

Alan Bryman, University of Leicester

Like the editors, I have been asked this question countless times. It reflects the fact that there is little definitive and unambiguous guidance in the qualitative research community regarding how large a sample should be. In both my book on social research methods (Bryman 2012) and in offering guidance to students, I tend to prefer to point to a number of factors they should consider (see also Morse 2000). I know this tactic is sometimes disappointing to students, but given the lack of agreement on this issue among practitioners and methodologists, it is the only responsible guidance that can be supplied. In this short note, I will refer to five factors.

First, there is the issue of *saturation*. As is well known, the notion of theoretical saturation derives from Glaser and Strauss's (1967) influential account of grounded theory. There, theoretical saturation is described as a process in which the researcher continues to sample relevant cases until no new theoretical insights are being gleaned from the data. Once saturation is achieved, the researcher would move on to a research question arising from the data collected and then sampling theoretically in relation to that question. As such, the answer to the question 'How large should my sample be?' would be a glib and unhelpful 'Whatever it takes [to saturate your theoretical categories]'. Such an approach to sampling is very demanding because it forces the researcher to combine sampling, data collection, and data analysis, rather than treating them as separate stages in a linear process. It also means that the researcher cannot possibly know at the outset how many cases he or she will need to collect data from, which causes problems when trying to formulate a research proposal or plan or when creating a budget. It is probably this pressure on the researcher that results in the common observation that saturation is often claimed when there is little evidence that it has been employed as a criterion for deciding when to stop sampling (Bryman 2012; Guest et al. 2006; O'Reilly and Parker in press). Guest et al. (2006) conducted an experiment on a corpus of transcripts from interviews with women in two West African countries and found that saturation was attained after twelve interviews. This might appear quite a low figure but the sample was quite homogeneous (women at high risk of HIV) and the research was tightly focused on how the women discussed sex. Further, there have been few guidelines on how to establish whether one has in fact achieved saturation. Bowen et al. (2010) have provided some useful guidance in this regard. They propose two stages which they employed in relation to two health-related projects: an initial sample of around ten cases followed by a further three cases to determine if any new themes emerge. This criterion is consistent with the findings of Guest et al.

A second factor is that it is sometimes suggested that there are *minimum requirements* for sample size in qualitative studies. For example, in *Social Research Methods*, I cite Warren's (2002) suggestion that the minimum number of interviews needs to be between twenty and thirty for an interview-based qualitative study to be published (Bryman 2012: 425). However, I also cite Gerson and Horowitz (2002: 223) as suggesting that 'fewer than 60 interviews cannot support convincing conclusions and more than 150 produce too much material to analyse effectively and expeditiously'. Contrasting these figures (20-30 versus 60-150) strongly suggests that there is quite a lot of variety in what is believed to be the minimum requirement, so that it is unsurprising to find that actual sample sizes vary considerable in qualitative research. For example, Mason (2010) reports that when he looked at the abstracts of doctoral thesis abstracts relating to interview-based qualitative studies in Great Britain and Ireland, he found that the range was 1 to 95 (the mean was 31 and the median 28). Mason also refers to an online article (the link no longer works and I was unable to track it down) which examined 50 articles based on grounded theory and found sample sizes to vary between 5 and 350 (see Bryman 2012 for more information).

Moreover, it is likely that what these figures conceal is that sample sizes will be significantly influenced by a third influence on sample size – the *style or theoretical underpinnings* of the study. Life story research or a study based on Interpretive Phenomenological Analysis is likely to entail a much smaller sample size because of the fine-grained analysis that is often involved. It is simply not necessary to generate a large

corpus of data for such research (much the same applies to conversation analysis – O’Reilly and Parker, in press). Thus, on the one hand, researchers need to be aware that there is a view that there are expectations about minimum sample size in order to be able to publish one’s results; on the other hand, there is very little agreement about what that minimum sample size is! What is almost certainly crucial is that the researcher must be prepared to justify the sample size with which he or she has ended up. One of the advantages of the saturation concept is that it can be used to justify the size of one’s sample.

A fourth factor that is likely to influence sample size is the *heterogeneity of the population* from which the sample is drawn. For some research questions, the population may be quite heterogeneous with a good deal of sub-group variability. It is possible, if not likely, that a researcher will want to capture at least some of this variability in view of the likelihood that it will be associated with significant variability in experiences and world views of participants.

