The Method Section

CHAPTER CONTENTS

Section 1: General Considerations 76
  Adapt the Material on Method to Your Study 76
  The Method Section Flows from the “Questions, Hypotheses, or Models” Section 76
  Operationalizing Terms May Result in New Conceptualizations 79
  Restrain the Design to Realistic Limits 80
  Eliminate Plausible Alternative Explanations 82
  Design Efficiency 86
  Give Special Care to Those Sections Critical to Your Research Method 87

Section 2: Developing the Subsections of Method 87
  Participants—Population and Sample 88
  Situation 90
  Focus of Action—Treatment(s), Independent and Dependent Variable(s) 90
  Records—Instrumentation and Observations 91
  Comparison and Contrast—The Basis for Sensing Attributes or Changes 95
  Time Schedule—The Specification of the Procedure 96
  Problems in Data Collection 96
  Analysis 98
  Expected End Product 98

An Alternative Dissertation Format: Articles Ready for Publication 99
Worksheet 5.1: Study Methods Review 100

The method section, which describes the procedures that will be used, translates the problem section developed in the previous chapter into project activities. This is usually the most carefully read section of the whole proposal. Up to this point, you may have told in glowing terms and appealing generalities what
you hope to do and what this will mean to your field. The section on method brings this down to earth in operational terms. Frequently, proposals that sound as though they will revolutionize a field appear much more mundane in the method section; the techniques proposed for attacking the problem may fall far short of what was implied when the earlier sections were written. Obviously, the method section should fulfill the expectations created by the foregoing sections.

The following material assumes that you have a reasonably clear idea of what you wish to study and how you wish to study it. The term design is used to describe the latter, both here and in the chain of reasoning. So the study's design is described in the method section. Obviously, that term design applies quite loosely to studies with emergent topics.

Our discussion of the method section is in two parts: In Section 1 we discuss general points to take into consideration in developing and describing method. Section 2 is a detailed discussion of each of the subsections of design that together, typically, are required to describe the method.

SECTION 1: GENERAL CONSIDERATIONS

Adapt the Material on Method to Your Study

Of all the proposal parts, the method section is the most dependent on the nature of your study. Some of the material in this chapter may be irrelevant to your kind of study. Consider what is important in your study and adapt it. For example:

If the study is a sample survey, elaborate on the sampling section. If it is an emergent study, collapse the sections on the links in design into a single description of how and where data will be gathered and the initial focus of attention. Suppose you plan to examine the educational and medical records of late-nineteenth-century Italian immigrants to determine how they differed from nonimmigrants. Describe where you are going to study the records, what you expect to find there, how you will get access, what information will be gathered and how, and what analyses will be performed.

Further discussion of how to adapt this section is given in chapters 7, 8, and 9, addressing the special requirements of various kinds of proposals.

The Method Section Flows from the "Questions, Hypotheses, or Models" Section

Considering a project as a chain of reasoning, the "design" link in the chain is logically derived from the previous links, specifically the "questions, hypotheses, or models" link. The design specifies the operations by which you will investigate whatever you chose at the questions-hypotheses-or-models link. If it is a question, it will indicate where, how, and when you will seek an answer to it. If it is a hypothesis or a model, it will describe how you will provide evidence in support of the prediction and relate it to the underlying explanation of the phenomena that resulted in the prediction. Also, it will indicate how relevant alternative explanations can be ruled out.

The process is usually one of a direct translation of the concepts in the question, hypothesis, or model into the choice of:

1. participants,
2. situation,
3. focus of action—the core variables such as treatment and effect,
4. records—measures and observations,
5. comparison and contrast (basis for sensing attributes and changes)—the basis on which the change due to an independent variable or experimental treatment or whatever happens at the focus of attention will be sensed (e.g., pre/post comparison or comparison with another group), and
6. time schedule—the study's procedures and the schedule of the various activities involved. These may involve observations, sensing the presence and strength of the independent variable, administering an experimental treatment, and/or measuring and observing any effect.

For example:

Consider this hypothesis: "Up to some reasonable point, the more time African American students spend studying African American history, the stronger their self-concept." This hypothesis suggests that increasing levels of study of African American history will result in gains in self-concept up to some point. To develop the design, decisions will have to be made about how to translate into operational terms all the concepts in the hypothesis:

The researcher must specify what is meant by "African American students" in terms of age, grade, and whether such variables as socioeconomic class or urban/rural background are important.

What does "study African American history" mean? Will any African American history curriculum do, or does it need to be one that stresses African American accomplishments?

Is there a measure of self-concept that is valid for African American students at the age chosen? The easily available measures for college students are not appropriate for elementary pupils.

What design will determine whether the variables change together as hypothesized? For example, one design might involve groups of students who are comparable except
ADVICE COMMON TO MOST PROPOSALS

for the time they have studied African American history. You could compare self-concept change for those who have studied it for short vs. longer times. Alternatively, a longitudinal design might be developed that follows the changes in a group exposed to a lengthy study of African American history.

For many of the decisions (e.g., choice of age and grade or use of comparable groups vs. longitudinal designs) you must choose among alternative translations. Thus, a variety of interpretations can result when a hypothesis is translated into operational terms. The same process is involved in translating questions and models.

Some terms seem to immediately translate into design features. Here are some examples:

Long-term retention vs. immediate recall Requires multiple posttests. Note that even here there are alternatives: the same group can be tested several times, or, to eliminate the effect of retesting, use different groups—one tested immediately, others for different lengths of retention.

Cumulative treatment effect Requires multiple posttests. Again these could all be of the same group or of different groups, each tested after a different length of treatment.

Anticipatory effect (e.g., effect of studying sculptures or enjoying paintings) Pre—and posttesting. One group before exposure to sculpture as well as before and after paintings exposure; one group before and after paintings exposure only.

Enhancing or interactive effect of a variable with treatment (e.g., printing the words that are key to understanding a text in a contrasting color in a reading test) Separate treatment groups with and without the presence of the interactive variable conditions.

