5

The Method Chapter

Describing Your Research Plan

The Method chapter of a dissertation, article, or proposal describes the exact steps that will be undertaken to address your hypotheses or research questions. For this reason, the Method section follows logically from the statement of the problem in much the same way as research questions follow from the Review of the Literature. The goal of this chapter is to provide a clear and complete description of the specific steps to be followed. It is necessary to describe these steps in sufficient detail to help a naïve reader to replicate your study.

In Chapter 3, we suggested that students select a suitable problem before selecting the appropriate method with which to study that problem. In some dissertations, however, the method may in fact be the topic of study. A student may become excited about the possibilities of a particular data collection technique or method prior to generating a suitable problem. In most cases, though, the appropriate method of study is generated by careful consideration of the research questions and the applicable method by which those questions may be studied. The following material is divided into two major sections, one that focuses primarily on the quantitative dissertation and one that focuses on the qualitative dissertation; we strongly recommend that you read both sections. The issues to be considered are not independent of each other, and there is considerable overlap in the content of this chapter, regardless of the method being employed.

The Method Chapter in a Quantitative Dissertation

The Method chapter is the place in which the exact steps you will be following to test your questions are enumerated. The Method chapter typically contains the following three subsections: Subjects or Participants, Instrumentation or Measures, and Procedures. In addition, the Method chapter of a dissertation proposal often contains a Statistical Analysis or Data Analysis section, in which procedures for approaching the data are outlined. Research that uses special equipment frequently contains an Apparatus section, in which the nature and type of equipment are described.

A frequent mistake contained in first drafts of Method chapters involves overuse of the opportunity to rehash the material contained in the Review of the Literature or Statement of the Problem, which were presented in another chapter. The Method chapter should be viewed primarily as a set of directions for conducting a specific piece of research. We follow this perspective in describing each of the subsections of a Method chapter.

How to Begin a Method Chapter

A reasonable way to begin a Method chapter is to develop an introductory paragraph that describes both the design of the study and the organization of the chapter. This prepares the reader for what is to follow and provides a framework within which to incorporate the materials. This paragraph says to the reader, "This is the Method chapter, this is how it is organized, and this is the type of design I used." The most difficult component of the introductory paragraph may appear to be the design statement. There are dozens of books that take research design as their sole topic. Our purpose here is not to review this extensive material. In this context, we are suggesting only a short sentence or two that informs the reader of the general type of design and any suitable label that may be applied to it.

For example, imagine a study in which you plan to mail 700 surveys to retired male executives in an attempt to assess their adjustment to retirement. It is probably sufficient simply to state something like the following: "This study used a cross-sectional survey design to assess adjustment to retirement of a sample of retired male executives." Similarly, imagine an experimental design in which two dichotomous variables will be used as independent variables. It may be sufficient to describe the design as "an experimental study using a 2 × 2 factorial

design." An example of such a design is described in a dissertation by Sangster (1991):

> An experimental, 2 × 2 factorial, pretest-posttest design was used to test the hypothesis. The independent variables were (a) training type and (b) management experience. Respective levels of independent variables were descriptive and prescriptive training, and high and low management experience. (p. 102)

Determining exactly which type of design description best fits your study can be accomplished by referencing one of the many texts that describe designs (e.g., Campbell & Stanley, 2005; Kerlinger & Lee, 1999), but a specific label may not be as important as a clear explanation of the structure of your study. For example, a dissertation by Macdonald (1990) examined the relationships among empathy, personal maturity, and emotional articulation. Macdonald described her design as follows:

> The research design was a correlational design utilizing cross-sectional survey methodology and includes a number of survey instruments. The purpose of the design was to correlate the scores of the personality tests listed below with the scores on responses to the short stories, as well as measure the interrelationship of the responses to the short stories. (p. 115)

Describing Your Sample

The Subjects (or Participants) section of the Method chapter describes the source and number of subjects you will be using (in a proposal) or the source and number of subjects actually obtained (in a completed project). Note that we use the term *subjects* in a very broad context. The subjects may be informants or participants in the research, organizations or events, documents or segments of commercials edited from television programs, or even an entire society, as in the case of anthropological research. In studies of human beings, the term *participants* is generally preferable to the term *subjects*. One goal of this section of the Method chapter is to describe why and how the particular unit of analysis was selected.

A frequent problem we have observed is that students often confound a description of the sampling of subjects with the procedures used to collect data from these subjects. The former belongs in the Subjects section, whereas the latter belongs in the Procedures section. Of particular importance in the Subjects section are the specific sampling procedures

used, the rationale for selecting a given number of subjects, and the source of the subjects. Which type of sample will you draw: random, stratified, purposive? Where will your subjects be located? How many subjects are necessary for your particular design? Each of these questions is important and should be dealt with in the Subjects section. We discuss each in turn in this section.

The Sampling Design. As is the case with research design in general, the issue of sampling is a complex one. Our purpose, however, and your purpose as researcher, is not to discuss the knotty problems of sampling. Your job is to describe how you will accomplish this task in your study, taking all the theoretical and practical issues into account. Knowledge of the various types of sampling design is a prerequisite to developing an appropriate description of your particular sampling problem.

Two examples may be of assistance in developing your approach. In a dissertation examining gender differences in professional identity among practicing psychotherapists, New (1989) obtained a mailing list of all licensed psychologists in the state of California. Beginning at a randomly selected start point, she then sampled every *n*th name until she had exhausted the list. Her dissertation described this process as follows:

> A systematic random sample was drawn from a list of all psychologists who have been licensed in the state of California for five or more years. The sample was obtained from a research center specializing in the distribution of selected sampling lists. A table of random numbers was used to locate the initial sampling point. Every 20th person was selected such that an initial group of 400 was obtained. These persons were mailed the study materials.

In a second instance, illustrating a nonrandom procedure, Caddell (1989) selected students from classes at three universities to examine moral education in the college environment. He described his sampling procedure as follows:

> The questionnaire was administered to selected undergraduate and graduate level classes at California State University–Fullerton, California Baptist College and Pacific Christian College. The sample was divided into a subset from the morally nondirective institution (Cal State–Fullerton, $N = 178$) and a subset from morally directive institutions (Cal Baptist College, $N = 104$ and Pacific Christian College, $N = 71$).

Locating Participants. Of the two examples presented in the previous section, only one describes both where and how the participants were

obtained as well as the sampling design used. In the study using university students, Caddell (1989) failed to describe how the "selected undergraduate and graduate level classes" were obtained. In fact, the procedure was one of asking professors at each institution for permission. This decidedly nonrandom procedure could probably best be described as a convenience sample.

