

**This item is provided to support UOB courses.**

**Its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission.**

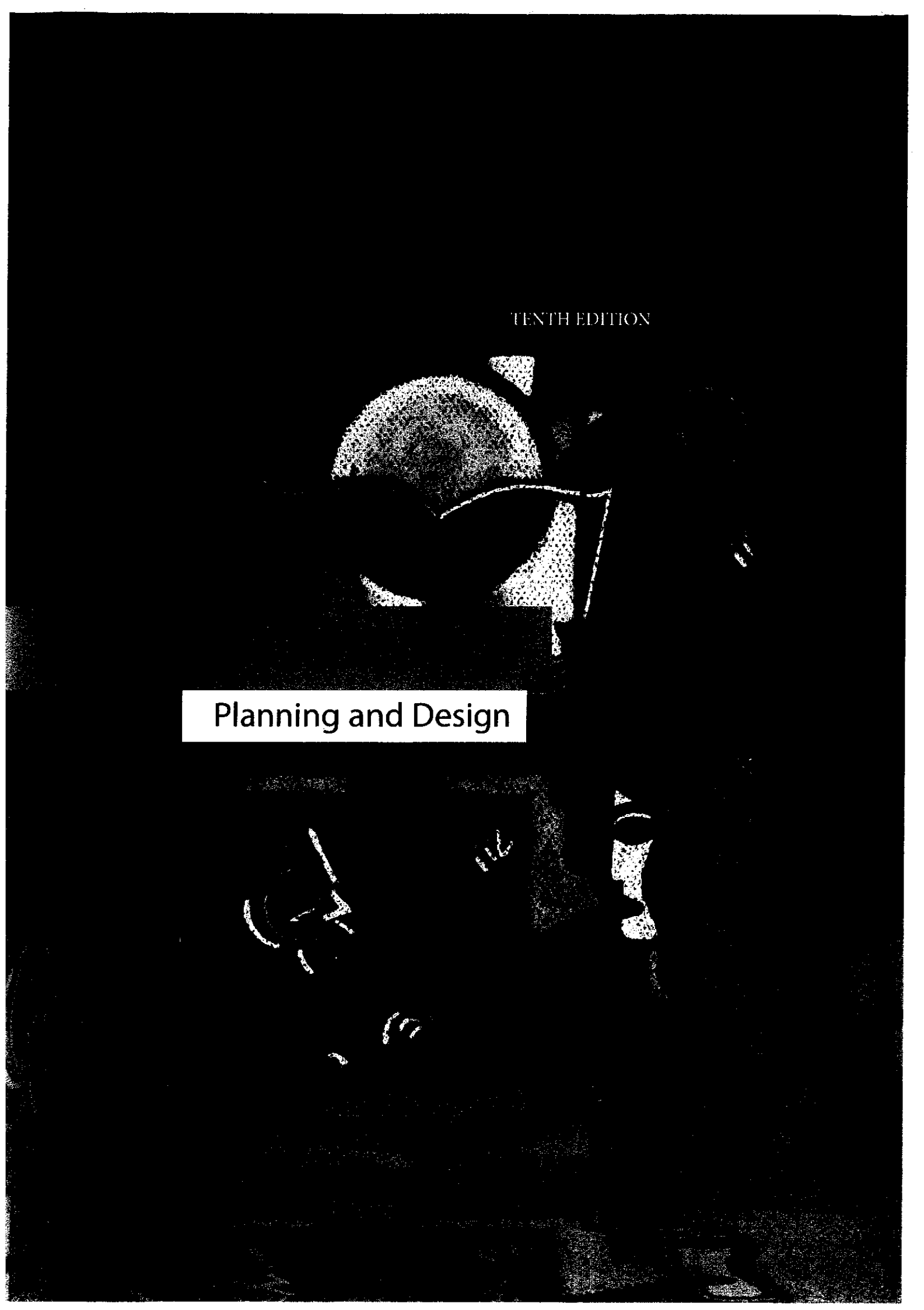
**However, users may print, download, or email it for individual use for learning and research purposes only.**

هذه الوثيقة متوفرة لمساندة مقرارات الجامعة.

ويمنع منعاً باتاً نسخها في نسخ متعددة أو إرسالها بالبريد الإلكتروني إلى قائمة تعميم بدون الحصول على إذن مسبق من صاحب الحق القانوني للملكية الفكرية لكن يمكن للمستفيد أن يطبع أو يحفظ نسخة منها لاستخدام الشخصي لأغراض التعلم والبحث العلمي فقط.

TENTH EDITION

Planning and Design



## Pearson's MyLab™

### PROVEN RESULTS

For over 10 years, instructors and students have reported better grades through increased engagement and real-time insights into progress.

### ENGAGING EXPERIENCES

MyLab is designed to reach students in a personal way. Engaging learning and practice opportunities lead to assessments that create a personalized study plan.

### A TRUSTED PARTNERSHIP

With millions of students registered annually, MyLab is the most effective and reliable learning solution available today.

To order this text with **MyEducationLab™**  
use ISBN 0-13-289361-4

ISBN-13: 978-0-13-269324-0  
ISBN-10: 0-13-269324-0



[www.pearsonhighered.com](http://www.pearsonhighered.com)

PEARSON

ALWAYS LEARNING

# 4

## Planning Your Research Project

Before the construction of a building, an architect develops a detailed set of plans. These plans enable the builder to erect a strong, well-designed structure. Researchers should be no less detailed and precise in the planning of a research design. Plans, specifications, and design: All of these serve well the architect, the builder, and the researcher alike.

Architectural planning and research planning have much in common. Each requires a conceptualization of the overall organization of a project and a detailed specification of the steps to be carried out; only after such meticulous planning can work on the project actually begin. For successful completion, a building requires plans that are clearly conceived and accurately drawn. A research project should be no less completely conceptualized and precisely detailed.

### Planning a General Approach

When we talk about a general strategy for solving a research problem, we are talking about a **research design**. The research design provides the overall structure for the procedures the researcher follows, the data the researcher collects, and the data analyses the researcher conducts. Simply put, research design is *planning*.

Nothing helps a research effort be successful so much as carefully planning the overall design. More research effort is wasted by going off half-prepared—with only a vague set of ideas and procedures—than in any other way. You will be much more efficient and effective as a researcher if you identify your resources, your procedures, and the forms that your data will take—always with the central goal of solving your research problem in mind—at the very beginning of your project.

### The Basic Format of All Research

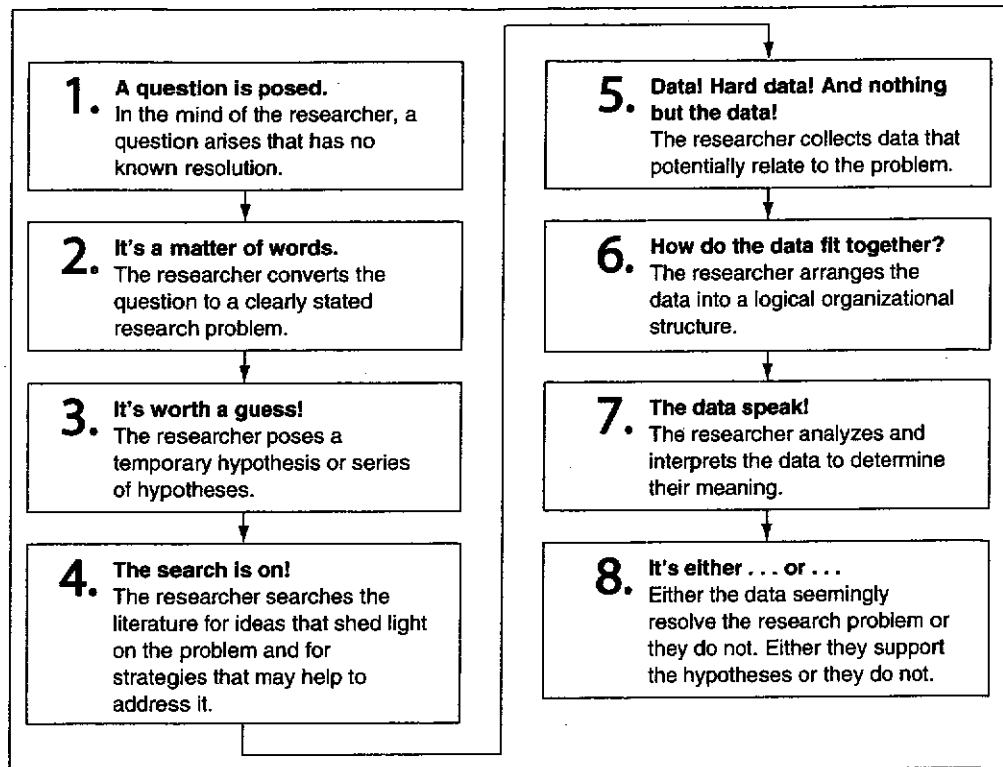
The research process follows a basic format. No matter which academic discipline gives rise to the research endeavor, the general research procedure is fundamentally the same. Although you will sometimes hear people refer to “social science research,” “nursing research,” or “marketing research,” in fact the search for information to solve a problem or to answer a question seldom fits squarely within the bounds of a single academic discipline.

