

## **Sampling in interview-based qualitative research: A theoretical and practical guide**

### **Abstract**

Sampling is central to the practice of qualitative methods, but compared with data collection and analysis, its processes are discussed relatively little. A four-point approach to sampling in qualitative interview-based research is presented and critically discussed in this article, which integrates theory and process for the following: (1) Defining a *sample universe*, by way of specifying inclusion and exclusion criteria for potential participation; (2) Deciding upon a *sample size*, through the conjoint consideration of epistemological and practical concerns; (3) Selecting a *sampling strategy*, such as random sampling, convenience sampling, stratified sampling, cell sampling, quota sampling or a single-case selection strategy; and (4) *Sample sourcing*, which includes matters of advertising, incentivising, avoidance of bias, and ethical concerns pertaining to informed consent. The extent to which these four concerns are met and made explicit in a qualitative study has implications for its coherence, transparency, impact and trustworthiness.

### **Keywords**

Sampling, purposive sampling, random sampling, theoretical sampling, case study, stratified sampling, quota sampling, sample size, recruitment

## ***Qualitative Research in Psychology, in press***

Sampling is an important component of qualitative research design that has been given less attention in methodological textbooks and journals than its centrality to the process warrants (Mason, 2002). In order to help fill this void, the current article aims to provide academics, students and practitioners in Psychology with a theoretically-informed and practical guide to sampling for use in research that employs interviewing as data collection. Recognised methods in qualitative psychology that commonly use interviews as a data source include Interpretative Phenomenological Analysis (IPA), Grounded Theory, Thematic Analysis, Content Analysis, and some forms of Narrative Analysis. This article presents theoretical and practical concerns within the framework of four ‘pan-paradigmatic’ points: (1) setting a *sample universe*, (2) selecting a *sample size*, (3) devising a *sample strategy* and (4) *sample sourcing*. Table 1 summarises the principle features of these. All of the aforementioned methods can be used in conjunction with this four-point approach to sampling.

*(Insert Table 1)*

### **Point 1: Defining a sample universe**

The first key concern in the four-point approach is defining the sample universe (also called ‘study population’ or ‘target population’). This is the totality of persons from which cases may legitimately be sampled in an interview study. To delineate a sample universe, a set of inclusion criteria or exclusion criteria, or a combination of both, must be specified for the study (Luborsky & Rubinstein, 1995; Patton, 1990). *Inclusion* criteria should specify an attribute that cases must possess to qualify for the study (e.g. a study on domestic violence that specifies that participants must be women who have suffered partner violence that was reported to the police or social services), while *exclusion* criteria must stipulate attributes that disqualify a case from the study (e.g. a study on exercise that stipulates that participants must *not* be smokers). Together, these criteria draw a boundary around the sample universe, as illustrated in Figure 1.

*(Insert Figure 1)*

### ***Homogeneity and heterogeneity in the sample universe***

The more inclusion and exclusion criteria that are used to define a sample universe, and the more specific these criteria are, the more *homogenous* the sample universe becomes. Sample universe homogeneity can be sought along a variety of parameters, such as *demographic homogeneity*, *geographical homogeneity*, *physical homogeneity*, *psychological homogeneity* or *life history homogeneity* (see Table 2 for descriptions of these). The addition of exclusion or inclusion criteria in these different domains increases sample homogeneity.

## ***Qualitative Research in Psychology, in press***

One of these forms of homogeneity - *psychological homogeneity* - is established if a criterion for case inclusion is a particular mental ability, attitude or trait. In order to make case selection possible based on this kind of criterion, quantitative data from questionnaires or tests can be used as a sampling tool (Coleman, Williams & Martin, 1996). For example, Querstret and Robinson (2013) gained quantitative data on the extent to which individuals self-report having a personality that varies across different social contexts, and used this data to select individuals who were one standard deviation or more above the mean for 'cross-context variability'. These persons were then interviewed for a qualitative study about the motivations for, and experiences of, varying behaviour and personality according to social context.

*(Insert Table 2)*

The extent of sample universe homogeneity that a research study aims at is influenced by both theoretical and practical factors. Theoretically, certain qualitative methods have a preference for homogenous samples; for example *Interpretative Phenomenological Analysis* is explicit that homogenous samples work best in conjunction with its philosophical foundations and analytical processes (Smith, Flowers & Larkin, 2009). By maintaining a measure of sample homogeneity, IPA studies remain contextualised within a defined setting, and any generalisation from the study is made cautiously to that localised sample universe, and not beyond at any more speculative or abstract level.

Conversely, there are approaches that aim to gain samples that are intentionally heterogeneous, for example the *variation sampling* technique of Grounded Theory (Strauss and Corbin, 1998), or the *cross-contextual* approach described by Mason (2002). The rationale for gaining a heterogeneous sample is that any commonality found across a diverse group of cases is more likely to be a widely generalizable phenomenon than a commonality found in a homogenous group of cases. Therefore, heterogeneity of sample helps provide evidence that findings are not solely the preserve a particular group, time or place, which can help establish whether a theory developed within one particular context applies to other contexts (Mason, 2002).

Cross-cultural qualitative research is another instance where a demographically and geographically heterogeneous sample may be called for. Such research selects individuals from different cultures in order to compare them and search for similarities and differences (see Table 2). An example of qualitative research conducted at such a scale was the EURO CARE study – the sample universe was comprised of persons caring for co-resident spouses with Alzheimer's in 14 European countries (Murray, Schneider, Banerjee, & Mann, 1999; Schneider, Murraray, Banerjee, & Mann, 1999). This

## ***Qualitative Research in Psychology, in press***

influential piece of research shows that cross-cultural qualitative research can be successfully conducted with a culturally heterogeneous sample universe, if resources are available.