In many studies, it may not be practical to obtain a truly random sample. Even college sophomores, who constitute the classic population of many university-based studies, are a select group with their own idiosyncracies that exert limitations on the ability to generalize findings to people in general. Students who choose to study relatively unusual phenomena, such as men who continue to live with their parents into and beyond midlife or adults who grew up in foster homes, may end up relying on a snowball technique to obtain a sufficient number of participants. This technique consists of other individuals, including previously identified participants, knowing of and recommending additional potential participants. This approach to sampling may or may not be suitable, depending on the tolerance of the study to absorb the potential bias inherent in the method. One relevant example is Carol Crane's (2005) dissertation, which required her to locate an adequate number of participants with developmental synesthesia, a rare neuropsychological condition. A snowball technique for identifying participants is more common in qualitative studies. Take, for example, Diane Armstrong's (1994) fascinating dissertation on the nature of dreams of 18 congenitally and adventitiously blind men and women or Devon Jersild's (2006) exploration of women who grew up with particularly successful and accomplished mothers.

An increasingly popular way of identifying participants and collecting data is through the Internet. Our students have used two different approaches. One is to design or hire an individual or company that designs Web sites for research data collection. All of the questionnaires and measures for the study are available on the Web site and presented in a clear, systematic way that enables eligible participants to complete them on their home computers at a time of their choosing. Each participant's responses are downloaded into a database so that the results can easily be manipulated and analyzed statistically. Sometimes the participants are recruited on the same Web site or via Internet discussion lists that appeal to individuals who have the required characteristics to qualify for the study. Tracy Zemansky (2005) helped design her own Web site to complete a recent dissertation on the recovery process of long-term members of Alcoholics Anonymous. She obtained 164 qualified participants in a

short period of time. A slightly different approach is to use an existing service that allows researchers to post their measures on a secure domain within a public Web site that is available exclusively for such survey and data-gathering activities. Potential participants are recruited by the researcher through e-mails, electronic discussion lists, letters, or personal contact and directed to the Web site, invited to create a secure online identification code and password, and asked to type in the appropriate survey number to access the study. The site also provides the necessary informed consent forms. Corinne Goodwin (2006) recently used this kind of online research service (www.psychdata.com) to complete her dissertation on the quality of the supervisory relationship in the training of clinical psychologists. She too was gratified to obtain more participants than she anticipated within a very short period of time using this strategy.

What are the advantages and disadvantages of selecting participants based on their willingness to access an Internet Web site and complete surveys electronically? We do not yet know enough about the comparability of these samples, but what we do know is encouraging (Wright, 2005). Though there may be bias in terms of computer access and computer savvy, there is also the possibility of obtaining geographically heterogeneous samples that may not be available when using traditional data collection strategies. Also, many participants seem to trust and appreciate the anonymity they have by addressing a machine rather than the researcher directly. Of course, the accompanying risk is that when might be more likely for participants to fudge their responses using this more impersonal format. It does seem clear that researchers can obtain a large number of responses very efficiently over the Internet and collect data in a form that allows for relatively painless analysis. The previous discussion does not exhaust the varieties of Internet-based data collection. This is so critical for the dissertation process that we have devoted all of Chapter 11 to a thorough discussion of the many sources of data and strategies for data collection through the Internet.

The Appropriate Number of Participants. Determining the appropriate number of participants for a given design is one of the most difficult sampling problems. Generally, given considerations of cost and time, students wish to obtain as few as their committee will allow; however, this decision is not arbitrary. Most students tend to underestimate the number of participants necessary to draw meaningful conclusions from the data. For example, imagine a study from the field of education in which the researcher wishes to compare the level of creativity of students who are sent to public schools to those who receive home schooling. If in fact there

is a difference, the student-researcher must collect enough data for this difference to appear in the results. The smaller the difference, the more data the student-researcher must collect.

The best method to approximate the number of participants is to conduct a power analysis. A power analysis advises the researcher regarding how many subjects are necessary to detect any effects that result from the independent variables, given (a) the size of the effect of these variables in the population, (b) the type of statistical tests to be used, and (c) the level of significance (or alpha level) of the study. The level of power, expressed as a probability, lets the researcher know how likely he or she is to avoid a Type II error. A Type II error occurs when one fails to reject the null hypothesis, even though it is false. Failing to reject a false null hypothesis means that an effect existed but was not detected by the study. As the probability of a Type II error increases, the power of the study decreases. In fact, power is equal to 1 minus the probability of a Type II error. Thus, if a probability of a Type II error is .15, power is .85 (1 – .15 = .85). Said less technically, an underpowered study is likely to obtain nonsignificant findings.

Historically, power calculations have been difficult to conduct, and students and their committees frequently relied on general rules of thumb to determine the appropriate number of subjects. Computer programs now simplify these calculations, and we strongly recommend that they be used in dissertation planning. Most power analysis software exists as programs that are either individually purchased or purchased as add-ons to existing software. An example of an individually purchased program is nQuery Advisor 6.0 (http://www.statsolusa.com). nQuery Advisor is an easy-to-use program that helps researchers determine the effect sizes and sample sizes necessary to conduct a sufficiently powered study for a wide variety of statistical procedures and designs. An example of a program that is an add-on to an existing program is the SPSS module, Sample Power (www.spss.com). One problem with these programs is that they may be quite expensive; however, cooperative arrangements with the academic community often permit deep discounts for students. Check with your college or university software vendor to determine the exact pricing of this software. An alternative is to calculate power by hand. Though the calculations can be confusing for the novice, with a little help a student can often successfully estimate the appropriate sample size using formulas and tables. Three sources that we recommend are Cohen's (1988) *Statistical Power Analysis for the Behavioral Sciences;* Kraemer's (1987) *How Many Subjects? Statistical Power Analysis in Research;* and Wilcox's (2001) *Modern Statistical Methods: Substantially Improving Power and Accuracy*.

Finally, for more detail on the concepts of power and effect size, we recommend Chapter 4 of our own text *Your Statistical Consultant: Answers to Your Data Analysis Questions* (Newton & Rudestam, 1999). This chapter contains an extended discussion of the issues raised here and alternative formulations of the logic of statistical hypothesis testing.

In our earlier example of the student interested in comparing homeschooled with traditionally schooled youngsters, a *t* test might be used to compare levels of creativity between these two groups. Table 5.1 gives an example of a power analysis designed to answer the question "How many participants will I need to test the difference between the means of two groups, if I use a level of significance of .05 and desire a power of .80?" The table shows the number of subjects necessary to obtain the specified power level of .80 (a standard level, generally considered acceptable) using an alpha level of .05 (also a generally acceptable criterion) if the size of the effects (i.e., the mean differences relative to the standard deviation) are small, medium, or large.