THE METHOD SECTION

Operationalizing Terms May Result in New Conceptualizations

As terms are operationalized, you often come to a different understanding of the study from when it was initially conceptualized. Terms take on new meaning, and often the initial conceptualization has to be sharpened and modified as the problem becomes better understood.

Suppose that you start out to study the relation of per pupil expenditures to achievement across a set of public school districts. In operationally defining per pupil expenditures (determining their dollar value), you find that different districts include different costs.

In an effort to get comparable data across districts, you adjust each district to include a common set of basic costs. But at that point, the study begins changing. There isn't much variability in these basic costs across districts; the variability is in the nonbasics, the discretionary money available to a school's principal to improve instruction. So that becomes the focus of the study, forcing you to go back and change the whole front end of the proposal to fit this new conceptualization of the problem.

Some researchers argue that you really come to understand the problem only in operationalizing the study. However, operationalization may never be completely satisfactory when you are dealing with constructs that can’t be concretized so as to satisfy everyone (e.g., personality characteristics such as likableness, monetary estimates of the value of good health, etc.). Remember this if you are dissatisfied with your study and/or the redevelopment of the method section seems never-ending. A compromise operationalization may be the only way to study your problem but, also, the source of your dissatisfaction.

Sometimes, when the questions, hypotheses, or models are given operational translations, it becomes immediately apparent that the problem is too large or too complex. In the per pupil expenditure example above, an attempt to estimate all the discretionary resources available in a given classroom might put the project beyond the realm of feasibility (parent volunteer time, laptop computers brought into the class by students, etc.). Yet these might be important inputs to the classroom in certain circumstances. First attempts at problem definition are particularly susceptible to impracticality where the student insists on doing “something significant.”

Refocusing and delimiting the problem to restore feasibility are the answers. But, sometimes, even after the problem has been refocused, certain requirements may still be too great. Consider whether these may be handled by alternative design choices. For example, if there are too few cases to establish both a control and an experimental group, the participants may be used as their own control with pre—and posttests.

Development of the design is an iterative process. The researcher sets an initial set of pieces in place, but changing one sets off a cycle of resulting changes. That may in
turn result in a reconceptualization, further changes, and so on, until all the pieces fit together and are feasible. The new curriculum takes too long, so it must be reduced to its essentials. That is likely to result in weaker learning, which in turn requires the size of the sample be increased in order to detect it. To get a larger sample you must include atypical persons. Now you must find a comparison group that includes similar atypical persons. And so it goes.

Often you must go all the way back to the beginning and redesign the study on a different basis. Many cycles may take place before a satisfactory solution is reached.

Restrain the Design to Realistic Limits

Even as the design is first being considered, you must make tentative decisions on what level of resources you can practically employ. Take into account your own time and what access and cooperation you can expect from other institutions, participants, etc. These estimates are important for making methodological decisions: the possible number of participants, number and location of study settings, and so on. Indeed, the limits may rule out certain methods that take too long, such as a longitudinal study. Getting parents' consent to test children may be difficult or impossible. The most desirable and cooperative institutions may be too distant.

Some of the limits are easy to estimate, others more difficult, but some reasonable determination must be made for all of them if development of the design is to proceed realistically. Further, just as other parts of the design are successively adapted, so initial limits may have to be adjusted as the plan develops. Since many of these judgments are based on practical experience, seek the advice of your chairperson, committee members, and other researchers who have conducted similar studies and learned what is realistic through hard experience.

Resource Limits

As soon as you begin to translate the study into operational terms, the question immediately arises, "How big shall I make it?" Although it need not be answered precisely at the outset, some working estimates must be set. A key one is how long you can afford to work on your dissertation. Here again, seek the opinions of others: ask more senior graduate students how long the various pieces of their dissertation research are taking them.

Institutional Limits

When other institutions or agencies are involved, either as collaborators or as sources of data collection, support, etc., their perspective must be considered to ensure that requests made of them are reasonable. Most institutions operate by trying to do too much, for too many, with too little time and resources, and so may limit access to participants, facilities, equipment, or personnel. Further, they tend to resist changes in routines that interfere with "business as usual."

Ethical Limits

Ethical limits must be considered in developing the design. Any federally funded research involving human participants must be approved by a Committee on the Protection of Human Subjects concerned with the ethical implications of the study. Federal regulations prescribe the composition of the committee, which includes individuals outside the university. Although approval is required only for federally funded projects, nearly every university extends that requirement to all other research involving human participants, including dissertations. If clearance has already been routinely obtained, note it in the discussion of method. If not yet obtained, or expected not to be routine, it may require a section of its own along with other assurances. Committee on the Protection of Human Subjects clearances are further discussed in the next chapter (p. 110-11).

Time Limits on Proposal Development

It seems you ought to be able to control your schedule. But pressures to get your degree in a reasonable time, to gather data before certain natural breaks in institutional schedules, faculty unavailability due to trips and sabbaticals, and other scheduling difficulties may impinge on your timetable. Circumstances may, for example, require completion and approval of your proposal by an early date, enforce a particular schedule on data collection, or compel use of nonpreferred sites for data collection.

Consider the trade-offs involved in rushing to meet the immediate deadline or, if there is one, waiting for a later one when some of these problems could be more successfully resolved. A several-month delay in proposal approval might pay handsome dividends in more cooperative site conditions as the staff of these institutions and agencies are given a chance to contribute to the research plan and feel it is partly theirs. Considering that this may make for a more cooperative milieu and possibly better data, the delay may be worthwhile. But other considerations such as the availability of your own or a key person's time may be overriding.

