2

Sampling in Qualitative Research

WHAT IS SAMPLING?

Sampling is one of the most important aspects of research design. In a research context, sampling refers to the process of selecting a subset of items from a defined population for inclusion into a study. We say *items* because the subset of things to include in your study may not always be people. Individuals are certainly the most common sampling unit in social and behavioral research, but sampling units can also be groups, events, places, and points in or periods of time. Whatever your sampling unit, you'll need to consider carefully *how* items are to be chosen.

The degree of generalizability (or not) of findings and their representativeness relative to the larger population (or not) are dependent on (or constrained by) one's sampling strategy. Likewise, the validity of study findings is related to sampling. If, for example, the "key informants" selected for a study are not very knowledgeable about the topic being investigated, the resulting data will not be particularly informative. Because sampling is so important, we devote an entire chapter to the subject. The chapter is divided into the three major components of selecting participants for a research study—deciding whom to include, choosing how to include them (i.e., sampling strategy), and determining how to bring them into your study (recruiting techniques).

Defining Key Terms

Study Population

The entire group of elements that you would like to study. This is often a group of individuals (e.g., firefighters in New York City, migrant workers in North Carolina, female sex workers in Bangkok), but it can also be composed of larger units such as families, institutions, communities, cities, and so on. The study population is entirely defined by the researcher, based on the study objectives. Note that a single study can include multiple populations, depending on how your research problem is defined.

Sampling Frame

A sampling frame is a list of all the elements in a study population. It is always defined by your study population. If, for example, your study population is vendors in market X, then your frame would be a list of all vendors in that market. Note that the sampling frame is the largest possible sample of a study population. It is often not easy to obtain in field research.

Sampling Unit

This is the thing (person, place, event, etc.) that is selected for inclusion in the study. For example, if you are sampling high school students, your sampling unit would be one student.

Study Sample

The group of sampling units chosen from the larger population for inclusion in the study becomes your study sample. Note that any given sample can be part of more than one population.

SELECTING WHOM TO SAMPLE

For highly targeted research studies, selecting whom to sample may not be a difficult task. Your study population may be precisely defined by your research objectives. If, for example, your research objective is to explore the personal experiences of women who participated in a particular clinical trial or programmatic intervention, your primary inclusion criterion is already determined by your research objective. But not all research questions are this precise, especially in qualitative research, which is often exploratory in design and purpose. For most research initiatives, you will need

to think carefully about whom to include in your research as you design your study, and, in some cases, as data collection progresses.

Whom you include in your study will be based on several criteria, but first and foremost, the rationale behind your design should directly reflect your research objective(s) or question(s). Ask yourself: What specific individuals or types of individuals (i.e., roles, occupations) might know a lot about my research topic in the study site(s)? This is the basis of the key informant concept. Some individuals are simply more knowledgeable than others about certain subjects. If you are investigating medicinal plant use within an indigenous community, for example, traditional healers would likely be at the top of your list of whom to include.

One of the hallmarks of qualitative research is the ability to account for context and gain a more holistic picture of a research topic. Including the most knowledgeable individuals in a qualitative study is essential, but it is often not enough to fully understand the social, cultural, and contextual complexities associated with a research question, particularly if it's on the complicated end of the spectrum. Seeking and documenting multiple viewpoints is part of any good research study, even more so for qualitative inquiry. To this end, think about who else in your study community has knowledge about the subject and who may have a different perspective. Referring back to the medicinal plant use example, you might decide to also sample adult women if they use medicinal plants in the home, or local pharmacists if they sell herbal remedies in their shops. Each of these groups will likely have different types of knowledge and use plants in different ways. Capturing this range of opinions and experiences would be essential if your interest is at the community level.

Perhaps your research is intended to inform policy. If this is the case, you'll need to consider including (a) individuals who would be most affected by any policy change to emerge from your study findings and (b) individuals who have control over policy pertaining to your research topic and study site(s). In other words, identify the key stakeholders. Including stakeholders in your sample provides not only important information, but also it can facilitate their cooperation if and when recommendations are developed.

Unless you have unlimited funding and time (i.e., you're independently wealthy and have nothing else to do), you will need to prioritize whom you include in your study. Talking to everyone in your study community and documenting every voice is usually not possible. But defining inclusion criteria and narrowing the scope of your study is not an easy task. In fact, when we teach research methods, the most common type of questions we receive pertains to choosing participants. Students often ask questions such as, Do I need to include both men and women or both wealthy and impoverished individuals in my study? or If I have focus groups with patients, do I need to do focus groups with doctors too? We have two general responses to these types of sampling questions:

It is always a good idea to elicit multiple perspectives on a given research question. Men and women, for example, often have different experiences and take on different

social roles in many cultures, so including both in research studies is common practice. But think beyond gender. How would you expect knowledge, behavior, beliefs, and experience regarding your research topic to vary within your study population? Would you expect, for example, older individuals to have different or more knowledge than younger individuals, or wealthy people to have different perceptions than those more impoverished? Is there a group of individuals that share has a unique experience who might view the topic differently? You would ask the same types of questions if sampling events or places. Places and events often have temporal variability associated with them. Do the activities or locations you wish to sample vary by season? by day of the week? time of day? or some other temporal parameter?

Answering these types of questions in an informed manner requires a certain degree of knowledge about your research topic and study community. If you've never been to the study site(s) or worked within the study communities, consult other researchers or local residents who are knowledgeable about the area of study.

The decision to include or exclude particular groups or strata also depends on how your research question is framed and how your study population is defined. When conceptualizing your research question, you should have at least a sense of the study's scope, which is partially determined by budget and time available. Investigating an entire community's perspective on something is different (and much more difficult) than focusing your study on a very specific and targeted group of individuals (e.g., 2012 high school students in X county, or religious leaders in Baghdad). Another factor bearing on sampling decisions is the degree to which you want to explicitly compare data between groups. If comparison is a key element in your study design, you need to think about which groups you wish to compare. For each comparative group you add, the total sample size for your study will at least double (more so if your groups are segmented into more than two categories), so be mindful of budgetary and time constraints when making your selection.

It's often helpful to take an open, brainstorming approach to designing your sampling parameters. List all of the groups, people, events, or places that could provide relevant information on your research topic; then, rank or prioritize the list based on the salience of the information that you would expect to obtain from each. Let budgetary and timing issues help you reduce your sample(s) from the ideal to the feasible. Creating a title for a research study is another incredibly useful exercise in defining the boundaries of your research. It forces you to think about who you really want to include in your study and why. A good title should tell your audience who your study population is. Referring back to the example of medicinal plant usage, think how different your sampling strategy would be for each of the following three study titles:

Title 1: "Medicinal Plant Usage in Village X"

Title 2: "Medicinal Plant Usage Among Traditional Healers in Village X"

Title 3: "Medicinal Plant Usage Among Traditional Healers and Mothers in Village X"

In Title 1, your study population is defined as everyone in community X. This, at least theoretically, includes men, women, children, young, old, traditional healers, merchants, and so forth. With such a long list of subpopulations, you will likely need to refine your question to exclude less relevant groups, or stratify your sample so that at least a few members of key groups are represented in your sample. Title 1 suggests a large study and will require a relatively complicated sampling strategy for an in-depth qualitative study. In contrast, Title 2 requires that you identify traditional healers in the community and include at least some of them in your study, a much easier proposition than Title 1. Finally, Title 3 suggests a comparative design, in which usage is compared between traditional healers and mothers. In terms of complexity, this study design falls between Title 1 and Title 2. The title suggests inclusion of two subpopulations within the village, both of which should be relatively easy to identify.

Inductive Sampling

Much applied research is relatively focused in scope, with clear objectives from the outset. The range of stakeholders is known up front, as are the cultural and social contexts in which the study is embedded. But a good deal of qualitative research is inductive and exploratory in nature. Sometimes you may not know who the stakeholders are when you begin your research, or who in the study community is the most knowledgeable about your research topic. In this case you would gather as much information as you could from existing resources. If you still don't have enough information to make an informed decision, your research design will be, by necessity, exploratory in design. You will likely need to employ an inductive sampling strategy.

Inductive, or emergent, sampling allows for the inclusion of groups and types of cases not originally specified or included in the study design. The flexibility to follow new leads during fieldwork and to take advantage of new information as it is collected and reviewed is a major strength of inductive sampling, and of qualitative research in general. Consider one of the author's studies on contraceptive decision making in India. The original sampling design included married women, married men, and health care providers. After a few interviews it was apparent that mothers and mothers-in-law play an important role in contraceptive decisions of their children in Indian culture. These two groups were subsequently incorporated into the sampling plan.