There are however challenges inherent in using a heterogeneous sample. The first is that findings will be relatively removed from real-life settings, and the second is that the sheer diversity of data may lessen the likelihood meaningful core cross-case themes being found during analysis. Therefore, all researchers must consider the homogeneity/heterogeneity trade-off for themselves and delineate a sample universe that is coherent with their research aims and questions, and with the amount of research resources they have at their disposal.

The sample universe is not only a practical boundary that aids the process of sampling, it also provides an important theoretical role in the analysis and interpretation process, by specifying what a sample is a sample *of*, and thus defining who or what a study is about. The level of generality to which a study's findings is relevant and logically inferable *is* the sample universe (Mason, 2002), thus the more clearly and explicitly a sample universe is described, the more valid and transparent any generalisation can be. If a study does not define a sample universe, *or* makes claims beyond its own sample universe, this undermines its credibility and coherence.

### **Point 2: Deciding on a sample size**

The size of a sample used for a qualitative project is influenced by both theoretical and practical considerations. The practical reality of research is that most studies require a provisional decision on sample size at the initial design stage. Without a provisional number at the design stage, the duration and required resource-allocation of the project cannot be ascertained, that makes planning all but impossible. However *a priori* sample specification need not imply inflexibility – instead of a fixed number, an approximate sample size range can be given, with a minimum and a maximum.

Interview studies that have a *nomothetic* aim to develop or test general theory are to a degree reliant on sample size to generalise (Robinson, 2012). Sample size is by no means the only factor influencing generalisability, but it is part of the picture. O'Connor and Wolfe's grounded theory of midlife transition, which was based on interviews with a sample of 64 adults between the ages of 35 and 50 (O'Connor and Wolfe, 1987) illustrates this point. A way of working with larger sample sizes in qualitative research, which prevents analytical overload, is to combine separate studies together into larger syntheses. For example, I recently combined findings from a series of three studies on the topic of early adult crisis into a single analytical synthesis and single article. One contributing study had a sample of 16 cases, the second had a sample of 8 cases, and the third employed a sample of

## ***Qualitative Research in Psychology, in press***

26 cases. These were analysed and reported as separate studies originally, before being combined into the synthesis paper with a total sample of N=50 (Robinson, Wright & Smith, 2013).

Very large-scale qualitative interview projects include hundreds of individuals in their sample. For example the aforementioned EURO CARE project employed a sample size of approximately 280 (20 persons of for each of 14 countries) (Murray, Schneider, Banerjee, & Mann, 1999), and the MIDUS study (The Midlife in the United States Study) is a study that has involved over 700 structured interviews (Wethington, 2000). While such projects do require time, money, many researchers and a robust purposive sampling strategy (see below), they are achieved by breaking up the research into smaller sub-studies that are initially analysed on their own terms before being aggregated together.

Interview research that has an *idiographic* aim typically seeks a sample size that is sufficiently small for individual cases to have a locatable voice within the study, and for an intensive analysis of each case to be conducted. For these reasons, researchers using *Interpretative Phenomenological Analysis* are given a guideline of 3 to 16 participants for a single study, with the lower end of that spectrum suggested for undergraduate projects and the upper end for larger-scale funded projects (Smith, Flowers & Larkin, 2009). This sample size range provides scope for developing cross-case generalities, while preventing the researcher being bogged down in data, and permitting individuals within the sample to be given a defined identity, rather than being subsumed into an anonymous part of a larger whole (Robinson & Smith, 2010).

Case study design is often referred to as a distinct kind of method that is separable from standard qualitative method (e.g. Yin, 2009). In relation to interview-based case-studies, a more integrative view is taken here, in which the decision to do a N=1 case study is a sample size decision to be taken as part of the four-point rubric set out in this guide. The resulting case study can then be analysed using an idiographic interview-focused method such as IPA. There are a number of different reasons for choosing a sample size of 1, and Table 3 lists six of these; *psychobiography, theoretical or hermeneutic insight, theory-testing or construct-problematizing, demonstration of possibility, illustration of best practice* and *theory-exemplification*. All of these warrant a sample size of 1 and also require associated sample strategies, which are discussed later in this article.

These case study objectives are not mutually exclusive. An example of a paper that evidences multiple aims is Sparke's narrative analysis of the autobiography of cyclist Lance Armstrong (Sparke, 2004). It includes aspects of psychobiography, hermeneutic insight and construct problematizing.

*(Insert Table 3)*

***Pragmatic and theoretical justifications for altering sample size during interview-based research***

In all qualitative studies there are strong grounds for monitoring data collection as it progresses and altering sample size within agreed parameters if need be on theoretical or practical grounds (Silverman, 2010). Indeed, monitoring and being responsive to the practical realities of research is a key skill for the qualitative researcher, as collecting in-depth data leads to challenges that are *never* entirely predictable at the outset of a project. Mason (2002) refers to this skill as ‘organic’ sampling. For example, recruiting participants, the final and fourth concern discussed in this article, is an unpredictable business and if it proves to be more difficult than anticipated, a reduction in target sample size may be required. Conversely, recruitment may lead to more potential cases than was anticipated, so the researcher may consider at this point expanding the target sample size, if logistically manageable. The other major practical reason for changing sample size is if the availability of resources, funding, time or researcher manpower lessens or increases during the course of a project.

Of all qualitative methodologies, Grounded Theory puts most emphasis on being flexible about sample size as a project progresses (Glaser, 1978). According to Grounded Theory, as the researcher collects data, analysis should proceed *at the same time*, not be left till later. Simultaneous analysis permits a researcher to make real-time judgements about whether further data collection is likely to produce any additional or novel contribution to the theory-development process and therefore whether further sample acquisition would be appropriate or not (Strauss & Corbin, 1998). Sample size may be increased if ongoing data analysis leads the researcher to realise that he/she has omitted an important group or type of person from the original sample universe, who should be added to the sample in order to enhance the validity or transferability of the findings or theory (Silverman, 2010). Alternatively, if the researcher judges that ‘theoretical saturation’ has been reached, it is assumed that further data collection will not bring incremental benefit to the theory-development process (Strauss & Corbin, 1998), and data collection will be halted. Guest, Bunce and Johnson (2006) provide a useful set of guidelines for determining theoretical saturation when using interviews.