Sort out those things that can be done satisfactorily in the time available for proposal development from those that are unwise to attempt or, perhaps, cannot be done even if tried. Attempting too much usually results in a proposal that shows it, as does similar haste in data collection and analysis in the dissertation. As in sewing, "find a pattern that fits the cloth available," or as in sports, "find a league in which you can comfortably play."
Eliminate Plausible Alternative Explanations

Whatever the methodology, all studies concerned with setting forth an explanation for a relationship must be concerned with the elimination of competing plausible explanations. Whether experimental, qualitative, survey, or whatever, if the design is not adequate to ensure the integrity of the study's chain of reasoning against plausible alternative explanations, readers may prefer an alternative to the explanation that the study is intended to support. The proposal should describe how the design so structures the study that plausible alternative explanations are ruled out as significant explanatory factors. Here's an example.

In a study comparing the effect of two different curricula, the researcher would be concerned that any initial differences in the groups might be reflected in their after-treatment performance. Otherwise, such after-treatment differences might be attributed as well to the initial differences as to the effect of the curricula. In this situation, the researcher might be expected to control such potentially contaminating factors as the beginning level of competence, general academic ability, and/or motivation.

The term design seems to go with experimental as in experimental design. You might be tempted, therefore, to assume that this discussion is of little importance to other than experimental studies, to a qualitative study, for example. Nothing could be further from the truth! A qualitative study observer, for example, must protect against a variety of alternative explanations; to name just a few:

- the possible effect of the observer's prevailing attitudes and values as they affect observations,
- the possible choice of individuals and times to observe which are "atypical" samples,
- the possible effect of "dropouts"—persons present at the start of the observations but not as they progress (usually referred to by the name mortality),
- the possibility of going "native" and perceiving things differently as observations progress.

Clearly, when we refer to design in qualitative studies, we are using the term to refer to such investigator decisions as whom to study, what persons or situations to contrast, what instances in time to compare, and similar judgments.

Control of Alternative Explanations

How do you control for possible alternative explanations that might equally plausibly be considered the cause of what you are studying? We have three ways: (1) elimination, (2) adjusting, and (3) spreading their effect equally across whatever groups or individuals are being compared (if all units are equally affected, then differences must be due to something else, presumably what you are studying). This last method is used in the two-curricula comparison. Individuals are ranked on achievement (assuming that achievement is a reflection of both ability and motivation so it controls for both), and persons with even-numbered ranks are assigned to one curriculum group, those with odd-numbered ranks the other.

Clearly, one of your tasks in writing the proposal is to identify potentially serious plausible alternative explanations and discuss which ones to control and how to control them. That is, you must:

- decide which alternatives are the most serious threats to the study,
- decide how they can be controlled,
- determine how controls for the set of the most serious threats can be combined into a design, and finally
- determine whether the design is feasible, adjusting it until it is.

This requires asking such questions as:

- How likely is each of these alternative explanations to appear?
- In your estimation, therefore, how critical is it that each of them be controlled?
- If there are several requiring control, how will you prioritize their relative importance?
- Will your chairperson, committee, and intended audience likely agree with your priority order?
- How, taking your own and these other opinions into account, shall these alternatives be prioritized in their claim on your resources?
- Which design best controls the top-priority alternative explanations?
- Is that design feasible? If not, how can it be modified so it is?
- Given the other claims on resources, what design is preferable?

The final decision must depend on the particular circumstances of each study, but a general principle is to find the design configuration that provides the best possible use of available resources at the same time that it:

1. Gives priority to the most serious alternative causes of the effect, taking into account their likelihood, and
2. Control by elimination if that is possible; by adjustment as a second choice if a good method is available; or, where it is not, as a third choice, by building them equally into the groups being compared.

Control for as many variables as are important and as can feasibly be accommodated. This is one of many areas subject to your good judgment for
which no set of foolproof rules can be provided. Every study is a compromise between what it is realistically possible to control and those variables that would be nice to control in the most perfect of all possible worlds.

Unfortunately, not all judges will weigh the desirability of controlling possible contaminating factors the same way. Their "most acceptable compromise" may differ from yours. Once again, this is a place to demonstrate your mastery of the problem. Nobody knows better than you do the multiple sources of contamination that might affect your study. Therefore, in your proposal, convincingly indicate:

- The nature and basis of the particular compromise being proposed,
- The reasons for accepting it,
- The reasons for choosing to control the variables selected,
- The reasons for ignoring certain others, and
- How the design realistically controls the critical variables without sacrificing the integrity of the study.

Where do you place this explanation of elimination of alternative explanations in the proposal? Usually, you will find it in the discussion of one of the links of the design, especially the comparison and contrast—the basis for sensing-attributes-and-changes link. But it can be covered anywhere it fits; the important thing is be sure it is included.

It is possible that you can do the study only in a laboratory-like situation if your design becomes sufficiently complex. This markedly reduces the generality of the findings. Such a consideration is obviously more of a worry in an applied or developmental study than in one dealing with basic research. But even in doing basic research, the need for generality may force consideration of other design choices.

Avoid expediency as a reason for failing to control a factor if reasonable effort and/or expense would permit doing so. For less critical variables, experienced faculty will recognize the reality of expediency as a good and sufficient basis.

Which alternative explanations are likely to be most troublesome varies with methodology and the study's circumstances. The most thorough delineation of alternative explanations has been in the context of experimental studies where they are termed threats to validity (see Campbell and Stanley, 1963; Cook and Campbell, 1979; Krathwohl, 1998/2004, pp. 526–531; Shadish, Cook, and Campbell, 2002; thirty-three of them are listed in Wortman, 1994). But there are also lists for qualitative studies (see Krathwohl, 1998/2004, pp. 317–320).

Some Illustrative Common Alternative Explanations to Be Eliminated or Controlled

For purposes of illustration, some examples that plague a variety of types of studies are briefly described below.

Reactivity. The effect of special attention is a reaction to the perception that there is something special about this situation. It usually elicits 'I'd better do what is right' or 'I'd better be good' behavior. One looks for reactive effects where obtrusiveness tips the situation from normal to special:

- The presence of an observer can change normal behavior—a teacher better controls her temper; the children are on their "good behavior."
- The treatment obtrusively stands out from the normal sequence of events—the experimental group is taken from the classroom to the computer cluster.
- Measurement of effect is obtrusive—students spot the video camera that is recording their use of reference books in the library.

