Sampling Designs

Sampling is a complex and technical topic. Yet at the same time, sampling is familiar to us all. In the course of our daily activities, we get information, make decisions, and develop predictions through sampling. A nursing student may select an elective course by sampling two or three classes on the first day of the semester. Patients may generalize about nurses' friendliness in a particular hospital based on the care they received from a sample of nurses. We all come to conclusions about phenomena based on exposure to a limited portion of those phenomena.

Researchers, too, usually obtain data from samples. For example, in testing the efficacy of a new asthma medication, researchers reach conclusions without administering the drug to all asthmatic patients. Researchers, however, cannot afford to draw conclusions about the effectiveness of interventions or the validity of relationships based on a sample of only three or four subjects. The consequences of making erroneous decisions are more momentous in disciplined inquiries than in private decision making.

Quantitative and qualitative researchers have different approaches to sampling. Quantitative researchers seek to select samples that will allow them to generalize their results to broader groups. They therefore develop a **sampling plan** that specifies in advance how study participants are to be selected and how many to include. Qualitative researchers are not concerned with issues of generalizability but rather with a holistic understanding of the phenomenon of interest. They make sampling decisions during the course of data collection based on informational and theoretical needs, and typically do not develop a formal sampling plan in advance. This chapter discusses sampling issues for both quantitative and qualitative research.

BASIC SAMPLING CONCEPTS IN QUANTITATIVE STUDIES

Sampling is a critical part of the design of quantitative research. Let us first consider some terms associated with sampling—terms that are used primarily (but not exclusively) with quantitative studies.

Populations

A **population** is the entire aggregation of cases in which a researcher is interested. For instance, if a nurse researcher were studying American nurses with doctoral degrees, the population could be defined as all U.S. citizens who are registered nurses (RNs) and who have acquired a Ph.D., D.N.Sc., Ed.D., or other doctoral-level degree. Other possible populations might be all male patients who underwent cardiac surgery in St. Peter's Hospital during 2002, all women currently in treatment for breast cancer in Boston, or all children in Canada with cystic fibrosis. As this list illustrates, a population may be broadly defined, involving thousands of individuals, or may be narrowly specified to include only several hundred people.

Populations are not restricted to human subjects. A population might consist of all the hospital records on file in a particular hospital, all the blood samples taken from clients of a health maintenance organization, or all the high schools in the United States with a school-based clinic that dispenses contraceptives. Whatever the basic unit, the population always comprises the entire aggregate of elements in which the researcher is interested.

As noted in Chapter 9, it is sometimes useful to make a distinction between target and accessible populations. The **accessible population** is the aggregate of cases that conform to the designated criteria *and* that are accessible as a pool of subjects for a study. The **target population** is the aggregate of cases about which the researcher would like to make generalizations. A target population might consist of all diabetic people in the United States, but the accessible population might consist of all diabetic people who are members of a particular health plan. Researchers usually sample from an accessible population and hope to generalize to a target population.

TIP: A serious issue for the development of an evidence-based practice is information about the populations on whom research has been conducted. Many quantitative researchers fail to identify their target population, or discuss the issue of the generalizability of the results. The population of interest needs to be carefully considered in planning and reporting a study.

Eligibility Criteria

Researchers should be specific about the criteria that define who is included in the population. Consider the population of American nursing students. Does this population include students in all types of nursing programs? How about RNs returning to school for a bachelor's degree? Or students who took a leave of absence for a semester? Do foreign students enrolled in American nursing programs qualify? Insofar as possible, the researcher must consider the exact criteria by which it could be decided whether an individual would or would not be classified as a member of the population. The criteria that specify population characteristics are referred to as **eligibility criteria** or **inclusion criteria**. Sometimes, a population is defined in terms of characteristics that people must *not* possess (i.e., stipulating the **exclusion criteria**). For example, the population may be defined to exclude people who cannot speak English.

Inclusion or exclusion criteria for a study often reflect considerations other than substantive or theoretical interests. The eligibility criteria may reflect one or more of the following issues:

- *Costs*. Some criteria result from cost constraints. For example, when non–English-speaking people are excluded, this does not necessarily mean researchers are not interested in non–English speakers, but may mean that they cannot afford to hire translators and multilingual data collectors.
- *Practical concerns*. Sometimes, there are other practical constraints, such as difficulty in including people from rural areas, people who are hearing impaired, and so on.
- *People's ability to participate in a study*. The health condition of some people may preclude their participation. For example, people with mental impairments, who are in a coma, or who are in an unstable medical condition may need to be excluded.
- *Design considerations*. As noted in Chapter 9, it is sometimes advantageous to define a fairly homogeneous population as a means of controlling extraneous variables.

The criteria used to define a population for a research project have implications for both the interpretation of the results and the generalizability of the findings.

Example of inclusion and exclusion criteria: Keele-Smith and Price-Daniel (2001) used an experimental design to examine the effect of crossing legs on blood pressure measurements. Study participants were seniors, and could be either normotensive or hypertensive. People were excluded if they were taking antihypertensive medications and had not taken their medication that day; had a diagnosis of peripheral vascular disease; had lower leg amputations; had had surgery within the 2 prior weeks; or could not cross their legs.

Samples and Sampling

Sampling is the process of selecting a portion of the population to represent the entire population. A **sample**, then, is a subset of population elements. An **element** is the most basic unit about which information is collected. In nursing research, the elements are usually humans.

Samples and sampling plans vary in quality. *The overriding consideration in assessing a sample in a quantitative study is its representativeness*. A **representative sample** is one whose key characteristics closely approximate those of the population. If the population in a study of blood donors is 50% male and 50% female, then a representative sample would have a similar gender distribution. If the sample is not representative of the population, the external validity (generalizability) of the study is at risk.

Unfortunately, there is no way to make sure that a sample is representative without obtaining the information from the population. Certain sampling procedures are less likely to result in biased samples than others, but a representative sample can never be guaranteed. This may sound discouraging, but it must be remembered that researchers operate under conditions in which error is possible. Quantitative researchers strive to minimize those errors and, if possible, to estimate their magnitude.

Sampling designs are classified as either probability sampling or nonprobability sampling. **Probability sampling** involves random selection in choosing the elements. The hallmark of a probability sample is that researchers can specify the probability that each element of the population will be included in the sample. Probability sampling is the more respected of the two approaches because greater confidence can be placed in the representativeness of probability samples. In **nonprobability samples**, elements are selected by nonrandom methods. There is no way to estimate the probability that each element has of being included in a nonprobability sample, and every element usually does *not* have a chance for inclusion.

Strata

Sometimes, it is useful to think of populations as consisting of two or more subpopulations, or **strata**. A stratum is a mutually exclusive segment of a population, established by one or more characteristics. For instance, suppose our population was all RNs currently employed in the United States. This population could be divided into two strata based on gender. Alternatively, we could specify three strata consisting of nurses younger than 30 years of age, nurses aged 30 to 45 years, and nurses 46 years or older. Strata are often used in the sample selection process to enhance the sample's representativeness.

Sampling Bias

Researchers work with samples rather than with populations because it is more cost-effective to do so. Researchers typically have neither the time nor the resources to study all members of a population. Furthermore, it is unnecessary to gather data from an entire population; it is usually possible to obtain reasonably accurate information from a sample.

Still, data from samples *can* lead to erroneous conclusions. Finding 100 people willing to participate in a study seldom poses difficulty. It is considerably more problematic to select 100 subjects who are not a biased subset of the population. **Sampling bias** refers to the systematic over-representation or under-representation of some segment of the population in terms of a characteristic relevant to the research question.

As an example of consciously biased selection, suppose we were investigating patients' responsiveness to nurses' touch and decide to use as our sample the first 50 patients meeting eligibility criteria in a specific hospital unit. We decide to omit Mr. Z from the sample because he has shown hostility to nurses. Mrs. X, who has just lost a spouse, is also excluded from the study because she is under stress. We have made conscious decisions to exclude certain individuals, and the decisions do not reflect bona fide eligibility criteria. This can lead to bias because responsiveness to nurses' touch (the dependent variable) may be affected by patients' feelings about nurses or their emotional state.

Sampling bias usually occurs unconsciously, however. If we were studying nursing students and systematically interviewed every 10th student who entered the nursing library, the sample of students would be biased in favor of library-goers, even if we are conscientious about including every 10th entrant regardless of the person's appearance, gender, or other traits.

Sampling bias is partly a function of population homogeneity. If population elements were all identical with respect to key attributes, then any sample would be as good as any other. Indeed, if the population were completely homogeneous, that is, exhibited no variability at all, then a *single* element would be a sufficient sample to draw conclusions about the population. For many physiologic attributes, it may be safe to assume a reasonably high degree of homogeneity. For example, the blood in a person's veins is relatively homogeneous and so a single blood sample chosen haphazardly is adequate. For most human attributes, however, homogeneity is the exception rather than the rule. Age, health condition, stress, attitudes, habits—all these attributes reflect the heterogeneity of humans. When variation occurs in the population, then similar variation ideally should be reflected in a sample.

TIP: One straightforward way to increase the generalizability of a study is to select study participants from two or more sites, such as from different hospitals, nursing homes, communities, and so on. Ideally, the two different sites would be sufficiently divergent that broader representation of the population would be obtained.