In planning a research design, the researcher in quest of new knowledge cannot be shackled by discipline-specific methodological restraints. The course of a research project will frequently lead the investigator into new and unfamiliar territories that have historically been associated with other content areas. The sociologist trying to resolve a problem in sociology may come face to face with problems that are psychological or economic. The educational researcher exploring the causes of a learning disability may need to consider the domains of neurophysiology, psychopathology, endocrinology, and family counseling. On the way to finding a solution for a problem in criminology, the student in criminal justice may venture into the realms of abnormal psychology and behavioral genetics. Any good researcher must be *eclectic*, willing to draw on whatever sources seem to offer productive methods or evidence for resolving the research problem.

Figure 4.1 presents the basic format for the research process. It is important to note, however, that a research project does not always follow these steps in the exact sequence depicted

**FIGURE 4.1**

The basic format of the research process



here. Especially in the case of qualitative research—an approach we describe later in the chapter—a researcher may order the steps somewhat differently. For instance, in some forms of qualitative research, the researcher formulates a hypothesis and reads much of the literature only *after* collecting a substantial amount of data (more on this point in Chapter 6).

Instead of limiting their thinking to departmentalized knowledge, researchers might do much better to think of problems as arising out of broad generic areas within whose boundaries all research falls: people, things, records, thoughts and ideas, and dynamics and energy. Let's briefly consider some research problems that may fall within each of these areas.

- *People.* In this category are found research problems relating to children, senior citizens, families, communities, cultural groups, ancestors, employees, management, mental and physiological processes, learning, motivation, social and educational problems, crime, rehabilitation, medical treatments, nutrition, language, religion.
- *Things.* In this category are found research problems relating to animal and vegetable life, viruses and bacteria, inanimate objects (rocks, soil, buildings, machines), matter (molecules, atoms, subatomic matter), stars, galaxies.
- *Records.* In this category are found research problems relating to newspapers, journals, letters, manuscripts, registers, speeches, interviews, minutes, legal documents, financial and corporate statements, census reports, archeological remains, sketches, drawings, paintings, music.
- *Thoughts and ideas.* In this category are found research problems relating to concepts, theories, perceptions, opinions, beliefs, reactions, issues, semantics, poetry, cartoons, caricatures.
- *Dynamics and energy.* In this category are found research problems relating to human interactions, metabolism, chemical reactions, radiation, radio and microwave transmissions, quantum mechanics, thermodynamics, hydrodynamics, hydrologic cycles, atomic and nuclear energy, wave mechanics, atmospheric and oceanic energy systems, solar energy, black holes.

We do not intend the preceding lists to be mutually exclusive or all-inclusive. We merely present them to give you an idea of the many research possibilities that each category suggests.

## Research Planning versus Research Methodology

Do not confuse overall research planning with research methodology. Whereas the general approach to *planning* a research study may be similar across disciplines, the techniques one uses to collect and analyze data—that is, the *methodology*—may be specific to a particular academic discipline. Such is the case because data vary so widely in nature. You cannot deal with a blood cell in the same way that you deal with a historical document, and the problem of finding the sources of Coleridge's "Kubla Khan" is entirely different from the problem of finding the sources of radio signals from extragalactic space. You cannot study chromosomes with a questionnaire, and you cannot study attitudes with a microscope.

In planning the research design, therefore, it is extremely important for the researcher not only to choose a viable research problem but also to consider the kinds of data that an investigation of the problem will require, as well as reasonable means of collecting and interpreting those data. Many beginning researchers become so entranced with the glamour of the problem that they fail to consider practical issues related to data availability, collection, and interpretation.

Comparing the brain wave patterns of children who are gifted versus those of average ability may be an engaging project for research, but consider the following issues:

- Will you be able to find a sufficient number of children who are willing to participate in the study and whose parents will grant permission for their children to participate?
- Do you have an electroencephalograph at your disposal?
- If so, do you have the technical skills to use it?
- Are you sufficiently knowledgeable to interpret the electroencephalographic tracings you obtain?
- If so, do you know how you would interpret the data and organize your findings so that you could draw conclusions from them?

Unless the answer to all of these questions is *yes*, it probably is better that you give up this project in favor of one for which you have the knowledge, skills, and resources to carry through to completion. Your research should be *practical* research, built on precise and realistic *planning* and executed within the framework of a clearly conceived and feasible *design*.

## General Criteria for a Research Project

A good research project has four important qualities you can use as criteria in evaluating your own research plan. Here, briefly, are those qualities.

**Universality.** The research project should be one that might be carried out by *any* competent person. As a researcher, you are a director and producer—an agent whose function is to collect, organize, and report what the collected data seem to indicate—but another, equally knowledgeable individual might take your place and complete the project with essentially the same results. (We will find an exception to this criterion later in the chapter, when we discuss the nature of data collection in qualitative research.)

**Replication.** The research should be repeatable. Any other competent researcher should be able to take your problem and, collecting data under the same circumstances and within the same parameters you have used, achieve results comparable to those you have obtained.

**Control.** The researcher must in some way isolate, or *control*, those factors that are central to the research problem. Control is important for replication: An experiment should be repeated under the identical conditions and in the identical way in which it was first carried out. Control is also important for consistency within the research design. For instance, if you want to compare the effects of two different treatments on some other factor, then you should keep everything else (aside from the specific treatments you are studying) as similar as possible.

Control is more easily achieved in some areas than in others. In the physical sciences, for example, such factors as temperature, pressure, electrical potential, humidity, and the like are

highly controllable. Control is more challenging—but certainly possible—in research areas concerned with human behavior.

Control is especially important in experimental research designs. Accordingly, we talk more about this issue in Chapter 9.

**Measurement.** The data should be able to be measured in some way. This, again, is often easily accomplished in the physical sciences. Measurement is typically less precise and less accurate in the humanities and social sciences. Later in the chapter we discuss strategies for enhancing measurement procedures in these areas.

---

## The Nature and Role of Data in Research

Research is a viable approach to a problem only when data exist to support it. The term *data* is plural (singular is *datum*) and comes from the past participle of the Latin verb *dare*, which means “to give.” Data are those pieces of information that any particular situation *gives* to an observer.

### What Are Data?

Researchers must always remember that data are not absolute reality—the pure, naked Truth that underlies all the phenomena we observe. Rather, data are *manifestations* of that reality. No one has ever looked upon Truth itself. We are like those individuals who live in a dungeon, across the floor of which a beam of sunlight passes. That light gives us an idea of what the sun must be like, but if we can never behold the sun, we shall never know the difference between it and the shaft of light on the dungeon floor.<sup>1</sup>

The researcher is in a factual dungeon. He or she will never be able to see the original source of the data. For instance, we often see what other people do—the behaviors they exhibit, the things they create, and the effects of their actions on others. But the actual people “inside”—those individuals we will never know!

Research seeks, through data, to discover underlying truths. Yet such is probably an endless pursuit. Experienced researchers are constantly aware that the Truth they most ardently seek is forever just beyond what is represented by the data and, hence, just beyond human grasp. For instance, the scientist probing the nature of subatomic matter may detect tiny entities that make up larger bits of matter. Yet the scientist knows, too, that such entities are probably made up of sub-sub-subentities that beckon yet evade measurement and investigation.

The mind yearns to understand the Truth. To pursue that goal, we have chosen the path of research. But the path always ends at the farthest reaches of the data, which are at the brink of the canyon in whose depths lies the inaccessible, ultimate Truth.

### Data Are Transient and Ever Changing

Whenever we look at data analytically, we gain new insights. But at the same time, we also discern new problems that demand further research.

Data are not only elusive but also transient. Data that the researcher is permitted to glimpse may exist for only a split second. Consider, as an example, a sociologist who plans to conduct a survey in order to learn about people’s attitudes and opinions in a certain city. The sociologist’s research assistants begin by administering the survey in a particular city block. By the time they move on to the next block, the data they have collected are already out of date. Some people in the previous block who indicated that they held a particular opinion have now changed their minds and have a somewhat different opinion. They may have seen a television program or heard a discussion that has changed it. Some people have moved away, and others have moved in; some

---

<sup>1</sup>For readers interested in philosophy, our dungeon analogy is based loosely on Plato’s Analogy of the Cave, which he used in Book VII of *The Republic*.

have died, and others have been born. Tomorrow, next week, next year—what we thought we had “discovered” may have changed completely.

Thus is the transient nature of data. We catch merely a fleeting glance of what seems to be true at one point in time but is not necessarily true the next. Even the most carefully collected data may have an elusive quality about them; at a later point in time they may have no counterpart in reality whatsoever. Data are volatile: They evaporate quickly.

## Primary Data versus Secondary Data

The researcher's only perceptions of Truth are various layers of truth-revealing facts. In the layer closest to the Truth are **primary data**; these are often the most valid, the most illuminating, the most truth-manifesting. Farther away is a layer consisting of **secondary data**, which are derived not from the Truth itself, but from the primary data.

Earlier we used the analogy of the researcher as someone who sits in a dungeon and tries to understand the sun only by looking at a shaft of sunlight that falls on the floor. This direct beam of sunlight represents the *primary data*. Although the shaft is not the sun itself, it has come directly from the sun. But now imagine that the imprisoned researcher sees the sunlight, not as a direct beam, but as a pattern of shimmering light on the floor. The sunlight (primary data) has fallen onto a shiny surface and then been reflected—distorted by the imperfections of the shiny surface—to an image that is in some ways similar but in other ways dissimilar to the original shaft of light. This reflection of light is secondary data.