Obviously, reactivity is eliminated or at least reduced when things proceed naturally, or as much so as possible. Concealing the observer by providing him a social role in the group being observed and allowing time for him to become a normal part of the situation may control for reactivity. Having the usual classroom teacher, social worker, or other professional administer a special treatment, instead of the researcher, may do so as well. Further, that person may be the best one to decide how and when to introduce a treatment into the situation. Where measurement is a problem, see Webb, Campbell, Schwartz, and Sechrest (1981) and Lee (2000), books on unobtrusive measurement and methods as ways of reducing reactive behavior.

Researcher Expectancy and Placebo Effects. A very closely related influence is the expectation of the researcher that influences result in the direction the researcher hopes to see. It refers both to the elicitation of such behavior from those studied as well as to faulty or self-deceiving perceptions by those recording the study's results. Researchers or their assistants may inadvertently tip the scales in favor of preferred results in a variety of ways: Participant observers may give inadvertent cues to desired behavior. Participants typically try to fathom the purpose of the study and give the responses they perceive as wanted. Ambiguous situations may be recorded as instances of the study's expected outcome. Errors in recording, observation, or measurement procedure may unintentionally favor the expected outcome (when totaling your checking account register, why do errors usually favor you instead of the bank?).

Use of "double-blind" procedures, where neither observer, measurer, nor subject knows the intended outcome of the study, eliminate expectancy effect. Treatments appear as identical as possible, but participants are coded so some uninvolved party can separate comparison groups after treatment. The control treatment is referred to as a "placebo" or "placebo treatment" after the inert pill that is used to mimic an experimental drug. Double-blind procedures cannot be used when: (1) the participant's knowledge of treatment is part of the treatment itself; (2) it is obvious which treatment is to be favored from merely ob-
serving the treatment or being exposed to it, (3) the treatment can be readily identified from side effects, or (4) withholding a more favorable treatment would have ethical consequences.

Selection and Mortality Effects. Selection and mortality are opposite sides of a coin. Selection adds an alternative explanation by affecting the composition of the group through the nature of the persons selected for study, mortality by those leaving the study as it progresses.

The alternative explanation comes about because the persons selected are distinguished from those not studied by a factor that may also cause the desired effect. If unrecognized, it can lead to the wrong conclusion, like assuming that, generally, bottles are discarded in the ocean with their caps on, because they predominate along the shoreline. Rather, the others sank—these are the survivors. Similarly, high school graduates and college and graduate students are “survivors.” Volunteering is a common selective factor. Alternative explanations arise from the fact that those who volunteer are different from non-volunteers (usually brighter, better educated, higher in social status, more sociable, have a higher need for social approval, etc. [Rosenthal and Rosnow, 1975]).

“Mortality” as an alternative explanation is not the death of an individual, but the change in group composition resulting from their leaving. “Leavers” depart for a reason—uncomfortable, bored, afraid to fail, etc. Because this modifies the average characteristics of the study group, their leaving should be noted and taken into account. Their leaving can be easily overlooked when concentrating on others in the group.

Instrument Decay. Changes in the measuring instrument over time might cause one to conclude an effect occurred when it was the recording standards that were changing. In qualitative studies, since the observer is the “instrument,” changes may occur as she becomes more familiar with those studied and/or the situation. Where measuring equipment is concerned, as in hearing or other discrimination tests, lack of calibration may cause the effect. Where essay tests are involved, the first ones graded may be held to a different standard than the last. Instrument decay can take many forms.

A General Strategy. The important thing is to be aware of possible alternative explanations, to describe the likely ones in the proposal, to tell how they will be handled, or if they won’t, why not. Show in the description how well you have analyzed the design and how familiar you are with the literature on this topic so that you have recognized and adequately taken into account the relevant threats to your study.

Design Efficiency

At some point, determine whether the design is maximally efficient. For example, can better use be made of participants or informants; data collection points be reduced; more data collected at each visit, measurement, or observation; persons be scheduled more efficiently; and so on? The development of the work plan, discussed in the next chapter, is especially helpful in showing where economies can be made in scheduling (your time is the major resource to allocate for a dissertation). If feasible, considerable savings in resources can result from combining your study with that of other researchers so as to use the same participants, situations, and/or data.

Give Special Care to Those Sections Critical to Your Research Method

As noted at the beginning of this section, the nature of your study affects what parts of the proposal are critical and therefore need special attention. But, regardless of the kind of study, the design aspects of the method section deserve special attention because within any research method, there are a variety of ways to proceed.

Choice of design is still an art. A design’s strengths in one aspect may result in a weakness elsewhere. Choice requires assessing the gains and losses involved in various alternatives. Unfortunately, they are rarely known accurately in advance; good estimates come from knowing one’s field and having worked with it long enough to have learned which options yield gains, which losses, and their frequency and seriousness. As a new researcher you may not have the experience to weigh all these variables as your chairperson and committee will, but use all the resources you can to develop the best possible design. Talk with other graduate students and particularly with those faculty members who frequently serve as design consultants. Then rely on your chairperson and committee to point out problems and solutions that you may have overlooked.

Because choice of design is an art, reasonable persons may differ as to the best design for a given problem. Your initial choice may not be that which springs to the mind of your chairperson or committee members. But they may be thinking in stereotypes, and your approach may indeed be best. Help readers follow your line of reasoning so that they, too, may see your design choice rationale—your reasons for so choosing and why this choice over alternatives. Creating a strong proposal is also a matter of knowing your audiences and being able to adequately anticipate and meet their concerns. There is more on this point in the material that follows.