NONPROBABILITY SAMPLING

Nonprobability sampling is less likely than probability sampling to produce accurate and representative samples. Despite this fact, most research samples in nursing and other disciplines are nonprobability samples. Three primary methods of nonprobability sampling are convenience, quota, and purposive.

Convenience Sampling

Convenience sampling entails using the most conveniently available people as study participants. A faculty member who distributes questionnaires to nursing students in a class is using a convenience sample, or an **accidental sample**, as it is sometimes called. The nurse who conducts an observational study of women delivering twins at the local hospital is also relying on a convenience sample. The problem with convenience sampling is that available subjects might be atypical of the population of interest with regard to critical variables.

Convenience samples do not necessarily comprise individuals known to the researchers. Stopping people at a street corner to conduct an interview is sampling by convenience. Sometimes, researchers seeking people with certain characteristics place an advertisement in a newspaper, put up signs in clinics or supermarkets, or post messages in chat rooms on the Internet. These approaches are subject to bias because people select themselves as pedestrians on certain streets or as volunteers in response to posted notices.

Snowball sampling (also called **network sampling** or **chain sampling**) is a variant of convenience sampling. With this approach, early sample members are asked to identify and refer other people who meet the eligibility criteria. This method of sampling is often used when the research population is people with specific traits who might otherwise be difficult to identify (e.g., people who are afraid of hospitals). Snowballing begins with a few eligible study participants and then continues on the basis of referrals from those participants until the desired sample size has been obtained.

Convenience sampling is the weakest form of sampling. It is also the most commonly used sampling method in many disciplines. In heterogeneous populations, there is no other sampling approach in which the risk of sampling bias is greater.

Example of a convenience sample: Board and Ryan-Wenger (2002) prospectively examined the long-term effects of the pediatric intensive care unit experience on parents and on family adaptation. The researchers used convenience sampling to recruit three groups of parents: those with a hospitalized child in the pediatric intensive care unit, those with a child in a general care unit, and those with nonhospitalized ill children.

Quota Sampling

A **quota sample** is one in which the researcher identifies population strata and determines how many participants are needed from each stratum. By using information about population characteristics, researchers can ensure that diverse segments are represented in the sample, preferably in the proportion in which they occur in the population.

Suppose we were interested in studying nursing students' attitude toward working with AIDS patients. The accessible population is a school of nursing with an undergraduate enrollment of 1000 students; a sample of 200 students is desired. The easiest procedure would be to use a convenience sample by distributing questionnaires in classrooms or catching students as they enter or leave the library. Suppose, however, we suspect that male and female students have different attitudes toward working with AIDS victims. A convenience sample might result in too many men or women. Table 13-1 presents fictitious data showing the gender distribution for the population and for a convenience sample in the first two columns. In this example, the convenience sample over-represents women and under-represents men. We can, however, guide the selection of study participants so that the sample includes the correct number of cases from both strata. The far-right panel of Table 13-1 shows the number of men and women required for a quota sample for this example.

If we pursue this same example a bit further, you may better appreciate the dangers of a biased sample. Suppose that one of the key questions in this study was, "Would you be willing to work on a unit that cared exclusively for AIDS patients?" The percentage of students in the population who would respond "yes" to this inquiry is shown in the first column of Table 13-2. Of course, we would not know these values; they are displayed to illustrate a point. Within the population, men are more likely than women to be willing to work on a unit with AIDS patients, yet men were seriously under-represented in the convenience sample. As a result, there is a discrepancy between the population and sample values on the outcome variable: about 27% more students in the population are favorable toward working with AIDS victims (14.0%) than we would conclude based on results from the convenience sample (11.0%). The quota sample, on the other hand, does a better job of reflecting the viewpoint of the population (14.5%). In actual research situations, the distortions from a convenience sample may be smaller than in this example, but could be larger as well.

Quota sampling does not require sophisticated skills or a lot of effort—and it is surprising that so

TABLE 13.1 Numbers and Percentages of Students in Strata of a Population, Convenience Sample, and Quota Sample

STRATA	POPULATION	CONVENIENCE SAMPLE	QUOTA SAMPLE
Male	200 (20%)	10 (5%)	40 (20%)
Female	800 (80%)	190 (95%)	160 (80%)
Total	1000 (100%)	200 (100%)	200 (100%)

TABLE 13.2 Students Willing to Work on an AIDS Unit, in the Population, Convenience Sample, and Quota Sample

	NUMBER IN POPULATION	NUMBER IN CONVENIENCE SAMPLE	NUMBER IN QUOTA SAMPLE
Male	55	3	12
Female	85	19	17
Total Number of Willing Students	140	22	29
Total Number of All Students	1000	200	200
Percentage Willing	14.0%	11.0%	14.5%

few researchers use this strategy. Many researchers who use a convenience sample could probably design a quota sampling plan, and it would be advantageous to do so. Stratification should be based on one or more variables that would reflect important differences in the dependent variable under study. Such variables as age, gender, ethnicity, educational attainment, and medical diagnosis are often good stratifying variables.

Except for identifying the strata and the desired representation for each, quota sampling is procedurally similar to convenience sampling. The subjects in any particular cell constitute, in essence, a convenience sample from that stratum of the population. Referring back to the example in Table 13-1, the initial sample of 200 students constituted a convenience sample from the population of 1000. In the quota sample, the 40 men constitute a convenience sample of the 200 men in the population. Because of this fact, quota sampling shares many of the same weaknesses as convenience sampling. For instance, if a researcher is required by a quota sampling plan to interview 10 men between the ages of 65 and 80 years, a trip to a nursing home might be the most convenient method of obtaining those subjects. Yet this approach would fail to represent the many senior citizens who live independently in the community. Despite its problems, quota sampling represents an important improvement over convenience sampling and should be considered by quantitative researchers whose resources prevent the use of probability sampling.

Example of a quota sample:
Williams, Soetjiningsih, and Williams (2000) studied Balinese mothers' expectations for children's development. The researchers used quota sampling to ensure an equal number of urban and rural Balinese mothers, and an equal number of male and female children.

Purposive Sampling

Purposive sampling or **judgmental sampling** is based on the belief that researchers' knowledge about the population can be used to hand-pick sample members. Researchers might decide purposely to select subjects who are judged to be typical of the population or particularly knowledgeable about the issues under study. Sampling in this subjective manner, however, provides no external, objective method for assessing the typicalness of the selected subjects. Nevertheless, this method can be used to advantage in certain situations. Newly developed instruments can be effectively pretested and evaluated with a purposive sample of diverse types of people. Purposive sampling is often used when researchers want a sample of experts, as in the case of a needs assessment using the key informant approach or in Delphi surveys. Also, as discussed later in this chapter, purposive sampling is frequently used by qualitative researchers.

Example of purposive sampling: Friedemann, Montgomery, Rice, and Farrell (1999) studied family members' involvement in the nursing home. The first stage of their sampling plan involved purposively sampling 24 nursing homes with a diversity of policies related to family involvement, based on a survey of 208 nursing homes in southern Michigan. In the second stage, all family members of residents admitted to these nursing homes during a 20-month window were invited to participate.

Evaluation of Nonprobability Sampling

Although a nonprobability sample is often acceptable for pilot, exploratory, or in-depth qualitative research, for most quantitative studies, the use of nonprobability samples is problematic. Nonprobability samples are rarely representative of the population. When every element in the population does not have a chance of being included in the sample, it is likely that some segment of it will be systematically underrepresented.

Why, then, are nonprobability samples used in most nursing studies? Clearly, the advantage of these sampling designs lies in their convenience and economy. Probability sampling, discussed next, requires skill and resources. There is often no option but to use a nonprobability approach or to abandon the project altogether. Even hard-nosed research methodologists would hesitate to advocate the abandonment of an idea in the absence of a random sample. Quantitative researchers using nonprobability samples out of necessity must be cautious about the inferences and conclusions drawn from the data. With care in the selection of the sample, a conservative interpretation of the results, and replication of the study with new samples, researchers may find that nonprobability samples work reasonably well.

TIP: If you use a convenience sample, you can still take steps to enhance the sample's representativeness. First, identify important extraneous variables—factors that affect variation in the dependent variable. For example, in a study of the effect of stress on health, family income would be an important extraneous variable because poor people tend to be less healthy (*and* more stressed) than more affluent ones. Then, decide how to account for this source of variation in the sampling design. One solution is to eliminate variation from extraneous variables, as discussed in Chapter 9. In the stress and health example, we might restrict the population to middle-class people. Alternatively, we could select the convenience sample from two communities known to differ socioeconomically so that our sample would reflect the experiences of both lower- and middle-class subjects. This approaches using a quota sampling method. In other words, if the population is known to be heterogeneous, you should take steps either to make it more homogeneous (thereby redefining the population) or to capture the full variation in the sample.

PROBABILITY SAMPLING

Probability sampling involves the random selection of elements from a population. Random selection should not be (although it often is) confused with random assignment, which was described in connection with experimental designs in Chapter 8. Random assignment refers to the process of allocating subjects to different treatment conditions at random. Random *assignment* has no bearing on how subjects in an experiment were selected in the first place. **Random sampling** involves a selection process in which each element in the population has an equal, independent chance of being selected. The four most commonly used probability sampling methods are simple random, stratified random, cluster, and systematic sampling.

Simple Random Sampling

Simple random sampling is the most basic probability sampling design. Because the more complex probability sampling designs incorporate features of simple random sampling, the procedures involved are described here in some detail.