As another example, consider the following incident: You see a car veer off the highway and into a ditch. You have witnessed the entire event. Afterward, the driver says he had no realization that an accident might occur until the car went out of control. Neither you nor the driver will ever be able to determine the absolute truth underlying the accident. Did the driver have a momentary seizure of which he was unaware? Did the car have an imperfection that the damage from the accident obscured? Were other factors involved that neither of you noticed? The answers lie beyond an impenetrable barrier. The true cause of the accident may never be known, but the things you witnessed, incomplete as they may be, are primary data that emanated directly from the accident itself.

Now along comes a newspaper reporter who interviews both you and the driver. The reporter writes an account of the accident for the local paper. When your sister reads the account in the paper the next morning, she gets, as it were, the sunlight-reflection-on-the-floor version of the event. The newspaper article provides secondary data. The data are, of necessity, distorted—perhaps only a little, perhaps quite a bit—by the channels of communication through which they must pass to her. The reporter's writing skills, your sister's reading skills, and the inability of language to reproduce every nuance of detail that a firsthand observation can provide—all of these factors distort what you actually observed.

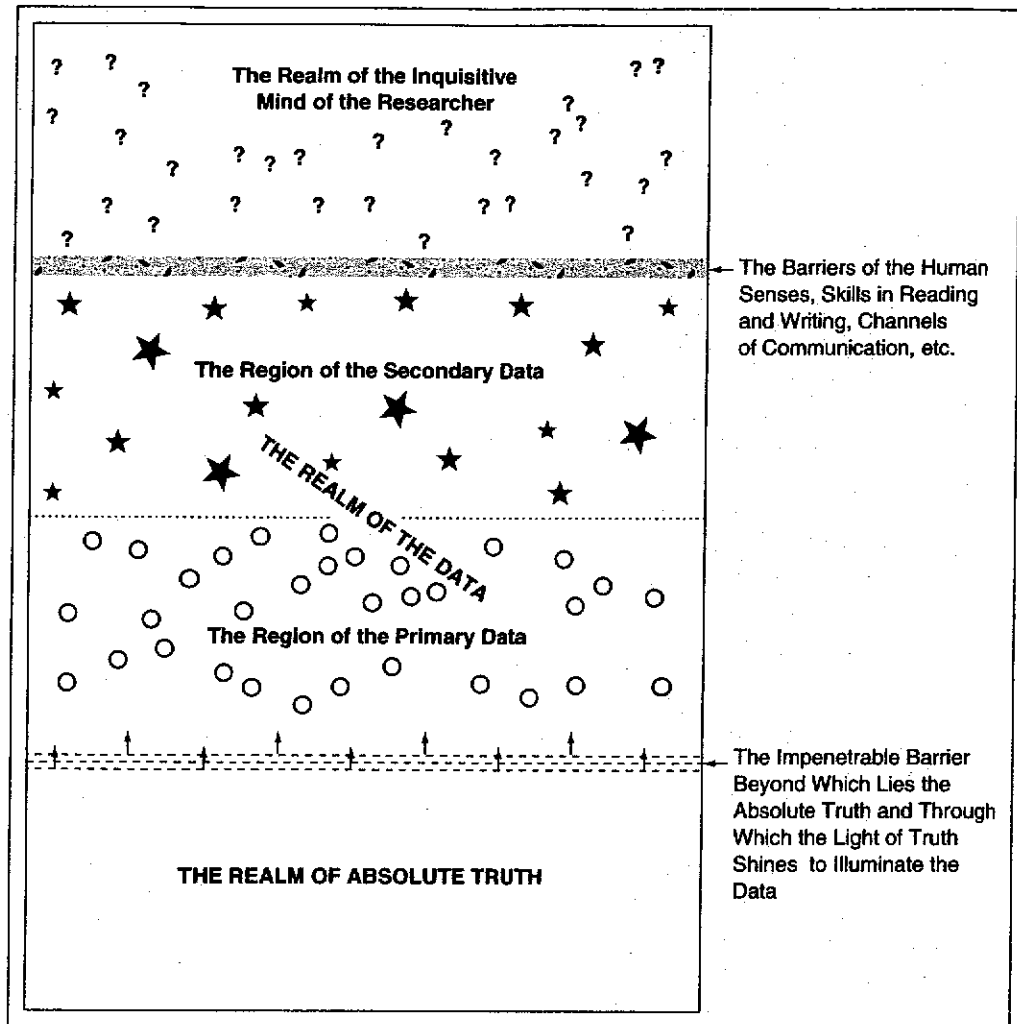
What we have been saying about data and their relation to Truth is represented by Figure 4.2. Lying farthest away from the researcher—and, hence, least accessible—is The Realm of Absolute Truth. It can be approached by the researcher only by passing through two intermediate areas that we have labeled The Realm of the Data. Notice that a barrier exists between The Realm of Absolute Truth and The Region of the Primary Data. Small bits of information leak through the barrier and manifest themselves as data. Notice, too, the foggy barrier between The Realm of the Data and The Realm of the Inquisitive Mind of the Researcher. This barrier is comprised of many things: the distortions and insensitivities of the human senses, the weaknesses of instrumentation, the inability of language to communicate thoughts exactly, the inability of two human beings to witness the same event and report it in precisely the same way, and so on.

Researchers must never forget the overall idea underlying Figure 4.2. Keeping it in mind can prevent them from making exaggerated claims or drawing unwarranted conclusions. No researcher can ever glimpse Absolute Truth, nor can anyone even perceive the data that reflect that Truth except through imperfect senses and imprecise channels of communication. Such a humiliating awareness helps researchers be cautious in the interpretation and reporting of research findings—for instance, by using such words and phrases as *perhaps*, *it seems*, *one might conclude*, *it would appear to be the case*, and *the data are consistent with the hypothesis that*.



**FIGURE 4.2**

The relation between  
data and Truth



## Criteria for the Admissibility of Data

Not all data that come to the researcher's attention are acceptable for use in a research project. Data can be defective. If they are, they may affect the validity of the researcher's conclusions. The imperfections in the data stem from the imperfections and irregularities of nature. If researchers include in the mass of data those that are imperfect or irregular—those that are distorted, inaccurate, hopelessly entangled with irrelevant variables, and so on—they corrupt the entire body of data.

As previously mentioned, any research effort should be *replicable*; that is, it should be able to be repeated by any other researcher at any other time under precisely the same conditions. To ensure such precision, certain criteria must be adopted, certain limits established, and certain standards set up that all data must meet in order to be admitted for study.

For example, imagine that an agronomist wants to determine the effect of ultraviolet light on growing plants. *Ultraviolet* is a vague term: It encompasses a range of light waves that vary considerably in nanometers. The agronomist must narrow the parameters of the data so that they will fall within certain specified limits. What precisely does she or he mean by *ultraviolet light*? Within what nanometer range will ultraviolet emission be acceptable? At what intensity? For what length of time? At what distance from the growing plants? What precisely does the researcher mean by the phrase "effect of ultraviolet light on growing plants"? All plants? A specific genus? A particular species? The agronomist must be quite specific about all of these things so that another researcher can achieve very similar conditions in replicating the study.

By prescribing such criteria and insisting on standards, we can control the types of data admitted and regulate the conditions under which the research effort proceeds. Any data not meeting the criteria are excluded from the study. It is, of course, much easier to control data in the physical sciences, where we can measure a phenomenon fairly objectively, but such control is also possible in the humanities, the social sciences, and the arts. In these latter disciplines, the criteria are usually established by definition, and admissible data must meet the definitive parameters.

We hem in the data on all sides, placing on them the restrictions of criterion after criterion, so that we are able to isolate only those data acceptable for our use. The rest of the data are inadmissible. The restrictions we identify are the **criteria for the admissibility of data**.

When we standardize the data, admitting only those that comply with our criteria, we can better control the research effort and conclude with greater certainty what appears to be true. Therefore, to ensure the integrity of the research, we must set forth *beforehand* precisely what standards the data must meet. These criteria must be set forth clearly in both the research proposal and the final research report. Only in this way can the researcher and the consumer of the research come to a common, intelligent understanding of the investigation at hand.

## Planning for Data Collection

Basic to any research project are four fundamental questions about the data. To avoid serious trouble later on, the researcher must answer these questions specifically and concretely. Clear answers can help bring any research planning and design into focus.

1. *What data are needed?* This question may seem like an overly simple one, but in fact a specific, definitive answer is fundamental to any research effort. On a sheet of paper, write the answers to the following questions: To resolve the problem, what data are mandatory? What is their nature? Are they historical documents? Interview excerpts? Questionnaire responses? Observations? Measurements made before and after an experimental intervention? Specifically, what data do you need, and what are their characteristics?
2. *Where are the data located?* Those of us who have taught courses in research methodology are constantly awed by the fascinating problems that students identify for research projects. But then we ask a basic question: "Where will you get the data to resolve the problem?" Some students either look bewildered and remain speechless or else mutter something such as, "Well, they must be available *somewhere*." Not *somewhere*, but *precisely where*? If you are doing a study of documents, where are the documents you need? At exactly which library and in what collection will you find them? What society or what organization has the files you must examine? Where are these organizations located? Specify geographically—by town, street address, and postal code! Suppose a nurse or a nutritionist is doing a research study about Walter Olin Atwater, whose work has been instrumental in establishing the science of human nutrition in the United States. Where are the data on Atwater located? The researcher can go no further until that basic question is answered.
3. *How will the data be obtained?* To know where the data are located is not enough; you need to know how you might acquire them. With privacy laws, confidentiality agreements, and so on, obtaining the information you need might not be as easy as you think. You may indeed know what data you need and where you can find them, but an equally important question is, How will you get them? Careful attention to this question marks the difference between a viable research project and a pipe dream.
4. *How will the data be interpreted?* This is perhaps the most important question of all. The three former hurdles have been overcome. You have the data in hand. But you must also spell out precisely what you intend to do with the data to solve the research problem or one of its subproblems.