SECTION 2: DEVELOPING THE SUBSECTIONS OF METHOD

The method section describes the structure of the investigation: the way participants or situations will be studied; how groups will be organized; if there is a treatment, when and how it will be administered; when observations will be made, of whom, when, and, if known, of what; the protection against alternative explanations; and the like. Begin the write-up of this section with a one-paragraph summary or overview of the method to be used.
Then, in whatever order seems most appropriate for what you plan to do, cover the six links of design in the chain of reasoning so as to describe these various aspects of method:

1. Participants—population and sample
2. Situation
3. Focus of action—the core variables such as treatment and effect,
4. Records—instrumentation and data collection,
5. Comparison and contrast (basis for sensing attributes and changes)—the basis on which the change due to an independent variable or experimental treatment or whatever happens at the focus of the study will be sensed (e.g., pre/post comparison or comparison with another group), and
6. Time schedule—the procedure.

Having described your data gathering plan, next describe your

• analysis plan and
• expected end product.

Although one can use the six links in design as an organizing framework, most proposals will not have a subsection for each of them. This is apparent in the annotated proposals in part 5 of this book, and they are typical. But the information describing all six links is somewhere accounted for. Be sure to adapt your proposal format to best describe your study. The following discussion specifies what is typically included to describe each link, examples of where this appears in the proposal, suggestions for writing it, and some of the common, and/or most serious, errors.

Participants—Population and Sample

For all studies involving gathering data from people, a description of who they are is essential to determining the potential generalizability of the study findings. The characteristics of the population to which the sample studied belongs define the group to whom the study’s results may transfer. Obviously, this generality should be consistent with the generality claimed in the problem statement and objectives sections. The representativeness of the sample indicates how confidently we can generalize from sample to population.

While random sampling provides on average a sample that is representative in every respect—even some characteristics we don’t care about, like length of one’s little toe—there is no guarantee that any given sample will be representative of those characteristics crucial to our study. Therefore, we often take steps to ensure that the sample is representative with respect to key variables. There are a number of ways of doing this such as stratified and cluster sampling with random selection within strata or clusters.

For studies concerned mainly with description, characterizing the nature of the participants allows readers to determine what, if any, parallels exist to their own experiences, thus allowing a determination of whether the results “ring true” and, if they do, what, if any, implications the study might have.

Therefore, regardless of how those studied were or will be selected, be sure to describe that process in detail, giving a rationale for why that process is the best of those available. If you seek findings that generalize, indicate the variables that will be used as the basis for ensuring representativeness—e.g., the basis of stratified and cluster sampling, the significance of those variables for the study, and why they were chosen over others. Indicate where the data on the variables used to stratify or cluster individuals will be obtained. If there is any reason to believe the database from which they are to be selected is not error free, give some indication of the anticipated error’s extent and its likely impact on the study.

To study a proposition that is presumed to be universally applicable, you can use anyone or any situation except where the choice of participant or situation might favor or disfavor it. Any random sample of the world’s population will do. We often substitute a convenience sample such as graduate or undergraduate students for such an unbiased sample. But if one or more characteristics of university students would normally be expected to affect the study’s outcome, you must explain why you believe this will not be so for your study. It is important to anticipate such concerns.

Sample size is another important decision. Giving a good rationale is more impressive than picking an arbitrary number or using whatever size convenience sample is available. Power analyses provide such a rationale by providing a design basis such that if the expected result does appear, the study will be sensitive enough to show it as statistically significant. Increasingly, studies intended for publication must be designed using such analyses. They require making some decisions about:

1. How precise must the estimate be? Put another way, how small a difference is to be sensed? Other things being equal, the greater the precision required and the smaller the difference to be sensed, the larger the sample required.

2. How different are individuals with respect to the characteristic being estimated; how much variability is there? If everybody is about the same, other things being equal, you can estimate from a few cases. But if people differ greatly, that is, there is high variability from person to person, more cases will be needed.

3. How much certainty is required of the estimate? This is another way of asking whether you want to use the 1 percent level of significance, 5 percent level, 10 percent level, and so on. At the 5 percent level, your confidence that the
population value is bracketed by the confidence interval is expressed by odds of 19 to 1. Again, other things being equal, the greater the certainty required (e.g., the 1 percent rather than 5 percent level, the smaller the confidence interval), the larger the sample required.

Where does the information come from to determine sample size? For questions 1 and 2 above, from your pilot studies, from other researchers’ use of the same instruments with comparable participants, or, failing these, from “guesstimates” made on the best basis you can command. With such estimates and a decision on question 3, any good statistics book or Cohen (1988) or Lipsey (1989) will show how to calculate a sample size such that if an event of interest occurs, the odds heavily favor that it will be statistically significant. For an interactive statistical power analysis site on the Web, try http://www.stat.uiowa.edu/~rlenth/Power/index.html (accessed October 1, 2004). To create the tilde (~), press “Shift” and the key to the left of “1.” Alternatively, after “link” copy and paste its URL into Google.com to find similar sites (e.g., linkwww.stat.uiowa.edu/~rlenth/Power/index.html).

Occasionally a student will propose using the total population, which, though large, is presumably manageable. In such instances, even though feasible, it may be preferable to work more intensively or carefully with a sample than to use the same resources trying to cover the entire population. Indeed, if a power analysis indicates a sample instead of a census can satisfactorily be employed, the resources required to canvass an entire population when concentrated on a sample may result in better and deeper information or the same information obtained more cheaply. However, if your research is intended to convince lay policy makers, there may be no substitute for a census. How best to employ your resources is determined by the sophistication of your intended audience and the purposes of your study.

Situation

In many instances, the situation or the setting in which you will gather your data is determined by the sample, so one has already described it in the previous section. But where that is not the case, such description indicates where the design will be implemented. It helps readers determine the possible applicability of findings to comparable situations. Though typically covered in the population and sample section, description of the situation may be covered in other sections of the proposal.

See for example, Beissner’s paragraph 48, which is in the “Procedure” section.