**Point 3: Selecting a sample strategy**

Once a sample universe is defined and an approximate or exact sample number decided upon, a researcher must then ask themselves the question – how do I select cases for inclusion in the sample? The strategic options available at this point can be categorised into (a) random/convenience sampling strategies and (b) purposive sampling strategies.

## ***Qualitative Research in Psychology, in press***

### ***Random and convenience sampling strategies***

Random sampling is the process of selecting cases from a list of all (or most) cases within the sample universe population using some kind of random selection procedure. This process is used in opinion polls and social research surveys – typical methods include random selection of numbers from a phone book or of addresses from the electoral roll. Quantitative studies in psychology often claim to use a random sampling procedure, when they do not. Instead they typically locate a nearby source of potential participants who are convenient in their proximity and willingness to participate (i.e. Psychology students) and are in all likelihood *not* a random cross-section of the sample universe. This is called *convenience sampling*. It is used in quantitative research and sometimes in qualitative research too. It proceeds by way of locating any convenient cases who meet the required criteria and then selecting those who respond on a first-come-first-served basis until the sample size quotient is full. The problem of using this approach in quantitative research is that statistics function on the basis that samples are random, when they are typically not. For qualitative research, the danger of convenience sampling is that if the sample universe is broad, unwarranted generalisations may be attempted from a convenience sample. The best way of justifying the use of convenience samples in qualitative research is by defining the sample universe as demographically and geographically local and thus restricting generalisation to that local level, rather than attempting decontextualized abstract claims. For example, if the convenience sample is psychology students at a particular university in the UK, then by making the sample universe ‘young university-educated adults in the UK’ rather than ‘people in general’ the link between sample and target population is enhanced, while potential generalisation is narrowed and thus made more logically justifiable.

### ***Purposive sampling strategies***

Purposive sampling strategies are non-random ways of ensuring that particular categories of cases within a sampling universe are represented in the final sample of a project. The rationale for employing a purposive strategy is that the researcher assumes, based on their *a-priori* theoretical understanding of the topic being studied, that certain categories of individuals may have a unique, different or important perspective on the phenomenon in question and their presence in the sample should be ensured (Mason, 2002; Trost, 1986). Summarised below are *stratified*, *cell*, *quota* and *theoretical* sampling, which are all purposive strategies used in studies that employ multiple cases. Following this I describe *significant case*, *intensity*, *deviant case*, *extreme case* and *typical case* sampling, which are purposive strategies that are best employed when selecting a single case study.

## ***Qualitative Research in Psychology, in press***

All of these are processes for ensuring that certain types of individuals within a sample universe definitely end up in a final sample.

### *Stratified sampling*

In a stratified sample, the researcher first selects the particular categories or groups of cases that he/she considers should be purposively included in the final sample. The sample is then divided up or 'stratified' according to these categories and a target number of participants is allocated to each one. Stratification categories can be geographical, demographic, socio-economic, physical or psychological – the only requirement is that there is a clear theoretical rationale for assuming that the resulting groups will differ in some meaningful way.

If there are just two stratification criteria in a study, the resulting framework can be illustrated as a simple cross-tabulated table, as shown in Figure 2.a. In this table, gender and age provide the basis for the sample stratification of a hypothetical study on the experience of life following divorce. If more than two variables are used in a sampling framework, an alternative way of illustrating the stratification is using a 'nested table', as shown in Figure 2.b (Troost, 1986). Here, the variable of 'with children/without children' is added to the divorce study sampling framework. It should be born in mind from a practical view that the more stratification criteria one includes in a sample frame, the more complicated recruitment becomes and the longer the process of finding participants. Therefore researchers should devise a sample strategy that takes into account how much time they have and the resources at their disposal.

*(Insert Figure 2)*

As previously mentioned, to include a purposive sampling stratification there must be clear theoretical grounds for the categories used. For example, in this hypothetical study on post-divorce experiences that Figure 2 refers to, the theoretical grounds for sampling for men and women could be that women are more likely to get custody of children than men in the UK and thus a systematic difference between sexes would be justifiably expected. Age could be justified as a sampling criterion on the basis that younger adults typically find it easy to re-partner than older adults, meaning the post-divorce experience may differ by age. The presence or absence of dependent children could be included because issues of child custody add a great deal of complexity and potential stress to post-divorce proceedings so those with and without children could be expected to differ. In a real study such theoretical rationales for purposive criteria would ideally have referenced sources.

### *Cell sampling*



Cell sampling is like stratified sampling insofar as it provides a series of a priori categories that must be filled when gaining sample. The difference between cell sampling and stratified sampling is that the latter employs categories that are discrete and non-overlapping, while in the former, cells can overlap like a Venn diagram (Miles & Huberman, 1994). As a hypothetical example, a study on popular phobias may choose to purposively select individuals who a) have a phobia of a certain animal, or b) have a phobia of heights, or c) both types of phobia. This example is illustrated in Figure 3.

*(Insert Figure 3)*

### *Quota sampling*

The process of quota sampling is a more flexible strategy than stratified or cell sampling. Instead of requiring fixed numbers of cases in particular categories, quota sampling sets out a series of categories and a *minimum number* of cases required for each one (Mason, 2002). As the sample is gathered, these quotas are monitored to establish whether they are being met. As a hypothetical example, imagine a study run in London in which 30 couples are to be interviewed about the challenges of moving an elderly relative into a care home. A quota sample might list the following:

1. At least 10 couples whose experience pertains to moving their father into supported accommodation, and at least 10 whose experience pertains to their mother.
2. At least 10 couples from low socio-economic groups (C2, D or E).
3. At least 15 couples recruited from outside South-East England.
4. At least 5 unmarried couples.
5. At least 5 couples who live 100 miles or more away from the parent in question .

