



Quantitative Methods

We turn now from the introduction, the purpose, and the questions and hypotheses to the methods section of a proposal. This chapter presents essential steps in designing quantitative methods for a research proposal or study, with specific focus on survey and experimental designs. These designs reflect postpositivist philosophical assumptions, as discussed in [Chapter 1](#). For example, determinism suggests that examining the relationships between and among variables is central to answering questions and hypotheses through surveys and experiments. The reduction to a parsimonious set of variables, tightly controlled through design or statistical analysis, provides measures or observations for testing a theory. Objective data result from empirical observations and measures. Validity and reliability of scores on instruments lead to meaningful interpretations of data.

In relating these assumptions and the procedures that implement them, this discussion does not exhaustively treat quantitative research methods, such as correlational and causal comparative approaches so that the focus can be on surveys and experiments. Excellent, detailed texts provide information about survey research (e.g., see Babbie, 2007; Creswell, 2012; Fink, 2002; Salant & Dillman, 1994). For experimental procedures, some traditional books (e.g., Campbell & Stanley, 1963; Cook & Campbell, 1979), as well as some newer texts, extend the ideas presented here (e.g., Boruch, 1998; Field & Hole, 2003; Keppel & Wickens, 2003; Lipsey, 1990; Thompson, 2006). In this chapter, the focus is on the essential components of a method section in proposals for a survey and an experiment.

DEFINING SURVEYS AND EXPERIMENTS

A **survey design** provides a quantitative or numeric description of trends, attitudes, or opinions of a population by studying a sample of that population. From sample results, the researcher generalizes or draws inferences to the population. In an *experiment*, investigators may also identify a sample and generalize to a population; however, the basic intent of an **experimental design** is to test the impact of a treatment (or an intervention) on an outcome, controlling for all other factors that might influence that outcome. As one form of control, researchers randomly assign individuals to groups. When one group receives a treatment and the other group does not, the experimenter can isolate whether it is the treatment and not other factors that influence the outcome.

COMPONENTS OF A SURVEY METHOD PLAN

The design of a survey method section follows a standard format. Numerous examples of this format appear in scholarly journals, and these examples provide useful models. The following sections detail typical components. In preparing to design these components into a proposal, consider the questions on the checklist shown in [Table 8.1](#) as a general guide.

Table 8.1 A Checklist of Questions for Designing a Survey Method

	Is the purpose of a survey design stated?
	Are the reasons for choosing the design mentioned?
	Is the nature of the survey (cross-sectional vs. longitudinal) identified?
	Is the population and its size mentioned?
	Will the population be stratified? If so, how?
	How many people will be in the sample? On what basis was this size chosen?
	What will be the procedure for sampling these individuals (e.g., random, nonrandom)?
	What instrument will be used in the survey? Who developed the instrument?
	What are the content areas addressed in the survey? The scales?
	What procedure will be used to pilot or field-test the survey?
	What is the timeline for administering the survey?
	What are the variables in the study?
	How do these variables cross-reference with the research questions and items on the survey?
	What specific steps will be taken in data analysis to do the following:
(a) _____	Analyze returns?
(b) _____	Check for response bias?
(c) _____	Conduct a descriptive analysis?
(d) _____	Collapse items into scales?
(e) _____	Check for reliability of scales?
(f) _____	Run inferential statistics to answer the research questions or assess practical implications of the results?
	How will the results be interpreted?

The Survey Design

In a proposal or plan, the first parts of the method section can introduce readers to the basic purpose and rationale for survey research. Begin the discussion by reviewing the purpose of a survey and the rationale for its selection for the proposed study. This discussion can do the following:

- Identify the purpose of survey research. This purpose is to generalize from a sample to a population so that inferences can be made about some characteristic, attitude, or behavior of this population. Provide a reference to this purpose from one of the survey method texts (several are identified in this chapter).
- Indicate why a survey is the preferred type of data collection procedure for the study. In this rationale, consider the advantages of survey designs, such as the economy of the design and the rapid turnaround in data collection. Discuss the advantage of identifying attributes of a large population from a small group of individuals (Fowler, 2009).
- Indicate whether the survey will be cross-sectional—with the data collected at one point in time—or whether it will be longitudinal—with data collected over time.
- Specify the form of data collection. Fowler (2009) identified the following types: mail, telephone, the Internet, personal interviews, or group administration (see also Fink, 2012; Krueger & Casey, 2009). Using an Internet survey and administering it online has been discussed extensively in the literature (Nesbary, 2000; Sue & Ritter, 2012). Regardless of the form of data collection, provide a rationale for the procedure, using arguments based on its strengths and weaknesses, costs, data availability, and convenience.

The Population and Sample

In the methods section, follow the type of design with characteristics of the population and the sampling procedure. Methodologists have written excellent discussions about the underlying logic of sampling theory (e.g., Babbie, 2007; Fowler, 2009). Here are essential aspects of the population and sample to describe in a research plan:

- Identify the population in the study. Also state the size of this population, if size can be determined, and the means of identifying individuals in the population. Questions of access arise here, and the researcher might refer to availability of sampling frames—mail or published lists—of potential respondents in the population.
- Identify whether the sampling design for this population is single stage or multistage (called clustering). Cluster sampling is ideal when it is impossible or impractical to compile a list of the elements composing the population (Babbie, 2007). A single-stage sampling procedure is one in which the researcher has access to names in the population and can sample the people (or other elements) directly. In a multistage or clustering procedure, the researcher first identifies clusters (groups or organizations), obtains names of individuals within those clusters, and then samples within them.
- Identify the selection process for individuals. I recommend selecting a *random sample*, in which each individual in the population has an equal probability of being selected (a systematic or probabilistic sample). With randomization, a representative sample from a population provides the ability to generalize to a population. If the list of individuals is long, drawing a random sample may be difficult. Alternatively, a *systematic sample* can have precision equivalent **random sampling** (Fowler, 2009). In this approach, the researcher chooses a random start on a list and selects every X numbered people on the list. The X number is based on a fraction determined by the number of people on a list and the number that are to be selected on the list (e.g., 1 out of every 80th person). Finally, less desirable is a nonprobability sample (or *convenience sample*), in which respondents are chosen based on their convenience and availability.
- Identify whether the study will involve *stratification* of the population before selecting the sample. This requires that characteristics of the population members be known so that the population can be stratified first before selecting the sample (Fowler, 2009). Stratification means that specific characteristics of individuals (e.g., gender—females and males) are represented in the sample and the sample reflects the true proportion in the population of individuals with certain characteristics. When randomly selecting people from a population, these characteristics may or may not be present in the sample in the same proportions as in the population; stratification ensures their representation. Also identify the characteristics used in stratifying the population (e.g., gender, income levels, education). Within each stratum, identify whether the sample contains individuals with the characteristic in the same proportion as the characteristic appears in the entire population.
- Discuss the procedures for selecting the sample from available lists. The most rigorous method for selecting the sample is to choose individuals using a random sampling, a topic discussed in many introductory statistics texts (e.g., Gravetter & Wallnau, 2009).
- Indicate the number of people in the sample and the procedures used to compute this number. In

