, their refusal may, in part, be a result of the recruitment strategy, which may need revision. Information about non-contacts is important to inform the recruitment strategy, or to check the accuracy of contact details on the frame.

Information about those contacted but who were ineligible for the purposes of the study is useful for checking the effectiveness of the sampling frame. Their agreement to be contacted again in the future could also be sought, providing further valuable information.

3. Office for National Statistics’ Qualitative Respondent Register

At the ONS, respondent recruitment is conducted by the researchers themselves and not by an outside agency. A major part of the work carried out by the ONS qualitative research unit relates to testing social survey questionnaires, including the Census, administered to the general population. We have been investigating ways of constructing sampling frames for qualitative social studies.

The ONS continuous social surveys can provide details of survey respondents but, particularly where the survey is run annually, the information is often difficult to retrieve quickly and on an ad hoc basis. Furthermore, providing a list of potential respondents is not a survey manager’s priority.

We have therefore looked to the National Statistics Omnibus survey to provide a list of potential respondents and their details. The Omnibus is a multi-purpose survey, run on a monthly basis, covers the whole of Great Britain and selects one person living in a private household, aged 16 or over, for interview. Data are available within a few weeks of fieldwork completion.

The Omnibus Survey uses the Royal Mail’s small user Postcode Address File as its sampling frame, which aims to include all private household addresses in Great Britain. A new sample of 100 postal sectors is selected at random each month. The postal sectors are selected with probability proportionate to size and, within each sector, 30 addresses are selected randomly. Around 1800 individuals respond to the survey each month.¹

A potential respondent register is being built from details of those who responded to the Omnibus survey and gave their contact details and permission to be contacted again in the future by the ONS. Every month around 1470 (82%) individual contact details are added to the register. Sampling frames for specific projects can then be constructed from the register.

¹ It should be noted that since April 2005 the number of sampled postcode sectors has been reduced to 67. From this date onwards just over 1200 interviews are achieved each month.
The register is currently stored as an SPSS database. Details of potential respondents remain on the database for a period of four years, after which time they are deleted. This is because around 1/3 of respondents from the original monthly Omnibus sample, are likely to have moved over that time period (a higher proportion among the private renter population), and details may then be out of date. Furthermore, when respondents have taken part in a maximum of two qualitative studies they are not contacted again; so respondent burden is minimised. Contact is usually made by telephone followed by a face-to-face interview.

As discussed earlier, it is important for the researcher to be aware of any potential deficiencies in the register or frames produced from it, particularly if those deficiencies relate to the research question or subject under investigation. The advantages and disadvantages of the register in this respect are detailed in section 3.2.

The ONS respondent register contains the following information about those who agreed to be contacted:

- Around fifty classificatory variables including both household and individual level information such as household type and composition; tenure; number of cars; respondent age and sex; marital, education, health and employment status; socio-economic classification; ethnic group and income band.
- Geographical information such as postcode and postcode sector, Local Authority and Government Office Region.
- Comments from Omnibus interviewers about the household such as how to find the address.
- A note of the Omnibus topic areas covered each month.
- Address details (100% coverage of Omnibus sample agreeing to recall).
- Respondent name (100% coverage).
- Telephone number (96% coverage).
- Email address (19% coverage).
- Number of times respondent contacted via the register.
- Dates when contacted.
- Times of contact.
- Contact outcomes.
- Number of times respondent participated.
- Details of project participated in.
- Dates of participation.
- Name of researcher conducting interview.

The register is kept and maintained in a secure environment, access procedures have been verified and agreed with those responsible for data protection at ONS. Respondent confidentiality is paramount. Only two members of staff have access to the database.