By using minimum quotas such as these, this strategy ensures that key groups are represented in the sample, while providing flexibility in the final sample composition, thus making it an easier job for recruitment than the more exacting approaches of stratified and cell sampling. It is therefore popular in applied forms of qualitative research, such as market research.

### *Theoretical sampling*

Theoretical sampling differs from the aforementioned purposive strategies, for it takes place *during* the collection and analysis of data, following provisional sampling and analysis of some data (Coyne, 1997; Strauss, 1987). The term theoretical sampling was originally associated with Grounded Theory, but its principles apply across other methods too (Mason, 2002). The process involves either (a) locating cases from new groups of participants (for example a 'comparison' group to provide

## ***Qualitative Research in Psychology, in press***

contrast with existing participants) or from new locations (Strauss, 1987), to build extra heterogeneity into the sample, or (b) re-structuring an already-gathered sample into a new set of categories that have emerged from analysis, and replacing any stratification/cells/quotas that were chosen a-priori (Draucker, Martsof, Ross & Rusk, 2007). For example, Sandelowski et al. (1992) conducted a study of couples who were experiencing parental fertility problems. Initially a stratified frame of medical and social criteria was used. Then after analysis of the initial data, an alternative, more grounded way of grouping couples was found based on the explanations of fertility that participants gave. As a result, the researchers changed to an emergent sample categorisation, and purposively sought new participants to fill this new typology. Another example of theoretical sampling was the process of sampling used to develop a 'theory of academic change' in higher education institutions (Conrad, 1978). The researcher selected individuals to interview from positions in four higher education establishments according to emerging theory, without any a-priori purposive sample strategy at all. Each interviewee was selected based on the provisional analysis of previous interviews, followed by the researcher's judgment as to what kind of person would help fill out the developing theory and explore its internal structure and external generalisability.

### *Purposive sampling strategies for selecting case study participants*

If a decision has been taken by a researcher to do an N=1 case study, a strategy is then required to select the case in question, which fits with its overall objective (see Table 3). If the objective is psychobiography, an individual could be chosen due to his or her historical or theoretical significance – this is called the *significant sample* approach (Simonton, 1999). If the objective is theoretical insight, a strategy of *intensity sampling* can be used – this aims to locate an *information-rich* case, chosen specifically to be insightful, comprehensive, articulate and/or honest (Miles & Huberman, 1994). For example in my own research, an intensity sampling strategy was taken for the purposes of selecting a case study on early adult crisis; from a previous interview sample, individuals were selected who had provided a comprehensive account of both the inner and outer dimensions of their crisis experience, following which one person was selected randomly from this information-rich group (Robinson & Smith, 2010b).

If the objective of a case study is theory-testing or construct-problematism, a *deviant case* strategy may be used, whereby an individual is sought intentionally to explore the limits or problems inherent in a theory (Eysenck, 1976). Alternatively, if the aim is to demonstrate the *possibility* of a phenomenon, an *extreme case* strategy may be used, which locates who is not representative of others, but instead shows an extreme or unusual behaviour, ability or characteristic (Patton, 1990). Finally, if the research objective is illustration of best practice or theory-exemplification, a *typical*

## ***Qualitative Research in Psychology, in press***

*case* strategy (also called emblematic case or paradigmatic case) could be used (Miles & Huberman, 1994) in which the case is chosen precisely because it is a typical example of a theory or therapeutic application. For example, Hodgins and Peden (2000) describe a case study of using cognitive behavioural therapy for treating kleptomania, in which an individual was selected whose presentation was typical of the challenges involved.

### **Point 4: Sourcing sample**

Once a researcher has decided on a sample universe, a provisional sample number and a sample strategy for their interview study, then comes the hands-on part – they need to go and source their participants from the real world. As well as practical and organisational skills, this stage of sampling requires ethical skills and sensitivity; all potential interviewees should be informed of the study's aims, of what participation entails, of its voluntary nature, of how anonymity is protected and any other information that will help them reach an informed, consensual decision to participate, prior to any agreement to participate.

An aspect of voluntary participation is that individuals who consent to be involved in interviews may be different to those that do not, in ways that are not related to sampling criteria - this is called the *self-selection bias* (Costigan & Cox, 2001). In interviews, extensive intimate self-disclosure is sometimes required and this is likely to lead to a sample containing individuals who are more open, more patient and more interested in the topic than the general sample universe. My own research has found that women are more likely to put themselves forward to participate in qualitative research than men, which concurs with research showing a higher tendency towards self-disclosure in females than males (Dindia & Allen, 1992). This female bias can be easily counteracted in a mixed-gender purposive sampling frame that ensures male and female representation, but more subtle systematic biases in differences between participants and non-participants are harder to deal with. For example, Abrams (2010) refers to a study on 'vulnerable youth', which required individuals to respond to a flier and call the researcher to participate, but thereby was limited to those youth who were proactive and confident enough to actively call up a stranger to volunteer in a study and therefore in all probability not as psychologically vulnerable as those who did not call up. The self-selection bias is not possible to circumvent in interview-based research, as voluntary participation is central to ethical good practice, therefore all a researcher can do is be aware of the possibility for bias and consider it's possible impact on findings and generalisability.