Now go back and look carefully at how you have worded your research problem. Will you be able to get data that might adequately provide a solution to the problem? And if so, might they reasonably lend themselves to interpretations that shed light on the problem? If the answer to either of these questions is *no*, you must obviously rethink the nature of your problem. If, instead, both answers are *yes*, a next important step is to consider how you might best *measure* what you observe.

---

## Identifying Good Measurement Instruments

As you zero in on the nature of the data necessary for solving your research problem, you will probably discover that you must pin down your observations by measuring them in some way. In some cases you will be able to use one or more existing instruments—perhaps an oscilloscope to measure patterns of sound, a published personality test to measure a person's tendency to be either shy or outgoing, or a rating scale that a previous researcher has developed to assess parents' childrearing practices. In other situations you may have to develop your *own* measurement instruments—perhaps a survey to assess people's opinions about welfare reform, a paper-and-pencil test to measure what students have learned from a particular instructional unit, or a checklist to evaluate the quality of a new product.

Measurement instruments provide a basis on which the entire research effort rests. Just as a building with a questionable foundation is unlikely to be safe for habitation, so, too, will a research effort employing faulty measurement tools provide little of value in solving the problem under investigation.

## Defining Measurement

What exactly *is* measurement? Typically we think of measurement in terms of such objects as rulers, scales, gauges, and thermometers. In research, **measurement** takes on a somewhat different meaning:

Measurement is limiting the data of any phenomenon—substantial or insubstantial—so that those data may be interpreted and, ultimately, compared to a particular qualitative or quantitative standard.

Let's look more closely at this definition. The first five words are *measurement is limiting the data*. When we measure something, we set a limit that constrains the data in some way. We erect a barrier beyond which those data cannot go. What is a foot, a mile, a pound? Each is a unit of measure governed by a numerical constraint: 12 inches constrain a foot; 5,280 feet, a mile; and 16 ounces, a pound.

Now let's look at the next six words: *of any phenomenon—substantial or insubstantial*. This phrase is all-inclusive. Nothing exists that the researcher cannot measure in some way. In some cases, observable physical entities are measured. These are **substantial** measurements; that is, the things being measured have physical substance, an obvious basis in the physical world. An astronomer measures patterns and luminosity of light in the night sky; a neurologist measures intensity and location of activity in the brain; a chemist measures the mass of a compound both before and after transforming it in some way. All of these are attempts to measure substantial phenomena. Some devices designed to measure substantial phenomena, such as high-powered telescopes and MRI machines, are highly specialized and used only in particular disciplines. Others, such as balance scales and tape measures, are applicable to many fields of inquiry.

We can also measure those things—if “things” they be—that are **insubstantial**, that exist only as concepts, ideas, opinions, feelings, or other intangible entities. For example, we might attempt to measure the economic “health” of business, the degree to which students have

“learned,” or the extent to which people “value” physical exercise. We seek to measure these intangibles, not with tape measures or scales, but with the Dow-Jones index, achievement tests, questionnaires, or interviews.<sup>2</sup>

We continue with the next seven words of our definition of measurement: *so that those data may be interpreted*. We cannot emphasize this point enough: Research involves not only the collection but also the interpretation of data—the transformation of data into new discoveries, revelations, and enlightenments.

Now we finish our definition: *and, ultimately, compared to a particular qualitative or quantitative standard*. A researcher must have a goalpost, a true north, a point of orientation. In research, we call these standards *norms, averages, conformity to expected statistical distributions, goodness of fit, accuracy of description*, and the like.

Measurement is ultimately a comparison: a thing or concept measured against a point of limitation. We compare the length of an object with the scale of a ruler or a measuring tape. We “measure” an ideology against the meaning of it as articulated by its originator. For example, the essence of a philosophy arises from the writings and teachings of its founder: Platonism from Plato, Marxism from Karl Marx, and romanticism, perhaps, from Jean Jacques Rousseau. The essence of a religious belief lies in its sacred writings, in the precepts of its great teachers, and in its creed. The meaning of freedom is articulated in many political documents—for instance, in the Declaration of Independence and the Constitution of the United States. Against these original sources, it is possible to measure the thoughts and ideas of others and to approximate their similarity to or deviance from those sources.

Data examined statistically are constantly being interpreted in comparison with statistical norms: the normal curve, probability tables, and other accepted statistical standards. Data analyzed in a nonnumerical, qualitative manner are compared across data sources, across methods, and across time.

As you can see, then, our definition of measurement implies much more than an everyday understanding of measurement might suggest. Measurement provides an important tool with which data may be inspected, analyzed, and interpreted so that the researcher may probe the meaning that lies below the surface.

## Measuring Insubstantial Phenomena: An Example

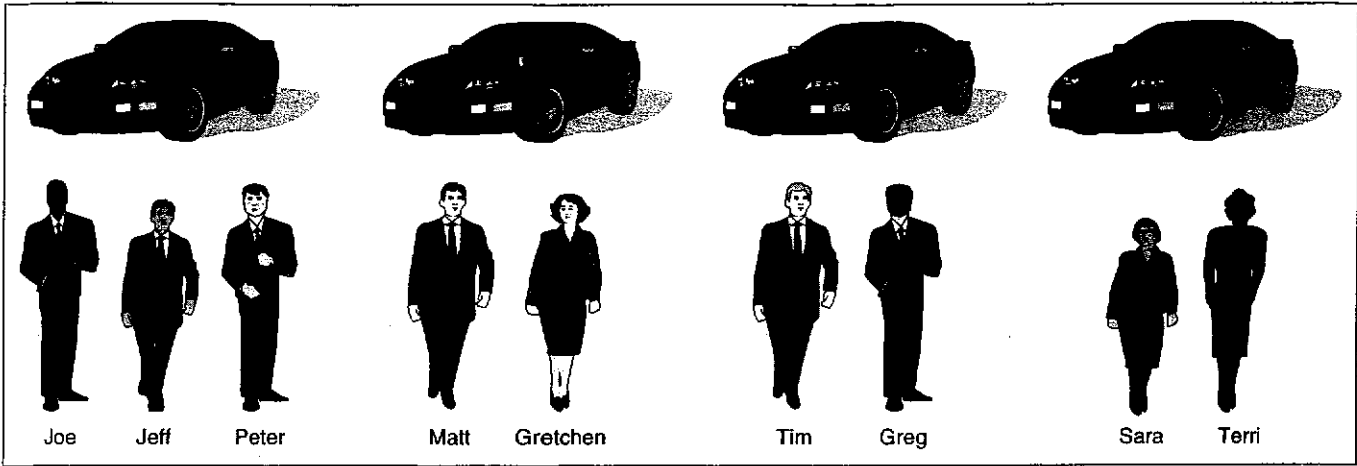
Measuring insubstantial phenomena—those phenomena that have no obvious, concrete basis in the physical world—can sometimes involve considerable creativity on the part of the researcher. For example, imagine that we want to examine—and hence also to *measure*—the interpersonal dynamics among a small group of people. Let’s take a group of nine people, shown in Figure 4.3, who work together in the human resources department of a large corporation. They are to attend a recognition dinner at an exclusive hotel.

They arrive in four cars. They enter the hotel in the following order: Terri, Sara, Greg, Tim, Gretchen, Matt, Peter, Jeff, and Joe. They greet one another and have time for a brief conversation before dinner. Most of them position themselves in conversation groups, as shown in Figure 4.4.

To the perceptive observer, the interpersonal dynamics within the group soon become apparent. Who greets whom with enthusiasm or with indifference? Who joins in conversation with whom? Who seems to be a relative outsider? However, *to merely observe the behavior of individuals in a particular situation is not to measure it*.

One possible approach to measuring the group’s interpersonal dynamics is to give each person in the group a slip of paper on which to record three choices: (a) one or more individuals in

<sup>2</sup>You may sometimes see the substantial–insubstantial distinction referred to as *manifest variables* (which can be directly observed and measured) versus *latent variables* (which lie below the surface and can be measured only indirectly through their effects on another, observable entity; e.g., see Bartholomew, 2004).



**FIGURE 4.3**

Recognition dinner participants

the group whom the person likes most, (b) one or more individuals whom the person likes least, and (c) one or more individuals for whom the person has no strong feeling one way or the other. When using this method, we should poll each person in the group individually and guarantee that every response will be kept confidential.