In simple random sampling, researchers establish a **sampling frame,** the technical name for the list of the elements from which the sample will be chosen. If nursing students at the University of Connecticut were the accessible population, then a roster of those students would be the sampling frame. If the sampling unit were 500-bed (or larger) hospitals in Canada, then a list of all such hospitals would be the sampling frame. In practice, a population may be defined in terms of an existing sampling frame rather than starting with a population and developing a list of elements. For example, if we wanted to use a telephone directory as a sampling frame, we would have to define the population as community residents who are customers of the telephone company *and* who had a number listed at the time the directory was published. Because not all members of a community own a telephone and others do not have listed numbers, it would not be appropriate to consider a telephone directory as the sampling frame for the entire population.

Once a sampling frame has been developed, elements are numbered consecutively. A table of random numbers would then be used to draw a sample of the desired size. An example of a sampling frame for a population of 50 people is presented in Table 13-3. Let us assume we want to select randomly a sample of 20 people. As in the case of random assignment, we would find a starting place in a table of random numbers by blindly placing our finger at some point on the page. To include all numbers between 1 and 50, two-digit combinations would be read. Suppose, for the sake of the example, that we began random selection with the first number in the random number table of Table 8-2 (p. 171), which is 46. The person corresponding to that number, D. Abraham, is the first subject selected to participate in the study. Number 05, C. Eldred, is the second selection, and number 23, R. Yarinsky, is the third. This process would continue until the 20 subjects were chosen. The selected elements are circled in Table 13-3.

TABLE 13.3 Sampling Frame for Simple Random Sampling Example

1. N. Alexander	26. G. Berlin
2. T. Brock	27. C. Coulton
3. H. Collado	28. R. De los Santos
4. F. Doolittle	29. D. Edelstein
5. C. Eldred	30. B. Fink
6. R. Fellerath	31. J. Gueron
7. B. Goldman	32. J. Hunter
8. G. Hamilton	33. R. Joyce
9. R. Ivry	34. Y. Kim
10. S. James	35. A. London
11. V. Knox	36. J. Martinez
12. S. Lynn	37. C. Nicholson
13. C. Michalopoulos	38. R. Ortega
14. L. Nelson	39. K. Paget
15. J. O'Brien	40. G. Queto
16. M. Price	41. J. Riccio
17. J. Quint	42. E. Scott
18. D. Romm	43. L. Traeger
19. R. Seupersad	44. E. Vallejo
20. P. Tang	45. J. Wallace
21. N. Verma	46. D. Abraham
22. R. Widom	47. D. Butler
23. R. Yarinsky	48. O. Cardenas
24. M. Zaslow	49. F. Derocher
25. M. Agudelo	50. K. Edin

It should be clear that a sample selected randomly in this fashion is not subject to researchers' biases. Although there is no guarantee that a randomly drawn sample will be representative, random selection does ensure that differences in the attributes of the sample and the population are purely a function of chance. The probability of selecting a markedly deviant sample is low, and this probability decreases as the size of the sample increases.

Simple random sampling tends to be laborious. Developing the sampling frame, numbering all the elements, and selecting sample elements are time-consuming chores, particularly if the population is large. Imagine enumerating all the telephone subscribers listed in the New York City telephone directory! If the elements can be arranged in computer-readable form, then the computer can be programmed to select the sample automatically. In actual practice, simple random sampling is not used frequently because it is a relatively inefficient procedure. Furthermore, it is not always possible to get a listing of every element in the population, so other methods may be required.

Example of a simple random sample: Yoon and Horne (2001) studied the use of herbal products for medicinal purposes in a sample of older women. A random sample of 86 women aged 65 or older who lived independently in a Florida County was selected, using a sampling frame compiled from information from the state motor vehicle agency.

Stratified Random Sampling

In **stratified random sampling,** the population is first divided into two or more strata. As with quota sampling, the aim of stratified sampling is to enhance representativeness. Stratified sampling designs subdivide the population into homogeneous subsets from which an appropriate number of elements are selected at random.

Stratification is often based on such demographic attributes as age, gender, and income level. One difficulty is that the stratifying attributes must be known in advance and may not be readily discernible. If you were working with a telephone directory, it would be risky to guess a person's gender, and age, ethnicity, or other personal information could not be used as stratifying variables. Patient listings, student rosters, or organizational directories might contain information for a meaningful stratification. Quota sampling does not have the same problem because researchers can ask prospective subjects questions that determine their eligibility for a particular stratum. In stratified sampling, however, a person's status in a stratum must be known before random selection.

The most common procedure for drawing a stratified sample is to group together elements belonging to a stratum and to select randomly the desired number of elements. Researchers can either select an equal number of elements from each stratum or select unequal numbers, for reasons discussed later. To illustrate the procedure used in the simplest case, suppose that the list in Table 13-3 consisted of 25 men (numbers 1 through 25) and 25 women (numbers 26 through 50). Using gender as the stratifying variable, we could guarantee a sample of 10 men and 10 women by randomly sampling 10 numbers from the first half of the list and 10 from the second half. As it turns out, our simple random sampling did result in 10 elements being chosen from each half of the list, but this was purely by chance. It would not have been unusual to draw, say, 8 names from one half and 12 from the other. Stratified sampling can guarantee the appropriate representation of different segments of the population.

Stratifying variables usually divide the population into unequal subpopulations. For example, if the person's race were used to stratify the population of U. S. citizens, the subpopulation of white people would be larger than that of African-American and other nonwhite people. The researcher might decide to select subjects in proportion to the size of the stratum in the population, using **proportionate stratified sampling**. If the population was students in a nursing school that had 10% African-American students, 10% Hispanic students, and 80% white students, then a proportionate stratified sample of 100 students, with racial/ethnic background as the stratifying variable, would consist of 10, 10, and 80 students from the respective strata.

When researchers are interested in understanding differences among strata, proportionate sampling may result in insufficient numbers for making comparisons. In the previous example, would the researcher be justified in drawing conclusions about the characteristics of Hispanic nursing students based on only 10 cases? It would be unwise to do so. For this reason, researchers often adopt a **disproportionate sampling design** when comparisons are sought between strata of greatly unequal size. In the example, the sampling proportions might be altered to select 20 African-American students, 20 Hispanic students, and 60 white students. This design would ensure a more adequate representation of the two racial/ethnic minorities. When disproportionate sampling is used, however, it is necessary to make an adjustment to the data to arrive at the best estimate of *overall* population values. This adjustment process, known as **weighting**, is a simple mathematic computation described in textbooks on sampling.

Stratified random sampling enables researchers to sharpen the precision and representativeness of the final sample. When it is desirable to obtain reliable information about subpopulations whose memberships are relatively small, stratification provides a means of including a sufficient number of cases in the sample by oversampling for that stratum. Stratified sampling, however, may be impossible if information on the critical variables is unavailable. Furthermore, a stratified sample requires even more labor and effort than simple random sampling because the sample must be drawn from multiple enumerated listings.

Example of stratified random sampling: Bath, Singleton, Strikas, Stevenson, McDonald, and Williams (2000) conducted a survey to determine the extent to which hospitals with labor and delivery services had policies about screening pregnant women for hepatitis B. A stratified random sample of 968 hospitals (stratified by number of beds and affiliation with a medical school) was selected.

Cluster Sampling

For many populations, it is impossible to obtain a listing of all elements. For example, the population of full-time nursing students in the United States would be difficult to list and enumerate for the purpose of drawing a simple or stratified random sample. It might also be prohibitively expensive to sample students in this way because the resulting sample would include only one or two students per institution. If personal interviews were involved, the interviewers would have to travel to students scattered throughout the country. Large-scale surveys almost never use simple or stratified random sampling; they usually rely on cluster sampling.

In **cluster sampling**, there is a successive random sampling of units. The first unit is large groupings, or clusters. In drawing a sample of nursing students, we might first draw a random sample of nursing schools and then draw a sample of students from the selected schools. The usual procedure for selecting samples from a general population is to sample successively such administrative units as states, cities, census tracts, and then households. Because of the successive stages in cluster sampling, this approach is often called **multistage sampling**. The resulting design is usually described in terms of the number of stages (e.g., three-stage cluster sampling).

The clusters can be selected either by simple or stratified methods. For instance, in selecting clusters of nursing schools, it may be advisable to stratify on program type. The final selection from within a cluster may also be performed by simple or stratified random sampling.

For a specified number of cases, cluster sampling tends to be less accurate than simple or stratified random sampling. Despite this disadvantage, cluster sampling is more economical and practical than other types of probability sampling, particularly when the population is large and widely dispersed.

Example of cluster/multistage sampling: Trinkoff, Zhou, Storr, and Soeken (2000) studied nurses' substance abuse, using data from a two-stage cluster sample. In the first stage, 10 states in the United States were selected using a complex stratification procedure. In the second stage, RNs were selected from each state (a total sample of 3600) by simple random sampling.

Systematic Sampling

The final sampling design can be either probability or nonprobability sampling, depending on the exact procedure used. **Systematic sampling** involves the selection of every *k*th case from a list or group, such as every 10th person on a patient list or every 100th person in a directory of American Nurses Association members. Systematic sampling is sometimes used to sample every *k*th person entering a bookstore, or passing down the street, or leaving a hospital, and so forth. In such situations, unless the population is narrowly defined as all those people entering, passing by, or leaving, the sampling is nonprobability in nature.