survey research, investigators often choose a sample size based on selecting a fraction of the population (say, 10%), select the size that is unusual or typical based on past studies, or base the sample size simply on the margin of error they are willing to tolerate. Instead, Fowler (2009) suggested that these approaches are all misguided. Instead, he recommended that sample size determination relates to the analysis plan for a study. One needs to first determine the subgroups to be analyzed in study. Then, he suggested going to a table found in many survey books (see Fowler, 2009) to look up the appropriate sample size. These tables require three elements. First, determine the margin of error you are willing to tolerate (say $\pm 4\%$ **confidence interval**). This is a + or – figure that represents how accurate the answers given by your sample correlate to answers given by the entire population. Second, determine the confidence level for this margin of error (say 95 out of 100 times, or a 5% chance). Third, estimate the percentage of your sample that will respond in a given way (50% with 50/50 being the most conservative because people could respond either way). From here, you can then determine the sample size needed for each group. Using Fowler's (2009) table, for example, with a margin of error of $\pm 4\%$, a confidence error of 95%, and a 50/50 chance that the sample contains our characteristic, we arrive at a sample size of 500.

Instrumentation

As part of rigorous data collection, the proposal developer also provides detailed information about the actual survey instrument to be used in the proposed study. Consider the following:

- Name the survey instrument used to collect data. Discuss whether it is an instrument designed for this research, a modified instrument, or an intact instrument developed by someone else. If it is a modified instrument, indicate whether the developer has provided appropriate permission to use it. In some survey projects, the researcher assembles an instrument from components of several instruments. Again, permission to use any part of other instruments needs to be obtained. In addition, instruments are being increasingly designed through online surveys products (see Sue & Ritter, 2012, for a discussion of products such as Survey Monkey and Zoomerang and important criteria to consider when choosing software and a survey host). Using products such as these, researchers can create their own surveys quickly using custom templates and post them on websites or e-mail them for participants to complete. The software program then can generate results and report them back to the researcher as descriptive statistics or as graphed information. The results can be downloaded into a spreadsheet or a database for further analysis.

- To use an existing instrument, describe the established validity of scores obtained from past use of the instrument. This means reporting efforts by authors to establish **validity in quantitative research**—whether one can draw meaningful and useful inferences from scores on the instruments. The three traditional forms of validity to look for are (a) content validity (do the items measure the content they were intended to measure?), (b) predictive or concurrent validity (do scores predict a criterion measure? Do results correlate with other results?), and (c) **construct validity** (do items measure hypothetical constructs or concepts?). In more recent studies, construct validity has become the overriding objective in validity, and it has focused on whether the scores serve a useful purpose and have positive consequences when they are used in practice (Humbley & Zumbo, 1996). Establishing the validity of the scores in a survey helps to identify whether an instrument might be a good one to use in survey research. This form of validity is different than identifying the threats to validity in experimental research, as discussed later in this chapter.

- Also mention whether scores resulting from past use of the instrument demonstrate **reliability**. Look for whether authors report measures of internal consistency (Are the items' responses consistent across constructs?) and test-retest correlations (Are scores stable over time when the instrument is administered a second time?). Also determine whether there was consistency in test administration and scoring (Were errors caused by carelessness in administration or scoring? See Borg & Gall, 2006).

- When one modifies an instrument or combines instruments in a study, the original validity and reliability may not hold for the new instrument, and it becomes important to reestablish validity and reliability during data analysis.

- Include sample items from the instrument so that readers can see the actual items used. In an appendix to the proposal, attach sample items or the entire instrument.

- Indicate the major content sections in the instrument, such as the cover letter (Dillman, 2007, provides a useful list of items to include in cover letters), the items (e.g., demographics, attitudinal

items, behavioral items, factual items), and the closing instructions. Also mention the type of scales used to measure the items on the instrument, such as continuous scales (e.g., *strongly agree* to *strongly disagree*) and categorical scales (e.g., yes/no, rank from highest to lowest importance).

- Discuss plans for pilot testing or field-testing the survey and provide a rationale for these plans. This testing is important to establish the content validity of scores on an instrument and to improve questions, format, and scales. Indicate the number of people who will test the instrument and the plans to incorporate their comments into final instrument revisions.

- For a mailed survey, identify steps for administering the survey and for following up to ensure a high response rate. Salant and Dillman (1994) suggested a four-phase administration process (see Dillman, 2007, for a similar three-phase process). The first mail-out is a short advance-notice letter to all members of the sample, and the second mail-out is the actual mail survey, distributed about 1 week after the advance-notice letter. The third mail-out consists of a postcard follow-up sent to all members of the sample 4 to 8 days after the initial questionnaire. The fourth mail-out, sent to all nonrespondents, consists of a personalized cover letter with a handwritten signature, the questionnaire, and a pread-dressed return envelope with postage. Researchers send this fourth mail-out 3 weeks after the second mail-out. Thus, in total, the researcher concludes the administration period 4 weeks after its start, providing the returns meet project objectives.