Inductive sampling—also called "theoretical sampling" in the grounded theory literature (Charmaz, 2006; Glaser & Strauss, 1967)—is more commonly employed in longer-term research such as ethnographic studies. The iterative and inductive process of collecting data, analyzing them, and formulating new sampling and recruiting procedures takes time. In applied research, which generally has more targeted research questions and shorter time horizons, inductive sampling methods are

logistically more difficult (though the study of Indian contraceptive use just cited was an applied project, so it is possible).

Inductive sampling can also be problematic at the regulatory level. You may have been required in a funding proposal and/or ethics review application to state how many individuals you will sample from which populations, as well as your sampling strategy and method of recruitment for each (in applied research this is typically the case). If sampling procedures or characteristics of your sample are altered significantly, most ethical oversight bodies require a protocol amendment and subsequent review. So if you know up front you will be sampling inductively, it's best to incorporate that flexibility into your proposal and protocol from the beginning.

SAMPLING APPROACHES

In the previous section, we talked about choosing whom to include or not include in your study. Once this has been figured out (or at least partially so), the next step is to think about *how* you will sample the units (usually people) in your study population(s). Essentially, there are three general sampling approaches to choose from—censuses, non-probability sampling, and probability sampling—each with its own unique strengths and limitations.

Census

A census is the process of collecting information from or about *every item* in a study population. It is technically not a sampling method since no selective process is involved; i.e., you include everyone in your sampling frame. Censuses are often associated with large quantitative data collection efforts, such as the U.S. Census, but they can also be employed in qualitative research. Theoretically, a census is always preferable to any kind of sample. It obviates the need to make inferences from the sample to a larger population, and avoids criticisms about the representativeness of your "sample." Unfortunately, censuses are rarely employed in qualitative field research due to logistical constraints. Most study populations are large, and including everyone from a population in a research study is usually prohibitively time consuming and costly. Moreover, once you reach a large enough sample size—in either qualitative or quantitative research—the law of diminishing returns kicks in, and the effort-to-information ratio becomes lopsided; each new item in your sample adds very little information to the data that has already been collected.

Nonetheless, there are certainly situations in which the population of interest is small enough to permit a census. Some examples might include all of the senior

executives in a large company, every caregiver in a public health clinic, or all elders in a small village. We recommend using a census if possible, as it eliminates the potential for criticisms pertaining to generalizability of findings (at least relative to your study population). In most cases, some sort of strategy will be required to select a sample from a larger study population.

Non-Probability Sampling

Non-probability sampling is typically defined in relation to its probability-based counterpart. There are two main points to this comparison:

- Non-probability sampling does not involve random selection.
- Relatedly, since non-probabilistic samples cannot depend upon probability theory, we can't know the odds or probability with which the chosen sample represents the population.

This does not necessarily mean that non-probability samples are not representative of the larger study population or that we can't conceptually generalize from our sample to a larger population in some cases. It does mean, though, that we can't statistically extend our findings to a larger population or place confidence intervals around our findings. Fortunately, qualitative research is not intended, nor expected, to generate these metrics.

Non-probabilistic sampling is the norm in qualitative research, for several reasons:

- Statistical analyses (inferential)—most of which require probability-based samples—of qualitative data are not common.
- Qualitative inquiry is not intended or designed for statistical generalizability.
- Qualitative sample sizes are generally too small to be subject to probability theory.

Qualitative studies seek to generate rich, contextually laden, explanatory data and are therefore not concerned with generating population-based estimates and *p*-values. Probability samples provide little benefit to this end, and the effort needed to acquire a proper sampling frame, as is necessary for most probability samples, is not justified. And, in many cases, employing a probability sample in a qualitative research study can be fundamentally detrimental to the validity of one's findings because it may gloss over specific individuals who have unique attributes (e.g., knowledge, experiences, social position) relative to your research question(s).

The suitability of non-probability samples for qualitative research is also related to the types of objectives that drive much qualitative inquiry. In many cases, qualitative studies are designed to understand common processes, shared experiences and understandings, or to identify shared cultural knowledge and norms. It doesn't take a probabilistic sample to gain insight into these types of topics. In fact, as we discuss in more detail below, sample sizes as small as six individuals are often adequate in this regard. How many of your neighbors would you need to ask to find out what time the local supermarket opens? How many colleagues would you need to interview to get a general picture of what it's like to visit a dentist in your country? Much experience and knowledge is so widely shared or processes so standardized within a culture or community that only a few individuals are needed to address a research question. In cases where we are interested in variability—as opposed to commonalities and patterns—qualitative inquiry is extremely useful in establishing the *range of* attributes associated with the phenomena of interest. If it is a *measure* of variability you're after, however, don't bother with qualitative methods. Structured inquiry with a probabilistic sample is the way to go.

Purposive Sampling

The most commonly employed non-probabilistic sampling approach is purposive, or purposeful, sampling (also sometimes called "judgment" sampling). Perhaps the most intuitive way to think of purposive sampling is that you choose study participants based on the purpose of their involvement in the study. In Bernard's words, "you decide the purpose you want your informants (or communities) to serve, and you go out and find some" (Bernard, 2000, p. 176). The logic and power of purposive sampling, argues Patton, "lie in selecting information-rich cases for study in depth . . . those [cases] from which one can learn a great deal about issues of central importance to the purpose of the inquiry" (Patton, 2002, p. 230). Operationally, this means establishing one or more eligibility criteria for inclusion into a study. Your criteria could be as simple and broad as "adult males in community x." Or they may include multiple criteria exhibited by a specific group of people. If, for example, you wanted to understand how a niche market might react to the introduction of a new beer bottle shape, you might screen for sex (males), marital status (single), annual income (> $80K), and consumptive behavior (drink more than 12 bottles/week). Or perhaps your research is aimed at understanding a very specific experience, such as traveling through security checkpoints in a conflict zone. Your criterion might be individuals who have traveled through X number of checkpoints in the past month. Table 2.1, extracted and adapted from Patton's *Qualitative Research and Evaluation Methods* (2002) provides a framework from which to start thinking about the various dimensions of purposive sampling.

Table 2.1 Types of Purposive Sampling Approaches

Sampling Type	Characteristics	Example
Homogeneous	• Sample is similar on one or more dimensions • Often used in focus group discussions • Useful if the population to which the results will be inferred is also homogeneous, as it simplifies analysis • One of the most common sampling approaches • Data are often collected from multiple homogeneous groups for a more holistic perspective and/or comparative purposes	For a U.S.-based marketing study of female motorcycle riders, you might sample female motorcycle owners from four different regions of the country: for example, Los Angeles, New York City, Atlanta, and Chicago. If you were interested in assessing the introduction of a new sex education curriculum in secondary schools, you would likely interview various homogeneous stakeholder groups: teachers, parents, students.
Extreme or Deviant Case	• Focus on unusual manifestations of the phenomenon of interest • Often used to find key factors associated with extreme behavior • Such cases are easily identified with quantitative data (e.g., the *nth* percentile)	You may use this approach if you want to identify outstanding successes or notable failures, such as the best or worst students in a class. In reproductive health studies, participants might be selected based on extreme (reported) condom usage (i.e., all the time versus never).
Intensity	• Information-rich cases that manifest the phenomenon of interest intensely, but not extremely • Differ from extreme/deviant cases in that they are in a normal range and not on the extreme end of the curve	You may be interested in sampling good students and poor students or individuals who are above average/below average on a given dimension (rather than the "extreme" cases of brilliance and abject failure).
Typical Case	• Illustrate or highlight what is typical, normal, average • If selected well, a few individuals can provide accurate insight into general patterns and processes across a larger population • Illustrative, not definitive • Particularly good for understanding phenomena that are widely shared across a culture or community	If your research objective is to understand the experience of purchasing a new vehicle from a GM dealership, you could simply interview individuals who recently purchased a new, average priced car from a GM dealership (i.e., not a Corvette or a Volt). The assumption is that if the process and experience are fairly typical across dealerships, relatively few participants need to be sampled.