Systematic sampling can be applied so that an essentially random sample is drawn. If we had a list, or sampling frame, the following procedure could be adopted. The desired sample size is established at some number (n). The size of the population must be known or estimated (N). By dividing N by n, the sampling interval width (k) is established. The **sampling interval** is the standard distance between elements chosen for the sample. For instance, if we were seeking a sample of 200 from a population of 40,000, then our sampling interval would be as follows:

$$k = \frac{40{,}000}{200} = 200$$

In other words, every 200th element on the list would be sampled. The first element should be selected randomly, using a table of random numbers. Let us say that we randomly selected number 73 from a table. The people corresponding to numbers 73, 273, 473, 673, and so forth would be sampled. Alternatively, we could randomly select a number from 1 to the number of elements listed on a page, and then randomly select every *k*th unit on all pages (e.g., number 38 on every page).

Systematic sampling conducted in this manner yields essentially the same results as simple random sampling, but involves far less work. Problems would arise if the list were arranged in such a way that a certain type of element is listed at intervals coinciding with the sampling interval. For instance, if every 10th nurse listed in a nursing personnel roster were a head nurse and the sampling interval was 10, then head nurses would either always or never be included in the sample. Problems of this type are rare, fortunately. In most cases, systematic sampling is preferable to simple random sampling because the same results are obtained in a more efficient manner. Systematic sampling can also be applied to lists that have been stratified.

Example of a systematic sample: Tolle, Tilden, Rosenfeld, and Hickman (2000) explored barriers to optimal care of the dying by surveying family members of decedents. Their sampling frame was 24,074 death certificates in Oregon, from which they sampled, through systematic sampling, 1458 certificates. They then traced as many family members of the decedents as possible and conducted telephone interviews.

Evaluation of Probability Sampling

Probability sampling is the only viable method of obtaining representative samples. If all the elements in the population have an equal probability of being selected, then the resulting sample is likely to do a good job of representing the population. A further advantage is that probability sampling allows researchers to estimate the magnitude of sampling error. **Sampling error** refers to differences between population values (such as the average age of the population) and sample values (such as the average age of the sample). It is a rare sample that is perfectly representative of a population; probability sampling permits estimates of the degree of error. Advanced textbooks on sampling elaborate on procedures for making such estimates.

The great drawbacks of probability sampling are its inconvenience and complexity. It is usually beyond the scope of most researchers to sample using a probability design, unless the population is narrowly defined—and if it *is* narrowly defined, probability sampling may seem like “overkill.” Probability sampling is the preferred and most respected method of obtaining sample elements, but it may in some cases be impractical.

TIP: Whenever possible, it is useful to compare sample characteristics with population characteristics. Published information about the characteristics of many groups of interest to nurses may be available to help provide a context for evaluating sampling bias. For example, if you were studying low-income children in Detroit, you could obtain information on the Internet about salient characteristics (e.g., race/ethnicity, age distribution) of low-income American children from the U. S. Bureau of the Census. Population characteristics could then be compared with sample characteristics, and differences taken into account in interpreting the findings.

SAMPLE SIZE IN QUANTITATIVE STUDIES

Quantitative researchers need to pay careful attention to the number of subjects needed to test research hypotheses adequately. A sophisticated procedure known as **power analysis** can be used to estimate sample size needs, but some statistical knowledge is needed before this procedure can be explained. In this section we offer guidelines to beginning researchers; advanced students can read about power analysis in Chapter 20 or consult a sampling or statistical textbook.

There are no simple formulas that can tell you how large a sample is needed in a given quantitative study, but we can offer a simple piece of advice: You should use the largest sample possible. The larger the sample, the more representative of the population it is likely to be. Every time researchers calculate a percentage or an average based on sample data, they are estimating a population value. Smaller samples tend to produce less accurate estimates than larger ones. In other words, the larger the sample, the smaller the sampling error.

Let us illustrate this with an example of monthly aspirin consumption in a nursing home facility (Table 13-4). The population consists of 15 residents whose aspirin consumption averages 16 aspirins per month, as shown in the top row of the table. Eight simple random samples—two each with sample sizes of 2,

TABLE 13.4 Comparison of Population and Sample Values and Averages: Nursing Home Aspirin Consumption Example

NUMBER OF PEOPLE IN GROUP	GROUP	INDIVIDUAL DATA VALUES (NUMBER OF ASPIRINS CONSUMED, PRIOR MONTH)	AVERAGE
15	Population	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30	16.0
2	Sample 1A	6, 14	10.0
2	Sample 1B	20, 28	24.0
3	Sample 2A	16, 18, 8	14.0
3	Sample 2B	20, 14, 26	20.0
5	Sample 3A	26, 14, 18, 2, 28	17.6
5	Sample 3B	30, 2, 26, 10, 4	14.4
10	Sample 4A	22, 16, 24, 20, 2, 8, 14, 28, 20, 4	15.8
10	Sample 4B	12, 18, 8, 10, 16, 6, 28, 14, 30, 22	16.4

3, 5, and 10—have been drawn. Each sample average represents an estimate of the population average (16). Under ordinary circumstances, of course, the population value would be unknown, and we would draw only one sample. With a sample size of two, our estimate might have been wrong by as many as eight aspirins (sample 1B, average of 24), which is a 50% error. As the sample size increases, the averages get closer to the true population value, *and* the differences in the estimates between samples A and B get smaller as well. As the sample size increases, the probability of getting a markedly deviant sample diminishes. Large samples provide an opportunity to counterbalance atypical values. Unless a power analysis can be done, the safest procedure is to obtain data from as large a sample as is practically feasible.

Large samples are no assurance of accuracy, however. When nonprobability sampling methods are used, even a large sample can harbor extensive bias. The famous example illustrating this point is the 1936 presidential poll conducted by the magazine *Literary Digest,* which predicted that Alfred M. Landon would defeat Franklin D. Roosevelt by a landslide. About 2.5 million individuals participated in this poll—a substantial sample. Biases resulted from the fact that the sample was drawn from telephone directories and automobile registrations during a depression year when only the well-to-do (who preferred Landon) had a car or telephone. Thus, a large sample cannot correct for a faulty sampling design.

Because practical constraints such as time, subject availability, and resources often limit sample size, many nursing studies are based on relatively small samples. In a survey of nursing studies published over four decades (the 1950s to the 1980s), Brown, Tanner and Padrick (1984) found that the average sample size was under 100 subjects in all four decades, and similar results were reported in a more recent analysis (Moody, Wilson, Smyth, Schwartz, Tittle, & VanCott, 1988). In many cases, a small sample can lead to misleading or inconclusive results. Below we discuss some considerations that affect sample size requirements in quantitative studies.

Homogeneity of the Population

If there is reason to believe that the population is relatively homogeneous, a small sample may be adequate. Let us demonstrate that this is so. The top half of Table 13-5 presents hypothetical population values for three different populations, with only 10 people in each population. These values could

TABLE 13.5 Three Populations of Different Homogeneity

GROUP	INDIVIDUAL DATA VALUES	LOWEST VALUE	HIGHEST VALUE	AVERAGE
Population A	100 110 105 95 90 110 105 95 90 100	90	110	100.0
Population B	110 120 105 85 80 120 115 85 80 100	80	120	100.0
Population C	100 130 125 75 70 130 125 75 70 100	70	130	100.0
Sample A	110 90 95	90	110	98.3
Sample B	120 80 85	80	120	95.0
Sample C	125 70 75	70	125	90.0

reflect, for example, scores on a measure of anxiety. In all three populations, the average anxiety score is 100. In population A, however, the individuals have similar anxiety scores, ranging from a low of 90 to a high of 110. In population B, the scores are more variable, and in population C, the scores are more variable still, ranging from 70 to 130.

The second half of Table 13-5 presents three sample values from the three populations. In the most homogeneous population (A), the average anxiety score for the sample is 98.3, which is close to the population average of 100. As the population becomes less homogeneous, the average sample values less accurately reflect population values. In other words, there is greater sampling error when the population is heterogeneous on the key variable. By increasing the sample size, the risk of sampling error would be reduced. For example, if sample C consisted of five values rather than three (say, all the even-numbered population values), then the sample average would be closer to the population average (i.e., 102 rather than 90).

For clinical studies that deal with biophysiologic processes in which variation is limited, a small sample may adequately represent the population. For most nursing studies, however, it is safer to assume a fair degree of heterogeneity, unless there is evidence from prior research to the contrary.

Effect Size

Power analysis builds on the concept of an **effect size**, which expresses the strength of relationships among research variables. If there is reason to expect that the independent and dependent variables will be strongly related, then a relatively small sample should be adequate to demonstrate the relationship statistically. For example, if we were testing a powerful new drug to treat AIDS, it might be possible to demonstrate its effectiveness with a small sample. Typically, however, interventions have modest effects, and variables are usually only moderately correlated with one another. When there is no *a priori* reason for believing that relationships will be strong (i.e., when the effect size is expected to be modest), then small samples are risky.

Attrition

In longitudinal studies, the number of subjects usually declines over time. This is most likely to occur if the time lag between data collection points is great; if the population is mobile or hard to locate; or if the population is a vulnerable one at risk of death or disability. If resources are devoted to tracing subjects, or if the researcher has an ongoing relationship with them (as might be true in clinical studies), then the rate of attrition might be low. It is a rare longitudinal study, however, that maintains the full research sample. Therefore, in estimating sample size needs, researchers should factor in anticipated loss of subjects over time.