Table 5.1 shows that if medium effects were predicted, it would require 64 participants per group to achieve a power of .80 when using an alpha level of .05. Even if the effects are large, the researcher would need at least 26 participants per group to achieve the same power level (.80). Another way to say this is that even if the research hypothesis is true, the researcher may fail to support this hypothesis because of insufficient sample size. As the sample size increases, it becomes more likely that the statistical tests will detect any effects that exist in the data.

Finally, it is important to make three critical points about the use of power analysis. First, in many dissertations the use of power analysis may be unrealistic. Enough subjects to meet the requirements of a purely mathematical procedure may not exist in some cases, and qualitative dissertations, case studies, oral histories, and intensive interviews may rely

Table 5.1 Power Analysis for a *t* Test

Effect Size	*N Per Group*	*Total N*
Small	393	786
Medium	64	128
Large	26	52

NOTE: To achieve a power level of .80, while setting the level of significance (alpha) at .05.

more on what the student and committee deem reasonable to develop a convincing argument, independent of statistical testing. Second, in cases where quantitative analysis is essential, students often use a multivariate framework with many variables. The number of subjects necessary in these contexts is likely to be much higher than that suggested earlier for a simple two-group *t* test. Third, even in highly quantitative studies, the use of statistical inference may not be appropriate, and thus power analysis would also be meaningless. In these cases, it is unreasonable to critique the study for failure to use significance testing or for lack of power. Not all quantitative studies rest solely on finding significance.

Instrumentation: Describing Your Research Tools

The Instrumentation (or Measures) section of a Method chapter describes the particular measures you will employ and how they will measure the variables specified in your research questions and hypotheses. This section makes your case for the use of particular measures as the best and most appropriate to your specific research environment.

If you are using instruments that have been used previously, especially standardized and widely recognized scales, then you should consider the following information as particularly relevant to your description: information about the appropriateness of the use of the instrument with the population and setting described in the proposal, information about the measurement characteristics of the instrument, and information about the administration and scoring of the scales. Each of these three areas is discussed separately.

Is the Measure Appropriate? A number of scales may exist that measure the same phenomenon. How is the researcher to select the "best"? (Or how does one guard against the claim that he or she should have used something else?)

The first step is to support the use of a given instrument with the population selected for your study. Measures that work well with adults may not be particularly effective with adolescents or children. Measures designed for use in one culture may work poorly in another. The writer's goal is to locate empirical studies that demonstrate the use of an instrument with a population as closely representative as possible of the population to be used in his or her study.

The second step is to demonstrate that the authors of the measure you select conceive of the phenomenon in terms similar to the manner in which you have conceptualized the same phenomenon. For example, all

measures that assess depression are not the same because those who originally designed the measures viewed depression from different theoretical positions. The contents of the measures reflect these different positions. Thus, it may be important to indicate that you have chosen a particular instrument because it reflects the conceptualization of the phenomenon in a manner that is consistent with your perspective.

What Are the Measurement Characteristics of the Instrument? By *measurement characteristics,* we mean the reliability, validity, and structure of the measure. Reliability refers to the ability of a measure to produce consistent results. Validity indicates that a measure in fact measures what it purports to measure. Structure refers to the number and meaning of subscales contained in a given instrument.

It is important to realize that in a proposal the only information likely to be available regarding an instrument's reliability and validity is that contained in the existing literature. After the student collects his or her data, it is important to add to this body of literature by reporting the reliability and validity of the instrument as evidenced in the new sample. The reliability of an instrument is a characteristic of the population in which that instrument is used. Thus, an instrument that achieves high reliability in one sample will not necessarily receive that same level of reliability in another sample representing a different population.

How Does One Administer and Score the Measures? It is important for the reader to understand how an instrument is administered and scored. Some measures are self-administered and may simply be mailed or passed out with instructions to, for example, "check the box that most represents your feelings at the present time." Others, such as the Rorschach, require extensive training to administer and score. For others, the scoring methods are a well-guarded secret, and completed protocols must be sent to a central location for computerized scoring (for a fee, of course). In any case, we recommend that you pilot test any instruments you use, whether you make them up yourself or whether they are standard research tools. It may help to ask your pilot participants which instructions and items they found difficult or confusing.

If possible, include a copy of each instrument, with its instructions to the respondent, in an appendix to the dissertation. However, this usually is not appropriate with copyrighted instruments.

The following example, drawn from a doctoral dissertation (Hardwick, 1990), provides what we consider to be a clearly written description of a rather complex instrument. The section we quote contains information

about the structure, scoring, and administration of the instrument. Information about the reliability and validity of the instrument is contained in materials we have not chosen to quote here.

> The Bell Object Relations Reality Testing Inventory (BORRTI) is a 90-item "true-false" self-report inventory which was used to operationalize the dependent variables, object relations and reality testing. It yields four object relations (OR) subscales and three reality testing (RT) subscales. Each dimension, OR and RT, is measured by 45 items which are worded to reflect various levels of object relations and reality testing. The four OR subscales are: Alienation, Insecure Attachment, Egocentricity, and Social Incompetence. The three RT subscales are Reality Distortion, Uncertainty of Perception, and Hallucinations and Delusions. The test also yields a quasi-summary score for each overall OR and RT construct. This score is the sum of all items to which the participant responded in a pathological direction. Thus, a participant could conceivably have normal subscale scores while still responding to any number of pathological OR and/or RT items.

The author goes on to describe examples of test items and the meaning of each subscale, including the interpretation of high and low scores as well as the measurement characteristics of each subscale. As noted earlier, it is important to include a brief assessment of the reliability and validity of each measure, especially as they pertain to previous research conducted on populations as similar as possible to those in the proposed study. A dissertation by Gilbert (2007) provides the following description of the Levels of Emotional Awareness Scale (LEAS):

> The LEAS is a self-report instrument in which the respondent is asked to imagine 20 scenes described in the form of brief vignettes. A sample item is: "A neighbor asks you to repair a piece of furniture. As the neighbor looks on, you begin hammering the nail but then miss the nail and hit your finger. How would you feel? How would the neighbor feel?" Responses are scored separately for each scene according to a Glossary and Scoring Manual and summed to determine the total score. (p. 101, cited in Gilbert, 2007)

> The LEAS has demonstrated high interrater reliability ($r(20) = .84$). Internal consistency is also high (Cronbach's alpha = .81). Construct validity has been established by comparing the LEAS with scores on measures of emotion perception and of the ability to reason in terms of hypothetical emotions. The LEAS has been found to correlate with measures of ego development and the cognitive complexity of descriptions of parents. (Lane et al., 1990, p. 101, cited in Gilbert, 2007)

Copyrighted scales typically are not reproduced in a dissertation or research publication. To use copyrighted scales in a study, permission should be obtained in writing from the holder of the copyright.

What If I Design My Own Instruments?