Fifth and finally, the *breadth and scope of research questions* vary quite a lot in qualitative research and this too is likely to influence sample size. A fairly narrow research focus like the one involved in the research by Guest et al. (2006: 62) – ‘how women talk about sex and their perceptions of self-report accuracy’ in Nigeria and Ghana – can be contrasted with that of Butler and Robson (2001: 2146) – ‘the pattern of gentrification in inner London and in particular...the variability of the process’. However, breadth and scope are not entirely objective attributes of a research focus, so there is likely to be some disagreement about appropriate sample sizes along this dimension.

In this brief commentary, I have tried to sketch some considerations that might be taken into account when contemplating sample size for a qualitative study. I am aware that for some readers it may be a frustrating account, but it is better to give a candid point of view than provide numerical or other guidelines which are contentious and therefore likely to be misleading. As I have said, the most crucial thing is to be prepared to justify your sample size and in this briefing I’ve tried to suggest some reflections in this regard. The five factors that I have mentioned can be used to springboards for thinking about and justifying sample size. The other crucial issue to bear in mind is not to make inappropriate inferences from the kind and size of sample you end up with (Bryman 2012: 426-7; Onwugbuzie and Leech 2010).

References

- Bryman, A. (2012) *Social Research Methods*, fourth edition. Oxford: Oxford University Press.
- Butler, T., and Robson, G. (2001). ‘Social Capital, Gentrification and Neighbourhood Change in London: A Comparison of Three South London Neighbourhoods’, *Urban Studies*, 38: 2145–62.
- Francis, J.J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M.P., and Grimshaw, J.M. (2010). ‘What is Adequate Sample Size? Operationalising Data Saturation for Theory-Based Interview Studies’, *Psychology and Health*, 25: 1229-45.
- Gerson, K., and Horowitz, R. (2002). ‘Observation and Interviewing: Options and Choices’, in T. May (ed.), *Qualitative Research in Action*. London: Sage.
- Glaser, B. G., and Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- Guest, G., Bunce, A., and Johnson, L. (2006). ‘How Many Interviews are Enough? An Experiment with Data Saturation and Variability’, *Field Methods*, 18: 59–82.
- Mason, M. (2010). ‘Sample Size and Saturation in PhD Studies Using Qualitative Interviews’ [63 paragraphs], *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 11/3, art. 8, <http://nbn-resolving.de/urn:nbn:de:0114-fqs100387> (accessed 1 October 2010).
- Morse, J.M. (2000). ‘Editorial: Determining Sample Size’, *Qualitative Health Research*, 10: 3-5.
- O’Reilly, M. and Parker, N. (in press) ‘“Unsatisfactory Saturation”: A Critical Exploration of the Notion of Saturated Sample Sizes in Qualitative Research’, *Qualitative Research* (in press).

Onwuegbuzie, A. J., and Leech, N. L. (2010). 'Generalization Practices in Qualitative Research: A Mixed Methods Case Study', *Quality and Quantity*, 44: 881–92.

Warren, C. A. B. (2002). 'Qualitative Interviewing', in J. F. Gubrium and J. A. Holstein (eds), *Handbook of Interview Research: Context and Method*. Thousand Oaks, CA: Sage.

Kathy Charmaz, Sonoma State University*

When novices ask, "How many interviews do I need?" their question likely rests on three presuppositions. First, the question presupposes that the number of interviews answers a researcher's concern about performance, whether this concern is about meeting barely adequate, credible, or exemplary standards. Second, the question presupposes that experts can specify a concrete number of interviews and third, that they would agree on the same concrete number. All three presuppositions are problematic. Forming any answer to the question is more complex than it seems and raises a series of related questions. An answer based primarily on the topic, research purpose, disciplinary traditions, institutional human subjects' reviews, or the researcher's professional goals does not suffice, although such concerns figure in planning an interview research project. Fundamental questions about epistemology must be addressed. What do you seek to know? What do you need to learn? How can interviews inform these questions? A paradox arises: you may not know what you need to find out until you grapple with analyzing your data. Most qualitative interview research is an emergent process of learning about and interpreting research participants' views of their experience. Important foci often remain implicit. Planning solid interview studies entails allowing for following emergent ideas and directions.

A standard answer to the question of how many interviews is that it depends on your research purpose. Might you have multiple purposes that complement or supersede your research purpose? Do you intend to meet a course or doctoral requirement and later present and publish papers from your study? What are the norms of your discipline? Are you aiming for credibility within or across disciplines and professions?