*Study Advertising*

Ways of recruiting participants for interviews are only limited by a researcher's ingenuity in how to disseminate the message of his/her research study to the sample universe. Advertising is one way of making this happen; it can be done using print and face-to-face methods, but increasingly online advertising is becoming more and more popular. Hamilton and Bowers (2006) debate the pros and cons of using the internet to publicise research, and conclude that the benefits include wide outreach to different populations, while the disadvantages include a sample that is likely skewed towards higher income and education levels, and there is also the possibility of fraudulent claims about identity, if face-to-face contact during data collection is not required. Furthermore, disadvantaged groups may be less likely to have access to the internet. Sample sourcing can also benefit from the internet in another way; online video and audio call services such as Skype can be used to conduct the interviews, meaning that sample can be sourced from a geographically dispersed area without encountering major logistical challenges or costs.

Recruiting interviewees within organisations presents particular advertising challenges. Typical advertising channels within an organisation may include group email, intranet, notice boards or internal mail. A researcher will first need to secure permission from an individual who has access to channels of communication used within the organisation. This person may be termed the recruitment gatekeeper (Devers & Frankel, 2000). Ideally, a gatekeeper can be turned into a research 'champion' within the organisation, who will actively help with publicising the study and encourage participation. For example, in a previous interview study on the experience of major change within a government department, the HR manager was identified as the gatekeeper to recruitment. Fortunately he supported the study and championed it for us, as well as advertising it through email circulars to all staff (Robinson & Griffiths, 2004). In such instances, offering to produce an executive report of the findings for the organisation is frequently essential to achieving sampling sourcing.

Advertising that is aimed at hard-to-reach or vulnerable populations can benefit from a face-to-face approach to build trust. Abrams (2010) describes a face-to-face strategy used in a study with youths in juvenile correctional institutions. She gave presentations to groups of 5 – 10 youths in a room within the institution, during which the purposes, procedures and voluntary nature of the study were made explicit, followed by an opportunity to ask questions. Following this, youths could confidentially put their name on a list to express their interest. Having gained a list of interested individuals, consent from parents/guardians was received prior to arranging interviews.

## ***Qualitative Research in Psychology, in press***

### *Referral processes: Snowball sampling*

An alternative recruitment strategy to advertising is *snowball sampling* (also called chain sampling, chain-referral sampling or referral sampling). This involves asking participants for recommendations of acquaintances who might qualify for participation, leading to 'referral chains'. This may be particularly useful when the population being studied is unlikely to respond to advertisements due to the stigmatising or illegal nature of the topic (Heckathorn, 2002). Such topic areas include drug use, crime, homelessness or stigmatizing illness such as HIV. Classic qualitative research that employed snowball sampling to investigate drug use includes the work by Lindesmith (1968) on opiate addiction, or Biernacki and Waldorf (1981) on recovery from heroin addiction.

### *The question of incentives and 'respondent-driven sampling'*

When recruiting people into an interview study, the question of whether to offer them a financial incentive for participation is a key decision that the researcher must make. The benefits of incentives are that they increase the likelihood of participation by adding additional motivation, and also increase retention in longitudinal studies (Yancey, Ortega & Kumanyika, 2006). The downside is that they provide a motivation for fabricating information in the interview in order to gain the money. The danger of gaining 'dodgy data' is such that if a researcher is confident that he/she can source the required sample without offering financial incentives, this is preferable. This problem is greater in low-income populations. A further ethical issue pertaining to incentives is that they may motivate participation in an interview on a topic that may elicit distress. Alternatives to financial incentives exist that can provide other perceived benefits to participation. Firstly participants can be offered a copy of their interview transcript and/or a report of the study's findings, however researchers should be aware that this may not be an incentive to all participants and could in some cases lead to distress; therefore this decision should be made sensitively with a clear option to opt out. Secondly, making clear the potential applied benefits of the research can act as an incentive for many adults.

Incentives can also be used within a snowball sampling process as a way of stimulating recruitment.

*Respondent-driven sampling* is a form of snowball sampling that gives financial incentives for recruiting others into the study (Heckathorn, 1997; Johnston & Sabin, 2010; Wang et al., 2005). In the process, a researcher first recruits a small number of 'seed' subjects. These individuals are then given a number of 'recruitment coupons'. They are then told that if they pass the coupons on to other persons who end up participating, they will be paid a fixed amount per referral. All new recruits are then offered the same recruitment coupons to continue the process. This incentivised

## ***Qualitative Research in Psychology, in press***

recruitment can be done in conjunction with a financial incentive for doing an interview – a combination referred to as ‘dual incentives’.

### **The importance of sampling to qualitative research validity**

Addressing the four sampling issues that have been outlined in this article is central to enhancing the validity of any particular interview study. Yardley’s criteria for evaluating qualitative research are respected benchmarks for assessing the validity of a study (Robinson, 2008; Smith et al., 2009; Yardley, 2000). These are ‘sensitivity to context’, ‘rigour’, ‘transparency’, ‘coherence’, and ‘impact and importance’. Well-conducted sampling enhances all four of these.

With regards to ‘sensitivity to context’, a fully articulated, contextualised sample universe prevents unwarranted generalisation and helps to locate the study within a place, a time and a meaningful group. A sample universe can be given contextual richness by locating it clearly within a particular culture or subculture, and giving it a historical location if required. Historical sensitivity means that unwarranted generalisation of findings to the past or the future will be less likely. This is particularly relevant if a phenomenon is being studied that changes over historical time, for example marriage or retirement.

With regards to ‘rigour’, Yardley herself writes that sampling is integral to meeting this criterion: ‘Rigour...depends partly on the adequacy of the sample – not in terms of size but in terms of its ability to supply all the information needed for comprehensive analysis’ (2000, p.221). Key decisions to enhance rigour are the relationship of the sample to the sample universe, the appropriate choice of sampling strategy, the robustness of the sample sourcing approach, and the overall fit between research questions and total sample strategy.