In rare instances, a student may not be able to locate any existing measures that tap the construct the student wants to measure. Our first recommendation is to send the student back to the library with the instruction "keep looking." This advice reflects our strong belief that developing your own instrument is generally not a good idea. Research based on hastily thrown together instruments, which lack sufficient pretesting and are questionable in terms of reliability and validity, is of little scientific value. If the student is unable to locate a satisfactory instrument, he or she may wish to consider changing the focus of the research from one that attempts to relate the construct in question to other constructs, to one that attempts to design and assess a new instrument, a valid dissertation topic in and of itself. In this case, building the instrument becomes the central topic, and examining its relationships to other variables becomes part of the process of establishing the validity of the new instrument. Woodard (2001) set out to examine the relationship between the personality construct of hardiness and stress, coping, and physiological functioning in his dissertation. He ended up constructing and validating a measure of courage, which he proposed to be an integral component of the existential concept of authenticity, which is postulated to include the three components of hardiness: control, commitment, and challenge.

Research that concentrates on instrument development is a valuable enterprise and often makes greater contributions than research that attempts to relate existing measures to each other in some new and as yet untried fashion. Nevertheless, such research requires large numbers of subjects, frequent retesting, and sophisticated statistical models. Both the student and his or her committee must carefully weigh the pros and cons of undertaking such an enterprise. We provide six steps that are commonly involved in the process of instrument design:

1. Identify suitable content for an initial pool of items. This helps shape the content validity and face validity of the instrument and is facilitated by reading the theoretical and empirical literature and surveying the opinion of experts.

2. Create an initial pool of items, including items that may be parts of other instruments. This involves deciding on an item and response format

(e.g., 7-point Likert scale) and writing a large pool of logical, unambiguous, understandable items that reflect some dimension of the underlying construct being measured.

3. Obtain a group of expert judges to rate the pool of items for appropriateness (content validity) and clarity (wording). Eliminate poorly rated items based on systematic criteria.

4. Survey an initial pool of respondents with all of the remaining items. Perhaps include other measures to correlate with the items, such as a social desirability scale to control for the tendency to make a positive impression. The respondents should represent the population for which the instrument is being designed.

5. Determine the structure and reliability of the instrument and its subscales. This includes assessing the internal consistency of the scale(s) (i.e., how the items correlate with one another, using a reliability measure such as a coefficient alpha, and how each item correlates with the total score of the instrument or the individual subscales, or both) and may involve the use of exploratory and confirmatory factor analytic procedures. Reduce the item pool by discarding items that do not correlate sufficiently (either positively or negatively) with the total score as well as items that do not discriminate well (e.g., items that almost everyone endorses or rejects).

6. Survey a new pool of suitable respondents to validate the new measure. This includes correlating the responses to the scale with responses to other measures with which it is hypothesized to correlate positively (convergent validity) and negatively (discriminant validity). It might also include some form of predictive validity measure by testing to see if the scale effectively predicts a relevant criterion (e.g., a test of academic potential predicting subsequent school grades). Confirmatory factor analytic procedures are also likely to be used in this step.

What If I Modify an Established Instrument?

It is not uncommon for a student to modify questions or add questions to a validated instrument to facilitate its use. For example, when using some instruments with younger populations, it may be necessary to reword or eliminate questions regarding relations with the opposite sex. Phrases such as "My sex life is satisfactory" may not be appropriate for 7-year-old children.

Our position is that the modification of an existing instrument is perfectly acceptable but that such changes may make the norms invalid and may affect both the reliability and validity of the instrument. A close examination of many instruments currently in use reveals that there has been considerable borrowing among various authors over a lengthy time period. When borrowing or modification occurs, it becomes the

responsibility of the student to justify such changes and make his or her case for the reliability and validity of the instrument in its revised form.

It is often advisable to use an existing instrument in conjunction with a new instrument of the student's design. Even if the student feels that his or her instrument is better, the use of multiple measures of a single concept can be very useful when establishing the reliability and validity of a new instrument. Of course, when a new instrument fails, as it often does, the old standard can be used in its place. When Slanger (Slanger & Rudestam, 1997) designed her dissertation on participants in extreme risk sports, she knew that the application of the self-efficacy measurement to new domains was limited by the situational specificity of the measure. Thus, she supplemented a generalized self-efficacy scale with a new scale of her own design that focused on significant physical risk taking. To her surprise and delight, her previously invalidated scale was a better predictor of her dependent variable than previously existing measures.

Most dissertations contain some form of demographic data sheet. These sheets contain questions that assess gender, ethnicity, years of employment, marital status, education, and the like. Although such questions need to be thought through carefully, we do not consider adding such questions to a battery of existing instruments in the same context as scale development. We suggest examining surveys from the National Opinion Research Center General Social Survey (GSS) for examples of how to solicit demographic information. For GSS survey information, including the exact wording of hundreds of demographic questions, visit http://webapp.icpsr.umich.edu/GSS/.

What If I Want to Use Archival Data?

Questionnaires are not the only type of data collection instrument. Behavioral observations, extended interviews, and archival data all constitute valid sources of data for dissertation research. Students need to be cautioned about the potential problems of relying on archival data (i.e., data previously gathered and available from other sources prior to the design of the current study) for their dissertations. Here are three issues that must be considered by the student and his or her dissertation committee when considering this option:

1. In our experience, it is not uncommon for students to stumble on or be given a large data set prior to formulating a research question—or even to having a specific research interest. When this happens, students are

tempted to generate a suitable research question and set of hypotheses to conform to the available data. This is putting the cart before the horse: Whenever possible, the method follows from the question rather than precedes it. Research questions need to be justified on their own merits.

2. Archival data sets often suffer from missing, incomplete, or compromised data. This can take the form of an insufficient sample size, the absence of information on important variables that were not included in the original data collection, or the reliance on flawed or out of date measures. The dissertation student is responsible for the adequacy of the data set, and blame for any deficiencies in design or data should not be passed off on other agents.

3. Archival data are usually owned or controlled by others. Students who use them for their dissertation research must ascertain that they will have total access to the data and authorship and control over the resultant dissertation and, in most cases, associated publications.

In one telling example, a student we know who relied on someone else's data discovered, in the middle of data analysis, that the owner of the data had moved out of the country and taken the data with him. Although we have our misgivings about students relying on archived data, we are certainly aware of many excellent studies emanating from them. For example, Newton's (1991) dissertation examined the incidence of depression, substance abuse, and eating disorders in the families of anorexics. Data were obtained from existing medical records. Newton supported his data collection procedures as follows:

> A number of authors have studied the validity of data found in medical records (Harlow & Linet, 1989; Horwitz, 1986; Paganini-Hill & Ross, 1982; and Hewson & Bennett, 1987). Horwitz (1986) compared epidemiological data collected from medical records and by interview for 462 subjects who were part of a case-control study of chronic disease. Results showed that agreements between medical record and interview data are variable, and depend on the type of data examined and the strategy for handling incomplete or ambiguous responses. Horwitz did find a 93% agreement between medical records and interviews for family history of breast cancer. (p. 45)

In this quotation, the author let the readers know that he is aware of potential problems with his method, he has read the literature, and the literature supports the validity of this technique in some instances.