Attrition problems are not restricted to longitudinal studies. People who initially agree to cooperate in a study may be subsequently unable or unwilling to participate for various reasons, such as death, deteriorating health, early discharge, discontinued need for an intervention, or simply a change of heart. Researchers should expect a certain amount of subject loss and recruit accordingly.

Subgroup Analyses

Researchers are sometimes interested in testing hypotheses not only for an entire population but for subgroups. For example, we might be interested in determining whether a structured exercise program is effective in improving infants' motor skills. After testing the general hypothesis with a sample of infants, we might wish to test whether the intervention is more effective for certain infants (e.g., low-birthweight versus normal-birth-weight infants). When a sample is divided to test for **subgroup effects**, the sample must be large enough to support these divisions of the sample.

Sensitivity of the Measures

Instruments vary in their ability to measure key concepts precisely. Biophysiologic measures are usually very sensitive—they measure phenomena accurately, and can make fine discriminations in values. Psychosocial measures often contain a fair amount of error and lack precision. When measuring tools are imprecise and susceptible to errors, larger samples are needed to test hypotheses adequately.

IMPLEMENTING A SAMPLING PLAN IN QUANTITATIVE STUDIES

Once decisions are made about the sampling design and sample size, the plan must be implemented. This section provides some practical information about implementation of a sampling plan.

Steps in Sampling in Quantitative Studies

The steps to be undertaken in drawing a sample vary somewhat from one sampling design to the next, but a general outline of procedures can be described.

1. *Identify the population*. It is good to begin with a clear idea about the target population to which you would ideally like to be able to generalize your results. Unless you have extensive resources, you are unlikely to have access to the entire target population, and so you will also need to identify the portion of the target population that is accessible to you. Researchers often *begin* by identifying an accessible population, and then decide how best to define the target population.
2. *Specify the eligibility criteria*. The criteria for eligibility in the sample should then be spelled out. The criteria should be as specific as possible with respect to characteristics that might exclude potential subjects (e.g., extremes of poor health, inability to read English). The criteria might lead you to redefine your target population.
3. *Specify the sampling plan*. Once the accessible population has been identified, you must decide (a) the method of drawing the sample and (b) how large it will be. Sample size specifications should consider the aspects of the study discussed in the previous section. If you can perform a power analysis to determine the desired number of subjects, it is highly recommended that you do so. Similarly, if probability sampling is an option for selecting a sample, that option should be exercised. If you are not in a position to do either, we recommend using as large a sample as possible and taking steps to build representativeness into the design (e.g., by using quota sampling).
4. *Recruit the sample*. Once the sampling design has been specified, the next step is to recruit prospective study participants according to the plan (after any needed institutional permissions have been obtained) and ask for their cooperation. Issues relating to subject recruitment are discussed next.

Sample Recruitment

Recruiting subjects to participate in a study involves two major tasks: identifying eligible candidates and persuading them to participate. Researchers may in some cases need to spend time early in the project deciding the best sources for recruiting potential participants. Researchers must ask such questions as, Where do people with the characteristics I want live or obtain care in large numbers? Will I have direct access to subjects, or will I need administrative approval? Will there be sufficiently large numbers in one location, or will multiple sites be necessary? During the recruitment phase, it may be necessary to develop a **screening instrument**, which is a brief interview or form that allows researchers to determine whether a prospective subject meets all eligibility criteria for the study.

The next task involves actually gaining the cooperation of people who have been deemed eligible for the study. It is critical to have an effective recruitment strategy. Most people, given the right circumstances, will agree to cooperate, but some are hesitant. Researchers should ask themselves, What will make this research experience enjoyable, worthwhile, convenient, pleasant, and nonthreatening for subjects? Factors over which researchers have control that can influence the rate of cooperation include the following:

- *Recruitment method*. Face-to-face recruitment is usually more effective than solicitation by a telephone call or a letter.

- *Courtesy*. Successful recruitment depends on using recruiters who are pleasant, courteous, respectful, and nonthreatening. Cooperation sometimes is enhanced if characteristics of recruiters are similar to those of prospective subjects—particularly with regard to gender, race, and ethnicity.
- *Persistence*. Although high-pressure tactics are never acceptable, persistence may sometimes be needed. When prospective subjects are first approached, their initial reaction may be to decline participation, because they might be taken off guard. If a person hesitates or gives an equivocal answer at the first attempt, recruiters should ask if they could come back at a later time.
- *Incentives*. Gifts and monetary incentives have been found to increase participation rates.
- *Research benefits*. The benefits of participating to the individual and to society should be carefully explained, without exaggeration or misleading information.
- *Sharing results*. Sometimes it is useful to provide people with tangible evidence of their contribution to the study by offering to send them a brief summary of the study results.
- *Convenience*. Every effort should be made to collect data at a time and location that is convenient for subjects. In some cases, this may mean making arrangements for transportation or for the care of young children.
- *Endorsements*. It may be valuable to have the study endorsed or acknowledged by a person, group, or organization that has prospective subjects' confidence, and to communicate this to them. Endorsements might come from the institution serving as the research setting, from a funding agency, or from a respected community group or person, such as a church leader. Press releases in advance of recruitment are sometimes advantageous.
- *Assurances*. Prospective subjects should be told who will see the data, what use will be made of the data, and how confidentiality will be maintained.

TIP: Subject recruitment often proceeds at a slower pace than researchers anticipate. Once you have determined your sample size needs, it is a good idea to develop contingency plans for recruiting more subjects, should the initial plan prove overly optimistic. For example, a contingency plan might involve relaxing the eligibility criteria, identifying another institution through which participants could be recruited, offering incentives to make participation more attractive, or the lengthening the recruitment period. When contingency plans are developed at the outset, it reduces the likelihood that you will have to settle for a less-than-desirable sample size.

Generalizing From Samples

Ideally, the sample is representative of the accessible population, and the accessible population is representative of the target population. By using an appropriate sampling plan, researchers can be reasonably confident that the first part of this ideal has been realized. The second part of the ideal entails greater risk. Are the unemployed nurses in Atlanta representative of all unemployed nurses in the United States? Researchers must exercise judgment in assessing the degree of similarity.

The best advice is to be realistic and conservative, and to ask challenging questions: Is it reasonable to assume that the accessible population is representative of the target population? In what ways might they differ? How would such differences affect the conclusions? If differences are great, it would be prudent to specify a more restricted target population to which the findings could be meaningfully generalized.

TIP: As you recruit your sample, it is wise to document thoroughly. The more information you have about who the sample is and who it is not, the better able you will be to identify potential biases. Biases can occur even in probability sampling because the *selection* of elements that are representative of the population does not guarantee the *participation* of all those elements, and refusal to participate in a study is rarely random. Thus, you should calculate a **response rate** (the number of people participating in the study relative to the number of people sampled) and document the

nonresponse bias, that is, differences between the characteristics of participants and those of people who refused to participate in the study. Also, those who remain in a study should be compared with those who drop out to document any attrition biases. It may also be useful to document the reasons people give for not cooperating (or not continuing to cooperate) in a study.

SAMPLING IN QUALITATIVE RESEARCH

Qualitative studies almost always use small, nonrandom samples. This does not mean that qualitative researchers are unconcerned with the quality of their samples, but rather that they use different criteria for selecting study participants. This section examines considerations that apply to sampling in qualitative studies.

The Logic of Qualitative Sampling

Quantitative research is concerned with measuring attributes and relationships in a population, and therefore a representative sample is needed to ensure that the measurements accurately reflect and can be generalized to the population. The aim of most qualitative studies is to discover *meaning* and to uncover multiple realities, and so generalizability is not a guiding criterion.

Qualitative researchers begin with the following types of sampling question in mind: Who would be an information-rich data source for my study? Whom should I talk to, or what should I observe first, so as to maximize my understanding of the phenomenon? A critical first step in qualitative sampling is selecting settings with high potential for information richness.

As the study progresses, new sampling questions emerge, such as the following: Who can I talk to or observe that would confirm my understandings? Challenge or modify my understandings? Enrich my understandings? Thus, as with the overall design in qualitative studies, sampling design is an emergent one that capitalizes on early learning to guide subsequent direction.

Types of Qualitative Sampling

Qualitative researchers usually eschew probability samples. A random sample is not the best method of selecting people who will make good informants, that is, people who are knowledgeable, articulate, reflective, and willing to talk at length with researchers. Various nonprobability sampling designs have been used by qualitative researchers.

Convenience Sampling

Qualitative researchers sometimes use or begin with a convenience sample, which is sometimes referred to in qualitative studies as a **volunteer sample.** Volunteer samples are especially likely to be used when researchers need to have potential participants come forward and identify themselves. For example, if we wanted to study the experiences of people with frequent nightmares, we might have difficulty readily identifying a sufficient number of potential participants. In such a situation, we might recruit sample members by placing a notice on a bulletin board, in a newspaper, or on the Internet, requesting people with frequent nightmares to contact us. In this situation, we would be less interested in obtaining a representative sample of people with nightmares, than in obtaining a broad and diverse group representing various experiences with nightmares.