The third criterion, ‘transparency’, is enhanced by being explicit in a final research report as to how all four points – sample universe, sample size, sample strategy and sample sourcing – were met. Articles that stipulate one or two of these, but omit to mention others, will lack transparency, repeatability and auditability. Furthermore, the process of recruitment is often influenced by the researcher’s own background, location and connections, and if that is the case, appropriate reflexive acknowledgement of any conflict of interests or possible bias also aids transparency.

‘Coherence’ is maximised by systematically fitting the sampling processes with research aims, research questions, data collection and analysis, so that all are mutually supported and theoretically consistent, and by maintaining consistency between the reach of the sample universe and attempted generalisations. The 4-point approach to sampling is ultimately aimed at linking the

## ***Qualitative Research in Psychology, in press***

various processes of sampling to enhance coherence and prevent sampling theory becoming divorced from the practicalities of the process.

The final principle of 'impact and importance' is the extent to which a study contributes to theory or practice. This criterion requires the research to have theoretical or practical relevance *beyond* the sample used. For that to happen, the sample universe must have been specified with sufficient clarity, that those who read the research know where application or generalisation is appropriate, and thus understand *who* the study is important *for*.

### **Summary**

In summary, four points holistically encompass the challenge of sampling in interview-based qualitative studies: defining the *sample universe*, deciding on a *sample size*, selecting a *sample strategy* and *sourcing cases*. These are not sequential steps in a linear process, for decisions pertaining to each point can iteratively affect the other three and vice versa. Good sampling involves shuffling back and forth between each point to eventually devise a four-point strategy that is coherent, achievable and appropriate to research aims. The more explicitly and systematically all four points are addressed in a study, the greater the validity of that study and the stronger the quality of any corresponding write-up. The strategies and concepts described within each of the four points here are merely indicative as to the possibilities available – they can be re-combined or added to in countless new ways, and I strongly encourage all researchers to keep innovating.

## **Qualitative Research in Psychology, in press**

### **References**

- Abrams, L.S. (2010). Sampling 'Hard to Reach' Populations in Qualitative Research: The Case of Incarcerated Youth. *Qualitative Social Work, 9*, 536-550.
- Biernacki, P. & Waldorf, D. (1981). Snowball sampling: Problems and techniques of chain referral sampling. *Sociological Methods and Research, 10*, 141-163.
- Coleman, T., Williams, M. & Wilson, A. (1996). Sampling for qualitative research using quantitative methods. 1. Measuring GPs' attitudes towards discussing smoking with patients. *Family Practice, 13*, 526-530.
- Conrad, C.F. (1978). A grounded theory of academic change. *Sociology of Education, 51*, 101-112.
- Costigan, C.L. & Cox, M.J. (2001). Fathers' participation in family research: Is there a self-selection bias? *Journal of Family Psychology, 15f*, 706-720.
- Coyne, I.T. (1997). Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries? *Journal of Advanced Nursing, 26*, 623-630.
- Devers, K.J. & Frankel, R.M. (2000). Study Design in Qualitative Research—2: Sampling and Data Collection Strategies. *Education for Health, 13*, 263-271.
- Dindia, K. & Allen, M. (1992). Sex differences in self-disclosure: a meta-analysis. *Psychological Bulletin, 112*, 106-24.
- Draucker, C.B., Martsof, D.S., Ross, R., & Rusk, T.B. (2007). Theoretical sampling and category development in Grounded Theory. *Qualitative Health Research, 17*, 1137-1148.
- Erikson, E. (1969). *Ghandi's Truth*. New York: W. W. Norton.
- Erikson, E. (1993). *Young man Luther: A study in psychoanalysis and history*. New York: W. W. Norton.
- Eysenck, H. (1976). Introduction. In H.Eysenck (Ed.), *Case Studies in Behavioural Therapy* (pp.1-15). London: Routledge and Kegan Paul.
- Festinger, L., Riecken, H.W., & Schacter, S. (1956). *When prophecy fails: A social and psychological study of a modern group that predicted the end of the world*. University of Minnesota Press.
- Glaser, B. (1978). *Theoretical sensitivity: Advances in grounded theory*. Mill Valley, CA: The Sociology Press.
- Guest, G., Bunce, A. & Johnson, L. (2006), How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*, 59-82.
- Hamilton, R.J., & Bowers, B.J. (2006). Internet Recruitment and E-Mail Interviews in Qualitative Studies. *Qualitative Health Research, 16*, 821-835.
- Heckathorn, D.D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems, 44*, 174-199.