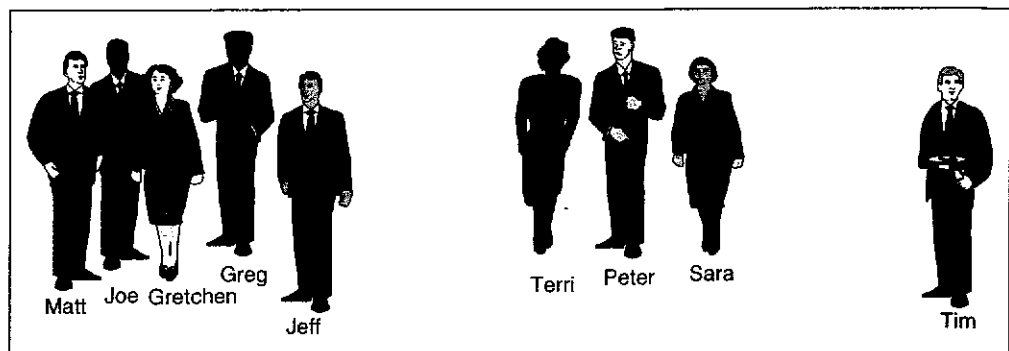
We can then draw a chart, or **sociogram**, of these interpersonal reactions, perhaps in the manner depicted in Figure 4.5. We might also assign “weights” that place the data into three numerical categories: +1 for a positive choice, 0 for indifference, and -1 for a negative reaction. Categorizing the data in this way, we can then construct a sociometric matrix. To create a matrix, we arrange the names of each person twice: vertically down the left side of a grid and horizontally across the top of the grid. The result is shown in Table 4.1. The dashes in the grid reflect the fact that the people can choose other individuals but cannot choose *themselves*.

Certain relationships begin to emerge. As we represent group dynamics in multiple forms, certain clusters of facts suggest the following conclusions:

- Jeff is the informal or popular leader (sometimes called the “star”) of the group. He received five choices and only one rejection (see the “Jeff” column in Table 4.1). The sociogram confirms Jeff’s popularity with his colleagues.
- Probably some factions and possible tension are present in this group. Notice that Peter, Sara, and Terri form a subclique, or “island,” that is separated from the larger clique that Jeff leads. The apparent liaison between these two groups is Joe, who has mutual choices with both Jeff and Peter.
- Friendship pairs may lend cohesion to the group. Notice the mutual choices: Matt and Gretchen, Gretchen and Jeff, Jeff and Joe, Joe and Peter, Peter and Terri, Terri and Sara. The sociogram clearly reveals these alliances.

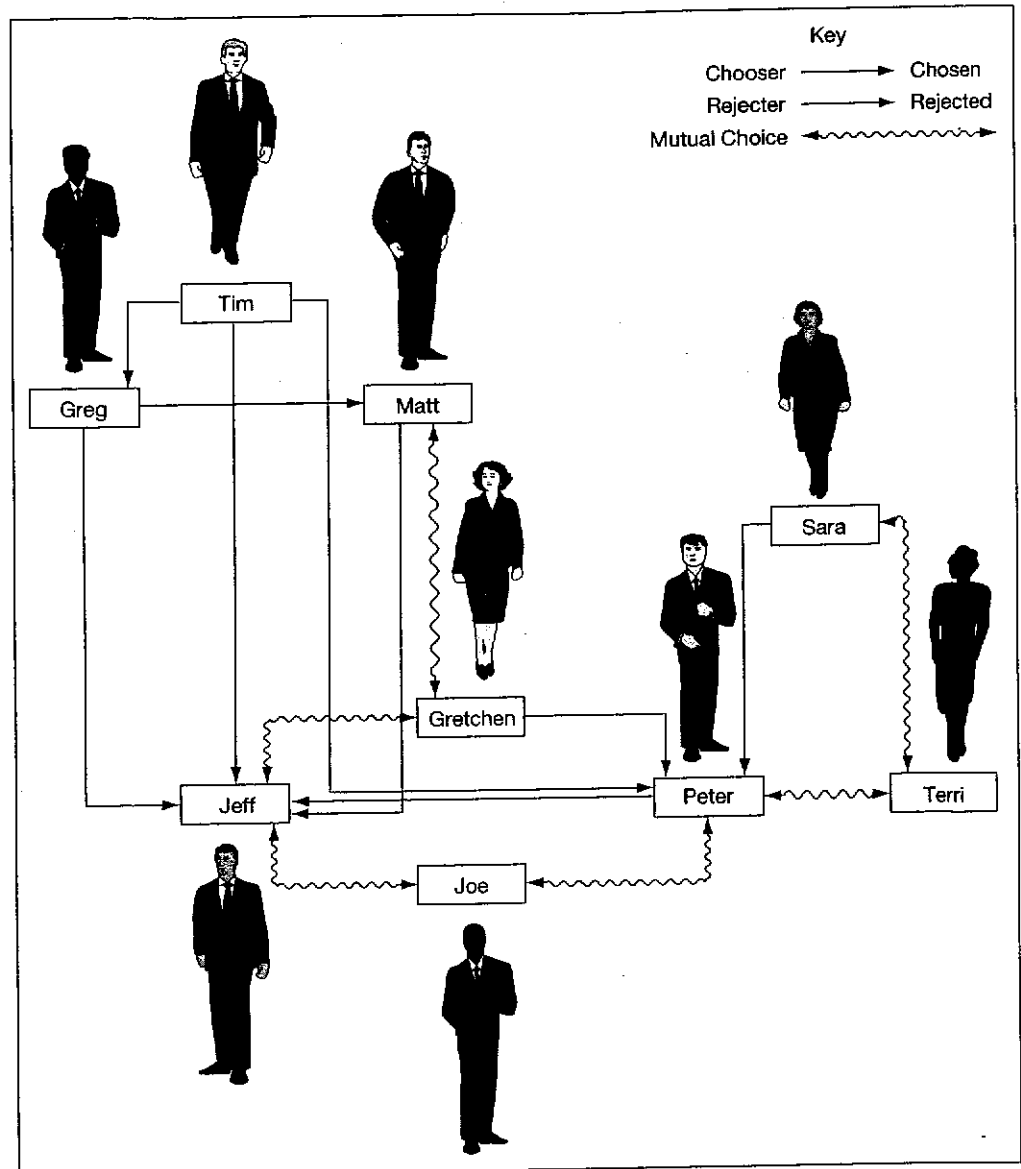
**FIGURE 4.4**

Conversation groups



**FIGURE 4.5**

Sociogram of interpersonal dynamics



■ Tim apparently is the isolate of the group. He received no choices; he is neither liked nor disliked. In such a position, he is probably the least influential member of the group.

With this example we have illustrated what it means to interpret data by measuring an insubstantial phenomenon and analyzing the resulting data. Notice that we did not just observe the behaviors of nine individuals at a social occasion, but we also looked below the surface to discover hidden social forces at play in their behaviors with one another. Our example is a simple one, to be sure. Measurement of interpersonal dynamics and social networks can certainly take more complex forms, including some that are especially helpful in studying social forces within large, extended groups (e.g., Chatterjee & Srivastava, 1982; Freeman, 2004; Wasserman & Faust, 1994).

## Types of Measurement Scales

Virtually any form of measurement falls into one of four categories, or scales: (a) nominal, (b) ordinal, (c) interval, and (d) ratio (Stevens, 1946). The scale of measurement will ultimately dictate the statistical procedures (if any) that can be used in processing the data.

**TABLE 4.1** Data from Figure 4.5 presented as a sociometric matrix

How Each Person Was Rated by the Others

	Gretchen	Joe	Greg	Sara	Peter	Jeff	Tim	Matt	Terri
Gretchen	—	0	0	0	-1	+1	0	+1	0
Joe	0	—	0	0	+1	+1	0	0	0
Greg	0	0	—	0	0	+1	0	+1	0
Sara	0	0	0	—	+1	0	0	0	+1
Peter	0	+1	0	0	—	-1	0	0	+1
Jeff	+1	+1	0	0	0	—	0	0	0
Tim	0	0	+1	0	-1	+1	—	0	0
Matt	+1	0	0	0	0	+1	0	—	0
Terri	0	0	0	+1	+1	0	0	0	—
Totals	2	2	1	1	1	4	0	2	2

### Nominal Scales

The word *nominal* comes from the Latin *nomen*, meaning “name.” Hence we can “measure” data to some degree by assigning names to them. Recall that the definition of measurement presented earlier includes the phrase *limiting the data*. That is what a **nominal scale** does—it limits the data—and just about all that it does. Assign a specific name to anything, and you have restricted that thing to the meaning of its name. For example, we can measure a group of children by dividing it into two groups: girls and boys. Each subgroup is thereby measured—restricted—by virtue of gender as belonging to a particular category.

Things can be measured nominally in an infinite number of ways. We can further measure girls and boys according to the home site of each child. Imagine that the town in which the children live is divided into two sections by Main Street, which runs from east to west. Those children who live north of Main Street are “the Northerners”; those who live south of it are “the Southerners.” In one period of U.S. history, we measured the population of the entire nation in just such a manner.

Nominal measurement is quite simplistic, but it does divide data into discrete categories that can be compared with one another. Let’s take an example. Imagine that we have six children: Zahra, Paul, Kathy, Binh, Ginger, and Nicky. They can be divided into six units of one child each. They can also form two groups: Zahra, Kathy, and Ginger (the girls) in one group and Paul, Binh, and Nicky (the boys) in the other. Perhaps all six children are students in a class that meets in Room 12 at Thompson’s Corner School. By assigning a room number, we have provided the class with a name, even though that “name” is a number. In this case, the number has no quantitative meaning: Room 12 is not necessarily bigger or better than Room 11, nor is it inferior to Room 13.