(Continued)

Table 2.1 (Continued)

Sampling Type	Characteristics	Example
Critical Case	• Permits logical generalization and maximum application of information to other cases because if it is true of this one case, it's likely to be true of all other cases	If you interview highly educated health professionals and they have a problem understanding a health message, you can be fairly certain that lay people will also have problems.
Confirming and Disconfirming Cases	• Helpful for elaborating and deepening initial analysis, seeking exceptions, and testing variation • Often used to validate (confirming) or expand/negate (disconfirming) data-driven models • Often sought out near the end of a research project to establish boundaries around one's interpretation of a dataset	If your data reveal a widely shared theme, take a second look through the same dataset in search of data that diverge or contradict that theme. Imagine, for example, 28 out of 30 participants have positive views of a product or program you're evaluating. The deviant two cases are analyzed more closely to understand why they have such different views. Often these views can be highly informative. If your research is flexible in design and you have enough time, you can follow up with those participants who expressed these divergent or contradictory viewpoints to gain further insight.
Politically Important Cases	• Purposely eliminating or including politically sensitive cases • Sample individuals to attract attention to the study or to avoid attracting undesired attention	Local leaders or political activists, for example, are often included in a study to increase political cooperation.
Criterion	• Picking all cases that meet some predetermined criterion • Often used in quality assurance processes	You may use criterion sampling to identify cases that fail set standards, such as sampling all children abused in a treatment facility. Or on the positive side, you may wish to identify all substance abuse program graduates who stay clean for *x* amount of time.
Maximum Variation	• Purposely picking a wide range of cases for a sample to get variation on dimensions of interests • Document unique variations that have emerged in relation to different conditions	You're interested in documenting the diversity of opinions regarding a new product or program that is scheduled for roll out. Seek diversity within your target population and select accordingly. Common parameters

(Continued)

Sampling Type	Characteristics	Example
Maximum Variation (continued)	• Identify important common patterns that cut across variation (cut through the noise of variation) • Also useful for generating a wide range of responses • Used when time/resource constraints prohibit sampling multiple groups	include age, sex, ethnicity, education, income level, geographic location, and occupation.
Stratified Purposeful	• Study population is broken down into strata that have some theoretical importance to the study objectives • Typically based on illustrative characteristics of particular subgroups of interest • Frequently used strategy, facilitates comparisons	Examples of commonly used strata include gender, ethnicity, age, education, and income level. Or it can be much more specific such as the amount of experience using a product or degree of exposure to an intervention.
Quota	• Variation of a stratified sample • Relative size of substrata are proportional to their relative size in the larger population	A general quota sample of men and women with an n of 20 would break down to 10 men and 10 women, since the sex ratio in general populations is 50/50. If, however, your study population was injection drug users, your quota sample of 20 would be composed of 18 men and 2 women, if 90% of injection drug users are men.
Theory-Based	• Finding real-world manifestations of a theoretical construct of interest so as to elaborate and examine the construct and its variations • Cases are chosen based on their ability to inform parts of the theoretical model	Models can come from various sources: for example, literature, primary quantitative data, primary qualitative data. In one of the author's studies, for example, quantitative data showed a significant inverse correlation between pornography viewing and subsequent condom use. We specifically searched the qualitative data from the study for any references that would inform this association (i.e., our theory that was derived from the quantitative analysis).

(Continued)

Table 2.1 (Continued)

Sampling Type	Characteristics	Example
Snowball or Chain Referral	• Utilizes participants' social networks to identify other participants • Ask existing participants to refer others based on certain criteria (often attributes similar to the referring participant) • Very useful for sampling hard-to-reach populations, but it is vulnerable to recruiting biases [Note: This is actually a recruiting technique, since chain referrals are used to *achieve* types of samples. However, since chain referral techniques are typically referred to as sampling methods in the literature (correct or not), we include them in this table.]	Chain referral techniques are often used with injection drug users since the activity is usually illicit, and users tend to have solid social connections based on the activity they share in common. Another example from marketing research comes from 3M. In their desire to better understand how surgical drapes are applied to the body, they initially contacted specialists they knew from various fields who used such devices, such as a veterinarian and a Broadway makeup artist. They then asked this initial group of experts to suggest other experts they knew who could offer more information.
Convenience or Haphazard	• Collect data from whatever cases present themselves • Convenience sampling has the lowest credibility of all sampling approaches—should be viewed as a last resort. Explicitly justify use of this technique when disseminating findings	Conducting research in conflict zones, for example, may permit only convenience samples due to security issues. In ethnographic research, initial interviews may be convenience based, until the researcher becomes more familiar with the study community.

Source: Adapted from Patton (2002, pp. 243–244).

Consider also what you intend to do and achieve with the data. Look at Table 2.1 again. Certainly, one's analytic aim would differ between, say, a homogeneous and a maximum variation sample. In the former, one is more likely looking for common themes across the group; in the latter, establishing the range of responses would be a more appropriate analytic aim. If, on the other hand, you're looking for unique individual attributes relative to your research question, extreme and critical case samples might be useful. Each of the variations of purposive sampling has a distinct function.

Probabilistic Sampling

As we mentioned earlier, non-probability sampling is the norm for qualitative inquiry, but this is not to the exclusion of probability sampling. A *probabilistic*, or *probability*, sample is derived using some form of *random selection*. Random selection means that the probability of each sampling unit being chosen is known. One of the defining features of probability sampling is that one can calculate the odds or probability that your sample represents the larger population well.

So when should a qualitative researcher use probability sampling? In our view, the main reason is that in some cases, it can enhance the representativeness of your sample. And, in cases where a probability sample may not actually enhance the representativeness of your sample, it can enhance the perception of such. As anyone in the research world is acutely aware, there is more to research than just data collection and analysis procedures. Logistics and political factors invariably play a role in determining how we carry out and report on our research. Even though one can't statistically generalize, say, from a small random sample of in-depth interviews to the larger study population—due to the small sample size and unstructured nature of the data—from a pragmatic perspective, it can often behoove a researcher to randomly choose a sample in qualitative research.

A personal experience from one of the authors' studies provides a poignant example. Guest was tasked with qualitatively assessing the procedures in a year-long clinical trial from the study participants' perspective. Twenty-four participants were selected from the study population of 400 (i.e., all participants enrolled in the clinical trial) using a simple random sampling procedure and were administered in-depth interviews. A thematic analysis of these data revealed several perceived shortcomings in the clinical trial procedures. When the data were presented to clinical trial staff, their immediate response was, "How do we know these data are representative of the larger clinical trial population?" The researcher's reply was simply, "Because the participants were chosen randomly." The criticism was curtailed right then and there. Whether or not a random sample of 24 open-ended interviews from a population of 400 is statistically generalizable is not the point here. Rather, the random sample was seen, at least theoretically, as being more representative of the population than a non-probabilistic sample. This example is not unique. The notion of a random sample is—rightly or wrongly—held in high esteem among many research audiences, whether corporate executives, funders, or other researchers.

The primary driver of your sampling strategy selection should be your study design and research objectives. For qualitative research, this often means some type of purposeful sampling strategy. Nonetheless, if you have a choice between choosing a homogeneous sample purposively or randomly from a larger population and you have a decent sampling frame, why not use a simple random sample? If it is possible to obtain a probabilistic sample (note that in the above example this was

uncharacteristically easy), and if there are no reasons to the contrary (that is, your use of a probability method would not capture your intended population, such as in extreme case sampling), random sampling approaches in qualitative research are perfectly acceptable.

Below we cover only two of the more common probabilistic sampling techniques—simple random and systematic. We do not discuss other methods, such as cluster, stratified, time and space, or respondent-driven sampling, as they are rarely employed in qualitative research.

Simple Random Sample (SRS)

Probably the most common probability sample, an SRS is also one of the most robust. In an SRS, each sampling unit has the same probability of being selected relative to all other units in the sampling frame. The procedure is straightforward and based on three fundamental steps.

1. Assign a number from 1 to *N* to each element in the population, with *N* being the total number of known elements in the population.
2. Obtain a list of n different numbers (*n* being the number of elements you want in your sample) within the 1 to *N* range, each one obtained using a random process. In contemporary research, random number generation is invariably done with computer software (Excel or any statistical program) or via the Internet, although hard copy random number tables are still available and used occasionally.
3. Choose the elements in the population corresponding to these numbers.

One of the main advantages of an SRS, compared to other probability methods, is that it is conceptually the simplest probabilistic sampling method and thus easy to analyze. In fact, many statistical analysis packages and standard statistical methods assume this type of sampling.