Sampling by convenience may be easy and efficient, but it is not in general a preferred sampling approach, even in qualitative studies. The key in qualitative studies is to extract the greatest possible information from the few cases in the sample, and a convenience sample may not provide the most information-rich sources. However, a convenience sample may be an economical and easy way to *begin* the sampling process, relying on other methods as data are collected.

Example of a convenience sample:
Young, Lynam, Valach, Novak, Brierton, and Christopher (2001) studied parent and adolescent conversations about health. Participants of Indo-Canadian and Euro-Canadian descent were recruited by posting notices in community centers, schools, health units, doctors' offices, and through

visits to community agencies. Thirty-five parent—adolescent dyads volunteered.

Snowball Sampling

Qualitative researchers also use snowball sampling, asking early informants to make referrals to other study participants. This method is sometimes referred to as **nominated sampling** because it relies on the nominations of others already in the sample. Researchers may use this method to gain access to people who are difficult to identify. Snowball sampling has distinct advantages over convenience sampling. The first is that it may be more cost-efficient and practical. Researchers may spend less time screening people to determine if they are appropriate for the study, for example. Furthermore, with an introduction from the referring person, researchers may have an easier time establishing a trusting relationship with new participants. Finally, researchers can more readily specify the characteristics that they want new participants to have. For example, in the study of people with nightmares, we could ask early respondents if they knew anyone else who had the same problem *and* who was articulate. We could also ask for referrals to people who would add other dimensions to the sample, such as people who vary in age, race, socioeconomic status, and so on.

A weakness of this approach is that the eventual sample might be restricted to a rather small network of acquaintances. Moreover, the quality of the referrals may be affected by whether the referring sample member trusted the researcher and truly wanted to cooperate.

Example of a snowball sample: Meadows, Thurston, and Berenson (2001) studied the messages that rural midlife women get about preventive health care. Study participants were recruited through convenience sampling at first, and subsequently through snowball sampling. A sample of 24 midlife women were interviewed.

Purposive Sampling

Qualitative sampling may begin with volunteer informants and may be supplemented with new participants through snowballing, but most qualitative studies eventually evolve to a purposive (or *purposeful*) sampling strategy—that is, hand-picking cases that will most benefit the study.

Example of a purposive sample: Gebbie, Wakefield, and Kerfoot (2000) purposefully selected 27 American nurses currently active in public health policy to describe their experiences in policy development.

In purposive sampling, several strategies have been identified (Patton, 2002), only some of which are mentioned here. Note that researchers themselves do not necessarily refer to their sampling plans with Patton's labels; his classification shows the kind of diverse strategies qualitative researchers have adopted to meet the theoretical needs of their research:

- **Maximum variation sampling** involves purposefully selecting cases with a wide range of variation on dimensions of interest. By selecting participants with diverse views and perspectives, researchers invite challenges to preconceived or emerging conceptualizations. Maximum variation sampling might involve ensuring that people with diverse backgrounds are represented in the sample (ensuring that there are men and women, poor and affluent people, and so on). It might also involve deliberate attempts to include people with different viewpoints about the phenomenon under study. For example, researchers might use snowballing to ask early participants for referrals to people who hold different points of view.
- **Homogeneous sampling** deliberately reduces variation and permits a more focused inquiry. Researchers may use this approach if they wish to understand a particular group of people especially well. Homogeneous sampling is often used to select people for group interviews.
- **Extreme (deviant) case sampling** provides opportunities for learning from the most unusual and extreme informants (e.g., outstanding successes and notable failures). The assumption underlying this approach is that extreme cases are rich in information because they are special

in some way. In some cases, more can be learned by intensively studying extreme cases, but extreme cases can also distort understanding of a phenomenon.

- **Intensity sampling** is similar to extreme case sampling, but with less emphasis on the extremes. Intensity samples involve information-rich cases that manifest the phenomenon of interest intensely, but not as extreme or potentially distorting manifestations. Thus, the goal in intensity sampling is to select rich cases that offer strong examples of the phenomenon.
- **Typical case sampling** involves the selection of participants who illustrate or highlight what is typical or average. The resulting information can be used to create a qualitative profile illustrating typical manifestations of the phenomenon being studied.
- **Critical case sampling** involves selecting important cases regarding the phenomenon of interest. With this approach, researchers look for the particularly good story that illuminates critical aspects of the phenomenon.
- **Criterion sampling** involves studying cases that meet a predetermined criterion of importance. Criterion sampling is sometimes used in multimethod studies in which data from the quantitative component are used to select cases meeting certain criteria for in-depth study. Sandelowski (2000) offers a number of helpful suggestions for combining sampling strategies in mixed-method research.
- **Theory-based sampling** involves the selection of people or incidents on the basis of their potential representation of important theoretical constructs. Theory-based sampling is a very focused approach that is usually based on an *a priori* theory that is being examined qualitatively.
- **Sampling confirming and disconfirming cases** is often used toward the end of data collection in qualitative studies. As researchers note trends and patterns in the data, emerging conceptualizations need to be checked. **Confirming cases** are additional cases that fit researchers' conceptualizations and offer enhanced credibility. **Disconfirming cases** are examples that do not fit and serve to challenge researchers' interpretations. These "negative" cases may offer new insights about how the original conceptualization needs to be revised or expanded.

It is important to note that almost all of these sampling strategies require that researchers have some knowledge about the setting in which the study is taking place. For example, to choose extreme cases, typical cases, or critical cases, researchers must have information about the range of variation of the phenomenon and how it manifests itself. Early participants may be helpful in implementing these sampling strategies.

Theoretical Sampling

The method of sampling used in grounded theory is called *theoretical sampling*. Glaser (1978, p. 36) defined this sampling as "the process of data collection for generating theory whereby the analyst jointly collects, codes, and analyzes his data and decides what data to collect next and where to find them, in order to develop his theory as it emerges." The process of theoretical sampling is controlled by the developing grounded theory. Theoretical sampling is not envisioned as a single, unidirectional line. This complex sampling technique requires researchers to be involved with multiple lines and directions as they go back and forth between data and categories as the theory emerges.

Glaser stressed that theoretical sampling is not the same as purposive sampling. Theoretical sampling's purpose is to discover categories and their properties and to offer interrelationships that occur in the substantive theory. "The basic question in theoretical sampling is: what groups or subgroups does one turn to next in data collection?" (Glaser, 1978, p. 36). These groups are not chosen before the research begins but only as they are needed for their theoretical relevance for developing further emerging categories.

Example of a theoretical sampling: Beck (2002) used theoretical sampling in her grounded theory study of mothering twins during the first year of life, in which 16 mothers of twins were interviewed in their homes. A specific example of theoretical sampling concerned what the mothers kept referring to as the "blur period"—the first few months of caring for the twins. Initially, Beck interviewed mothers whose twins were around 1 year of age. Her rationale was that these mothers would be able to reflect back over the entire first year of mothering the multiples. When these mothers referred to the "blur period," Beck asked them to describe this period more fully. The mothers said they could not provide many details about this period because it was "such a blur!" Beck then chose to interview mothers whose twins were 3 months of age or younger, to ensure that mothers were still immersed in the "blur period" and would be able to provide rich detail about what this phase of mothering twins was like.

Sample Size in Qualitative Research

There are no criteria or rules for sample size in qualitative research. Sample size is largely a function of the purpose of the inquiry, the quality of the informants, and the type of sampling strategy used. For example, a larger sample is likely to be needed with maximum variation sampling than with typical case sampling.

In qualitative studies, sample size should be determined based on informational needs. Hence, a guiding principle in sampling is **data saturation**—that is, sampling to the point at which no new information is obtained and redundancy is achieved. Morse (2000) has noted that the number of participants needed to reach saturation depends on a number of factors. For example, the broader the scope of the research question, the more participants will likely be needed. Data quality can also affect sample size. If participants are good informants who are able to reflect on their experiences and communicate effectively, saturation can be achieved with a relatively small sample. Also, if longitudinal data are collected, fewer participants may be needed, because each will provide a greater amount of information.

TIP: Sample size ambiguities sometimes create practical dilemmas when you are planning a study, or if you are seeking approval or funding for a project. Patton (2002) recommends specifying *minimum* samples that would reasonably be adequate for understanding the phenomenon. Additional cases can then be added, as necessary, to achieve saturation.

Evaluating Qualitative Samples

In a qualitative study, the sampling plan is evaluated in terms of *adequacy* and *appropriateness* (Morse, 1991). Adequacy refers to the sufficiency and quality of the data the sample yields. An adequate sample provides researchers with data without any "thin" spots. When researchers have truly attained saturation, informational adequacy has been achieved and the resulting description or theory is richly textured and complete. Appropriateness concerns the methods used to select the sample. An appropriate sample is one resulting from the identification and use of participants who can best supply information according to the conceptual requirements of the study. For example, a sampling plan that does not include disconfirming cases may not meet the information needs of the research.

TIP: No matter what type of qualitative sampling you use, you should keep a journal or notebook to jot down ideas and reminders regarding the sampling process (e.g., who you should interview next). Memos to yourself will help you remember valuable ideas about your sample.

Sampling in the Three Main Qualitative Traditions

There are similarities among the various qualitative traditions with regard to sampling: samples are usually small, probability sampling is almost never used, and final sampling decisions usually take place in the field during data collection. However, there are some differences as well.