The number of interviews depends on the analytic level to which the researcher aspires as well as these purposes. When researchers pursue straightforward research questions to resolve problems in local practice in applied fields, a small number of interviews may be enough. Guest, Bunce, and Johnson (2006) attempted to answer the question about how many interviews researchers (particularly those in applied fields) needed by conducting an experiment using their codebooks from an earlier qualitative interview study. They aimed to discover the point in data collection and analysis when new data does not alter themes in the code book. Guest, Bunce, and Johnson argue that twelve interviews suffice for most researchers when they aim to discern themes concerning common views and experiences among relatively homogeneous people. Twelve interviews may generate themes but not command respect.

Numerous thematic studies involve synthesizing data and sorting them into recognizable general categories. Subsequently such studies remain descriptive. Heterogeneity among the research participants, variation of experience and circumstance, comparative analytic methods, and development of an abstract conceptual analysis of the data all point to expanding the number of interviews. The nature of the research topic can also foster increasing the number of interviews. Opening secrets, silences, and liminal spaces likely increase the number of interviews needed, as does studying an area which does not come equipped with a widely-shared language.

Researchers sometimes claim that their method of choice such as discourse analysis or narrative inquiry leads to a small number of interviews. They reason that the intense scrutiny entailed in using this method precludes conducting a large number of interviews. Rationalization may serve as reason here. Similarly, some researchers mistake the efficiency of grounded theory as reason to shortcut data collection. Grounded theory is efficient but that does not mean a handful of interviews produces a respectable study. Conversely, having a substantial amount of data does not guarantee an original contribution.

Often the question of how many interviews assumes that conducting single interviews is the only method of gathering data. Is it? Not always. Sometimes researchers do not give themselves credit for observational, archival, and documentary research that they have done. Mixed qualitative methods can strengthen a study with a small number of interviews.

A very small sample can produce a study with depth and significance depending on the initial and emergent research questions and how the researcher conducted the study and constructed the analysis. In his classic study, Edward Speedling (1981) studied eight married men who had had heart attacks. The small sample belies a large effort. Speedling observed on the cardiac wards for several months, visited patients, talked with their family members, and subsequently interviewed the men and their wives during the hospital stays, after the husband's arrival at home, and remained involved in their lives for over three to four more months. If you conduct a study that relies only on interviews, the following guidelines may help. Increase your number of interviews when you: 1) pursue a controversial topic, 2) anticipate or discover surprising or provocative findings, 3) construct complex conceptual analyses, and 4) seek professional credibility. In short, my advice is to learn what constitutes excellence rather than adequacy in your field—and beyond, if your project portends of having larger import—and conduct as many interviews as needed to achieve it.

References

- Guest, G., Bunce, A., and Johnson, L. (2006) How many interviews are enough? : An experiment with data saturation and variability. *Field Methods*, 18, 59-82.
- Speedling, E. (1981) *Heart Attack: The Family Response at Home and in the Hospital*. New York: Tavistock.

* This article is adapted by permission from Sage Publications from the forthcoming second edition of Kathy Charmaz, *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*, 2nd ed., London: Sage Publications.

Linda Sandino, CCW Graduate School, University of the Arts/Victoria & Albert Museum

I first began to use interviews about twenty years ago when I was researching the history of a magazine for my MA in the history of design. As there was no archive beyond the magazine itself, oral history became a tool for examining the publication as field of cultural production. My list of must-do interviewees included three editors, five art directors, and two photographers for a dissertation of about 21,000 words. One of the art directors was constantly away. How would I have been able to write up a history of the magazine if I had not been able to record every one of its 'authors'? Does history demand completeness? This was a question I returned to in my PhD.

About ten years ago I began to undertake interviews for national and university archives focusing on life histories in the visual arts and design. Doing this kind of work makes you acutely aware of the partiality and contingency of all interviews, even those that have a thematic focus. No one interview, it seems, will ever be enough. Each interview generates others and so the archive grows. Who will listen to these interviews? What will they mean to those who 'overhear' them? Will they be of any use? I raise these questions because they reflect my preoccupations as an interviewer as well as a researcher who draws on interview based material. The question: 'how many interviews are enough?' is as relevant to archived interviews, as it is to ones we feel we need to conduct ourselves.

Nevertheless, the paradoxical condition of interviews, like that of most archive material, as both 'too much' information and 'too little', is, in my view, not a weakness but an opportunity. The role of historians, literary scholars, and narrative researchers (where I locate my work), is an interpretive one: what do these interviews represent? What do they mean? Depending on my research focus, one interview may suffice. For instance, using only one recording as a case study has served as a discussion for the different kinds of evidence an interview can produce; current research on museum curators as a group explores their individual, collective and institutional narrative identities. Poised at the intersection of auto/biography and history, each interview represents its own worldview that, nevertheless, contributes to the panorama of cultural institutions, art and design history. However, any one of these life histories could have formed a research project in itself, but as my work is about the history of the museum, one worldview is not enough. So, my answer to the question 'how many qualitative interviews is enough?' is that it all depends on the project and the discipline in which one might wish to situate the research.