Many large archives provide rich opportunities for further study, and there is probably no persuasive reason for re-collecting data that already

exist. One situation that merits using archived or secondary data is when important and relevant data exist that cannot effectively be duplicated by another researcher. The previously mentioned National Opinion Research Center General Social Survey is a good example. Another example is when archived data are critical to a new study, such as the student who wishes to discover how romantic relations have been expressed through the decades by analyzing popular song lyrics over the past 50 years.

Another situation calling for the use of secondary data occurs in academic departments or research centers when several (student) investigators collect or use the same large data set to ask and answer different research questions on the same general topic.

Procedures: Describing How You Did (or Will Do) It

The Procedures section provides a detailed description of the exact steps taken to contact your research participants, obtain their cooperation, and administer your instruments. After reading this section, one should know when, where, and how the data were collected. For example, in a mailed survey, one might describe the following steps: (1) mail precontact letter; (2) 1 week later, mail survey packet; (3) 2 weeks later, mail follow-up survey for nonrespondents. Exact copies of the precontact letter and cover letters accompanying the survey measures should be provided in appendices. Note that the sampling procedures have been described in the Subjects section and the measures described in the Instrumentation section. There is no need to repeat this information in the Procedures section. When procedures are complex and require the administration of multiple instruments over multiple time periods, a flowchart or table presenting the procedures visually becomes very helpful.

It is important that any information that might potentially affect the number of participants or their characteristics be included in this section. For example, much social research is conducted with college students. Were the students asked to volunteer? Were they given extra credit? Was participation mandatory as part of a course requirement? Similarly, when administering mailed surveys such simple information as whether the addresses were handwritten or mailing labels were used can affect response rates. Such information is critical to replicating the experiment and understanding the exact nature of the population sampled. It is also important to describe the procedures undertaken to gain access to the population. For example, if working with high school students, was it necessary to obtain permission of the school board or the principal, or did

one simply sample all sixth-period classes? Finally, the procedures for obtaining informed consent should be described in detail and a copy of the informed consent statement, if any, provided in an appendix. This is particularly important when working with minors, who are legally not permitted to provide their own consent to participate in research. (Chapter 12 contains a more detailed discussion of informed consent and other ethical issues.)

Typically, instructions to participants, as well as ethical release forms, are included in the appendices. The ethical release forms are crucial because they inform participants about the potential hazards of participating in the study (e.g., emotional upset), limits to confidentiality, and use of the data, and they make it clear that participation is voluntary. (In some rare instances, participation may not be voluntary.) The purpose of the release forms is to protect both you and those who participate in your research. Release forms also force you to think about the implications of your study on the physical and emotional well-being of humans or animals employed as subjects. Most universities have stringent procedures for conducting ethical research, which generally include passing proposals through an institutional review board or human subjects committee. A sample copy of a typical human subjects release form can be found in Chapter 12. These forms should include a description of the study, the right of refusal, an explanation of risks and potential discomfort, an opportunity to withdraw without penalty, and the potential for feedback.

Data Analysis: How to Justify and Describe an Analysis

A research proposal often includes a statement that describes the statistical tests that will be used to address the hypotheses and research questions. This section is usually the one that scares students the most (because students learn statistics in the abstract, not as applied to specific research questions). The great benefit of including this statement is that it forces you to think through how you will treat the data from your dissertation at the time the proposal is generated, rather than after the data are collected. Time spent thinking about these issues at the proposal stage saves considerable time and pain later. In this way, one can avoid spending countless hours collecting data that ultimately are not analyzable because they are not in the correct format in the first place. In the next chapter, we describe in detail how to present the results of a study once the data have been analyzed. In this section, we discuss how to propose a particular sort of analysis prior to having actually undertaken it.

This section is particularly difficult for a number of reasons, beyond the fact that students may not be prepared to apply statistics to their own research. First, statistical analysis is virtually never a one-shot affair. Data may be analyzed and reanalyzed many times before the researchers are satisfied that the data have been given the correct treatment. A technique that initially seemed perfectly reasonable will later seem inappropriate because of the number of cases or distributions presented by the completed sample. Second, interesting questions often arise after initial analyses are completed. If a hypothesis is not supported, one may search for alternative variables that might help explain the lack of support. For example, a hypothesis may be supported only among the highly educated. Research with a large number of noncollege graduates in the sample may obscure the relationship until the college graduates are analyzed separately from the noncollege graduates (i.e., education becomes a moderator variable).

What, then, is the student to do? First, we strongly recommend that, in consultation with your advisors, and experts if necessary, you describe which technique seems most appropriate given the nature of your hypothesis, the number of independent and dependent variables, and the level of measurement of each of the variables. The following chapter contains recommendations regarding the use of statistical techniques. Computer programs also can assist with the selection of statistical techniques, and many statistics texts present flowcharts directing students to appropriate statistical methods. Do not turn your data analysis section into a treatise on a particular technique paraphrased from a statistics text; instead, present a short, reasoned statement that a particular technique seems most appropriate. An example of a straightforward data analysis proposal is provided by Connell's (1992) study of stress in psychotherapists:

> The primary hypothesis, that there will be a significant difference between Group 1 and Group 2 on level of stress, will be tested using a one-tailed *t* test. . . . The expectation that Group 2 will score differently than Group 1 in 12 of the subscales of the Essi System StressMap will be tested using a multivariate analysis of variance (MANOVA).

When analyses are complex and based on a multivariate theoretical model, a broader statement of analytical strategy may be required. For example, a dissertation proposal by one of our students, Lisa Stewart (2006), examined the relationships among perceived stress, self-efficacy and depression, hopelessness, and suicidal ideation. She described the analysis of her theoretical model as follows:

Path analysis techniques will be used to examine the direct and indirect effects between the variables of initial stress, self-efficacy, depression, hopelessness, and suicidal ideation. Path coefficients will be computed via a series of bivariate and multiple regression analyses based on the hypothesized model.

Five direct, nonmediated relationships will be tested using multiple regression. These include self-efficacy and initial stress, self-efficacy and depression, self-efficacy and hopelessness, depression and hopelessness, and depression/hopelessness and suicidal ideation.

Three mediated relationships will be tested using multiple regression analyses: the effect of self-efficacy on suicidal ideation, the effect of self-efficacy on depression/hopelessness, and the effect of initial stress on suicidal ideation. (p. 65)

Should You Discuss the Limitations of Your Research?