Sampling in Ethnography

Ethnographers may begin by initially adopting a "big net" approach—that is, mingling with and having conversations with as many members of the culture under study as possible. Although they may converse with many people (usually 25 to 50), they often rely heavily on a smaller number of **key informants**, who are highly knowledgeable about the culture and who develop special, ongoing relationships with the researcher. These key informants are often the researcher's main link to the "inside."

Key informants are chosen purposively, guided by the ethnographer's theoretically informed judgments. Developing a pool of potential key informants often depends on ethnographers' prior theoretical knowledge to construct a relevant framework. For example, an ethnographer might make decisions about different types of key informants to seek out based on roles (e.g., physicians, nurse practitioners) or on some other theoretically meaningful distinction. Once a pool of potential key informants is developed, key considerations for final selection are their level of knowledge about the culture, and how willing they are to collaborate with the ethnographer in revealing and interpreting the culture.

It might be noted that there is some controversy among ethnographers about the use of "stranger" versus "insider" samples. It has been argued that ethnographers should not sample people whom they know or in whom they have a vested interest. According to this argument, it is not possible to do a valid ethnography "in your own backyard" (Glesne & Peshkin, 1992), despite the obvious advantage of having access to a lot of information and being able to gain people's cooperation. The problem is that if ethnographers are part of the culture under study, it may be difficult to get a handle on ingrained norms and values. Moreover, ethnographers who study people known to them have established relationships that can interfere with objective questioning and observation.

Although ethnographies have traditionally entailed studies of cultures in which researchers are strangers, not everyone agrees that this is essential. Field (1991), for example, has described the unique issues involved in nurse researchers studying their own culture and settings.

Sampling in ethnography typically involves more than selecting informants because observation and other means of data collection play a big role in helping researchers understand a culture. Ethnographers have to decide not only *whom* to sample, but *what* to sample as well. For example, ethnographers have to make decisions about observing *events* and *activities*, about examining *records* and *artifacts*, and about exploring *places* that provide clues about the culture. Key informants can play an important role in helping ethnographers decide what to sample.

Example of an ethnographic sample: Hoga, Alcantara, and deLima (2001) explored the involvement of men in reproductive health in a low-income community in Brazil. These ethnographers used Leininger's ethnonursing research method to collect data. Their sample consisted of 15 adult men, 7 of whom were key informants. "The key informants were selected based on their full knowledge about the domain of inquiry and the observations during the observation-participation-reflection process that they dictate their norms, values, and beliefs during social and mainly in their familiar examples and conversations with children and relatives" (p. 110).

Sampling in Phenomenological Studies

Phenomenologists tend to rely on very small samples of participants—typically 10 or fewer. There is one guiding principle in selecting the sample for a phenomenological study: all participants must have experienced the phenomenon under study and must be able to articulate what it is like to have lived that experience. It might thus be said that phenomenologists use a criterion sampling method, the criterion being experience with the phenomenon under study. Although phenomenological researchers seek participants who have had the targeted experiences, they also want to explore diversity of individual experiences. Thus, as described by Porter (1999), they may specifically look for people with demographic or other differences who have shared a common experience.

Example of a sample in a phenomenological study:

Orne, Fishman, Manka, and Pagnozzi (2000) studied the lived experience of being a medically uninsured working person. They purposively sampled 12 people who were working but lacked health insurance. The participants varied in terms of gender, occupation, and income.

Sampling in Grounded Theory Studies

Grounded theory research is typically done with samples of about 20 to 30 people, using theoretical sampling. The goal in a grounded theory study is to select informants who can best contribute to the evolving theory. Sampling, data collection, data analysis, and theory construction occur concurrently, and so study participants are selected serially and contingently (i.e., contingent on the emerging conceptualization). Sampling might evolve as follows:

1. The researcher begins with a general notion of where and with whom to start. The first few cases may be solicited purposively, by convenience, or through snowballing.
2. In the early part of the study, a strategy such as maximum variation sampling might be used, to gain insights into the range and complexity of the phenomenon under study.
3. The sample is adjusted in an ongoing fashion. Emerging conceptualizations help to focus the sampling process.
4. Sampling continues until saturation is achieved.
5. Final sampling often includes a search for confirming and disconfirming cases to test, refine, and strengthen the theory.

Example of sampling in a grounded theory study:

Patterson and Thorne (2000) studied how people with long-standing type 1 diabetes make decisions in relation to unanticipated blood glucose levels. Initial participants were selected to ensure variation with regard to attributes known to influence self-care decision-making (e.g., cohabitation). Subsequently, recruitment was directed on the basis of concepts that had relevance to the evolving theory. A total of 22 diabetic patients participated in the study.

RESEARCH EXAMPLES

In the following sections, we describe in some detail the sampling plans of two nursing studies, one quantitative and the other qualitative.

Research Example From a Quantitative Study

Holland and Carruth (2001) conducted a telephone survey to examine the risk factors of farm women who engage in activities that could exposed them to tetanus, and to study the circumstances related to tetanus immunization.

The researchers first used a purposive sampling method to select 10 parishes (counties) in southeast Louisiana. The counties were hand-picked to reflect agricultural and geographic diversity. The researchers had access to a sampling frame of 4804 farm owners in these 10 parishes (a list maintained by Louisiana State University Agricultural Centers and Farm Service Agency).

Next, a stratified random sample of farm owners was drawn, with parish as the stratifying variable. That is, in each parish, a random sample of farm owners was selected. (The research report did not indicate whether proportionate or disproportionate sampling was used.) Sampled farm owners were screened to determine whether an eligible woman lived in the household. Women were deemed eligible if they were 18 years or older and were members of a family participating in a farming operation. If the household included two or more such women, the woman who had the greatest involvement with farming was invited to participate.

A total of 1141 farms were determined to have an eligible sample member. Interviews were completed with 657 women, for a response rate of 57.6% among known eligible farms. The report did not indicate whether an analysis was done to evaluate response bias—although the absence of such information does not mean that such an analysis was not undertaken.

The results indicated that only 54% of the women had had a tetanus booster within the prior 10 years.

Older women were much less likely than younger women to be up to date on their immunizations.

Research Example From a Qualitative Study

Rillstone and Hutchinson (2001) conducted a grounded theory study to examine parents' experiences and feelings when faced with a potentially stressful pregnancy: a pregnancy subsequent to a prior pregnancy which the parents decided to terminate due to a fetal abnormality.

The initial sample was obtained from an urban community in northeastern Florida, where parents were recruited through obstetricians and reproductive endocrinologists. Because only 4 local parents were recruited, Rillstone and Hutchinson sought additional participants through an Internet support network, which yielded 18 additional parents. Sampling continued for an 8-month period. The total sample consisted of 13 women and 9 of their partners from across the nation, plus 2 local health care providers. After 20 interviews, the researchers felt they had achieved data saturation. To assess the validity of their developing grounded theory, two additional women and two health care providers (a nurse and a physician) were interviewed, either by telephone or in person.

The parents who participated in the study had confronted a wide range of diagnoses with the pregnancies where they had to choose whether to terminate the pregnancy. These varied diagnoses included Down syndrome, spina bifida, trisomy 18, bilateral renal agenesis, Prader-Willi syndrome, and autosomal recessive polycystic kidney disease.

The researchers concluded that their sample was not representative of all parents confronting such circumstances. Parents in the sample were in higher-than-average socioeconomic circumstances, were older than 24 years of age, and had high educational levels. Such a bias is consistent with the fact that Internet use is higher among more affluent and better-educated families.

Data analysis revealed that the basic problem these parents had to contend with was the reemergence of mental anguish. Parents coped with this mental anguish by developing emotional armor, limiting disclosure to others about both their past and present pregnancies, delaying attachment to the baby, and becoming increasingly attached to their health care providers.

SUMMARY POINTS

- **Sampling** is the process of selecting a portion of the **population,** which is an entire aggregate of cases.
- An **element** is the basic unit about which information is collected—usually humans in nursing research.
- The criteria that specify population characteristics are the **eligibility criteria** (or **inclusion criteria**).
- Researchers usually sample from an **accessible population**, but should identify the **target population** to which they would like to generalize their results.
- The main consideration in assessing a sample in a quantitative study is its **representativeness**—the extent to which the sample is similar to the population and avoids bias. **Sampling bias** refers to the systematic over-representation or under-representation of some segment of the population.
- The principal types of **nonprobability sampling** (wherein elements are selected by non-random methods) are convenience, quota, and purposive sampling. Nonprobability sampling designs are convenient and economical; a major disadvantage is their potential for bias.
- **Convenience sampling** (or **accidental sampling**) uses the most readily available or most convenient group of people for the sample. **Snowball sampling** is a type of convenience sampling in which referrals for potential participants are made by those already in the sample.
- **Quota sampling** divides the population into homogeneous **strata** (subpopulations) to ensure representation of the subgroups in the sample; within each stratum, subjects are sampled by convenience.
- In **purposive** (or **judgmental**) **sampling,** participants are hand-picked to be included in the sample based on the researcher's knowledge about the population.
- **Probability sampling** designs, which involve the random selection of elements from the population, yield more representative samples than

nonprobability designs and permit estimates of the magnitude of **sampling error**.