Variables in the Study

Although readers of a proposal learn about the variables in purpose statements and research questions/hypotheses sections, it is useful in the method section to relate the variables to the specific questions or hypotheses on the instrument. One technique is to relate the variables, the research questions or hypotheses, and sample items on the survey instrument so that a reader can easily determine how the data collection connects to the variables and questions/hypotheses. Plan to include a table and a discussion that cross-reference the variables, the questions or hypotheses, and specific survey items. This procedure is especially helpful in dissertations in which investigators test large-scale models. Table 8.2 illustrates such a table using hypotheticalal data.

Table 8.2 Variables, Research Questions, and Items on a Survey

Variable Name	Research Question	Item on Survey
Independent Variable 1: Prior publications	Descriptive research Question 1: How many publications did the faculty member produce prior to receipt of the doctorate?	See Questions 11, 12, 13, 14, and 15: publication counts for journal articles, books, conference papers, book chapters published before receiving the doctorate
Dependent Variable 1: Grants funded	Descriptive research Question 2: How many grants has the faculty member received in the past 3 years?	See Questions 16, 17, and 18: grants from foundations, federal grants, state grants
Control Variable 1: Tenure status	Descriptive research Question 3: Is the faculty member tenured?	See Question 19: tenured (yes/no)
Relating the Independent Variable 1: Prior publications to the Dependent Variable: Grants funded	Inferential Question 4: Does prior productivity influence the number of grants received?	See Questions 11,12,13,14,15 to Questions 16, 17, 18

Data Analysis and Interpretation

In the proposal, present information about the steps involved in analyzing the data. I recommend the following research tip—presenting them as a series of steps so that a reader can see how one step leads to another for a complete discussion of the data analysis procedures.

Step 1. Report information about the number of members of the sample who did and did not return the survey. A table with numbers and percentages describing respondents and nonrespondents is a useful tool to present this information.

Step 2. Discuss the method by which **response bias** will be determined. Response bias is the effect of nonresponses on survey estimates (Fowler, 2009). *Bias* means that if nonrespondents had responded, their responses would have substantially changed the overall results. Mention the procedures used to check for response bias, such as wave analysis or a respondent/nonrespondent analysis. In wave analysis, the researcher examines returns on select items week by week to determine if average responses change (Leslie, 1972). Based on the assumption that those who return surveys in the final weeks of the response period are nearly all nonrespondents, if the responses begin to change, a potential exists for response bias. An alternative check for response bias is to contact a few nonrespondents by phone and determine if their responses differ substantially from respondents. This constitutes a respondent-nonrespondent check for response bias.

Step 3. Discuss a plan to provide a **descriptive analysis** of data for all independent and dependent variables in the study. This analysis should indicate the means, standard deviations, and range of scores for these variables. In some quantitative projects, the analysis stops here with descriptive analysis, especially if the number of participants is too small for more advanced, inferential analysis.

Step 4. Assuming that you proceed beyond descriptive approaches, if the proposal contains an instrument with scales or a plan to develop scales (combining items into scales), identify the statistical procedure (i.e., factor analysis) for accomplishing this. Also mention reliability checks for the internal consistency of the scales (i.e., the Cronbach alpha statistic).

Step 5. Identify the statistics and the statistical computer program for testing the major inferential research questions or hypotheses in the proposed study. The **inferential questions or hypotheses** relate variables or compare groups in terms of variables so that inferences can be drawn from the sample to a population. Provide a rationale for the choice of statistical test and mention the assumptions associated with the statistic. As shown in [Table 8.3](#), base this choice on the nature of the research question (e.g., relating variables or comparing groups as the most popular), the number of independent and dependent variables, and the number of variables controlled (e.g., see Rudestam & Newton, 2007). Further, consider whether the variables will be measured on an instrument as a continuous score (e.g., age from 18 to 36) or as a categorical score (e.g., women = 1, men = 2). Finally, consider whether the scores from the sample might be normally distributed in a bell-shaped curve if plotted out on a graph or non-normally distributed. There are additional ways to determine if the scores are normally distributed (see Creswell, 2012). These factors, in combination, enable a researcher to determine what statistical test will be suited for answering the research question or hypothesis. In [Table 8.3](#), I show how the factors, in combination, lead to the selection of a number of common statistical tests. For further types of statistical tests, readers are referred to statistics methods books, such as Gravetter and Wallnau (2009).

Step 6. A final step in the data analysis is to present the results in tables or figures and interpret the results from the statistical test. An **interpretation in quantitative research** means that the researcher draws conclusions from the results for the research questions, hypotheses, and the larger meaning of the results. This interpretation involves several steps.

Table 8.3 Criteria for Choosing Select Statistical Tests

Nature of Question	Number of Independent Variables	Number of Dependent Variables	Number of Control Variables (covariates)	Type of Score Independent/ Dependent Variables	Distribution of Scores	Statistical Test	What the Test Yields
Group comparison	1	1	0	Categorical/ continuous	Normal	t-test	A comparison of two groups in terms of outcomes
Group comparison	1 or more	1	0	Categorical/ continuous	Normal	Analysis of variance	A comparison of more than two groups in terms of outcomes
Group comparison	1 or more	1	1	Categorical/ continuous	Normal	Analysis of covariance (ANCOVA)	A comparison of more than two groups in terms of outcomes, controlling for covariates
Association between groups	1	1	0	Categorical/ categorical	Non-normal	Chi-square	An association between two variables measured by categories
Relate variables	1	1	0	Continuous/ continuous	Normal	Pearson product moment correlation	Tells you the magnitude and direction of association between two variables measured on an interval (or ratio) scale
Relate variables	2 or more	1	0	Continuous/ continuous	Normal	Multiple regression	Learn about the relationship between several predictor or independent variables and an outcome variable. It provides the relative prediction of one variable among many in terms of the outcome

- Report how the results answered the research question or hypothesis. The *Publication Manual of*

the American Psychological Association (American Psychological Association [APA], 2010) suggests that the most complete meaning of the results come from reporting extensive description, **statistical significance testing**, confidence intervals, and effect sizes. Thus, it is important to clarify the meaning of these last three reports of the results. The statistical significance testing reports an assessment as to whether the observed scores reflect a pattern other than chance. A statistical test is considered to be significant if the results are unlikely by chance to have occurred, and the null hypothesis of “no effect” can be rejected. The researcher sets a rejection level of “no effect,” such as $p = 0.001$, and then assesses whether the test statistic falls into this level of rejection. Typically results will be summarized as “the analysis of variance revealed a statistically significant difference between men and women in terms of attitudes toward banning smoking in restaurants $F(2; 6) = 8.55$, $p = 0.001$.” Two forms of *practical evidence* of the results should also be reported: (a) the effect size and (b) the confidence interval. A confidence interval is a range of values (an interval) that describes a level of uncertainty around an estimated observed score. A confidence interval shows how good an estimated score might be. A confidence interval of 95%, for example, indicates that 95 out of 100 times the observed score will fall in the range of values. An **effect size** identifies the strength of the conclusions about group differences or the relationships among variables in quantitative studies. It is a descriptive statistic that is not dependent on whether the relationship in the data represents the true population. The calculation of effect size varies for different statistical tests: it can be used to explain the variance between two or more variables or the differences among means for groups. It shows the practical significance of the results apart from inferences being applied to the population.