3.1 Geographical mapping

We have recently begun to investigate an enhancement to the register that is intended to make it more efficient as a frame for qualitative purposes. The use of geographical mapping software means that each potential respondent, with the required sample characteristics, can be mapped at postcode level and a located across the country. This means that the spread of potential respondents with the required characteristics can be visualised. Areas where potential respondents
cluster are then visible and can be easily targeted for recruitment. This is especially useful when recruiting for face-to-face focus groups since the groups need to come together in one convenient location. It is more difficult to encourage a dispersed group to meet due to the time and cost involved in getting together. In addition, outliers, that is potential respondents who do not cluster, are also important in assuring breadth of coverage. These people may be different from those who cluster and should also be included in the research, at least in the initial stages.

3.2 Strengths and weaknesses of the ONS respondent register

The major benefit of having a qualitative potential respondent register such as this is the considerable efficiency gains for the office both in terms of quality and resource. However, the decision to use the register will depend upon the research objectives and requires an understanding of the strengths and weaknesses of the register, which are detailed below.

3.2.1 Strengths

*Efficiency gains from a sampling perspective*

- Potential respondent contact details are up-to-date. Telephone numbers are available in 96% of cases which makes the initial respondent contact simple, timely and less costly than contact made by post or in person.

- A large number of demographic and classificatory variables means that selection criteria can be varied and relevant.

- The size of the register (around 47,000 cases) means that sub-sampling for frame construction is both possible and effective.

- There is diversity and breadth of coverage since the original Omnibus sample design was based on random probability.

- When implemented, the mapping software will be useful in identifying where people who meet the sample criteria cluster and additional sampling methods, such as focussed enumeration or snowballing, may achieve a greater success rate.

- Even if the Omnibus respondent is unable to take part in the subsequent qualitative study there may be others in the household who fit the selection criteria and are willing to be interviewed.

- The register can be used as a starting point for frame construction and prior to administering a screening questionnaire.

- The fact that the Omnibus has a random sample design and only those selected can be included for interview means that potential qualitative respondents are not likely to be so called ‘professional qualitative respondents’. Indeed, many have never taken part in social research prior to being interviewed on the Omnibus.

---

*It should also be borne in mind that the original Omnibus sample has a clustered design.*
• Although the register is depleted through use for various studies it is also refreshed each month as a new Omnibus sample is added.

• Outcomes of any recruitment process are recorded which inform subsequent studies and periodic assessments of the completeness of the register can be made by comparing distributions with other data sources. This means that the researcher is informed at the outset of any potential deficiencies in the frame.

Efficiency gains from a research perspective

• Ensures efficient use of research resource as less time is spent identifying or constructing a suitable sampling frame.

• Ensures efficient use of research resource in identifying regions or areas to either target or avoid for recruitment.

• If implemented the additional use of the mapping software will be able to inform the development of the sampling and recruitment methodology.

• The mapping software would also aid decision making related to cost and timeliness. A more dispersed sample will require greater resource input when interviews are to be conducted face-to-face as researchers will have to travel further to interviews.

• The issues that are raised in recruitment, including those related to the geographical mapping information, may reveal issues that can further feed in to understanding during the research analysis or project reporting.

• From a practical perspective the SPSS database is easy to access and to manipulate. Sub-sampling frames can be produced quickly, particularly bearing in mind the number of classificatory variables included on the file. The ease of manipulation also means that it is simple to quality assure the sampling and recruitment process∗.

Efficiency gains from an operational perspective

• The fact that respondents have already taken part in an Omnibus interview and have been visited by an ONS interviewer means that greater confidence can be placed in the personal safety of the qualitative researchers when visiting people in their own homes.

• Because respondents have already taken part in the Omnibus survey and have agreed to ONS recall, respondents feel less pressurised to take part and more confident in the research process.

• Again, because respondents have already been through the Omnibus interview process they generally require less explanation as to what to expect in relation to the practical elements of taking part in the research e.g. making appointments.

∗ It is important to conduct probity checking of qualitative recruitment and interviewing in the same way as is conducted for quantitative inquiry.
• The previous experience of the Omnibus interviewer and comments such as how to find an address are also useful for the qualitative researcher.