## **Qualitative Research in Psychology, in press**

- Hodgins, D.C., & Peden, N. (2008). Cognitive-behavioral treatment for impulse control disorders. *Revista Brasileira de Psiquiatria, 30*, S31-40.
- Johnstone, L.G., & Sabin, K. (2010). Sampling hard-to-reach populations with respondent driven sampling. *Methodological Innovations Online, 5*, 38-48.
- Kraepelin, E. (1895). *Psychologische Arbeiten*. Leipzig: Wm Engelmann.
- Leshan, L. (1990). *The dilemma of psychology*. London: Penguin Books Ltd.
- Lindesmith, A.R. (1968). *Addiction and Opiates*. Chicago: Aldine.
- Luborsky, M.R. & Rubinstein, R.L. (1995). Sampling in qualitative research: Rationale, Issues and Methods. *Research on Aging, 17*, 89-113.
- Lynd, R.S. & Lynd, H.M. (1929). *Middletown*. New York: Harcourt, Brace.
- Lynd, R.S. & Lynd, H.M. (1937). *Middletown in transition*. New York: Harcourt, Brace.
- Mason, J. (2002). *Qualitative researching, 2<sup>nd</sup> Edition*. London: Sage.
- McAdams, D., & Schultz, T. (2010). *George W. Bush and the Redemptive Dream: A Psychological Profile*. OUP USA.
- McAdams, D., & West, S. (1997). Introduction: Personality psychology and the case study. *Journal of Personality, 65*, 757-783.
- Miles, M.B. and Huberman, A.M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook, 2nd edition*. London: Sage.
- Murray, J., Schneider, J., Banerjee, S., & Mann, A. (1999). EURO CARE: A cross-national study of coresident spouse carers for people with Alzheimer's disease: II--A qualitative analysis of the experience of caregiving. *International Journal of Geriatric Psychiatry, 14*, 662-667.
- Neisser, U. (1981). John Dean's memory: A case study. *Cognition, 9*, 1-22.
- O'Connor, D.J. & Wolfe, D.M. (1987). On managing midlife transitions in career and family. *Human Relations, 40*, 799-816.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd Edition). Newbury Park, CA: Sage.
- Querstret, D. & Robinson, O.C. (2013). Person, persona and personality modification: An in-depth qualitative exploration of quantitative findings. *Qualitative Research in Psychology, 10*, 140-159. 10.1080/14780887.2011.586450
- Reicher, S. (1996). The 'Battle of Westminster': developing the social identity model of crowd behaviour in order to explain the initiative and development of collective conflict. *European Journal of Social Psychology, 26*, 115-134.
- Robinson, O. (2012). A war of words: the history of the idiographic/nomothetic debate. *The Psychologist, 25*, 164-167.

## **Qualitative Research in Psychology, in press**

- Robinson, O. C., & Smith, J. A. (2010a). Investigating the form and dynamics of crisis episodes in early adulthood: The application of a composite qualitative method. *Qualitative Research in Psychology*, 7, 170-191.
- Robinson, O.C & Smith, J.A. (2010b). A case study of pre-midlife crisis: An extraordinary transformational episode in an ordinary life. In D.Robinson et al. (Eds.) *Narrative, Memory and Ordinary Lives* (pp. 145-157). Huddersfield: University of Huddersfield Press.
- Robinson, O.C., Wright, G.R.T. & Smith, J.A. (2013). The holistic phase model of early adult crisis. *Journal of Adult Development*. 20, 27-37, 10.1007/s10804-013-9153-y
- Robinson, O.C. & Griffiths, A. (2004). Coping with the stress of transformation change in a government department: A qualitative study. *Journal of Applied Behavioral Science*, 41, 204-221.
- Sandelowski M., Holditch-Davis D. & Harris B.G. (1992) Using qualitative and quantitative methods: the transition to parenthood of infertile couples. In *Qualitative Methods in Family Research* (Gilgun J.F., Daly K. & Handel G. eds), Sage, Newbury Park, California.
- Schneider J, Murray J, Banerjee S, Mann A. (1999). EURO CARE: a cross-national study of co-resident spouse carers for people with Alzheimer's disease: Factors associated with carer burden. *International Journal of Geriatric Psychiatry*, 14, 651-61.
- Silverman, D. (2010). *Doing qualitative research, 3<sup>rd</sup> Edition*. London: Sage.
- Simonton, D.K. (1999). Significant samples: The psychological study of eminent individuals. *Psychological Methods*, 4, 425-451.
- Smith, JA, Flowers, P & Larkin, M 2009, *Interpretative phenomenological analysis: theory, method and research*, Sage, London.
- Sparkes, A. (2004) Bodies, narratives, selves and autobiography: the example of Lance Armstrong. *Journal of Sport and Social Issues*, 28, 397\_428.
- Strauss, A. L. (1987). *Qualitative analysis for social scientists*. New York: Cambridge University Press.
- Strauss, A.L. & Corbin, J 1998, *Basics of qualitative research: techniques and procedures for developing grounded theory*, Sage, London.
- Trost, J.A. (1986). Statistically non-representative stratified sampling: A sampling technique for qualitative studies. *Qualitative Sociology*, 9, 54-57.
- Wang, J., Carlson, R.G., Falck, R.S., Siegal, H.A., Rahman, A., & Li, L. (2005). Respondent-driven sampling to recruit MDMA users: a methodological assessment. *Drug and Alcohol Dependence*, 78, 147–157.
- Wethington, E. (2000). Expecting stress: Americans and the “midlife crisis.” *Motivation and*

**Qualitative Research in Psychology, in press**

*Emotion*, 24, 85–103.

Yancey, A.K., Ortega, A.N. and Kumanyika, S.K. (2006). *Effective recruitment and retention of minority research participants. Annual Review of Public Health*, 27, 1-28.

Yardley, L. (2000). Dilemmas in qualitative health psychology. *Psychology and Health*, 15, 215-228.

Yin, R.K. (2009). *Case study research: Design and methods* (4<sup>th</sup> Edition). London: Sage Publications.

**Qualitative Research in Psychology, in press**

Table 1. The four-point approach to qualitative sampling

	<i>Name</i>	<i>Definition</i>	<i>Key decisional issues</i>
<i>Point 1</i>	<i>Define a sample universe</i>	Establish a sample universe, by way of specifically a set of inclusion and/or exclusion criteria.	Homogeneity vs. heterogeneity Inclusion and exclusion criteria.
<i>Point 2</i>	<i>Decide on a sample size</i>	Choose a sample size or sample size range, by taking into account what is ideal <i>and</i> what is practical.	Idiographic (small) vs. nomothetic (large)
<i>Point 3</i>	<i>Devise a sample strategy</i>	Select a purposive sampling strategy to specify categories of person to be included in the sample.	Stratified, cell, quota, theoretical strategies
<i>Point 4</i>	<i>Source the sample</i>	Recruit participants from the target population.	Incentives vs. non-incentivised. Snowball sampling varieties. Advertising study.