- Discuss the implications of the results for practice or for future research on the topic. This will require drawing inferences and conclusions from the results. It may involve discussing theoretical and practical consequences of the results. Focus should also be on whether or not the research questions/hypotheses were supported.

Example 8.1 A Survey Method Section

An example follows of a survey method section that illustrates many of the steps just mentioned. This excerpt (used with permission) comes from a journal article reporting a study of factors affecting student attrition in one small liberal arts college (Bean & Creswell, 1980, pp. 321–322).

Methodology

The site of this study was a small (enrollment 1,000), religious, coeducational, liberal arts college in a Midwestern city with a population of 175,000 people. [*Authors identified the research site and population.*]

The dropout rate the previous year was 25%. Dropout rates tend to be highest among freshmen and sophomores, so an attempt was made to reach as many freshmen and sophomores as possible by distribution of the questionnaire through classes. Research on attrition indicates that males and females drop out of college for different reasons (Bean, 1978, in press; Spady, 1971). Therefore, only women were analyzed in this study.

During April 1979, 169 women returned questionnaires. A homogeneous sample of 135 women who were 25 years old or younger, unmarried, full-time U.S. citizens, and Caucasian was selected for this analysis to exclude some possible confounding variables (Kerlinger, 1973).

Of these women, 71 were freshmen, 55 were sophomores, and 9 were juniors. Of the students, 95% were between the ages of 18 and 21. This sample is biased toward higher-ability students as indicated by scores on the ACT test. *[Authors presented descriptive information about the sample.]*

Data were collected by means of a questionnaire containing 116 items. The majority of these were Likert-like items based on a scale from “a very small extent” to “a very great extent.” Other questions asked for factual information, such as ACT scores, high school grades, and parents’ educational level. All information used in this analysis was derived from questionnaire data. This questionnaire had been developed and tested at three other institutions before its use at this college. *[Authors discussed the instrument.]*

Concurrent and convergent validity (Campbell & Fiske, 1959) of these measures was established through factor analysis, and was found to be at an adequate level. Reliability of the factors was established through the coefficient alpha. The constructs were represented by 25 measures—multiple items combined on the basis of factor analysis to make indices—and 27 measures were single item indicators. *[Validity and reliability were addressed.]*

Multiple regression and path analysis (Heise, 1969; Kerlinger & Pedhazur, 1973) were used to analyze the data. In the causal model ..., intent to leave was regressed on all variables which preceded it in the causal sequence. Intervening variables significantly related to intent to leave were then regressed on organizational variables, personal variables, environmental variables, and background variables. *[Data analysis steps were presented.]*

COMPONENTS OF AN EXPERIMENTAL METHOD PLAN

An experimental method discussion follows a standard form: (a) participants, (b) materials, (c) procedures, and (d) measures. These four topics generally are sufficient. In this section of the chapter, I review these components as well as information about the experimental design and statistical analysis. As with the section on surveys, the intent here is to highlight key topics to be addressed in an experimental methods section of a proposal. An overall guide to these topics is found by answering the questions on the checklist shown in [Table 8.4](#).

Table 8.4 A Checklist of Questions for Designing an Experimental Procedure

	Who are the participants in the study?
	What is the population to which the results of the participants will be generalized?
	How were the participants selected? Was a random selection method used?
	How will the participants be randomly assigned? Will they be matched? How?
	How many participants will be in the experimental and control group(s)?
	What is the dependent variable or variables (i.e., outcome variable) in the study? How will it be measured? Will it be measured before and after the experiment?
	What is the treatment condition(s)? How was it operationalized?
	Will variables be covaried in the experiment? How will they be measured?
	What experimental research design will be used? What would a visual model of this design look like?
	What instrument(s) will be used to measure the outcome in the study? Why was it chosen? Who developed it? Does it have established validity and reliability? Has permission been sought to use it?
	What are the steps in the procedure (e.g., random assignment of participants to groups, collection of demographic information, administration of pretest, administration of treatment(s), administration of posttest)?
	What are potential threats to internal and external validity for the experimental design and procedure? How will they be addressed?
	Will a pilot test of the experiment be conducted?
	What statistics will be used to analyze the data (e.g., descriptive and inferential)?
	How will the results be interpreted?

Participants

Readers need to know about the selection, assignment, and number of participants who will take part in the experiment. Consider the following suggestions when writing the method section for an experiment:

- Describe the selection process for participants as either random or nonrandom (e.g., conveniently selected). Researchers can select participants by random selection or random sampling. With random selection or random sampling, each individual has an equal probability of being selected from the population, ensuring that the sample will be representative of the population (Keppel & Wickens, 2003). In many experiments, however, only a convenience sample is possible because the investigator must use naturally formed groups (e.g., a classroom, an organization, a family unit) or volunteers. When individuals are not randomly assigned, the procedure is called a **quasi-experiment**.

- When individuals can be randomly assigned to groups, the procedure is called a **true experiment**. If a random assignment is made, discuss how the project will *randomly assign* individuals to the treatment groups. This means that of the pool of participants, Individual 1 goes to Group 1, Individual 2 to Group 2, and so forth so that there is no systematic bias in assigning the individuals. This procedure eliminates the possibility of systematic differences among characteristics of the participants that could affect the outcomes so that any differences in outcomes can be attributed to the experimental treatment (Keppel & Wickens, 2003).