Focus of Action—Treatment(s), Independent and Dependent Variable(s)

Here is where you describe what it is you are studying—the effect of one or more treatments, the effect of one or more independent variables, or whatever one is focusing attention on, such as what results when certain conditions occur.

Except in emergent studies, description of the treatment and variables is most often put in cause and effect terms where the causes are treatments or independent variables and the effects are the dependent variables—that is, they are dependent upon the presence, and often the strength, of the independent variable. In some instances, this information is covered in the instrumentation or measurement section.

Beissner, for example, describes the independent (paragraphs 40–47) and dependent (paragraphs 28–39) variables within the section titled “Instruments.” Note especially in Beissner that she also describes independent variables that might cause the same effects as the treatment—level of factual knowledge, critical thinking ability, and the processes by which one relates new knowledge to old (measured by her “Inventory of Learning Processes”). If one is to claim an effect has a particular cause, one must eliminate such alternative causes as plausible.

In an experimental study, one normally finds a careful description of the treatment.

Beissner assumes the readers are familiar with concept mapping and with Novak and Gowin (1984) (see her paragraph 14) and so describes only the scores that will be produced (see paragraph 46, the section “Instruments”). In contrast, although omitted in this cut-down version, Phelan provided detail regarding his treatment—a workshop on self-directed learning—in an appendix (see paragraph 26). Placement in the appendix is a common practice.

So it is clear, the specifications of the treatment and variables in the study, the focuses of action, need to be explicitly spelled out somewhere in the proposal, but they need not have a section of their own. They are frequently specified and described in the course of completing other sections of the proposal.

Records—Instrumentation and Observations

In this section, records—the measures and observations—to be made in gathering data should be detailed and their appropriateness for the task convincingly described. Instruments may be unnecessary in a case study with few individuals. They may be inappropriate for exploratory and emergent studies, where to start by using instruments would presuppose you already knew what you were seeking to study. But they are both appropriate and necessary for the many studies that are confirmatory in nature, highly structured in their approach, concerned with cross-case comparisons (individuals, programs, sites, etc.), or combinations of structured and/or exploratory-confirmatory designs.

In all of these instances, some forethought about instrumentation at the pro-
propositional stage will help reduce data collection and analysis problems and facilitate and enhance comparisons within cases and across them.

Instrumentation may range from those with very light structure—categories of behavior or phenomena to count—through increasingly structured data gathering: observation scales, rating scales, interview guides, interview schedules, conventional questionnaires, computer-adapted branched questionnaires, individually administered tests, and group tests. Which of these, if any, are appropriate will depend on the study, what is already available, and the trade-off of expending time and energy early in instrument construction in order to save time and energy at the analysis end. Whatever your choice, indicate it and describe its supporting rationale in the proposal.

The Observer as Instrument

As noted under "Instrument Decay," studies using observation have their own set of problems. The discussion of them and safeguards against them will typically appear under a section title like "Research Method" rather than "Instrumentation." This is discussed more extensively in part 3.

Measures

Some variables—time or distance measures, for instance—present little problem. But most studies in the behavioral sciences involve constructs that must be translated into behaviors that can be sensed in order for us to assess them. When we meet people, we cannot directly sense their intelligence, for instance, but we judge it by their behavior. Sometimes we do this by exposing them to a standard set of problems, an intelligence test. This permits comparison of their behavior with other persons on a common scale. Psychological, sociological, and economic constructs such as anxiety, socioeconomic class, and marginal utility require interpretation into characteristics that can be sensed and measured. The instrumentation section is where that translation is described and the case made for its adequacy.

This is another section in which the expansive rhetoric of the problem description may be reduced to mundane terms when the reader sees what the problem has become in measurement terms. If the realities of measurement are modest, keep the early rhetoric modest too.

Often the translation process helps to sharpen your understanding of the study’s constructs as you are forced to choose among alternatives that represent different operational definitions.

"Anxiety" may be undefined in the hypothesis, but one will find many possibilities when one comes to choosing a measure. They range from self-report of one’s "state" of anxiety to self-perceptions of it as a persistent "trait" to physiological measures (galvanic skin response or heart and breathing rate). Are these interchangeable definitions of the same characteristic? The problem definition and explanation should provide sufficient guidance to choose among the possibilities; if they do not, they need further refining.

When it is impossible to find totally satisfactory measures, describe the problem and justify, as well as possible, the measure that comes closest, indicating why it will be adequate for your study.

Be sure that all the terms critical to the questions, hypotheses, or models are discussed in this section. Variables mentioned earlier and then dropped leave ends dangling that are sure to be noticed by your chairperson or committee. And it is hoped that they do, since attending to them at the proposal stage may save you a real crisis later. In addition, dropping variables leaves the impression that you are not paying sufficient attention to important details and suggests there may be other carelessness in the proposal.

The interpretability of commonly used instruments may be well established for the purposes you intend. For new or experimental tests, however, your audience will expect empirical evidence of the test’s quality and meaning. If it is not available from use of the test by others, make provisions in the proposal for establishing that the test has appropriate characteristics (if possible, before the data are collected). For an example, see Beissner’s paragraph 39, regarding her plans to gather evidence of validity for the test she developed. Here, as elsewhere, do not assume that the reader will rush to the library to look up missing reliability and validity information. If there is any doubt that the reader is likely to know it, supply it.

Validity. Construct validity provides evidence that forms the basis for intended score interpretation and serves as a unifying framework for other validity evidence. Evidence based on relations with other variables shows, for instance, that the test correlates with an already accepted measure of the variable. It correlates with measures it ought to be related to and does not correlate with those it should not. Look at an example in paragraphs 28–31 and 45 of Beissner’s proposal.

Validity evidence based on content, also called "content" or "curricular validity," provides a comparison of the test items with specifications of what subject matter content and skills the test is supposed to cover. Predictive and concurrent validity evidence shows the measure predicts or is correlated with work or academic performance. Evidence of validity is usually found in a test’s manual or in such references as Buros’s Mental Measurements Yearbooks. Cite evidence for those kinds of validity needed for the problem posed.