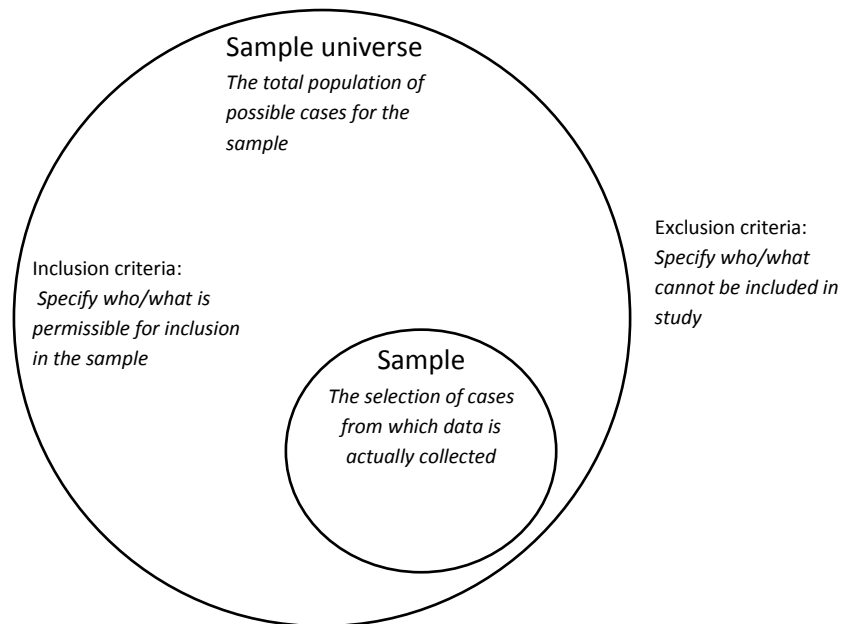
Table 2. Five types of sample homogeneity

<i>Source of homogeneity</i>	Description	Example hypothetical study and sample requirement
<i>Demographic homogeneity</i>	Homogeneity imparted by a demographic commonality such as a specific age range, gender, ethnic or socio-economic group	A study on menopause that requires participants to be women between the ages of 50 and 55
<i>Geographical homogeneity</i>	Refers to sample that is all drawn from the same location	A study that evaluates Cognitive Behavioural Therapy provision in Birmingham
<i>Physical homogeneity</i>	Occurs in a sample who must share a common physical characteristic	A study on coping with cystic fibrosis that requires all participants to currently suffer from the disease
<i>Psychological homogeneity</i>	Similarity within a sample imparted when participants are selected based on the possession of a particular trait or ability	A study into gifted children that requires participants to have an IQ of over 150
<i>Life history homogeneity</i>	Occurs in a sample if individuals share a past life experience in common	A study on motivations for migration that requires participants to have moved as a migrant to the UK between the ages of 20 and 40

Table 3. Six functions of an N=1 sample size

<b>Case-Study Function</b>	<b>Description</b>
<i>Psychobiography</i>	Providing information about an individual of particular interest or significance, by using existing theory and concepts allied to rich description to illuminate and interpret their behaviour and thoughts. Examples of this include the psychobiographical analysis of George W. Bush by McAdams and Schultz (2010), or Erikson's work on Gandhi (1969) and Martin Luther (1993).
<i>Theoretical or hermeneutic insight</i>	Theoretical insight gained from intensive examination of a single case can be put forward for validation in other samples. An example of this is the research by Neisser (1981) on the transcripts from one person in the Watergate Scandal, which was used to challenge and extend models of memory, or the work by Reicher (1996) on riots as a way of stimulating theory development in social and crowd processes.
<i>Theory-testing or construct-problematizing</i>	A general theory or hypothesis can be tested or challenged by recourse to evidence from one case, based on the logic that the assertion 'all swans are white' can be disproved by finding just <i>one</i> black swan (Eysenck, 1976). Such deductive logic is avoided by many qualitative psychologists as it is linked to objectivist epistemologies, but in fact quantitative research avoids this kind of logic too, for it provides strong justification for N=1 research, which is almost non-existent in orthodox quantitative journals. In fact, case-specific theory-testing works well within a qualitative rubric, as evidenced by Festinger's test of cognitive dissonance theory through a case study of a UFO cult (Festinger, Riecken & Schacter, 1956).
<i>Demonstration of possibility</i>	A case study can demonstrate that a phenomenon is possible. For example, if just one person was found to always show extrasensory abilities on repeated occasions, this would be an important finding in parapsychology, with or without corroboration from other cases. It would show that such abilities were <i>possible</i> and therefore that human beings have the requisite apparatus for ESP (Leshan, 1990).
<i>Illustration of best practice</i>	Providing information on the process of providing counselling, therapy or other one-to-one help. Through a comprehensively formulated case study, best-practice can be communicated, while the real-life context and challenges of delivering therapy can be conveyed.
<i>Theory-exemplification</i>	<i>Exemplifying</i> a particular theory or construct that is purported to explain or describe behaviour at the individual level, acting to bring the abstract theory into the concrete particular and so illuminate its descriptive or explanatory nature (McAdams & West, 1997).

Figure 1. Sample universe, inclusion/exclusion criteria and sample



**Qualitative Research in Psychology, in press**

Figure 2. Types of table used for illustrating stratified sampling

a) Cross-tabulated table illustrating a sample stratified by two variables: gender and age

	Male	Female
Ages 30-45	4	4
Ages 46-60	4	4

b) Nested table illustrating a stratified sample with three typological variables: gender, age and presence of dependent children

Male				Female			
30-45		46-60		30-45		46-60	
Dependent children	No dependent children	Dependent children	No dependent children	Dependent children	No dependent children	Dependent children	No dependent children
2	2	2	2	2	2	2	2



Figure 3. A hypothetical example of cell sampling showing a three-cell sample frame with 5 persons to be recruited who have phobias of animals, 5 who have phobias of heights, and 5 who have phobias of both.