- Identify other features in the experimental design that will systematically control the variables that might influence the outcome. One approach is *equating* the groups at the outset of the experiment so that participation in one group or the other does not influence the outcome. For example, researchers **match participants** in terms of a certain trait or characteristic and then assign one individual from each matched set to each group. For example, scores on a pretest might be obtained. Individuals might then be assigned to groups, with each group having the same numbers of high, medium, and low scorers on the pretest. Alternatively, the criteria for matching might be ability levels or demographic variables. A researcher may decide not to match, however, because it is expensive, takes time (Salkind, 1990), and leads to incomparable groups if participants leave the experiment (Rosenthal & Rosnow, 1991). Other procedures to place control into experiments involve using covariates (e.g., pretest scores) as moderating variables and controlling for their effects statistically, selecting homogeneous samples, or blocking the participants into subgroups or categories and analyzing the impact of each subgroup on the outcome (Creswell, 2012).

- Tell the reader about the number of participants in each group and the systematic procedures for determining the size of each group. For experimental research, investigators use a power analysis (Lipsey, 1990) to identify the appropriate sample size for groups. This calculation involves the following:

- A consideration of the level of statistical significance for the experiment, or alpha
- The amount of power desired in a study—typically presented as high, medium, or low—for the statistical test of the null hypothesis with sample data when the null hypothesis is, in fact, false
- The effect size, the expected differences in the means between the control and experimental groups expressed in standard deviation units

- Researchers set values for these three factors (e.g., $\alpha = 0.05$, power = 0.80, and effect size = 0.50) and can look up in a table the size needed for each group (see Cohen, 1977; Lipsey, 1990). In this way, the experiment is planned so that the size of each treatment group provides the greatest sensitivity that the effect on the outcome actually is due to the experimental manipulation in the study.

Variables

The variables need to be specified in an experiment so that it is clear to readers what groups are receiving the experimental treatment and what outcomes are being measured. Here are some suggestions for developing ideas about variables in a proposal:

- Clearly identify the independent variables in the experiment (recall the discussion of variables in [Chapter 3](#)). One independent variable must be the treatment variable. One or more groups receive the experimental manipulation, or treatment, from the researcher. Other independent variables may simply be measured variables in which no manipulation occurs (e.g., attitudes or personal characteristics of participants). Still other independent variables can be statistically controlled, such as demographics (e.g., gender or age). The method section must list and clearly identify all the independent variables in an experiment.
- Identify the dependent variable or variables (i.e., the outcomes) in the experiment. The dependent variable is the response or the criterion variable presumed to be caused by or influenced by the independent treatment conditions and any other independent variables. Rosenthal and Rosnow (1991) advanced three prototypic outcomes measures: (a) the direction of observed change, (b) the amount of this change, and (c) the ease with which the participant changes (e.g., the participant reacquires the correct response as in a single-subject design).

Instrumentation and Materials

During an experiment, one makes observations or obtains measures using instruments at a pretest or posttest (or both) stage of the procedures. A sound research plan calls for a thorough discussion about the instrument or instruments—their development, their items, their scales, and reports of reliability and validity of scores on past uses. The researcher also should report on the materials used for the experimental treatment (e.g., the special program or specific activities given to the experimental group).

- Describe the instrument or instruments participants complete in the experiment, typically filled out before the experiment begins and at its end. Indicate the established validity and reliability of the scores on instruments, the individuals who developed them, and any permissions needed to use them.
- Thoroughly discuss the materials used for the experimental treatment. One group, for example, may participate in a special computer-assisted learning plan used by a teacher in a classroom. This plan might involve handouts, lessons, and special written instructions to help students in this experimental group learn how to study a subject using computers. A pilot test of these materials may also be discussed, as well as any training required to administer the materials in a standard way. The intent of this pilot test is to ensure that materials can be administered without variability to the experimental group.

Experimental Procedures

The specific experimental design procedures also need to be identified. This discussion involves indicating the overall experiment type, citing reasons for the design, and advancing a visual model to help the reader understand the procedures.

- Identify the type of experimental design to be used in the proposed study. The types available in experiments are pre-experimental designs, quasi-experiments, true experiments, and single-subject designs. With pre-experimental designs, the researcher studies a single group and provides an intervention during the experiment. This design does not have a control group to compare with the experimental group. In quasi-experiments, the investigator uses control and experimental groups but does not randomly assign participants to groups (e.g., they may be intact groups available to the researcher). In a *true experiment*, the investigator randomly assigns the participants to treatment groups. A **single-subject design** or *N* of 1 design involves observing the behavior of a single individual (or a small number of individuals) over time.

- Identify what is being compared in the experiment. In many experiments, those of a type called between-subject designs, the investigator compares two or more groups (Keppel & Wickens, 2003; Rosenthal & Rosnow, 1991). For example, a factorial design experiment, a variation on the between-group design, involves using two or more treatment variables to examine the independent and simultaneous effects of these treatment variables on an outcome (Vogt, 2011). This widely used behavioral research design explores the effects of each treatment separately and also the effects of variables used in combination, thereby providing a rich and revealing multidimensional view. In other experiments, the researcher studies only one group in what is called a within-group design. For example, in a repeated measures design, participants are assigned to different treatments at different times during the experiment. Another example of a within-group design would be a study of the behavior of a single individual over time in which the experimenter provides and withholds a treatment at different times in the experiment to determine its impact.

- Provide a diagram or a figure to illustrate the specific research design to be used. A standard notation system needs to be used in this figure. A research tip I recommend is to use a classic notation system provided by Campbell and Stanley (1963, p. 6):

- *X* represents an exposure of a group to an experimental variable or event, the effects of which are to be measured.
- *O* represents an observation or measurement recorded on an instrument.
- *X*s and *O*s in a given row are applied to the same specific persons. *X*s and *O*s in the same column, or placed vertically relative to each other, are simultaneous.
 - The left-to-right dimension indicates the temporal order of procedures in the experiment (sometimes indicated with an arrow).
- The symbol *R* indicates random assignment.
- Separation of parallel rows by a horizontal line indicates that comparison groups are not equal (or equated) by random assignment. No horizontal line between the groups displays random assignment of individuals to treatment groups.