A final section of the Method chapter that we encourage students to include in a dissertation proposal is a statement on limitations and delimitations of the study. Delimitations imply limitations on the research design that you have imposed deliberately. These delimitations usually restrict the populations to which the results of the study can be generalized. For example, you may decide to study only males, either because the theory on which your hypotheses are based has not been studied in females or because you have a readily accessible population of males but not females. Limitations, on the other hand, refer to restrictions in the study over which you have no control. For example, you may be limited to only a narrow segment of the total population you wish to study, or you may be limited by the method you elect to use.

The Method Chapter in a Qualitative Dissertation

Too often students view the relatively unstructured nature of qualitative research designs as license to omit clarity and specificity in the Method chapter of their dissertations. Qualitative studies should not be considered as opportunities to ignore the planning process in research. Issues of identifying and soliciting participants, selecting and preparing research materials and data collection tools, and formulating procedures pertain here as in all studies. The reader needs to understand what you did and how you thought about it in order to appreciate the links among the research problem, the method, and the results.

The organization of the Method chapter, as well as the content that goes into it, may depend on your research model as well as the conventions of your discipline and preferences of your committee members. Josselson and Lieblich (2003) prefer the term *Plan of Inquiry* to *Method* because the word *method* seems to focus on the procedure rather than on how one thinks about the question. We find that students who conduct qualitative studies are prone to spend an inordinate amount of space explaining the philosophy of science supporting the design. Such long-winded explanations seem unnecessarily apologetic; after all, a quantitative dissertation doesn't require a treatise on logical positivism. Basically, the reader needs to know about your research strategy and how your data will be generated with as much specificity as you are able to offer at this stage of the project.

Qualitative research adopts views of sampling, instrumentation, and data analysis that are often directly contrary to views held by those conducting more traditional "rationalistic" inquiry. Lincoln and Guba (1985) called this "the paradox of designing a naturalistic inquiry" and argued that "the design specifications of the conventional paradigm form a procrustean bed of such a nature as to make it impossible for the naturalist to lie in it—not only uncomfortably, but at all" (p. 225). Nonetheless, there are some practical distinctions between these issues in quantitative and qualitative studies that are worthy of consideration.

Sampling and Sample Size in Qualitative Studies

Determining where and from whom data will be collected is directly analogous to our consideration of sampling. Quantitative studies generally rely on random or representative sampling to generalize findings from the sample to the population. The qualitative researcher is more apt to elect purposive or theoretical sampling to increase the scope or range of data exposed (random or representative sampling is likely to suppress more deviant cases) as well as to uncover the full array of multiple perspectives (Lincoln & Guba, 1985, p. 40).

A phenomenological study usually involves identifying and locating participants who have experienced or are experiencing the phenomenon that is being explored:

> Phenomenological research uses sampling which is idiographic, focusing on the individual or case study in order to understand the full complexity of the individual's experience. From this perspective, there is no attempt to claim an ability to generalize to a specific population, but instead, the findings are relevant from the perspective of the user of the findings. (Bailey, 1992, p. 30)

The participants, if you will, are the experiential experts on the phenomenon being studied. This means that the sample probably would not be randomly drawn from a group of college sophomores. Rather, the researcher uses criterion sampling, selecting participants who closely match the criteria of the study. Katz's (1995) participants, for instance, had to meet both inclusionary and exclusionary criteria: a deliberately diverse and representational sample of women who had experienced "ongoing, intractable symptoms of discomfort in the skin of the vulva and/or the vaginal vestibule, which may be diffuse or in specific loci, and either persistent or episodic" (p. 91) for a period of 1 year or more. Referral of participants came directly from medical providers. Moreover, most phenomenological studies engage a relatively small number of participants (10 or fewer might be appropriate) for a relatively long period of time (at least 2 hours). These factors should be carefully noted in the Method chapter.

Because the grounded theory study is inductive and theory evolves as the data are collected and explored, it may be neither possible nor advisable to establish the precise sample size beforehand. Strauss and Corbin (1998) stressed that several forms of sampling are appropriate at various stages of the study. The trick is to choose participants who can contribute to an evolving theory, participants whose main credential is experiential relevance. Thus, at the outset, open sampling might be most appropriate. This means selecting participants or observations without prejudice because no concepts have yet proven to be theoretically meaningful. In practice that might mean systematically choosing every *n*th name on a list, staying maximally flexible and open to discovery. Researchers who support this approach might regard 20 to 30 participants as a reasonable sample (e.g., Creswell, 1998). Other representatives of the grounded theory approach are leery of processing the amount of data accumulated from that many respondents. They recommend beginning with perhaps 5 or 6 participants who have been selected because they seem to have the phenomenon of interest in common, a process called homogeneous sampling (Glaser & Strauss, 1967). Participants who are different from the initial sample are added only if they represent some quality that emerges as significant for understanding (and perhaps generalizing) the phenomenon under study.

As the study proceeds, the chief criterion for sampling moves toward theoretical relevance. At this point, the researcher has begun to assimilate some early theoretical hunches and wishes to identify examples that demonstrate the range or variation of a concept in different situations and in relation to other concepts. The sampling is done to saturate a concept, to comprehensively explore it and its relationship to other concepts so that it

becomes theoretically meaningful. Because the theory emerges from the data, there is no viable way of determining these sampling dimensions beforehand. The researcher becomes increasingly selective in collecting a sample by adding to it based on the core variables that emerge as important to the theoretical understanding of the phenomenon under study. This process is sometimes known as discriminate sampling—choosing persons, sites, and documents that enhance the possibility of comparative analysis to saturate categories and complete the study. This might mean returning to previous interviews or sources of data as well as drawing on new ones. The study proceeds until there is theoretical saturation—gathering data until no new relevant data are discovered regarding a category and until the categories are well developed and validated. For the interested reader, Strauss and Corbin (1998) go into significant detail in describing the varieties of sampling relevant to grounded theory research.

Josselson and Lieblich (2003) agree that saturation—that is, stopping data collection when the results start to become redundant—is the key determinant of sample size. They caution, however, that real saturation never occurs because each new respondent has something unique to contribute to the study. They noted that it is usually the researcher who becomes saturated and that it is important to collect sufficient data to represent the breadth and depth of the phenomenon without becoming overwhelmed. Generally speaking, the longer, more detailed, and intensive the transcripts, the fewer the number of participants. In practice, this may mean specifying a range between 5 and 30 participants.

The ethnographer has different challenges to anticipate. The researcher's relationship with the group to be studied must be fully acknowledged and described. In some ethnographic studies, the researcher may already have membership status. More frequently, there is a gatekeeper or conduit for accessing the group. Once contact has been established, will everyone be approached to participate? Will sampling take place according to some thoughtful criteria? Will participants be selected opportunistically, according to convenience or eagerness to participate? The same rules apply in describing the use of artifacts and other sources of data that the ethnographer observes and records.