- **Simple random sampling** involves the random selection of elements from a **sampling frame** that enumerates all the elements; **stratified random sampling** divides the population into homogeneous subgroups from which elements are selected at random.
- **Cluster sampling** (or **multistage sampling**) involves the successive selection of random samples from larger to smaller units by either simple random or stratified random methods.
- **Systematic sampling** is the selection of every *k*th case from a list. By dividing the population size by the desired sample size, the researcher establishes the **sampling interval,** which is the standard distance between the selected elements.
- In quantitative studies, researchers should ideally use a **power analysis** to estimate **sample size** needs. Large samples are preferable to small ones because larger samples tend to be more representative, but even a large sample does not *guarantee* representativeness.
- Qualitative researchers use the theoretical demands of the study to select articulate and reflective informants with certain types of experience in an emergent way, capitalizing on early learning to guide subsequent sampling decisions.
- Qualitative researchers most often use purposive or, in grounded theory studies, **theoretical sampling** to guide them in selecting data sources that maximize information richness.
- Various purposive sampling strategies have been used by qualitative researchers. One strategy is **maximum variation sampling,** which entails purposely selecting cases with a wide range of variation.
- **Criterion sampling** involves studying cases that meet a predetermined criterion of importance.
- Another important strategy is **sampling confirming and disconfirming cases**, that is, selecting cases that enrich and challenge the researchers' conceptualizations.
- Other types of qualitative sampling include **homogeneous sampling** (deliberately reducing variation); **extreme case sampling** (selecting the most unusual or extreme cases); **intensity sampling** (selecting cases that are intense but not extreme); **typical case sampling** (selecting cases that illustrate what is typical); **critical case sampling** (selecting cases that are especially important or illustrative); and **theory-based sampling** (selecting cases on the basis of their representation of important constructs).
- Samples in qualitative studies are typically small and based on information needs. A guiding principle is **data saturation**, which involves sampling to the point at which no new information is obtained and redundancy is achieved.
- Criteria for evaluating qualitative sampling are informational adequacy and appropriateness.
- Ethnographers make numerous sampling decisions, including not only *whom* to sample but *what* to sample (e.g., activities, events, documents, artifacts); decision making is often aided by their **key informants** who serve as guides and interpreters of the culture.
- Phenomenologists typically work with a small sample of people (10 or fewer) who meet the criterion of having lived the experience under study.
- Grounded theory researchers typically use theoretical sampling and work with samples of about 20 to 30 people.

STUDY ACTIVITIES

Chapter 13 of the *Study Guide to Accompany Nursing Research*: *Principles and Methods*, *7th edition*, offers various exercises and study suggestions for reinforcing concepts presented in this chapter. In addition, the following study questions can be addressed:

1. Draw a simple random sample of 15 people from the sampling frame of Table 13-3, using the table of random numbers that appears in Table 8-2 on page 171. Begin your selection by blindly placing your finger at some point on the table.

2. Suppose you have decided to use a systematic sampling design for a research project. The known population size is 5000, and the sample size desired is 250. What is the sampling interval? If the first element selected is 23, what would be the second, third, and fourth elements selected?
3. Suppose you were interested in studying the attitude of clinical specialists toward autonomy in work situations. Suggest a possible target and accessible population. What strata might be identified if quota sampling were used?
4. What type of sampling design was used to obtain the following samples?
   a. 25 experts in critical care nursing
   b. 60 couples attending a particular prenatal class
   c. 100 nurses from a list of nurses registered in the state of Pennsylvania, using a table of random numbers
   d. 20 adult patients randomly selected from a random selection of 10 hospitals located in one state
   e. Every fifth article published in *Nursing Research* during the 1980s, beginning with the first article.
5. Suppose you wanted to study the experiences of nursing students during their first clinical assignment. Describe what you would need to do to select a sample using maximum variation sampling, critical case sampling, typical case sampling, and homogeneous sampling.

SUGGESTED READINGS

Methodologic References

Babbie, E. (1990). *Survey research methods* (2nd ed.). Belmont, CA: Wadsworth.

Brown, J. S., Tanner, C. A., & Padrick, K. P. (1984). Nursing's search for scientific knowledge. *Nursing Research, 33*, 26–32.

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). New York: John Wiley & Sons.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (rev. ed.). New York: Academic Press.

Diekmann, J. M., & Smith, J. M. (1989). Strategies for accessment and recruitment of subjects for nursing research. *Western Journal of Nursing Research, 11*, 418–430.

Field, P. A. (1991). Doing fieldwork in your own culture. In J. M. Morse (Ed.), *Qualitative nursing research: A contemporary dialogue*. Newbury Park, CA: Sage.

Glaser, B. G. (1978). *Theoretical sensitivity*. Mill Valley, CA: Sociology Press.

Glesne, C., & Peshkin, A. (1992). *Becoming qualitative researchers: An introduction*. White Plains, NY: Longman.

Kish, L. (1965). *Survey sampling*. New York: John Wiley & Sons.

Levey, P. S., & Lemeshow, S. (1980). *Sampling for health professionals*. New York: Lifetime Learning.

MacDougall, C., & Fudge, E. (2001). Planning and recruiting the sample for focus groups and in-depth interviews. *Qualitative Health Research, 11*, 117–126.

Moody, L. E., Wilson, M. E., Smyth, K., Schwartz, R., Tittle, M., & VanCott, M. L. (1988). Analysis of a decade of nursing practice research: 1977–1986. *Nursing Research, 37*, 374–379.

Morse, J. M. (1991). Strategies for sampling. In J. M. Morse (Ed.), *Qualitative nursing research: A contemporary dialogue*. Newbury Park, CA: Sage.

Morse, J. M. (2000). Determining sample size. *Qualitative Health Research, 10*, 3–5.

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

Polit, D. F., & Sherman, R. (1990). Statistical power analysis in nursing research. *Nursing Research, 39*, 365–369.

Porter, E. J. (1999). Defining the eligible, accessible population for a phenomenological study. *Western Journal of Nursing Research, 21*, 796–804.

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

Sandelowski, M. (2000). Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in Nursing & Health, 23*, 246–255.

Sudman, S. (1976). *Applied sampling*. New York: Academic Press.

Trinkoff, A. M., & Storr, C. L. (1997). Incorporating auxiliary variables into probability sampling designs. *Nursing Research, 46*, 182–185.

Williams, B. (1978). *A sampler on sampling*. New York: John Wiley & Sons.

Studies Cited in Chapter 13

Bath, S., Singleton, J., Strikas, R., Stevenson, J., McDonald, L., & Williams, W. (2000). Performance of U.S. hospitals on recommended screening and immunization practices for pregnant and postpartum women. *American Journal of Infection Control, 28*, 327–332.

Beck, C. T. (2002). Releasing the pause button: Mothering twins during the first year of life. *Qualitative Health Research 12*, 593–608.

Board, R., & Ryan-Wenger, N. (2002). Long-term effects of pediatric intensive care unit hospitalization on families with young children. *Heart & Lung, 31*, 53–66.

Friedemann, M., Montgomery, R. J., Rice, C., & Farrell, L. (1999). Family involvement in the nursing home. *Western Journal of Nursing Research, 21*, 549–567.

Gebbie, K. M., Wakefield, M., & Kerfoot, K. (2000). Nursing and health policy. *Journal of Nursing Scholarship, 32*, 307–315.

Hoga, L. A., Alcantara, A. C., & deLima, V. M. (2001). Adult male involvement in reproductive health: An ethnographic study in a community of Sao Paulo City, Brazil. *Journal of Transcultural Nursing, 12*, 107–114.

Holland, C., & Carruth, A. K. (2001). Exposure risks and tetanus immunization in women of family owned farms. *AAOHN Journal, 49*, 130–136.

Keele-Smith, R., & Price-Daniel, C. (2001). Effects of crossing legs on blood pressure measurement. *Clinical Nursing Research, 10*, 202–213.

Meadows, L. M., Thurston, W. E., & Berenson, C. A. (2001). Health promotion and preventive measures: Interpreting messages at midlife. *Qualitative Health Research, 11*, 450–463.

Orne, R. M., Fishman, S. J., Manka, M., & Pagnozzi, M. E. (2000). Living on the edge: A phenomenological study of medically uninsured working Americans. *Research in Nursing & Health, 23*, 204–212.

Patterson, B., & Thorne, S. (2000). Expert decision-making in relation to unanticipated blood glucose levels. *Research in Nursing & Health, 23*, 147–157.

Rillstone, P., & Hutchinson, S. A. (2001). Managing the reemergence of anguish: Pregnancy after a loss due to anomalies. *Journal of Obstetric, Gynecologic, and Neonatal Nursing, 30*, 291–298.

Tolle, S. W., Tilden, V. P., Rosenfeld, A. G., & Hickman, S. E. (2000). Family reports of barriers to optimal care of the dying. *Nursing Research, 49*, 310–317.

Trinkoff, A. M., Zhou, Q., Storr, C. L., & Soeken, K. L. (2000). Workplace access, negative proscriptions, job strain, and substance use in registered nurses. *Nursing Research, 49*, 83–90.

Williams, P. D., Soetjiningsih, & Williams, A. R. (2000). Balinese mothers' developmental timetables for young children. *Western Journal of Nursing Research, 22*, 717–735.

Yoon, S. L., & Horne, C. H. (2001). Herbal products and conventional medicines used by community-residing older women. *Journal of Advanced Nursing, 33*, 51–59.

Young, R. A., Lynam, M. J., Valach, L., Novak, H., Brierton, I., & Christopher, A. (2001). Joint actions of parents and adolescents in health conversations. *Qualitative Health Research, 11*, 40–57.