One of the challenges associated with an SRS, however, is the requirement that all the sampling units in the entire population be identified and numbered prior to selection (i.e., are enumerated). In other words, you need a sampling frame. In many research contexts, this is simply not possible. But in some contexts a little creativity can go a long way. For example, with the widespread accessibility of information and maps on the Internet, it's relatively easy to obtain an SRS of houses in a small town or village. Pull up a satellite photo of your study site (which will need to be validated on the ground), number the houses on your map, and randomly choose your sample from there. With a little ingenuity, you may be able to create a sampling frame where no "official" enumeration exists.

Another drawback to an SRS is that it may not adequately capture subpopulations if they are hidden or small in number, relative to the larger population. If, for example, for analytic purposes we wanted roughly an equal number of men and women in our study and only 10% of the study population was female, an SRS wouldn't likely produce enough women in the sample. We would instead use some type of stratified approach. The same principle holds if we are interested in sampling unique individuals or key informants within our study population. Using an SRS, or any probability sample for that matter, would not be very useful in this instance.

Systematic Sample

A systematic sample involves the selection of every *nth* unit within a study population. For example, we might select every *nth* name in a phone book, visit every 8th house in a village, or interview every 20th person entering a market or attending a health clinic. While theoretically systematic sampling is not true random sampling (because once the first unit is selected, the odds of selection for every other unit have been predetermined), it is as robust as an SRS for the vast majority of research contexts that require a probability sample. There are four basic steps to the process:

1. Define the sampling population.
2. Determine the desired sample size.
3. Estimate the size of sampling population.
4. Determine the sampling interval. This is calculated by dividing the estimated population size by the desired sample size.

Defining your population is trickier, and more important, than it might seem at first. If, for example, you're interested in an activity that has temporal variability—for example, people who visit a car dealership—inferences from your study's findings are constrained by your sampling parameters. Sampling on weekends will limit inferences from your findings to weekend customers; if you want to make broader inferences, include weekdays as well.

One of the main advantages of a systematic sample is that its use does not require a sampling frame, which makes it more versatile in many cases than a SRS. There are some limitations to the method, however. The first is that you need to be able to estimate the population size fairly well. If you underestimate, you will reach your sample size prematurely. Conversely, overestimation will result in having run through the entire population before obtaining your desired sample size.

Systematic sampling also works best if your sampling units are configured in an orderly or defined manner. So if you want to sample, say, people attending a sporting event, it helps if there are either defined entrances from which to count and sample

or orderly and fixed seating from which to count and select. If the sampling context is more fluid or less orderly, then systematic sampling becomes more difficult. Imagine, for example, trying to obtain a systematic sample at a large political rally in an open space, where people are milling about.

Work-arounds can be created for such amorphous contexts. One can, for example, create invisible transects through the crowd and, walking along a transect, select every *nth* person on that line. Multiple transects at different spatial orientations through an invisible center—for example, Corner A to Corner B, Corner C to Corner D, middle of Side A to middle of Side B, and so forth—can subsequently be created and followed if one transect is not enough to generate the desired sample size. In general, though, diffuse and mobile environments are not ideal for systematic sampling.

A less commonly encountered problem with systematic sampling is an order bias. If the list, event, or assemblage of units to be counted has some sort of regular patterning (e.g., people are seated in an alternating male, female configuration) this can introduce obvious, and significant, bias into a sample.

How Do I Choose a Sampling Strategy?

Above, we have described a number of sampling methods from which to choose. There are no prescriptive, mechanistic methods for selecting a sampling approach, but we can offer some criteria you can consider to help guide your efforts. The four steps below are intended as guidelines only. We cannot account for every possible permutation of contextual, theoretical, and logistical parameters of a given study. Every researcher will need to carefully consider ALL of the parameters of their particular research project to arrive at an informed decision.

1. Estimate the size of the population of interest. If it's small enough and ethically and logistically feasible to do, choose a census and proceed to Step 4. If a census is not possible or desirable, proceed to Step 2.
2. Consider (a) how much control you will have over your recruitment and sampling procedures and (b) how certain you are about *who, what, where* you need to sample for your study. If you have a good sense of what types of individuals (or events or places) will best inform your study and if you can access those individuals or places, a purposive sampling strategy is a good choice. Choose among the variations of purposive sample approaches outlined in Table 2.1 by identifying which variation best suits your research objectives and study population and which is feasible within your research context. If your population is "hidden" and socially connected somehow, consider augmenting your sampling strategy with a chain referral technique. Proceed to Step 4.

   If you have no or little control over who you can sample or if you are uncertain about who the most appropriate participants (or places and events) are for your study, carry on to Step 3.

3. If you're unsure of whom you should sample, do some more desk work (i.e., read more about your topic and the proposed study community). If you've exhausted this route, reach out to other researchers and local community members to help inform your research design and sampling strategy (this is good practice, even if you do have a well-informed and focused research design). If this doesn't provide enough information, you may have to initiate your study with a convenience sample, perhaps within a participant observation context. Proceed to Step 4.

   If you have no or little control over whom or what places you can access, a convenience sample is pretty much your only option.

4. Regardless of which sampling strategy you choose, consider how inductive and flexible your study procedures can be. The more room for iterative processes you have during data collection activities and the more time you have to carry out your study, the greater the consideration you can give to choosing an inductive sampling approach. If your sampling procedures have to be predetermined (due to funder or ethics committee requirements) or if you have substantial time constraints, inductive sampling is not your best option and may be logistically unfeasible. Unless you're employing a census, proceed to Step 5.

5. Once you have gone through Steps 1 through 4 and have selected a sampling method, determine whether or not adding a probability component to it would (a) enhance the representativeness of your sample and (b) not detract from the initial sampling strategy's purpose. If you answer in the affirmative to both these questions and it's logistically feasible to obtain an SRS or systematic sample, consider choosing one of these methods to enhance your study. Proceed to Step 6.

6. For each population and data collection method in your study, repeat Steps 1 through 5.

You likely will have noticed that sampling methods are not necessarily mutually exclusive. They also don't have to be employed in isolation relative to one another. You may, for example, use a convenience sample to identify people who are exceptionally knowledgeable about your study topic (i.e., *intensity* form of purposive sampling) and interview a sample of these key individuals. But imagine that you also want to make sure that you capture as diverse a range of viewpoints as possible from this population of knowledgeable individuals (*maximum variation* form of purposive sampling). As you complete the interviews with the initial sample of key informants, you ask them to refer others who are equally knowledgeable about the subject but who may have different perspectives (*chain referral* sampling). You've just utilized three sampling approaches to select one defined population (the convenience sample is, by its nature, undefined, although you would still record basic information about the individuals within the sample).

And if your study includes more than one population (which is common in field research), you may wish to sample one population using one method and another population using a different method. You may, for example, be able to get a census of

one group of individuals because their population size is extremely small, whereas another much larger group might have to be sampled purposively. There are no hard and fast rules about combining and integrating sampling techniques. Do what makes the most sense for your research and that can most effectively (and ethically) help you achieve your research objectives.

CHOOSING SAMPLE SIZES

In-Depth Interviews

Estimating probabilistic sample sizes is a fairly straightforward enterprise. They are estimated mathematically based on preselected parameters and objectives (i.e., *x* statistical power with *y* confidence intervals).[1] Non-probability samples are entirely different when it comes to estimating necessary sample sizes. Several years ago, in preparation for a sampling article, Guest, Bunce, and Johnson (2006) searched through the social, behavioral, and health science literature to see what general recommendations existed regarding non-probability sample sizes. The authors reviewed 24 research methods books and seven databases. They found that although numerous works explained how to select participants (e.g., Johnson, 1990; Trotter, 1991) or provided readers with factors to consider when determining non-probabilistic sample sizes (Flick, 2009; LeCompte & Schensul, 2010; Miles & Huberman, 1994; Morse, 1995; Patton, 2002; Rubin & Rubin, 2004), only seven sources provided guidelines for actual sample sizes.

Bertaux (1981), for example, argues that an *n* of 15 is the smallest acceptable sample size in qualitative research, while Bernard (2000, p. 178) observes that most ethnographic studies are based on 30 to 60 interviews. Morse (1994, p. 225) outlines more detailed guidelines. She recommends at least six participants for phenomenological studies; approximately 30 to 50 participants for ethnographies, grounded theory studies, and ethnoscience studies; and 100 to 200 sampling units in qualitative ethology.