Only a few statistical procedures are appropriate for analyzing nominal data. We can use the *mode* as an indicator of the most frequently occurring category within our data set; for example, we might determine that there are more boys than girls in Room 12 at Thompson’s Corner School. We can find the *percentage* of people in various subgroups within the total group; for example, we could calculate the percentage of boys in each classroom. We can use a *chi-square test* to compare the relative frequencies of people in various categories; for example, we might discover that more boys than girls live north of Main Street but that more girls than boys live south of Main Street. (We discuss these statistics, as well as the statistics listed in the following discussions of the other three scales, in Chapter 11.)

### Ordinal Scales

With an **ordinal scale** of measurement, we can think in terms of the symbols  $>$  (greater than) and  $<$  (less than). We can compare various pieces of data in terms of one being greater or higher than another. In essence, this scale allows us to *rank-order* our data—hence its name *ordinal*.

As an example, we can roughly measure level of education on an ordinal scale by classifying people as being unschooled or having completed an elementary, high school, college, or graduate education. Likewise, we can measure members of the workforce by grades of proficiency: unskilled, semiskilled, or skilled.

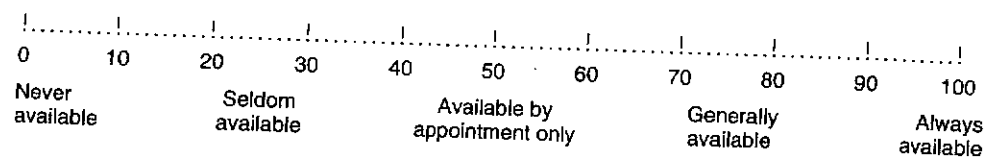
An ordinal scale expands the range of statistical techniques we can apply to our data. In addition to the statistics we can use with nominal data, we can also determine the *median*, or halfway point, in a set of data. We can use a *percentile rank* to identify the relative position of any item or individual in a group. We can determine the extent of the relationship between two characteristics by means of Spearman's *rank order correlation*.

### Interval Scales

An **interval scale** of measurement is characterized by two features: (a) It has equal units of measurement and (b) its zero point has been established arbitrarily. The Fahrenheit (F) and Celsius (C) scales for measuring temperature are examples of interval scales: The intervals between any two successive numbers of degrees reflect equal changes in temperature, but the zero point is not equivalent to a total absence of heat. For instance, when Gabriel Fahrenheit was developing his Fahrenheit scale, he first took as his zero point the coldest temperature he observed in Iceland. Later, he made it the lowest temperature obtainable with a mixture of salt and ice. This was purely an arbitrary decision. It placed the freezing point of water at  $32^{\circ}$  and the boiling point at  $212^{\circ}$  above zero.

The rating scales employed by many survey groups, businesses, and professional organizations are often assumed to be on interval scales. For instance, many universities ask students to use rating scales to evaluate the teaching effectiveness of various professors. Following is an example of an item from one university's teaching evaluation form:

Place an X on the scale below at the point where you would rate the availability of your professor for conferences.



Notice that the scale has 11 equidistant points ranging from 0 to 100. The equidistance creates what is presumed to be an interval scale for the measure. At five points along the scale are descriptive labels that can help students determine how they should rate their professor's availability. We might place descriptors at more places along the scale (perhaps at 10-point distances), thus potentially making the scale more sensitive or more accurate. For indicating the availability of a professor, such fineness of discrimination may not be either possible or desirable, but one can conceive of situations in which such a degree of discrimination may be necessary and appropriate.

Interval scales of measurement allow statistical analyses that are not possible with nominal or ordinal data. Because an interval scale reflects equal distances among adjacent points, any statistics that are calculated using addition or subtraction—for instance, *means*, *standard deviations*, and *Pearson product moment correlations*—can now be used.

### Ratio Scales

Two commonly used measurement instruments—a thermometer and a yardstick—might help you understand the difference between the interval and ratio scales. If we have a thermometer that measures temperature on the Fahrenheit scale, we *cannot* say that  $80^{\circ}\text{F}$  is twice as warm as



40°F. Why? Because this scale does not originate from a point of absolute zero; a substance may have some degree of heat even though its measured temperature falls *below zero*. With a yardstick, however, the beginning of linear measurement is absolutely the beginning. If we measure a desk from the left edge to the right edge, that's it. There is no more desk in either direction beyond those limits. A measurement of "zero" means there is no desk at all, and a "minus" desk width isn't even possible.

More generally, a **ratio scale** has two characteristics: (a) equal measurement units (similar to an interval scale) and (b) *an absolute zero point*, such that 0 on the scale reflects a total absence of the entity being measured.

Let's consider once again the "availability" scale presented earlier for measuring professor effectiveness. This scale could never be considered a ratio scale. Why? Because there is only one condition in which the professor would be absolutely unavailable—if the professor were dead!—in which case we would not be asking students to evaluate this individual.

What distinguishes the ratio scale from the other three scales is that *the ratio scale can express values in terms of multiples and fractional parts*, and the ratios are *true ratios*. A yardstick can do that: A yard is a *multiple* (by 36) of a 1-inch distance; an inch is *one-twelfth* (a *fractional part*) of a foot. The ratios are 36:1 and 1:12, respectively.

Ratio scales outside the physical sciences are relatively rare. And whenever we cannot measure a phenomenon in terms of a ratio scale, we must refrain from making comparisons such as "this thing is three times as great as that" or "we have only half as much of one thing as another." Only ratio scales allow us to make comparisons that involve multiplication or division.

We can summarize our description of the four scales this way:

If you can say that

- One object is different from another, you have a *nominal scale*;
- One object is bigger or better or more of anything than another, you have an *ordinal scale*;
- One object is so many units (degrees, inches) more than another, you have an *interval scale*;
- One object is so many times as big or bright or tall or heavy as another, you have a *ratio scale*.

(Senders, 1958, p. 51)

Table 4.2 provides a quick reference for the various types of scales, their distinguishing characteristics, and the statistical analysis possibilities for each scale. When we consider the statistical interpretation of data in later chapters (especially in Chapter 11), you may want to refer to this table to determine whether the type of measurement instrument you have used will support the statistical operation you are contemplating.

**TABLE 4.2** A summary of measurement scales, their characteristics, and their statistical implications

	Measurement Scale	Characteristics of the Scale	Statistical Possibilities of the Scale
Non-Interval Scales	Nominal scale	A scale that "measures" only in terms of names or designations of discrete units or categories	Enables one to determine the mode, percentage values, or chi-square
	Ordinal scale	A scale that measures in terms of such values as "more" or "less," "larger" or "smaller," but without specifying the size of the intervals	Enables one also to determine the median, percentile rank, and rank correlation
Interval Scales	Interval scale	A scale that measures in terms of equal intervals or degrees of difference, but with an arbitrarily established zero point that does not represent "nothing" of something	Enables one also to determine the mean, standard deviation, and product moment correlation; allows one to conduct most inferential statistical analyses
	Ratio scale	A scale that measures in terms of equal intervals and an absolute zero point	Enables one also to determine the geometric mean and the percentage variation; allows one to conduct virtually any inferential statistical analysis

## CONCEPTUAL ANALYSIS EXERCISE Identifying Scales of Measurement

Each of the following scenarios involves measuring one or more variables. Decide whether the various measurements reflect nominal, ordinal, interval, or ratio scales, and justify your choices. Be careful, as the answers are not always as obvious as they might initially appear. The answers are provided after the "For Further Reading" list at the end of the chapter.

1. An environmental scientist collects water samples from streams and rivers near large industrial plants and saves exactly one liter of water from each sample. Then, back at the lab, the researcher determines the amounts of certain health-jeopardizing bacteria in each sample. What measurement scale does the measurement of bacteria content reflect?
2. A tourism researcher is studying the relationship between (a) a country's average annual temperature and (b) the amount of tourist dollars that the country brings in every year. What scales underlie the two variables in this study?
3. A political science researcher wants to determine whether people's political party membership is correlated with the frequency with which they have voted in local elections in the past five years. The researcher can easily obtain information about people's party membership and voting records from the town clerks in several communities. To simplify data collection, the researcher uses the following coding scheme for party membership: 1 = Registered as Democrat, 2 = Registered as Republican, 3 = Registered as member of another party, 0 = No declared party affiliation. What measurement scale(s) underlie (a) political party membership and (b) voting frequency?
4. A marketing researcher in the United States wants to determine whether a certain product is more widely used in some parts of the country than others. The researcher separates the country into ten regions based on zip code; zip codes below 10000 are northeastern states, zip codes of 90000 and above are western states, and so on. What measurement scale does the researcher's coding scheme for the regions represent?
5. An economist is studying the home-buying behaviors of people of different income levels. The researcher puts people into four categories: Group A includes those earning up to \$20,000 per year, Group B includes those earning between \$20,001 and \$50,000 per year, Group C includes those earning between \$50,001 and \$100,000 per year, and Group D includes those earning more than \$100,000 per year. In this study, what kind of scale is income level?
6. A geographer is studying traffic patterns on four different types of roads that vary in quality: superhighways (i.e., roads accessible only by relatively infrequent on-off ramps), highways (i.e., roads that allow relatively high speeds for long distances but may have an occasional traffic light), secondary roads (i.e., well-paved two-lane roads), and tertiary roads (narrow, infrequently traveled roads; some may consist only of gravel). The type of road in this study reflects which type of measurement scale?
7. A psychologist is developing an instrument designed to measure college students' test anxiety. The instrument includes 25 statements—for example, "My heart starts to pound when I see the word *test* on a course syllabus" and "My palms get sweaty while I'm taking a multiple-choice test." Students must rate each of these statements on a 5-point scale, as follows:
  - 0 This is never true for me.
  - 1 This is rarely true for me.
  - 2 This is sometimes true for me.
  - 3 This is often true for me.
  - 4 This is always true for me.