Bindhi Shah, Southampton University

Recently I completed a qualitative research project, funded by the ESRC, which examined the role of religion and religious institutions in shaping identity, belonging and citizenship among second-generation Jains in the UK and USA. Since there is no social scientific literature on this group living in the 'west', this was an exploratory study aimed at uncovering the meaning of religion in the lives of non-Christian and non-Muslim young people in the two countries. Decisions about the type of methodological approach to adopt were guided by the following research questions: What makes a person Jain in the UK and USA? What notions of community are produced in the process of reconstructing Jainism? How are Jain religious identities further shaped by class, caste, gender, and migration histories? Does the Jain ethic of non-violence encourage involvement in concerns related to peace, justice and the environment in the wider society and promote a shared sense of citizenship? Given these broad interests, and individual and group level focus, I chose to conduct a multi-method qualitative research design that, nevertheless, was guided by the goal of generating rich and complex data that would illuminate important themes about the role of religion in the lives of young people in late modern societies. These methods included in-depth, semi-structured interviews with second-generation Jains, interviews with lay Jain leaders, observations of bi-annual conventions in each country as well as other small-scale social and religious events, and content analysis of magazines, newsletters, Internet sites and course/workshop materials produced and consumed by young Jains.

Within this multi-method and cross-cultural qualitative project, I decided to include interviews not only because they can add depth of understanding generated by the other methods but also to access young Jains' own views and meanings of their religious identities, practices and beliefs. Decisions about how many respondents to interview were guided by these concerns as well as by more pragmatic questions of time and funding and availability of sampling frame. I initially decided that twenty-five interviews in each country would be sufficient. As there are no established databases of Jain communities in either country I employed a mix of purposive and snowball sampling to select interviewees. Obviously, samples of interview subjects were not representative of Jain communities in both countries, however, I attempted to recruit a sample that was stratified in terms of gender and region of residence, and as far as was possible, of Jain sects. In the United States I had budgeted one-month of research time for conducting twenty-five interviews and participant-observations at a religious convention. This time frame demanded a very hectic schedule and enabled me to only find interviewees from the dominant Jain sects in three cities where there are large numbers of Jains, one on the west coast, one of the east coast and one in the mid-west. In the UK I had more time flexibility as I live in London and the majority of Jains in the UK also live in London. However, my time was divided between this project and other research commitments and so I was only able to locate a few interviewees outside of London. However, in both countries I was able to carry out an equal number of interviews with young Jain women and men. While this number of interviews may not support convincing conclusions on their own, I believed that this number was practical to conduct within given time and financial constraints. Additionally, given the exploratory nature of this research and the multi-scalar focus, the interview data generated illuminated important theoretical and substantive themes about the role and meaning of religion among young Jains as well as supplement data generated by the other qualitative methods employed. In the end I had the opportunity to conduct a total of thirty interviews in each country, but any more than that would have produced too much data to analyse adequately within the given time frames.

Conclusion

Our intention in putting this NCRM Methods Review together has been to answer the question ‘How many qualitative interviews is enough?’ To this end, we have gathered together a set of succinct ‘expert voice’ contributions from 14 prominent qualitative methodologists and five ‘early career reflections’ from those embarking on academic careers. These pieces range across epistemological and disciplinary positions, and across conversational and academic styles.

As we pointed out in our Introduction, the recurring answer to the question ‘how many’ is ‘it depends’. The usefulness of this resource for students, lecturers and researchers rests on the guidance offered by our contributors as to what it depends upon. These include epistemological and methodological questions about the nature and purpose of the research: whether the focus of the objectives and of analysis is on commonality or difference or uniqueness or complexity or comparison or instances. Practical issues to take into account include the level of degree, the time available, institutional committee requirements. And both philosophically and pragmatically, the judgment of the epistemic community in which a student or researcher wishes to be or is located, is another key consideration.

Whether or not we have ‘saturated’ the possible epistemological, methodological and pragmatic responses to the question of ‘how many qualitative interviews is enough?’ – the range of issues on which the answer ‘depends’ – is yet another question. What we have provided in this resource, though, is a very good starting point for anyone conducting qualitative research who is in need of advice and guidance on what to think about when it comes to ‘how many’ in sampling and case selection.