Creswell's (2006) ranges are somewhat different. He recommends between five and 25 interviews for a phenomenological study and 20 to 30 for a grounded theory

[1]Since the vast majority of inductive, qualitative field studies are not interested in statistical analyses of data, and subsequently probability sampling, we do not cover the subject of power analysis in this book. For those readers interested in pursuing the topic of probabilistic sample sizes in more detail, we suggest any basic statistical textbook, or the classic "Statistical Power Analysis for the Behavioral Sciences" (Cohen, 1988). Various sample size calculators are also available on the Internet. Though crude, they are simple to use and can provide ball park estimates (e.g., www.raosoft.com/samplesize.html).

study. Kuzel (1992, p. 41) ties his recommendations to sample heterogeneity and research objectives, recommending six to eight interviews for a homogeneous sample and 12 to 20 data sources "when looking for disconfirming evidence or trying to achieve maximum variation."

A key limitation of these recommendations is that, though derived from experience, none are based on empirical evidence (at least none of the authors cited above present evidence for their recommendations). During their review, Guest and colleagues (2006) observed that nearly all of the relevant literature recommended that purposive, non-probabilistic sample sizes be determined inductively: That is, sampling should continue until *theoretical saturation*—the point at which no or little new information is being extracted from the data—is reached (e.g., Bluff, 1997; Byrne, 2001; Fossey, Harvey, McDermott, & Davidson, 2002; Morse, 1995; Sandelowski, 1995). In fact, a good number of journals in the health sciences recognize theoretical saturation as the main criterion by which to justify adequate sample sizes in qualitative inquiry. Saturation is the status quo by which non-probability sample sizes should be determined in social and behavioral research, at least according to the methodological literature.

Guidelines and Qualitative Research

The issue of guidelines in qualitative research is a contentious one. Some scholars argue that we should not, or cannot, establish guidelines because practitioners of qualitative research do not agree about what these guidelines should be, or they disagree on the simple premise that qualitative research should follow guidelines at all (Chapple & Rogers, 1998; Sandelowski & Barroso, 2002). One argument is that guidelines are overly prescriptive and, in and of themselves, do not confer rigor (Barbour, 2001; Eakin & Mykhalovskiy, 2003). Another point critics make is that guidelines focus too much on procedure and methods and that their uncritical application "legitimize[s] substandard research" (Lambert & McKevitt, 2002). Some qualitative scholars posit that developing and/or adhering to predetermined structures or processes contradicts the inductive essence of qualitative inquiry (Denzin, 2009; Lambert & McKevitt, 2002).

For the most part, we don't disagree with the general intent behind any of these criticisms of guidelines. We do disagree, however, with the hyperbolic forms these arguments take. Using guidelines as templates to blindly design or evaluate qualitative research is indeed irresponsible. And, true, there will always be a disagreement over what constitutes best practice. Equally true, if one imposes *too much* structure on qualitative inquiry, it loses its inductive power. But "guidelines" are substantially different than "standards." As Guest

and MacQueen (2008) point out, guidelines are suggestions. Standards are prescriptive assertions. In other words, guidelines provide direction, not absolute rules, and they need to be applied intelligently and with good judgment. The suggestions, procedures, and templates we provide in this book are designed to help researchers think through their sampling, data collection, and data management procedures, not dictate one particular way of doing things. The suggestions, however, are based on the applied methodological literature and the combined experience of three applied qualitative researchers.

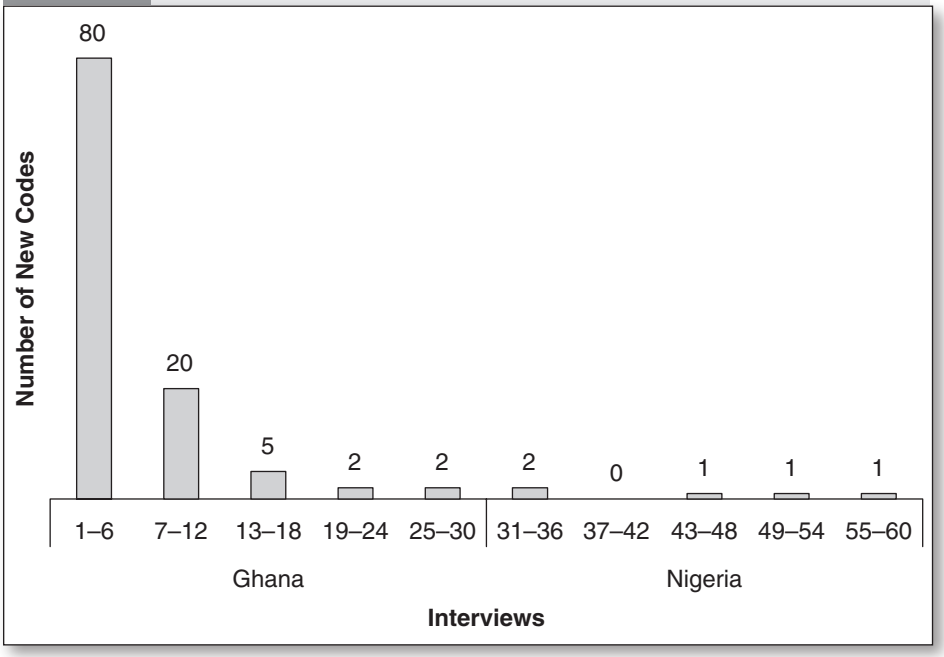
Research, including qualitative research, in an applied context is intended to address real-world problems and, if we're lucky enough, to inform policy. We do our best to understand a research problem by employing the most appropriate sampling and data collection method we can find (which are not always qualitative). Guidelines and best practices are landmarks to help us navigate the journey to our destination.

Guest and colleagues' review also revealed that the same literature did a poor job of operationalizing the concept of saturation, providing no description of how saturation might be determined. Using theoretical saturation as a determinant of sample size is problematic from another practical standpoint as well. As mentioned earlier, reviewers of research proposals and protocols (e.g., funders, institutional review boards [IRBs]) typically require researchers to state planned sample sizes for a study *before* it begins (Cheek, 2000). This requirement is antithetical to the very principles and procedures inherent in grounded theory and other inductively-based research designs (Charmaz, 2006, p. 30). Waiting to reach saturation in the field is, therefore, not often an option. Researchers must follow the a priori sampling plan they outlined in their proposal or protocol, regardless of the emergent situation on the ground.

So how does one go about estimating when saturation will occur before data are collected? An empirical study conducted by one of the book's authors sought to answer this question in the context of in-depth interviews (Guest et al., 2006). Using data from a study involving 60 in-depth interviews with female sex workers in two West African cities, the authors systematically documented the degree of thematic saturation and variability in the data over the course of their analysis. Based on their analysis, they posited that the dataset was relatively "thematically saturated" after only 12 interviews. After analyzing 12 of the 60 interviews, 100 of the 114 (88%) total codes applied to the entire dataset had been identified and developed. Seventy percent of all codes were identified within the first six interviews (Figure 2.1).

The magic number of six interviews is consistent with one other empirically-based study (Morgan, Fischoff, Bostrom, & Atman, 2002) and Morse's (1994) (albeit

Figure 2.1 Code Creation Over the Course of Data Analysis



Source: Guest, Bunce, and Johnson (2006).

unsubstantiated) recommendation for phenomenological studies. Similar evidence-based recommendations can be found for qualitative research in technology usability. Nielsen and Landauer (1993) created a mathematical model based on results of six different projects and demonstrated that six evaluators (participants) can uncover 80% of the major usability problems within a system and that after about 12 evaluators, this diagnostic number tends to level off at around 90%. While helpful, be aware that the numbers above come with two important caveats:

Caveat 1—All Studies Are Different. The studies from which the saturation data above were derived have three important factors in common. Researchers for each of these studies (a) used a fairly homogenous sample, (b) used a semi-structured data collection approach, and (c) were interested in finding patterns across the sample. Not every qualitative study exhibits any one, let alone all, of these characteristics. In general, the rapidity at which saturation is reached is related to five factors.

The degree of instrument structure. The more structure embodied in the instrument, the sooner saturation will be reached. Note that for studies using an

unstructured instrument, or no instrument at all, saturation may never be reached.

The degree of sample homogeneity. The more homogeneous the sample, the quicker saturation is achieved. Groups that are alike on various dimensions are more likely to think in similar ways and have similar experiences.

The complexity and focus of the study topic. For more complex and intricate topics, it will take longer to reach saturation than for simpler and more targeted topics.