In the following examples, this notation is used to illustrate pre-experimental, quasi-experimental,

true experimental, and single-subject designs.

Example 8.2 *Pre-Experimental Designs*

One-Shot Case Study

This design involves an exposure of a group to a treatment followed by a measure.

Group A X _____ O

One-Group Pretest-Posttest Design

This design includes a pretest measure followed by a treatment and a posttest for a single group.

Group A $O1$ _____ X _____ $O2$

Static Group Comparison or Posttest-Only With Nonequivalent Groups

Experimenters use this design after implementing a treatment. After the treatment, the researcher selects a comparison group and provides a posttest to both the experimental group(s) and the comparison group(s).

Group A X _____ O

Group B _____ O

Alternative Treatment Posttest-Only With Nonequivalent Groups Design

This design uses the same procedure as the Static Group Comparison, with the exception that the nonequivalent comparison group received a different treatment.

Group A $X1$ _____ O

Group B $X2$ _____ O

Example 8.3 *Quasi-Experimental Designs*

Nonequivalent (Pretest and Posttest) Control-Group Design

In this design, a popular approach to quasi-experiments, the experimental Group A and the control Group B are selected without random assignment. Both groups take a pretest and posttest. Only the experimental group receives the treatment.

Group A O _____ X _____ O

Group B O _____ O

Single-Group Interrupted Time-Series Design

In this design, the researcher records measures for a single group both before and after a treatment.

Group A $O-O-O-O-X-O-O-O-O$

Control-Group Interrupted Time-Series Design

This design is a modification of the Single-Group Interrupted Time-Series design in which two groups of participants, not randomly assigned, are observed over time. A treatment is administered to only one of the groups (i.e., Group A).

Group A $O-O-O-O-X-O-O-O-O$

Group B $O-O-O-O-O-O-O-O-O$

Example 8.4 True Experimental Designs

Pretest-Posttest Control-Group Design

A traditional, classical design, this procedure involves random assignment of participants to two groups. Both groups are administered both a pretest and a posttest, but the treatment is provided only to experimental Group A.

Group A $R-O-X-O$

Group B $R-O-O$

Posttest-Only Control-Group Design

This design controls for any confounding effects of a pretest and is a popular experimental design. The participants are randomly assigned to groups, a treatment is given only to the experimental group, and both groups are measured on the posttest.

Group A $R-X-O$

Group B $R-O$

Solomon Four-Group Design

A special case of a 2 X 2 factorial design, this procedure involves the random assignment of participants to four groups. Pretests and treatments are varied for the four groups. All groups receive a posttest.

Group A $R-O-X-O$

Group B $R-O-O-O$

Group C $R-X-O$

Group D $R-O-O$

Example 8.5 *Single-Subject Designs*

A-B-A Single-Subject Design

This design involves multiple observations of a single individual. The target behavior of a single individual is established over time and is referred to as a baseline behavior. The baseline behavior is assessed, the treatment provided, and then the treatment is withdrawn.

Baseline A Treatment B Baseline A

O-O-O-O-O-X-X-X-X-O-O-O-O-O

Threats to Validity

There are several threats to validity that will raise questions about an experimenter’s ability to conclude that the intervention affects an outcome and not some other factor. Experimental researchers need to identify potential threats to the internal validity of their experiments and design them so that these threats will not likely arise or are minimized. There are two types of threats to validity: (a) internal threats and (b) external threats. **Internal validity threats** are experimental procedures, treatments, or experiences of the participants that threaten the researcher’s ability to draw correct inferences from the data about the population in an experiment. **Table 8.5** displays these threats, provides a description of each one of them, and suggests potential responses by the researcher so that the threat may not occur. There are those involving participants (i.e., history, maturation, regression, selection, and mortality), those related to the use of an experimental treatment that the researcher manipulates (i.e., diffusion, compensatory and resentful demoralization, and compensatory rivalry), and those involving procedures used in the experiment (i.e., testing and instruments).

Table 8.5 Types of Threats to Internal Validity

Type of Threat to Internal Validity	Description of Threat	In Response, Actions the Researcher Can Take
History	Because time passes during an experiment, events can occur that unduly influence the outcome beyond the experimental treatment.	The researcher can have both the experimental and control groups experience the same external events.
Maturation	Participants in an experiment may mature or change during the experiment, thus influencing the results.	The researcher can select participants who mature or change at the same rate (e.g., same age) during the experiment.
Regression	Participants with extreme scores are selected for the experiment. Naturally, their scores will probably change during the experiment. Scores, over time, regress toward the mean.	A researcher can select participants who do not have extreme scores as entering characteristics for the experiment.
Selection	Participants can be selected who have certain characteristics that predispose them to have certain outcomes (e.g., they are brighter).	The researcher can select participants randomly so that characteristics have the probability of being equally distributed among the experimental groups.
Mortality	Participants drop out during an experiment due to many possible reasons. The outcomes are thus unknown for these individuals.	A researcher can recruit a large sample to account for dropouts or compare those who drop out with those who continue—in terms of the outcome.
Diffusion of treatment	Participants in the control and experimental groups communicate with each other. This communication can influence how both groups score on the outcomes.	The researcher can keep the two groups as separate as possible during the experiment.
Compensatory/Resentful demoralization	The benefits of an experiment may be unequal or resented when only the experimental group receives the treatment (e.g., experimental group receives therapy and the control group receives nothing).	The researcher can provide benefits to both groups, such as giving the control group the treatment after the experiment ends or giving the control group some different type of treatment during the experiment.
Compensatory rivalry	Participants in the control group feel that they are being devalued, as compared to the experimental group, because they do not experience the treatment.	The researcher can take steps to create equality between the two groups, such as reducing the expectations of the control group.
Testing	Participants become familiar with the outcome measure and remember responses for later testing.	The researcher can have a longer time interval between administrations of the outcome or use different items on a later test than were used in an

		earlier test.
Instrumentation	The instrument changes between a pretest and posttest, thus impacting the scores on the outcome.	The researcher can use the same instrument for the pretest and posttest measures.

SOURCE: Adapted from Creswell (2012).