Although not usually considered part of construct validity, “face validity”—that the test looks as though it measured what it was intended to measure—is very important when the study’s acceptance is determined by policy makers, parents, and others with little or no professional background. Treat it in the validity discussion if it is likely to be a factor in your study.
Reliability. Just as there are various kinds of evidence for validity, there are also for reliability: stability reliability—the test scores are stable over time; internal consistency reliability (homogeneity)—the various test items measure the same characteristic so the scores are interpretable; and equivalence reliability—different test forms are comparable. Which reliabilities are required depends on the design. For example, test results compared over a substantial period of time require evidence of stability reliability and internal consistency reliability (for an example, see paragraphs 41 and 44 of Beissner’s proposal). If the retest used a different form of the same test, equivalence reliability would be required as well. Again, such evidence is usually given in a test’s manual.

Objectivity. Observation scales, in particular, require that all observers in a study use them the same way so that they agree when rating the same phenomenon; this is objectivity. Observers often train by rating the same videotapes, continuing until all observers respond to events the same way. Describe any planned training and what level of agreement among observers will be sought. Remember that a correlation coefficient will show agreement on relative but not exact position on the score scale; it does not detect that one person is a tougher grader than another, for instance. Use the intraclass correlation to show exact correspondence of judgment.

Objectivity is also a problem for multiple raters of essay or similar material.

Beissner’s scoring of concept maps is an instance of this; see her paragraph 47 and the related annotation.

Sources of Instruments. If you are looking for available instrumentation, be sure to use the considerable resources for finding both established and experimental ones. At one time the sole source of information about tests was Oscar Buros’s Mental Measurements Yearbook, published at irregular intervals since 1938. Information on commercially available tests is available from the Buros Web site, http://www.unl.edu/buros (accessed October 1, 2004); reviews on the most heavily used tests can be downloaded for a fee. In addition, there are now a number of compilations of instruments (Backer, 1977; Goldman and Mitchell, 1995–2003; Fabiano and O’Brien, 1987; Educational Testing [ETS] Service’s TestLink—http://www.ets.org/testcoll/index.html [accessed October 1, 2004]).

The Internet is continually changing, so check for new sources, but these sources, which include ETS TestLink’s more than twenty thousand tests, should go a long way toward pointing you in possible directions. Pursue them in databases like the Social Science Citation Index, PsychINFO, or SociFile. In these, one may be able to find instances where a specific instrument has been used, its strengths and weaknesses noted, and sometimes an improved version.

Constructing and validating new instruments is both difficult and expensive. Established instruments may not be quite as close to the desired operational definition as new or experimental ones, but usually are better validated and more easily and widely understood by one’s audience. However, a specially designed instrument may result in a more on-target study. Which to choose? Consider such factors as: How much difference in validity would be gained with new construction? What are the odds the construction effort will be successful? How feasible is it? Must the results be accepted by a lay audience who might better accept the established instrument? Can data be obtained that document the new instrument to your audience’s satisfaction?

If you lean toward constructing a new instrument, consider it carefully with your chairperson and committee. Developing an instrument can be a dissertation in itself. They won’t want you to overcommit yourself by undertaking more than is reasonable in a dissertation, any more than you do. But they may have a better idea of the time and effort involved. So, lay it all out for them to consider. (Remember as indicated at the outset, approval of your proposal is a “shared decision-making situation”—see chapter 1, pp. 3–5. If given a green light, describe how the test will be developed and lay out a development plan. Display sample items in an appendix. Such plans will be found in many measurement books (e.g., Gronlund, 2001; Hopkins, 1998).

Comparison and Contrast—The Basis for Sensing Attributes or Changes

This link in the design serves two purposes that are most easily seen in experimental studies. The first purpose is providing a basis on which one can say that a treatment had an effect. This might be by comparison with an untreated experimental group as in the Beissner study. (Beissner describes the difference in treatment of the two groups in her “Procedure” section, paragraphs 48 and 49, and examines the effect of the treatment in her “Data Analysis” section, paragraph 50.) As with previous links, in those sections of the proposal where it is relevant—the procedure section being a common one—show how you will sense attributes or changes.

The second purpose is the elimination of alternative explanations, as discussed earlier in this chapter.

Beissner eliminates the alternative explanations of differences in prior knowledge, in critical thinking ability, and in learning style. The assessment of these variables is noted in Beissner’s “Instrument” section (paragraph 41 and 43) and their effect in the “Analysis” section (paragraph 50).

Although it is less obvious in nonexperimental research studies that seek explanations and generalizable findings, this link serves the same two purposes: noting attributes and changes and protecting against alternative explanations.
For instance, Warters is seeking attributes of effective programs. In paragraph 17, he notes how, using theory-based sampling procedures, he will select individuals for interview from three or four different treatment programs to allow "the opportunity for some comparison of treatment modalities" and to identify "specific aspects of group process thought useful by the men themselves in eliminating abuse."

In terms of protecting against alternative explanations, Warters notes in paragraph 31, that he is "currently functioning as an advocate of social intervention to reduce men's domestic violence," and that "This perspective will most certainly affect my interpretation of events and discussion during the course of my study." He then indicates in paragraph 32 how he hopes to overcome this concern on the part of his audience.

So, regardless of study method, if one is advancing a generalizable explanation, somewhere in the description of design, one needs to note how these two functions of this link will be attended.

**Time Schedule—The Specification of the Procedure**

The description of the procedure is a narration of the plan for data collection over time. It indicates what observations or measurements will be made, when, where, and of whom, and, if there is a treatment, how, where, when, and to whom it will be administered. It is usually the place where a reader can get the clearest idea of exactly what you plan to do and how and when you plan to do it. Usually, it is a verbal account, with the actual schedule with dates, which is described in the next chapter, forming a later section. It often contains a graphic of the work plan. Obviously, the procedural narrative and activities in the work plan or schedule should be coordinated so they tell the same story. Because of this relationship, you may wish to lay out the work plan or schedule section first and then describe it in the method or procedure section. Alternatively, constructing your work plan or schedule from the procedure section provides a test of its adequacy.