Study purpose. Finding high-level common themes across a sample will generally require fewer sampling units than identifying the maximum range of variation within a sample. If you're interested in finding the big issues, a small sample is often sufficient. Conversely, if your study objectives require the comprehensive documentation of all the idiosyncrasies exhibited within your sample, you'll need to sample substantially more than six or even 12 units.

Analyst categorization style. Some folks are "splitters." They tend to see detail in everything and create codebooks accordingly. On the other end of the continuum are "lumpers"—individuals who like to group things into a few large conceptual categories. Codebooks created by splitters will invariably include a lot more codes than codebooks created by analysts with a lumper bent. The smaller the codebook being used to code the data, the quicker saturation will be achieved.

So, if you have a heterogeneous sample, a less-structured instrument, a highly complex topic, are interested in the range of variation, and your analysts are all splitters, you may never reach saturation. The magic numbers of six and 12 per subgroup would be meaningless. They are minimum estimates. Use your best judgment to estimate your sample sizes by building off of the few existing empirical studies and considering the five factors above, rather than viewing recommendations as the ultimate authoritative source.

Caveat 2—Your Audience May Have Different Standards. We point out above that evidence-based studies, as well as the collective experience of qualitative researchers around the world, suggest that a sample size of six to 12 individuals per group is sufficient for many (but not all!) types of qualitative studies. For many audiences, however, this number is perceived as too small, regardless of the empirical or anecdotal evidence. Rarely, for example, will a journal accept a qualitative manuscript with such small sample sizes. Most will want an *n* of 25 or more. Likewise, we know of PhD committees who have insisted on increasing the sample size beyond 15 for a dissertation research project, even though 15 participants would have been more than sufficient for the particular research context.

Even in the corporate sector, where the cost of a single interview can be as much as several thousand dollars, we've experienced this propensity for wanting larger samples. Clients will often insist on an *n* of 20 to 30 per target group, in spite of the added cost and contrary to researcher recommendations. In sum, our advice is that while increasing one's sample size may not be necessary from a scientific perspective, you need to consider your audience's predilections. In most cases, it doesn't hurt to increase your sample size, and doing so provides the appearance of extra diligence (real or perceived).

Every research context is unique. And sampling decisions will need to be made within the parameters of a specific context. Notwithstanding, some general guidelines in selecting non-probabilistic sample sizes can be posited. Padgett (2008, p. 56) outlines four rules of thumb that we find helpful:

> *The smaller the sample size, the more intense and deep are the data being collected.* This is why, for example, case study analyses typically involve relatively small sample sizes.
>
> *Larger samples are needed for heterogeneity, smaller sizes for homogeneity.* This rule applies to both your research objectives and your study population. Finding shared, high-level themes across a group of individuals, for example, will require smaller sample sizes than those for revealing a large range of responses.
>
> *Avoid sacrificing depth for breadth.* Ideally, you'll have enough resources to cover both the depth and breadth of a particular topic, at least to a certain degree. But be careful not to stretch your resources too thin by including too many stakeholder groups, having too broadly defined a study population, or having too many research objectives. In qualitative inquiry, depth is your primary methodological objective.
>
> Larger numbers are preferred, as long as rule 3 is honored.

Focus Groups

To our knowledge, no empirical studies have been published that provide evidence-based guidance on how many focus groups are needed for a given research study.[2] In the focus group literature, as with the in-depth interview literature, scholars often cite theoretical saturation as the primary method for determining sample sizes. But since there are no supporting data as to how many groups this might actually entail, researchers must rely on general recommendations based on anecdotal experience. The

[2] There are, however, more specific recommendations for individual focus group size. We present these in Chapter 5.

recommendations in the literature range from as few as three groups per study population (or subpopulation) to as many as 50 groups (Greenbaum, 1997; Kitzinger & Barbour, 2001; Krueger & Casey, 2009/2010; Morgan, 1996; Powell & Single, 1996; Vaughn, Schumm, & Sinagub, 1996). A commonly agreed upon guideline is that saturation requires conducting at least two groups for each defining demographic, which means that the greater the variety of participants, the more focus groups are required (Carey, 1995; Knodel, 1993; Krueger & Casey, 2009/2010; Ulin, Robinson, & Tolley, 2005).

Based on the experience of others, as well as our own, we would recommend carrying out at least three focus groups per population or subpopulation. Even if one focus group for some reason goes awry (which does happen), you still have data from two groups that should provide enough insight for most research objectives. Note that focus groups are typically conducted with homogenous groups. So if you want to, say, include employees and managers from the same company in your study, this would mean three groups of employees and three groups of managers, if you were to use our minimum recommendation. Another point to keep in mind is that when analyzing focus group data, each *group* is the unit of analysis. One of the implications of this analytic property regards publishing. Focus groups are not often used as stand-alone data in peer reviewed publications, for two general reasons. The first is that response-dependence exists between participants within a group, which in turn affects validity (Guest, MacQueen, & Namey, 2012). The second is that you may need to conduct 15 or more focus groups (at the very least) for acceptance into most peer-reviewed journals, which entails recruiting 90 to 180 participants, depending on the size of your focus groups). That said, there are plenty of applications of focus group methodology for which publication is not the aim; we describe focus groups and their strengths in more detail in Chapter 5. The point is that it can be difficult to accurately estimate (and defend) an appropriate sample for focus groups, regardless of their purpose in the research design. At the time this book went to press, however, Guest received a grant to empirically study data saturation and variability in focus groups. We hope the results will comprise the establishment of an evidence base for focus group sample sizes upon which others can build.

Participant Observation

Because participant observation is so varied in its practice, there really aren't any recommendations as to the number of places, events, or behaviors you should observe. We can, however, offer several factors to consider:

The degree of structure in your participant observation procedures. If you are in the early exploratory stages of a study and are conducting participant observation to get a general lay of the land, you will probably need to carry out several

observation events to get a feel for what to focus on while observing and who to talk to. Conversely, the more you know about a community, its people, and your research topic, the more structure you can impose on observation activities. Templates and checklists might be used, for example, in the later stages of a research project. With more structure, you may need only one observation event to accomplish a particular data collection objective.

The degree of temporal variability inherent in your research topic. As a general rule, the more variability, the more participant observation events required. As mentioned earlier, you may need to observe at multiple times of the year, multiple days of the week, or multiple times during a day. It all depends on the activities, events, and behaviors in which you're interested. Imagine, for example, that you are interested in agricultural practices. These normally vary by season, so you would want to observe activities at different times of the year. Or maybe your research topic is drinking behavior. You would want to make sure you observe people and places during weekdays, weekends, and paydays, since you might expect drinking behavior to be different across these dimensions. If you can obtain some sort of census—that is, you or your research team is in the field every day for a particular event—even better.

The degree of spatial variability inherent in your research topic. The same principle for temporal variability holds for spatial variability. You want to make sure that you have as representative a set of observations as possible, within the logistical and budgetary constraints of your study. If you are observing behavior in shopping malls, for example, you might consider observing in malls within a range of geographic areas that are associated with different ethnic or socioeconomic groups. If there are few enough malls in your population, you can do even better and obtain a census of malls in your study area.

Being mindful of the above factors should help mitigate the possibility of missing a pivotally important and obvious piece of information in your observation activities. In addition to considering the three factors above, you can sample places, events, and times using any of the sampling methods discussed in this chapter. We mentioned a census, but you can choose observational venues randomly or based on specific criteria such as extreme cases. An example of the latter would be to observe drinking behavior on a specific holiday when certain locals are notoriously consumptive (think Carnival or Mardi Gras). In sum, as with sampling for other data collection methods, choosing what, when, and how to sample for participant observation activities should be guided by a combination of your research objectives, logistical parameters, and ethical considerations.

Sampling in Mixed Methods Research

Mixed methods research—defined as the integration of qualitative and quantitative methods in a single study—is becoming increasingly mainstream in social-behavioral studies. Commensurately, more is being written about sampling in mixed methods contexts (e.g., Collins, 2010; Teddlie & Yu, 2007). Of the various articles written, we find the work of Onwuegbuzie and Collins (2007) particularly insightful as it presents the mechanics of sample integration in a concise and practical way. They describe the combination of sampling efforts based on two underlying dimensions—timing of procedures and the type of relationship between sampling units in two or more samples. Onwuegbuzie and Collins (2007) combine these two dimensions with the various types of mixed methods designs to generate Figure 2.2.

Timing

Sequential. Samples are selected sequentially, at different times. A previous sample informs a subsequent sample. A subsequent sample is dependent in some way on a prior sample.