Elizabeth Moore's (1995) ethnographic study of organizational culture includes her list of 11 criteria for choosing an appropriate company to study, including issues of size, age, and tenure of employees; opportunities for observation; and access to company documents and employees. She received entry into a chemical company that met her criteria and gave her unlimited access to the organization. She offered all 29 employees the

opportunity to participate in confidential interviews and ended up interviewing a majority of the employees as well as the founders and their wives. The interviews were of four types:

1. Critical event interviews, which focused on critical periods in the evolution of the organization
2. Ethnographic interviews, which dealt with the artifacts and values within the organization
3. Spouse interviews
4. A customer interview

She also relied on a month of participant observation activities and analysis of organizational documents. The sampling procedures, interviews, observation activities, and documents were described in the Method chapter, and the interview protocols were placed in appendices to the dissertation.

Instrumentation (or Measures) in Qualitative Studies

The instrument of choice for the qualitative researcher is the human observer. Thus, qualitative researchers place particular emphasis on improving human observation and make no claims for the reliability and validity of the instrument in the rationalistic sense. (A discussion of reliability and validity in qualitative studies follows later in this chapter.) The task for the student proposing an observational study would be to place considerable emphasis on the training and practice of the observer(s).

Qualitative researchers also recognize the use of other, more traditional sorts of instrumentation, provided they are grounded in the central focus of the research and axioms underlying the method. Typically, our students have used interviews to generate discussion surrounding the major research questions. In these instances, the task of the instrumentation section is to describe the interview and demonstrate how the interview serves as a sufficient device to focus discussion on the research questions of the study.

Although the interview itself may be quite loosely structured and flexible, phenomenological researchers generally prepare some questions in advance, preferring to alter them if it seems appropriate as the interview progresses. Although the precise wording of questions may vary from time to time, certain types of questions are pro forma. Think of the questions as tools to draw out the participant to reflect on the experience and its implications in his or her life. Thus, one might request that a

participant relax and focus on the incident or the phenomenon and "describe the experience, how you felt, what you did, what you said, what thoughts you have about it." Further questions serve as probes to encourage the interviewee to dig deeper and reflect on the meaning of the experience:

> What aspects of the experience stand out for you?
>
> How has the experience affected you?
>
> What changes have you made in your life since the experience?

Grounded theory studies generally make primary use of interview techniques as well, although journals and other written records, as well as participant observation methods, may also be employed. Interviews may be either individual or group based, including what are popularly known as focus groups. LaPelle's (1997) dissertation provides a good example of a prepared opening statement to the interviewees:

> I would like you to think about performance review experiences that significantly affected your interest, motivation, or performance in your job or subsequent career decision. Please describe, in as much detail as you can remember, the circumstances surrounding these experiences. (p. 37)

Subsequent probes included possible circumstances, such as work responsibilities, interaction with supervisor, inclusion of rewards or recognition of work, skill development as a function of the review, and personal meaning of the review, among others. More objective questions were included to obtain clarification of work history and work environment after the narrative portion of the interview was completed. Note that the initial question is a question of discovery, to identify what is common among the experiences of an aggregate of individuals. In this light, David Rennie (1998), a well-known grounded theory researcher, cautioned researchers not to begin with a set of questions that are really categories for coding so that the categories don't emerge from the data and the data analysis becomes content analysis.

Whereas survey researchers typically insist on preparing a specific set of questions that need to be asked precisely and in a particular order, most qualitative interviewers start with a general plan of inquiry but not a formal set of questions. The Method chapter would include at least one opening question and then some detailed follow-up questions that the researcher may or may not use depending on the subsequent flow of the interview. A student who intends to take this route needs to obtain ample

training in conducting qualitative interviews prior to embarking on a dissertation. People construct meaning out of life events, and good interviewers learn to appreciate how they must listen patiently and sensitively and invite these stories by the way they frame their opening questions and follow-up interventions. In this spirit, some of our colleagues insist that students conduct a demonstration interview and include the analysis of it in a qualitative dissertation proposal (Josselson & Lieblich, 2003).

Several excellent texts are available for guidance in preparing and conducting qualitative research interviews. Among the more popular ones are Mishler (1991), Kvale (1996), Weiss (1994), and Rubin and Rubin (2004). Because these authors take somewhat different perspectives, it is a good idea to become familiar with more than one approach, then determine what best suits your own style and orientation.

Data Collection in Qualitative Studies

Issues of determining the successive phases of a study and planning data collection and logistics are subsumed under what we have described as Procedures. Regardless of the sort of study one is conducting, attention always must be paid to how the data are to be collected, independent of the form the data might take. Data recording may be described along two dimensions: fidelity and structure. An open-ended interview, when properly recorded, has high fidelity and little structure, whereas a standardized paper-and-pencil test has both high fidelity and high structure. We recommend the use of tape recorders to record interviews and place little reliance on the use of field notes (i.e., low fidelity and low structure). A diary or journal to record impressions, reactions, and other significant events that may occur during the data collection phase of research, however, is recommended as a useful source of supplementary information.

Data Analysis in Qualitative Dissertations

The issue of describing the data analysis in the Method chapter of the dissertation may be more problematic for those considering qualitative research than it is for those undertaking conventional quantitative research. One perspective comes from Lincoln and Guba (1985), who stated, "Not very much can be said about data analysis in advance of the study" (p. 241). This leaves the student in a difficult position, particularly if a committee is clamoring for a data analysis section in the proposal. Quite a bit has been said about data analysis within the context of qualitative studies, starting with an entire chapter of Lincoln and Guba's *Naturalistic Inquiry* (1985), which considers the issue of processing naturalistically obtained

data in considerable detail. Miles and Huberman (1994) devoted an entire book to the issue of qualitative data analysis, and many recent texts on qualitative research, such as Marshall and Rossman (2006) and Silverman (2006), discuss the topic extensively. We consider some of these works in more detail in the following chapter. The main point is that even though a student may not be able to refer to specific statistical procedures, we believe that the general framework of an analysis can be specified in advance. Students should expect that both qualitative and quantitative dissertations are likely to involve multiple phases of data analysis.

Validity and Reliability in Qualitative Dissertations

In traditional empirical research we are mindful of the importance of reliability, internal validity, and external validity of measures and procedures. Many qualitative researchers, on the other hand, forgo the use of the terms *validity* and *reliability* because of the historical link of these concepts to objectivist research and the argument that they are inappropriate to naturalistic inquiry. Nonetheless, all research carries the responsibility of convincing oneself and one's audience that the findings are based on critical investigation. The possibility that the results of a qualitative study rest on a few well-chosen examples has been pejoratively described as anecdotalism (Silverman, 2005). In contrast, the trustworthiness of a design becomes the standard on which it is likely to be judged, and the Method chapter is the place where evidence for methodological rigor is introduced.