Potential threats to external validity also must be identified and designs created to minimize these threats. **External validity threats** arise when experimenters draw incorrect inferences from the sample data to other persons, other settings, and past or future situations. As shown in [Table 8.6](#), these threats arise because of the characteristics of individuals selected for the sample, the uniqueness of the setting, and the timing of the experiment. For example, threats to external validity arise when the researcher generalizes beyond the groups in the experiment to other racial or social groups not under study, to settings not examined, or to past or future situations. Steps for addressing these potential issues are also presented in [Table 8.6](#).

Other threats that might be mentioned in the method section are the threats to **statistical conclusion validity** that arise when experimenters draw inaccurate inferences from the data because of inadequate statistical power or the violation of statistical assumptions. Threats to construct validity occur when investigators use inadequate definitions and measures of variables.

Table 8.6 Types of Threats to External Validity

Types of Threats to External Validity	Description of Threat	In Response, Actions the Researcher Can Take
Interaction of selection and treatment	Because of the narrow characteristics of participants in the experiment, the researcher cannot generalize to individuals who do not have the characteristics of participants.	The researcher restricts claims about groups to which the results cannot be generalized. The researcher conducts additional experiments with groups with different characteristics.
Interaction of setting and treatment	Because of the characteristics of the setting of participants in an experiment, a researcher cannot generalize to individuals in other settings.	The researcher needs to conduct additional experiments in new settings to see if the same results occur as in the initial setting.
Interaction of history and treatment	Because results of an experiment are time-bound, a researcher cannot generalize the results to past or future situations.	The researcher needs to replicate the study at later times to determine if the same results occur as in the earlier time.

SOURCE: Adapted from Creswell (2012).

- Practical research tips for proposal writers to address validity issues are as follows:
- Identify the potential threats to validity that may arise in your study. A separate section in a proposal may be composed to advance this threat.
 - Define the exact type of threat and what potential issue it presents to your study.
 - Discuss how you plan to address the threat in the design of your experiment.
 - Cite references to books that discuss the issue of threats to validity, such as Cook and Campbell (1979); Shadish, Cook, & Campbell (2001); and Tuckman (1999).

The Procedure

A proposal developer needs to describe in detail the procedure for conducting the experiment. A reader should be able to understand the design being used, the observations, the treatment, and the timeline of activities.

- Discuss a step-by-step approach for the procedure in the experiment. For example, Borg and Gall (2006) outlined steps typically used in the procedure for a pretest-posttest control group design with matching participants in the experimental and control groups:
 1. Administer measures of the dependent variable or a variable closely correlated with the dependent variable to the research participants.
 2. Assign participants to matched pairs on the basis of their scores on the measures described in Step 1.
 3. Randomly assign one member of each pair to the experimental group and the other member to the control group.
 4. Expose the experimental group to the experimental treatment and administer no treatment or an alternative treatment to the control group.
 5. Administer measures of the dependent variables to the experimental and control groups.
 6. Compare the performance of the experimental and control groups on the posttest(s) using tests of statistical significance.

Data Analysis

Tell the reader about the types of statistical analysis that will be used during the experiment.

- Report the descriptive statistics calculated for observations and measures at the pretest or posttest stage of experimental designs. This call for descriptive analysis is consistent with the recent *APA Publication Manual* (APA, 2010). These statistics are means, standard deviations, and ranges.
- Indicate the inferential statistical tests used to examine the hypotheses in the study. For experimental designs with categorical information (groups) on the independent variable and continuous information on the dependent variable, researchers use *t* tests or univariate analysis of variance (ANOVA), analysis of covariance (ANCOVA), or multivariate analysis of variance (MANOVA—multiple dependent measures). (Several of these tests are mentioned in [Table 8.3](#), which was presented earlier.) In factorial designs, both interaction and main effects of ANOVA are used. When data on a pretest or posttest show marked deviation from a normal distribution, use nonparametric statistical tests. Also, indicate the practical significance by reporting effect sizes and confidence intervals.
- For single-subject research designs, use line graphs for baseline and treatment observations for abscissa (horizontal axis) units of time and the ordinate (vertical axis) target behavior. Researchers plot each data point separately on the graph, and connect the data points with lines (e.g., see Neuman & McCormick, 1995). Occasionally, tests of statistical significance, such as *t* tests, are used to

compare the pooled mean of the baseline and the treatment phases, although such procedures may violate the assumption of independent measures (Borg & Gall, 2006).

Interpreting Results

The final step in an experiment is to interpret the findings in light of the hypotheses or research questions set forth in the beginning. In this interpretation, address whether the hypotheses or questions were supported or whether they were refuted. Consider whether the treatment that was implemented actually made a difference for the participants who experienced them. Suggest why or why not the results were significant, drawing on past literature that you reviewed ([Chapter 2](#)), the theory used in the study ([Chapter 3](#)), or persuasive logic that might explain the results. Address whether the results might have occurred because of inadequate experimental procedures, such as threats to internal validity, and indicate how the results might be generalized to certain people, settings, and times. Finally, indicate the implications of the results for the population studied or for future research.

Example 8.6 *An Experimental Method Section*

The following is a selected passage from a quasi-experimental study by Enns and Hackett (1990) that demonstrates many of the components in an experimental design. Their study addressed the general issue of matching client and counselor interests along the dimensions of attitudes toward feminism. They hypothesized that feminist participants would be more receptive to a radical feminist counselor than would nonfeminist participants and that nonfeminist participants would be more receptive to a nonsexist and liberal feminist counselor. Except for a limited discussion about data analysis and an interpretation section found in the discussion of their article, their approach contains the elements of a good method section for an experimental study.