Concurrent. Samples are drawn at the same (or close in) time. Samples are independent of one another.

Relationship Between Samples

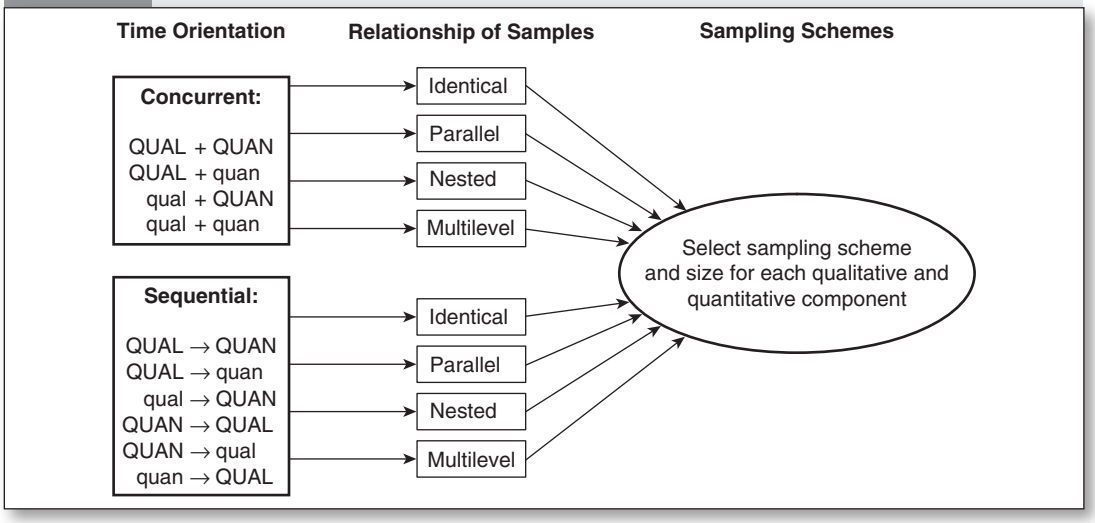
Identical. Sampling units and sample sizes are the same. Use if you want to directly compare qualitative to quantitative (QL/QT) responses at the level of measurement unit.

Parallel. Samples are drawn from the same population but are composed of different units. Use if there is no need to directly compare data at unit level or if identical sampling is too burdensome for participants.

Nested. A smaller sample is a subset of a larger sample. Use if you need to either directly compare between samples or infer from a smaller second sample to the larger sample.

Multilevel. Samples come from two or more different populations or two or more levels of analysis (definition depends on author). This feature is particularly useful if triangulating or comparing independent groups or levels of analysis.

Figure 2.2 Mixed Methods Sampling Approaches



Source: Onwuegbuzie and Collins (2007, p. 294).

RECRUITING

Recruiting methods are related to sampling strategies, but they are not the same thing. Recruitment of study participants refers to the process by which participants are first informed about and, if eligible, asked to join a study. Recruitment is the means used to obtain a desired sample. It is often an overlooked subject in the research design literature, which is why we provide a description of the more common recruitment techniques and their characteristics in Table 2.2 (see pages 68 and 69). As with every other dimension in research design, your method of recruitment will be determined by a combination of research objectives, logistical constraints and contextual parameters, and ethical considerations. More details on implementing recruiting strategies are provided in Chapters 4 and 5, in the context of in-depth interviews and focus groups, respectively.

WORKING INTERNATIONALLY

On a conceptual level, political borders and cultural practices have no bearing on sampling strategies. Sampling methods should be chosen, first and foremost, based on their appropriateness with respect to the proposed data analysis, their ability to

Table 2.2 Recruitment Techniques and Characteristics

Type	Technique	Advantages	Disadvantages
Media-based	Posters/ flyers	• Inexpensive • Targeted coverage • In many cases, the choice of venues acts as a prescreen for qualified interviewees.	• Dependent on visibility of the recruiting message • May miss many qualified interviewees
	Newspaper/ magazine	Relatively wide coverage	Expensive
	Radio/TV	Wide coverage	Expensive Coverage not highly targeted
	Internet	Inexpensive	• Little control over coverage • Dependent on visibility of the recruiting message • May miss many qualified interviewees
Investigator initiated	Door-to-door	Higher response rates (than approaches that are not face-to-face)	Labor intensive
	Facility-based	Good response rates	Generalizability limited to facility patrons
	Intercept	• If done at a venue relevant to the study topic or target interview population, can be an efficient way to locate qualified participants who might be hard to find through other methods. • May be linked to time/ space sampling.	• Involves approaching individuals in public places. • Subject to bias created by lack of representativeness of the intercept location(s) or by approach bias from the recruiter doing the intercepts
	E-mail	• Inexpensive • In many cases the choices of e-mail addresses acts as a prescreen for qualified interviewees	• Limits sample to Internet users • Potentially low response rate (dependent on population and topic) • Subject to self-selection bias

Type	Technique	Advantages	Disadvantages
	Phone	If randomly selected, useful for getting unbiased representation of residents of a particular area	• Limits sample to those with phones • Obtaining sampling frame may be difficult • Can be time consuming, especially if the screening criteria disqualify many from participation
	Mail	Inexpensive	Lower response rates
Socially-based	Chain referral (participants refer other participants)	Good for hard-to-reach populations	• Participants must have substantial social ties • Social connections between the interviewees may create bias in the research findings
	External referral (nonparticipants refer participants)	Can be effective, depending on who is referring (e.g., community leaders)	Possible ethical issues relating to participant volition
Panel/list-based	National panel	Prescreened lists, usually balanced to provide a nationally representative source of market research participants, can be easily screened to provide participants matching almost any set of demographic, behavioral, attitudinal, or other characteristics	• Expensive • Depending on how the panel is developed and maintained, panels may contain people who participate in studies simply to get the incentive money or who take on the personae of "professional respondents" causing skewed data • May miss particularly well-qualified interviewees who are not part of the database
	Research facility or professional/ affiliation-based panel	• Range in cost from free to expensive • Range in quality, depending on how the lists are developed and maintained	
		• Assuming the lists closely match the screening criteria, makes for fast, efficient recruiting of highly qualified interviewees • Useful for tough-to-locate categories of interviewees	• Dependent on the quality/ representativeness of the lists • Use of some lists may raise issues of privacy if those on the list have not preapproved being contacted by researchers

achieve the desired research objectives, and ethical considerations. But geography and culture can affect the logistics of implementing a particular sampling approach. As we discussed earlier, logistical issues can be a key factor, if not *the* primary determinant, of whom and how you sample. Reality has a way of impinging on the best of plans. Physical access to your study population, for example, can be impeded due to poor roads (whose conditions can be highly seasonal) or lack of transportation routes altogether. And helpful resources such as recruiting companies or reliable data to construct sampling frames are less likely to exist in developing countries. This is one of the areas where local collaborators and stakeholders can play a key role.

Involving local stakeholders (e.g., religious leaders, chiefs, politicians, institutional leaders) and/or local investigators before embarking on data collection is always good research practice and helps mitigate (though not eliminate) unwelcome surprises and barriers to implementation. It is especially critical when doing research in international settings. Not only is stakeholder cooperation necessary for the successful implementation of a study, but local experts can also tell you which recruitment techniques are most appropriate—and, more importantly, which ones are inappropriate—for your study population. For example, in one of the author's studies in Malawi, village chiefs helped the study team identify willing and eligible couples for the assessment of a family planning intervention (Shattuck, Kerner, Gilles, Hartmann, & Guest, 2011). Without the chiefs' help, recruitment would have taken twice as long at best, and may never have gotten off the ground at worst. Even experienced ethnographers, who may be very familiar with a particular community, rely on local community members to provide guidance for research processes and to gain entry into local communities. These same tenets hold true if conducting research among subpopulations within your own country, such as immigrant communities or socially isolated or deviant subpopulations.

SUMMING UP

The old adage that everything works well in theory is certainly true of sampling (and recruiting). A researcher may design *the* perfect sampling strategy for a study and have an airtight rationale for it. But the truth is, no matter how "perfect" a sampling strategy is, it can be quickly foiled by simple logistics on the ground. Reality has a way of keeping theory in check. Factors such as access to the study population and willingness of individuals to participate in research (or even to informally talk to field staff) are critical. Without either of these, there is no study. The study environment is another consideration. Working in conflict zones or other dangerous areas (e.g., high crime areas, red light districts) will seriously impede the ability to locate and talk freely with individuals.