Validating an argument or research process basically means showing that it is well founded and sound, whether or not the results generalize to a larger group. On the other hand, when a procedure or result is reliable, it means that we can depend on it (i.e., rely on it). As Richards (2005) pointed out, reliability is not the same as total consistency, anymore than the basic regularity of train schedules implies perfect predictability. She went on to note that the goal of employing standardized measures in a controlled setting is likely to be incompatible with naturalistic research.

It may not be necessary to use the traditional terms *reliability, internal validity,* and *external validity* in writing a qualitative dissertation. Guba and Lincoln (1985), for example, recommended the alternative constructs of credibility, transferability, dependability, and confirmability, and the interested reader is referred to their text for a more complete understanding of these terms. However, the Method chapter needs to attend to these issues in some convincing way. What follows is an overview of recommended approaches.

Reliability concerns the replication of the study under similar circumstances. It pertains to issues such as training interviewers and systematically recording and transcribing data. The naturalistic investigator derives consistency through coding the raw data in ways so that another person can understand the themes and arrive at similar conclusions. The researcher's coding scheme needs to be introduced in the Method chapter, with the understanding that the analysis is likely to be modified both during and after data collection.

Internal validity refers to the validity of a causal inference. From a social constructivist perspective, validation is the process of evaluating "the 'trustworthiness' of reported observations, interpretations, and generalizations" (Mishler, 1990, p. 419). In qualitative research, we ask about the extent to which the investigator's constructions are empirically grounded in those of the participants who are the focus of study (Flick, 2002). The credibility or truth value of findings might be ascertained by spending sufficient time with participants to check for distortions, exploring the participant's experience in sufficient detail, videotaping interviews for comparison with the recorded data, clarifying tentative findings with the participants, revising working hypotheses as more data become available, and checking multiple sources of data such as other investigators, written records, diaries, field notes, and so on. These procedures for adding to the credibility of the study need to be noted in the Method chapter, with the understanding that they might be amended as the study evolves.

External validity refers to the generalizability of the findings of the study. The qualitative study emphasizes the "thick description" of a relatively small number of participants within the context of a specific setting. The descriptions of the participants or setting under study are sufficiently detailed to allow for transferability to other settings. Samples can change as the study proceeds, but generalizations to other participants and situations are always modest and mindful of the context of individual lives. Moreover, generalization is the task of the reader rather than the author of qualitative studies.

Part of the challenge of completing a dissertation proposal using qualitative methods is to master the language of the qualitative paradigm of inquiry. The following procedures can be employed to enhance the trustworthiness of a qualitative research project. No study needs to include all of them, but whatever procedures are adopted should be described in the procedure section of the Method chapter.

Criteria of Adequacy and Appropriateness of Data (Morse, 1998). *Adequacy* pertains to the amount of data collected in a qualitative study, analogous to ensuring sufficient power by insisting on an adequate number of participants in a quantitative study. Adequacy is achieved when you have obtained enough data so that the previously collected data are confirmed (saturation) and understood. *Appropriateness* means that information has been sampled and chosen purposefully rather than randomly to meet the theoretical needs of the study. Multiple sources of data are obtained to provide saturation and confirmation of the emerging model.

Deviant Case Analysis. The constant comparative method implies that the researcher continues to build and test the completeness of a theory by reaching across participants to determine how the findings apply to cases that appear to be exceptions to the rule. By deliberately searching for these "deviant" cases, it becomes possible to test a provisional hypothesis and amend it to incorporate new and different data.

The Audit Trail. An audit trail refers to keeping a meticulous record of the process of the study so that others can recapture steps and reach the same conclusions. An audit trail includes not only the raw data but also evidence of how the data were reduced, analyzed, and synthesized, as well as process notes that reflect the ongoing inner thoughts, hunches, and reactions of the researcher. This critical self-reflection component (previously described as reflexivity) illuminates the researcher's potential biases and assumptions and how they might affect the research process. A further possible step, called an external audit, involves asking an external consultant who has no relationship to the study to review the materials and assess the findings and interpretations for consistency.

Member Checks. It is common in the qualitative literature for researchers to return to informants and present the entire written narrative, as well as the interpretations derived from the information, with the intention of confirming the accuracy and credibility of the findings. For some researchers, this is consistent with elevating the informant from the role of a participant in the study to the role of co-researcher. Others, such as Silverman (2005), advise caution when putting respondents in a "privileged" position by asking them to verify or validate the research findings.

Triangulation. Soliciting data from multiple and different sources as a means of cross-checking and corroborating evidence and illuminating a theme or a theory is known as triangulation. The different sources may

include additional participants, other methodologies, or previously conducted studies. Of course, it is equally likely with qualitative studies that different kinds of data will yield different interpretations.

Peer Review or Debriefing. Many qualitative researchers make use of peers or colleagues to play the role of devil's advocate, asking tough questions about data collection, data analysis, and data interpretation to keep the researcher honest. The other role of the peer reviewer is to provide professional and emotional support by being an empathic listener to the researcher along the way. Both the researcher and the debriefer typically keep written accounts of their sessions together.

This concludes the presentation of the basic elements of a Method chapter for a thesis or dissertation. When the study is completed and the Results and Discussion chapters have been added to the document, it is necessary to go back to the Method chapter and change the verb tense from future to past. We also recommend removing any detailed sections on data analysis from the Method chapter and incorporating that material into the Results chapter.

BOX 5.1 Student Suggestions

1. Pilot test any instruments you use, whether you make them up yourself or they are standard research tools. Every person can misread or misunderstand something. It also helps to ask your pilot participants specific questions, such as inquiring if a section was interesting or difficult.
2. I was amazed by the vague and ambiguous items often present in well-known instruments. You can't trust that other researchers have constructed scales that will be appropriate for your use. If you want to compare your results to scores reported in the literature, however, you can't tamper with the instrument.
3. If you decide to construct your own instrument, never start with demographic items. Most researchers put them at the end. They are boring. Be sure your questionnaire looks nice and aesthetic. It must be clear and easy to understand. Put your response alternatives in columns rather than rows. It is easier for the eye to follow.
4. It helps to give an incentive to participants. Include a stamped, self-addressed envelope. Most important is personal contact with participants. My committee thought I would get a 10% return rate from a mailing to American Psychological Association members, so I made sure to contact people who had access to subjects directly.
5. One relatively inexpensive way to encourage people to participate is to offer a lottery or drawing for money, a gift certificate, or some item of value. That way it will not cost an arm and a leg to pay every participant but will provide a worthwhile reward for one or a few participants.