Students who answer "never" to each of the 25 questions get the lowest possible score of 0 on the instrument. Students who answer "always" to each of the 25 questions get the highest possible score of 100 on the instrument. Thus, scores on the instrument range from 0 to 100. What kind of scale do the scores represent?

## Validity and Reliability in Measurement

Regardless of the type of scale a measurement instrument involves, the instrument must have both validity and reliability for its purpose. The validity and reliability of measurement instruments influence the extent to which a researcher can learn something about the phenomenon under investigation, the probability that the researcher will obtain statistical significance in any data analysis, and the extent to which the researcher can draw meaningful conclusions from the data.

In your research proposal and final research report, you should provide evidence that the instruments you use have a reasonable degree of validity and reliability for your purposes. But validity and reliability take different forms, depending on the nature of the research problem, the methodology being used to address the problem, and the nature of the data that are collected. In the following sections, we briefly discuss the kinds of validity and reliability you may need to consider and how you might determine and enhance them.

### Validity of Measurement Instruments

The validity of a measurement instrument is the extent to which the instrument measures what it is intended to measure. Certainly no one would question the premise that a yardstick is a valid means of measuring length. Nor would most people doubt that a thermometer measures temperature; for instance, in a mercury thermometer, the level to which the mercury rises is a function of how much it expands, which is a function of the degree to which it is hot or cold.

But to what extent does an intelligence test actually measure a person's intelligence? How accurately do people's annual incomes reflect their social class? And how well does a sociogram capture the interpersonal dynamics in a group of nine people? Especially when we are measuring *insubstantial* phenomena—phenomena without a direct basis in the physical world—our measurement instruments may be somewhat suspect in terms of validity.

Let's return to the rating-scale item presented earlier to assess a professor's availability to students (see p. 86) and consider its validity as such a measure. Notice how fuzzy some of the labels are. The professor is "always available." What does *always* mean? Twenty-four hours a day? Could you call the professor at 3:00 a.m. any day of the week, or only whenever the professor is on campus? If the latter is the case, could you call your professor out of a faculty meeting or out of a conference with the college president? We might have similar problems in interpreting "generally available," "seldom available," and "never available." On careful inspection, what seems at first glance to be a scale that anyone could understand has limitations *as a measurement instrument* for research purposes.

A paper-and-pencil test may be intended to measure a certain characteristic, and it may be *called* a measure of that characteristic, but these facts don't necessarily mean that the test actually measures what its creator says it does. For example, consider a paper-and-pencil test of personality traits in which, with a series of check marks, a person indicates his or her most representative characteristics or behaviors in given situations. The person's responses on the test are presumed to reveal relatively stable personality traits. The question that validity asks is: Does such a test, in fact, measure the person's personality traits, or does it measure something else altogether? The answer depends, at least in part, on the extent to which the person is or *can* be truthful in responding. If the person responds in terms of characteristics and behaviors that he or she believes to be socially desirable, the test results may reveal not the person's actual personality, but rather an idealized portrait of how he or she would like to be perceived by others.

The validity of a measurement instrument can take several different forms, each of which is important in different situations:

- **Face validity** is the extent to which, on the surface, an instrument *looks like* it is measuring a particular characteristic. Face validity is often useful for ensuring the cooperation of people who are participating in a research study. But because it relies entirely on subjective judgment, it is not, in and of itself, a terribly dependable indicator that an instrument is truly measuring what the researcher wants to measure.
- **Content validity** is the extent to which a measurement instrument is a representative sample of the content area (domain) being measured. Content validity is often a consideration when

a researcher wants to assess people's *achievement* in some area—for instance, the knowledge students have learned during classroom instruction or the job skills workers have acquired in a training program. A measurement instrument has high content validity if its items or questions reflect the various parts of the content domain in appropriate proportions and if it requires the particular behaviors and skills that are central to that domain.

- **Criterion validity** is the extent to which the results of an assessment instrument correlate with another, presumably related measure (the latter measure is, in this case, the *criterion*). For example, a personality test designed to assess a person's shyness or outgoingness has criterion validity if its scores correlate with other measures of a person's general sociability. An instrument designed to measure a salesperson's effectiveness on the job should correlate with the number of sales the individual actually makes during the course of a business week.
- **Construct validity** is the extent to which an instrument measures a characteristic that cannot be directly observed but is assumed to exist based on patterns in people's behavior (such a characteristic is a *construct*). Motivation, creativity, racial prejudice, love—all of these are constructs, in that none of them can be *directly* observed and measured. When researchers ask questions, present tasks, or observe behaviors as a way of assessing an underlying construct, they should obtain some kind of evidence that their approach does, in fact, measure the construct in question.

Sometimes there is universal agreement that a particular instrument provides a valid instrument for measuring a particular characteristic; such is the case for yardsticks, thermometers, barometers, oscilloscopes, and so on. But whenever we do *not* have such widespread agreement, we must provide evidence that an instrument we are using has validity for our purpose.

It is critical to note that the validity of any measurement instrument can vary considerably depending on the purpose for which it is being used. In other words, *the validity of an instrument is specific to the situation*. For example, a tape measure wrapped horizontally around a person's head is a valid measure of the person's head circumference but *not* a valid measure of the person's intelligence. Likewise, a widely used intelligence test might provide a reasonable estimation of children's general cognitive development but is *not* suitable for determining how well the children can perform in, say, a geometry class or interpersonal conflict.

**Determining the validity of a measurement instrument.** An in-depth discussion of how to determine validity is beyond the scope of this book; measurement textbooks, such as those listed in the "For Further Reading" section at the end of the chapter, provide more detailed information. But here we offer three examples of what researchers sometimes do to demonstrate that their measurement instruments have validity for their purposes:

- **Table of specifications.** To construct a measurement instrument that provides a representative sample of a particular content domain—in other words, to establish content validity—the researcher often constructs a two-dimensional grid (*table of specifications*) listing the specific topics and behaviors that reflect achievement in the domain. In each cell of the grid, the researcher indicates the relative importance of each topic-behavior combination. He or she then develops a series of tasks or test items that reflects the various topics and behaviors in appropriate proportions.
- **Multitrait-multimethod approach.** Two or more different characteristics are each measured using two or more different approaches (Campbell & Fiske, 1959). The different measures of the same characteristic should be highly correlated. The same ways of measuring different characteristics should *not* be highly correlated. For example, in a classroom situation, the constructs *achievement motivation* and *social motivation* might each be measured by both self-report questionnaires and teacher observation checklists. Statistical analyses should reveal that the two measures of achievement motivation are highly correlated and that the two measures of social motivation are also highly correlated. Results from the two self-report questionnaires—because they are intended to assess different and presumably unrelated characteristics—should *not* be highly correlated, nor should results from the two teacher checklists.

- **Judgment by a panel of experts.** Several experts in a particular area are asked to scrutinize an instrument and give an informed opinion about its validity for measuring the characteristic in question.

Although none of the approaches just described guarantees the validity of a measurement instrument, each one increases the likelihood of such validity.

## Reliability of Measurement Instruments

Imagine that you are concerned about your growing waistline and decide to go on a diet. Every day you put a tape measure around your waist and pull the two ends together snugly to get a measurement. But just how tight is “snug”? Quite possibly, the level of snugness might be different from one day to the next. In fact, you might even measure your waist with different degrees of snugness from one *minute* to the next. To the extent that you are not measuring your waist in a consistent fashion—even though you always use the same tape measure—you have a problem with reliability.

More generally, **reliability** is the consistency with which a measuring instrument yields a certain, consistent result when the entity being measured hasn't changed. As we have just seen in our waist-measuring situation, instruments that measure physical phenomena aren't necessarily completely reliable. As another example, think of a measuring cup that a baker might use while making a cake. When measuring a half-cup of flour, the baker won't always measure *exactly* the same amount of flour each time.

Instruments designed to measure social and psychological characteristics (insubstantial phenomena) tend to be even less reliable than those designed to measure physical (substantial) phenomena. For example, a student using the rating-scale item presented earlier for measuring professor availability might easily rate the professor as “70” one day and “90” the next, not because the professor's availability has changed overnight but because the student's interpretations of the phrases “generally available” and “always available” *have* changed. Similarly, if we asked the nine people portrayed in Figure 4.3 (Gretchen, Joe, Greg, etc.) to indicate the people they liked best and least among their colleagues, they wouldn't necessarily always give us the same answers they gave us previously, even if the interpersonal dynamics within the group have remained constant.

**Determining the reliability of a measurement instrument.** Like validity, reliability takes different forms in different situations. But in the case of reliability, its particular form is essentially equivalent to the procedure used to determine it. Following are four forms of reliability that are frequently of interest in research studies:

- **Interrater reliability** is the extent to which two or more individuals evaluating the same product or performance give identical judgments.
- **Test-retest reliability** is the extent to which a single instrument yields the same results for the same people on two different occasions.
- **Equivalent forms reliability** is the extent to which two different versions of the same instrument (e.g., “Form A” and “Form B” of a scholastic aptitude test) yield similar results.
- **Internal consistency reliability** is the extent to which all of the items within a single instrument yield similar results.