The best defense against "reality" is to be as prepared as you can be. If you're not familiar with the study site or population, visit it before finalizing the research design

and sampling strategy. If a personal visit isn't possible, consult local investigators or others familiar with the area and population. Get as much information as possible before deciding on a sampling plan.

Note that reality can occasionally be helpful and present an unexpected gift. In a recent study conducted by Guest in Kampala, Uganda, one of the populations in the study was female market vendors in a huge market (8,000 plus vendors, spread across a large area, much of which was undefined). When designing the study, Guest (the U.S. based investigator) developed an elaborate sampling plan—consisting of multiple integrated sampling approaches—in order to draw a probabilistic sample from the market's overwhelming chaos. On his first visit to the field site, Guest visited the market administrator with his local collaborator to get permission to conduct the study. During this brief conversation, the administrator mentioned that he had a list of all the market vendors and that the field team could copy it. Within minutes, the sampling strategy transformed from a cumbersome, multimethod approach to a simple random sample. The key is to be open to emergent possibilities when sampling. Be flexible when things don't work out as planned, and, always have a plan B in your back pocket.

REFERENCES

Barbour, R. S. (2001). Checklists for improving rigour in qualitative research: A case of the tail wagging the dog? *British Medical Journal, 322*, 1115–1117.

Bernard, H. (2000). *Social research methods: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Bertaux, D. (1981). From the life-history approach to the transformation of sociological practice. In D. Bertaux (Ed.), *Biography and society: The life history approach in the social sciences* (pp. 29–45). London, UK: Sage.

Bluff, R. (1997). Evaluating qualitative research. *British Journal of Midwifery, 5*(4), 232–235.

Byrne, M. (2001). Evaluating the findings of qualitative research. *Association of Operating Room Nurses Journal, 73*(3), 703–706.

Carey, M. (1995). Comment: Concerns in the analysis of focus group data. *Qualitative Health Research, 5*(4), 487–495.

Chapple, A., & Rogers, A. (1998). Explicit guidelines for qualitative research: A step in the right direction, a defense of the "soft" option, or a form of sociological imperialism? *Family Practice, 15*(6), 556–561.

Charmaz, K. (2006). *Grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.

Cheek, J. (2000). An untold story: Doing funded qualitative research. In N. Denzin & Y. Lincoln (Eds.), *Handbook for qualitative research* (2nd ed., pp. 401–420). Thousand Oaks, CA: Sage.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York, NY: Psychology Press.

Collins, K. (2010). Advanced sampling designs in mixed research: Current practices and emerging trends in the social and behavioral sciences. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (2nd ed., pp. 353–378). Thousand Oaks, CA: Sage.

Creswell, J. (2006). *Qualitative inquiry and research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.

Denzin, N. (2009). The elephant in the living room, or extending the conversation about the politics of evidence. *Qualitative Research, 9*, 139–160.

Eakin, J., & Mykhalovskiy, E. (2003). Reframing the evaluation of qualitative health research: Reflections on a review of appraisal guidelines in the health sciences. *Journal of Evaluation in Clinical Practice, 9*(2), 187–194.

Flick, U. (2009). *An introduction to qualitative research: Theory, method and applications* (4th ed.). Thousand Oaks, CA: Sage.

Fossey, E., Harvey, C., McDermott, F., & Davidson, L. (2002). Understanding and evaluating qualitative research. *Australian and New Zealand Journal of Psychiatry, 36*, 717–732.

Glaser, B., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New Brunswick, NJ: Aldine Transaction.

Greenbaum, T. L. (1997). *The handbook for focus group research.* Thousand Oaks, CA: Sage.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*, 59–82.

Guest, G., & MacQueen, K. (2008). Reevaluating guidelines for qualitative research. In G. Guest & K. MacQueen (Eds.), *Handbook for team-based qualitative research* (pp. 205–226). Lanham, MD: AltaMira.

Guest, G., MacQueen, K., & Namey, E. (2012). *Applied thematic analysis*. Thousand Oaks, CA: Sage.

Johnson, J. (1990). *Selecting ethnographic informants*. Newbury Park, CA: Sage.

Kitzinger, J., & Barbour, R. S. (2001). Introduction: The challenge and promise of focus groups. In J. Kitzinger & R. S. Barbour (Eds.), *Developing focus group research: Politics, theory and practice* (pp. 1–20). London, UK: Sage.

Knodel, J. (1993). The design and analysis of focus group studies: A practical approach. In D. L. Morgan (Ed.), *Successful focus groups: Advancing the state of the art* (pp. 35–50). Thousand Oaks, CA: Sage.

Krueger, R., & Casey, M. A. (2010). *Focus groups: A practical guide for applied research* (4th ed.). Thousand Oaks, CA: Sage. (Original work published 2009)

Kuzel, A. (1992). Sampling in qualitative inquiry. In B. Crabtree & W. Miller (Eds.), *Doing qualitative research* (pp. 31–44). Newbury Park, CA: Sage.

Lambert, H., & McKevitt, C. (2002). Anthropology in health research: From qualitative methods to multidisciplinary. *British Medical Journal, 325*, 210–213.

LeCompte, M., & Schensulm J. (2010). *Designing and conducting ethnographic research: An introduction* (2nd ed.). Lanham, MD: AltaMira.

Miles, M., & Huberman, A. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.

Morgan, D. L. (1996). Focus groups. *Annual Review of Sociology, 22*, 129–152.

Morgan, M., Fischoff, B., Bostrom, A., & Atman, C. (2002). *Risk communication: A mental models approach*. New York, NY: Cambridge University Press.

Morse, J. (1994). Designing funded qualitative research. In N. Denzin & Y. Lincoln (Eds.), *Handbook for qualitative research* (pp. 220–235). Thousand Oaks, CA: Sage.

Morse, J. (1995). The significance of saturation. *Qualitative Health Research, 5*, 147–149.

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. *Proceedings of InterCHI, 93*, 206–213.

Onwuegbuzie, A., & Collins, K. (2007). A typology of mixed methods sampling designs in social science research. *Qualitative Report, 12*, 281–316.

Padgett, D. (2008). *Qualitative methods in social work research* (2nd ed.). Thousand Oaks, CA: Sage.

Patton, M. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Powell, R. A., & Single, H. M. (1996). Focus groups. *International Journal of Qualitative Health Care, 8*, 499–504.

Rubin, H., & Rubin, I. (2004). *Qualitative interviewing: The art of hearing data* (2nd ed.). Thousand Oaks, CA: Sage.

Sandelowski, M. (1995). Sample size in qualitative research. *Research in Nursing and Health, 18*, 179–183.

Sandelowski, M., & Barroso, J. (2002). Reading qualitative studies. *International Journal of Qualitative Methods, 1*(1), Article 5, 74–108.

Shattuck, D., Kerner, B., Gilles, K., Hartmann, M., & Guest, G. (2011). Encouraging contraceptive uptake by motivating men to communicate about family planning: The Malawi Male Motivator Project. *American Journal of Public Health, 101*, 1089–1095.

Teddlie, C., & Yu, F. (2007). Mixed methods sampling. *Journal of Mixed Methods Research, 1*, 77–100.

Trotter, R., II. (1991). Ethnographic research methods for applied medical anthropology. In C. Hill (Ed.), *Training manual in applied medical anthropology* (pp. 180–212). Washington, DC: American Anthropological Association.

Ulin, P. R., Robinson, E. T., & Tolley, E. E. (2005). *Qualitative methods in public health: A field guide for applied research.* San Franciso, CA: Jossey-Bass.

Vaughn, S., Schumm, J. S., & Sinagub, J. (1996). *Focus group interviews in education and psychology.* Thousand Oaks, CA: Sage.

ADDITIONAL READING

Bernard, H. (2000). Sampling. In H. R. Bernard (Ed.), *Social research methods: Qualitative and quantitative approaches* (chap. 5, pp. 143–185). Thousand Oaks, CA: Sage.

Daniel, J. (2011). Sampling essentials: Practical guidelines for making sampling choices. Thousand Oaks, CA: Sage.

Johnson, J. (1990). *Selecting ethnographic informants.* Newbury Park, CA: Sage.

Onwuegbuzie, A., & Collins, K. (2007). A typology of mixed methods sampling designs in social science research. *Qualitative Report, 12*, 281–316.

Ritchie, J., & Lewis, J. (2003). Designing and selecting samples. In *Qualitative research practice* (chap. 4). Thousand Oaks, CA: Sage.