3.1.2 Weaknesses

Efficiency losses from a sampling perspective

• The register as a sampling frame is potentially biased in the sense that it only lists those who responded to the original Omnibus survey (65% on average) and then only those who then agreed to be contacted again in the future (80% on average).

• As the Omnibus is a survey of people aged 16 years or over, living in private households in Great Britain, the register can only list people in those circumstances and would not include people living in institutions. This means that the register alone may not be suitable for a study of the elderly or disabled for example.

• Despite the large number of cases listed on the register it may still not be suitable for studies of sub-groups of the general population where numbers could be very small, particular ethnic groups for example.

• Classificatory variables provided by the Omnibus are factual in nature. However, the sample criterion may be behavioural or attitudinal in nature. Therefore a further screening questionnaire would be required.

• Although refreshed each month and containing large numbers of potential respondents, the register may, in the longer term suffer from over use and resulting bias. This is because respondents are removed from the register after taking part in research on two occasions. So if a number of studies are carried out using the same selection criteria this could deplete the frame of certain types of respondent.

Efficiency losses from a research perspective

• Potential respondents have already taken part in an Omnibus interview and although agreed to recall, there is still an increase in respondent burden.

• Respondents contacted to take part in a qualitative study are usually offered a payment in the region of £15 to £30 depending on the nature of the study and data collection method used. Those who took part in an Omnibus interview were not provided with any incentive payment, other than a book of first class stamps. This may raise a question in the respondents mind about why they were not paid to take part in the original survey interview.

• The ease and accessibility of the register may mean that over time researchers become reliant on it when thinking about the most appropriate recruitment strategy, choosing to use the register as an easy option rather than it being the most appropriate option.
4. Conclusion

An effective sampling strategy is as important for qualitative as it is for quantitative research. The research objectives, sample population, resources available and reporting period, sampling and analysis techniques, data collection methodology, sample criteria, interview length sample size, sampling frame and recruitment method, all feed into the design of a qualitative sampling strategy.

Fundamental to the sampling strategy is the choice of sampling frame. Obtaining an appropriate sampling frame from existing lists or constructing a frame can be difficult and time consuming, although there may be no alternative. The decision is very much dependent on the nature of the study and the target population.

However, where studies involve members of the general public and sampling criteria are factual in nature the ONS qualitative respondent register provides a simple yet comprehensive and effective list containing around 47,000 potential respondents and 50 potential selection criteria, from which frames can be constructed. Furthermore, screening interviews can be conducted to aid frame construction, perhaps where relevant research criteria are not included.

The fact that the register contains such a large proportion of potential respondents while in contrast qualitative studies are small in scale, one might question why register bias should be considered at all. But misuse, overuse and poor documentation can certainly affect the quality of the register and sub-sampling of frames, and hence the quality of the sampling strategy employed.

The register is still under development and as such should currently be considered a prototype tool. However, it has proved effective as a source for sample frame construction for a number of social research projects already carried out by the ONS qualitative research unit, providing evidence of its value, encouraging further development and its formal operationalisation as a necessity.
Acknowledgements

Thanks are due to the ONS Omnibus research team, in particular Victoria Peacey, Laura Rainford and Sandra Short, without whose support the register could not have been constructed.

Thanks are also due to Zeeshan Rahman and Colin Hand in Data Collection Methodology Branch at ONS who helped to build and now manage the register.

Finally, thanks to Allen Gower from Statistics Canada for his input with regard to information about respondent recruitment specifications.

References and further reading


Elliot D. (1984), Conversion to the Postcode Address File as the main household survey sampling frame’ ONS Survey Methodology Bulletin No. 16.

Foster K., (1994), ‘The coverage of the Postcode Address File as a sampling frame’. ONS Survey Methodology Bulletin No. 34.


Patton M. Q. (1980), Qualitative Evaluation Methods, SAGE.

Ritchie J and J. Lewis (2003), Qualitative Research Practice: A Guide for Social Science Students and Researchers. SAGE.