Method

Participants

The participants were 150 undergraduate women enrolled in both lower-and upper-division courses in sociology, psychology, and communications at a mid-sized university and a community college, both on the west coast. [*The authors described the participants in this study.*]

Design and Experimental Manipulation

This study used a $3 \times 2 \times 2$ factorial design: Orientation of Counselor (nonsexist-humanistic, liberal feminist, or radical feminist) $\times$ Statement of Values (implicit or explicit) $\times$ Participants' Identification with Feminism (feminist or nonfeminist). Occasional missing data on particular items were handled by a pairwise deletion procedure. [*Authors identified the overall design.*]

The three counseling conditions, nonsexist-humanistic, liberal, and radical feminist, were depicted by 10 min videotape vignettes of a second counseling session between a female counselor and a female client. ... The implicit statement of values condition used the sample interview only; the counselor's values were therefore implicit in her responses. The explicit

statement of values condition was created by adding to each of the three counseling conditions a 2-min leader that portrayed the counselor describing to the client her counseling approach and associated values including for the two feminist conditions a description of her feminist philosophical orientation, liberal or radical. ... Three counseling scripts were initially developed on the basis of distinctions between nonsexist-humanistic, liberal, and radical feminist philosophies and attendant counseling implications. Client statements and the outcome of each interview were held constant, whereas counselor responses differed by approach. *[Authors described the three treatment conditions variables manipulated in the study.]*

Instruments

Manipulation checks. As a check on participants' perception of the experimental manipulation and as an assessment of participants' perceived similarity to the three counselors, two subscales of Berryman-Fink and Verderber's (1985) Attributions of the Term Feminist Scale were revised and used in this study as the Counselor Description Questionnaire (CDQ) and the Personal Description Questionnaire (PDQ). ... Berryman-Fink and Verderber (1985) reported internal consistency reliabilities of .86 and .89 for the original versions of these two subscales. *[Authors discussed the instruments and the reliability of the scales for the dependent variable in the study.]*

Procedure

All experimental sessions were conducted individually. The experimenter, an advanced doctoral student in counseling psychology, greeted each subject, explained the purpose of the study as assessing students' reactions to counseling, and administered the ATF. The ATF was then collected and scored while each subject completed a demographic data form and reviewed a set of instructions for viewing the videotape. The first half of the sample was randomly assigned to one of the twelve videotapes (3 Approaches $\times$ 2 Statements $\times$ 2 Counselors), and a median was obtained on the ATF. The median for the first half of the sample was then used to categorize the second half of the group as feminist or nonfeminist, and the remainder of the participants was randomly assigned to conditions separately from each feminist orientation group to ensure nearly equal cell sizes. The median on the final sample was checked and a few participants recategorized by the final median split, which resulted in 12 or 13 participants per cell.

After viewing the videotape that corresponded to their experimental assignment, participants completed the dependent measures and were debriefed. [pp. 35–36; *Authors described the procedure used in the experiment.*]

SUMMARY

This chapter identified essential components in designing a method section in a proposal for a survey or experimental study. The outline of steps for a survey study began with a discussion about the purpose, the identification of the population and sample, the survey instruments to be used, the relationship between the variables, the research questions, specific items on the survey, and steps to be taken in the analysis and the interpretation of the data from the survey. In the design of an experiment, the researcher identifies participants in the study, the variables—the treatment conditions and the outcome variables—and the instruments used for pretests and posttests and the materials to be used in the treatments. The design also includes the specific type of experiment, such as a pre-experimental, quasi-experimental, true experiment, or single-subject design. Then the researcher draws a figure to illustrate the design, using appropriate notation. This is followed by comments about potential threats to internal and external validity (and possibly statistical and construct validity) that relate to the experiment, the statistical analysis used to test the hypotheses or research questions, and the interpretation of the results.

Writing Exercises

1. Design a plan for the procedures to be used in a survey study. Review the checklist in [Table 8.1](#) after you write the section to determine if all components have been addressed.
2. Design a plan for procedures for an experimental study. Refer to [Table 8.4](#) after you complete your plan to determine if all questions have been addressed adequately.

ADDITIONAL READINGS

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 1–76). Chicago: Rand McNally.

This chapter in the Gage *Handbook* is the classical statement about experimental designs. Campbell and Stanley designed a notation system for experiments that is still used today; they also advanced the types of experimental designs, beginning with factors that jeopardize internal and external validity, the pre-experimental design types, true experiments, quasi-experimental designs, and correlational and ex post facto designs. The chapter presents an excellent summary of types of designs, their threats to validity, and statistical procedures to test the designs. This is an essential chapter for students beginning their study of experimental studies.

Fowler, F. J. (2009). *Survey research methods* (4th ed.). Thousand Oaks, CA: Sage.

Floyd Fowler provides a useful text about the decisions that go into the design of a survey research project. He addresses use of alternative sampling procedures, ways of reducing nonresponse rates, data collection, design of good questions, employing sound interviewing techniques, preparation of surveys for analysis, and ethical issues in survey designs.

Keppel, G. & Wickens, T. D. (2003). *Design and analysis: A researcher's handbook* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.

Geoffrey Keppel and Thomas Wickens provide a detailed, thorough treatment of the design of experiments from the principles of design to the statistical analysis of experimental data. Overall, this book is for the mid-level to advanced statistics student who seeks to understand the design and statistical analysis of experiments. The introductory chapter presents an informative overview of the components of experimental designs.

Lipsey, M. W. (1990). *Design sensitivity: Statistical power for experimental research*. Newbury Park, CA: Sage.

Mark Lipsey has authored a major book on the topics of experimental designs and statistical power of those designs. Its basic premise is that an experiment needs to have sufficient sensitivity to detect those effects it purports to investigate. The book explores statistical power and includes a table to help researchers identify the appropriate size of groups in an experiment.

Neuman, S. B., & McCormick, S. (Eds.). (1995). *Single-subject experimental research: Applications for literacy*. Newark, DE: International Reading Association.

Susan Neuman and Sandra McCormick have edited a useful, practical guide to the design of single-subject research. They present many examples of different types of designs, such as reversal designs and multiple-baseline designs, and they enumerate the statistical procedures that might be involved in analyzing the single-subject data. One chapter, for example, illustrates the conventions for displaying data on line graphs. Although this book cites many applications in literacy, it has broad application in the social and human sciences.

Thompson, B. (2006). *Foundations of behavioral statistics: An insight-based approach*. New York: The Guilford.

Bruce Thompson has organized a highly readable book about using statistics. He reviews the basics about descriptive statistics (location, dispersion, shape), about relationships among variables and statistical significance, about the practical significance of results, and about more advanced statistics such as regression, ANOVA, the general linear model, and logistic regression. Throughout the book, he brings in practical examples to illustrate his points.