Usually the account is labeled "Procedure," as it is in Beissner, where it begins with paragraph 48. But it sometimes appears, as in Warters, under the heading "Methods," where it starts with paragraph 11, or "Research Methods," as in Phelan, where it starts with paragraph 10. In all three instances, readers can get quite a clear picture of what the researcher plans to do from these sections.

**Problems in Data Collection**

Indicate your provisions for handling potential problems that may arise in the course of gathering data.

Beissner, for instance, fearing that students may communicate so that the control group may learn how to do concept mapping, indicates that "the participants will be advised not to discuss the content of their sessions with other study participants" (paragraph 49).
Analysis

The method of analysis must be consistent with the objectives and design. For instance, when the study calls for finding the extent of a relationship, some measure of the size of relations such as a correlation coefficient is in order. Too often, we find instead a contrast of high with low groups to compute a difference statistic such as a t test. A statistically significant t test would indicate that had a correlation been computed, it would be significantly different from zero. But the extent of statistical significance does not indicate the size of the relationship; it could be so low as to be, practically speaking, insignificant. That a difference is statistically significant at some extremely small percentage level may be testimony more to the statistical power of the study than to the strength of the relationship. Without knowing the size of the correlation, you don’t know whether the relationship is strong enough to permit any kind of reasonable prediction and, therefore, any practical application.

The assumptions of the statistics should fit the data. If they seem not to, describe the corrections that can be made. For instance, analysis of variance assumes normally distributed populations, but corrections in the level of significance can easily be made for nonnormal data.

A description of how missing data and/or unequal cell frequencies of a complex design are to be handled displays a sophistication that is comforting to the reader.

When new statistical techniques, computer programs, or other unfamiliar analytic tools are to be used, adequately describe them and show their advantages over current methods so the reader may be assured of their appropriateness.

It is not always possible to completely anticipate in advance the nature of the analysis that will be called for; it may depend on the nature of the data collected. This is especially true of content analysis procedures, but it may also be true of statistical methods. As is probably obvious by now, the best strategy is to reveal the depths to which these problems have been anticipated and describe the projected solution in sufficient detail as to clearly convey its nature. At the same time, show awareness of where anticipated departures from plan may occur. Before leaving this section, check to make sure the analytic procedures will handle all the relevant data that will be gathered and will yield evidence bearing on all the questions, hypotheses, and model aspects proposed for investigation.

Expected End Product

This section will not appear in all proposals, but is a good section to include if there are products in addition to the usual report of results or if the report of results is other than routine. Tests, evaluation instruments, curriculum materials, videotapes, audiotapes, films, pamphlets, and the like, even though they may be but by-products of the project, are sometimes more important and more en-

One possible very important end product is your dissertation formatted as an article ready to publish! Read how next to achieve that.

AN ALTERNATIVE DISSERTATION FORMAT: ARTICLES READY FOR PUBLICATION

If you are serious about obtaining a position in higher education, consider writing your dissertation as an article ready for publication in an appropriate journal. Nearly all institutions will allow this, even though it is a rarely chosen route in the social and behavioral sciences. Some encourage it; it is more common in the sciences than the social sciences, but ought to become more common over time. See Krathwohl (1994) and Duke and Beck (1999).

Having the dissertation ready to send off for publication will net you an early career publication and save you from trying to cut your dissertation to journal size at the same time that you are adjusting to your postdoctoral situation. After all, a new job, learning the ropes in a new institution, planning courses and writing teaching materials, starting a research program, and learning how to advise students, to say nothing of all the non-higher education problems of relocating, provide a full plate. You don’t need, in addition, a dissertation that you may possibly have grown tired of hanging around your neck.

Publication will require that you use the format of the particular journal to which you intend to submit. The length of their typical article will unquestionably be less than dissertation length, requiring a significant departure from dissertation format. Depending on your chairperson and committee, and the typical procedure at your institution, you may be required to put considerable work in an appendix so that your committee and the external readers at the final oral examination can understand what you did. The literature review may need to be prepared as a separate article for a different journal than the main body of the study.

If you plan to use the dissertation-as-article format, share this idea early with your chairperson and committee. Be sure they will cooperate with you in this endeavor. Then make your intent regarding writing format a part of the proposal compact. This makes clear to everyone what to expect.

If there is likely to be any question regarding “how much is enough,” that is, when your committee is willing to agree that you have fulfilled your “contract,” read the last section of chapter 7. It is discussed there because it is a common problem with qualitative proposals, but can be for other kinds of studies as well.

Finally, since the description of your study methods may be the most closely read section of your proposal, it is important to produce a carefully thought-out statement. Use Worksheet 5.1: Study Methods Review to assess and improve your methods statement.
**WORKSHEET 5.1**

**Study Methods Review**

*How Strong Are My Study’s Methods?*

As you develop your proposal, periodically review how strong your study's methods are, and where they need to be strengthened.

<table>
<thead>
<tr>
<th>In Describing My Study’s Methods Overall, How Well Have I...?</th>
<th>Strong</th>
<th>Acceptable</th>
<th>Weak Improvements Needed</th>
<th>Not Applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selected procedures that address the expectations created in prior chapters?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adapted my procedures statement to the specific methods of my study?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clearly shown the connection between my method and the question, hypotheses, or models I have chosen to study?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constructed a design that is relevant, feasible, and internally/coherent operationalization of the concepts being examined?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appropriately considered resource limits, institutional limits, ethical limits, and time limits?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identified and eliminated or controlled each major alternative explanation of the results I might obtain?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anticipated and accommodated possible study problems?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outlined appropriate analysis procedures?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued on next page)