For each of these forms, determining reliability involves two steps:

1. Getting two measures for each individual in a reasonably large group of individuals—in particular by doing one of the following:
  - a. Having two different raters evaluate the same performance for each individual (interrater reliability)
  - b. Administering the same instrument to the individuals at two different points in time—perhaps a day, a week, or a month apart (test-retest reliability)

- c. Giving each individual two parallel versions of the same instrument (equivalent forms reliability)
  - d. Administering only one instrument but calculating two subscores for the instrument—for instance, calculating one score for odd-numbered items and another score for even-numbered items (internal consistency reliability)
2. Calculating a correlation coefficient that expresses the degree to which the two measures are similar (see Chapter 11 for a discussion of correlation coefficients)

You can find more in-depth discussions about determining reliability in almost any general measurement textbook.

### Enhancing the Reliability and Validity of a Measurement Instrument

Both validity and reliability reflect the degree to which we may have *error* in our measurements. In many instances—and especially when we are measuring insubstantial phenomena—a measurement instrument may allow us to measure a characteristic only indirectly and so may be subject to a variety of biasing factors (e.g., people's responses on a rating scale are apt to be influenced by their interpretations, prejudices, memory lapses, etc.). In such cases, we have error due to the imperfect *validity* of the measurement instrument. Yet typically—even when we are measuring substantial phenomena—we may get slightly different measures from one time to the next simply because our measurement tool is imprecise (e.g., the waist or head size we measure may depend on how snugly we pull the tape measure). In such cases, we have error due to the imperfect *reliability* of the measure. Generally speaking, validity errors reflect biases in the instrument itself and are relatively constant sources of error. In contrast, reliability errors reflect *use* of the instrument and are apt to vary unpredictably from one occasion to the next.

We can measure something accurately only when we can also measure it consistently. Hence, by increasing the reliability of a measurement instrument, we might also increase its validity. A researcher can enhance the reliability of a measurement instrument in several ways. First, the instrument should always be administered in a consistent fashion. In other words, there should be *standardization* in use of the instrument from one situation or individual to the next. Second, to the extent that subjective judgments are required, specific *criteria* should be established that dictate the kinds of judgments the researcher makes. And third, any research assistants who are using the instrument should be *well trained* so that they obtain similar results for any single individual or situation being measured.

Yet even if we enhance the reliability of our measurements, we don't necessarily increase their accuracy. In other words, *reliability is a necessary but insufficient condition for validity*. For example, we could use a tape measure to measure a person's head circumference and claim that the result is a good reflection of intelligence. In this situation, we might have reasonable reliability—we are apt to get similar measures of an individual's head circumference on different occasions—but absolutely no validity. Head size is *not* a good indication of intelligence level.

Creative researchers use a wide variety of strategies to enhance the validity of their measurement instruments. One important strategy is to consult the literature in search of measurement techniques that other researchers have effectively used. Another is to show a first draft of an instrument to experienced colleagues and ask for their feedback and suggestions. Still another strategy is to conduct one or more *pilot studies* specifically to try out a particular instrument, carefully scrutinizing it for obvious or possible weaknesses, and then modify it in minor or major ways.

We cannot overemphasize the importance of determining and maximizing the validity and reliability of your measurement instruments. Without reasonably valid and reliable measures of the characteristics and phenomena under investigation, you cannot possibly obtain informative and useful data for addressing and solving your research problem.

As you plan your research project, you should clearly identify the nature of the measurement instruments you will use and carefully scrutinize them with respect to their potential validity and reliability. Furthermore, in your research proposal and final research report, you

should describe any instrument in explicit, concrete terms. For example, if you are using a particular piece of equipment to measure a certain physical characteristic or phenomenon, you should describe the equipment's specific nature (e.g., its manufacturer, model number, and level of precision). And if you are assessing some aspect of human thought or behavior, you should describe the questions asked or tasks administered, the overall length of the instrument (e.g., number of items, time required for administration), and the method of scoring responses.

## CONCEPTUAL ANALYSIS EXERCISE Identifying Problems with Validity and Reliability in Measurement

In each of the scenarios in this exercise, a researcher encounters a measurement problem. Some of the scenarios reflect a problem with the validity of a measure. Others reflect a problem with a measure's reliability—a problem that indirectly also affects the measure's validity. For each scenario, choose the most obvious problem from among the following alternatives:

- |                      |                                    |
|----------------------|------------------------------------|
| ■ Face validity      | ■ Interrater reliability           |
| ■ Content validity   | ■ Test-retest reliability          |
| ■ Criterion validity | ■ Equivalent forms reliability     |
| ■ Construct validity | ■ Internal consistency reliability |

The answers appear after the "For Further Reading" list at the end of this chapter.

1. After using two different methods for teaching basic tennis skills to non-tennis-playing adults, a researcher assesses the effectiveness of the two methods by administering a true-false test regarding the rules of the game (e.g., faults and double-faults, scoring procedures).
2. A researcher writes 120 multiple-choice questions to assess middle school students' general knowledge of basic world geography (e.g., what the equator is, where Africa is located). To minimize the likelihood that students will cheat on the test by copying one another's answers, the researcher divides the questions into three different sets to create three 40-item tests. In collecting data, the researcher distributes the three tests randomly to students in any single classroom. After administering the tests to students at many different middle schools, the researcher computes the students' test scores and discovers that students who answered one particular set of 40 questions scored an average of 3 points higher than students who answered either of the other two 40-question sets.
3. In order to determine what kinds of situations provoke aggression in gorillas, two researchers observe mountain gorillas in the Virunga Mountains of northwestern Rwanda. As they watch a particular gorilla family and take notes about family members' behaviors, the researchers often disagree about whether certain behaviors constitute "aggression" or, instead, reflect more peaceful "assertiveness."
4. A researcher uses a blood test to determine people's overall energy level after drinking or not drinking a can of a high-caffeine cola drink. Unfortunately, when two research assistants independently rate people's behaviors for energy level for a 4-hour period after drinking the cola, their results don't seem to have any correlation with the blood-test results.
5. In a two-week period during the semester, a researcher gains entry into several college classrooms in order to administer a short survey regarding college students' beliefs about climate change. The survey consists of 20 statements about climate change (e.g., "Devastating floods in recent years are partly the result of the earth's gradually rising overall temperature"), to which students must respond "Strongly disagree," "Disagree," "Agree," or "Strongly agree." Many of the students voluntarily put their names on their surveys. Thanks to the names on many survey forms, the researcher discovers that a few

students were in two of the classes surveyed and thus completed the survey twice. Curiously, however, these students sometimes gave different responses to particular statements on the two different occasions, and hence their overall scores were also different.

6. In order to get a sense of how harmonious most long-term marriages are, a researcher administers a questionnaire to married couples who have been married for at least 20 years. The questionnaire consists of 60 statements to which both spouses must individually respond either "This describes my marriage" or "This doesn't describe my marriage." All 60 statements describe a possible characteristic of a nonharmonious marriage (e.g., "We fight all the time," "We rarely agree about how to spend our money"), and the researcher has sequenced them in a random order on the questionnaire. Even so, the researcher discovers that respondents more frequently agree with the first 30 items than with the last 30 items. If one were to look only at responses to the first 30 items, then, one would think that married couples fight a lot. But if one were to look only at responses to the last 30 items, one would conclude that most long-term couples live in relative peace and harmony. (*Note:* We recommend that questionnaires *not* be slanted in a one-way direction, as this one is; see the "Constructing a Questionnaire" guidelines in Chapter 8).
7. A researcher develops and uses a questionnaire intended to measure the extent to which college students display tolerance toward a particular religious group. However, several experts in the researcher's field of study suggest that the questionnaire measures not how tolerant students actually *are*, but what students would like to *believe* about their tolerance for people of a particular religion.
8. Students in an introductory college psychology course must satisfy their "research methods" requirement in one of several ways; one option is to participate in a research study called "Intelligence and Motor Skill Learning." When students choosing this option report to the laboratory, one of their tasks is to respond as quickly as possible to a series of simple computer-generated questions. Afterward, the researcher debriefs the students about the nature of the study and tells them that the reaction-time measure was designed to be a simple measure of intelligence. Some of the students object, saying, "That's not a measure of intelligence! Intelligence isn't how quickly you can do something, it's how *well* you can do it."

---

## Linking Data and Research Methodology

Earlier in the chapter we described data as a link between Absolute Truth and the researcher's inquiring mind. Data are like ore: They contain pieces of the truth but are in a rather unrefined state. To extract meaning from the data, we employ what is commonly called *research methodology*.

Data and methodology are inextricably intertwined. For this reason, the methodology to be used for a particular research problem must always take into account the nature of the data that will be collected in the resolution of the problem.

An example may help clarify this point. Imagine that a man from a remote village decides to travel to the big city. While he is there, he takes his first ride on a commercial jetliner. No one in his village has ever flown before, so naturally when he returns home, they are curious about his trip. One afternoon two friends ask him about his experience, yet each one asks very different questions. The first friend asks, "How fast did you move?" "How far did you go?" and "How high did you fly?" The second friend asks, "How did you feel when you were moving so fast?" "What was it like being above the clouds?" and "What did the city look like from so high?" Both friends are asking questions that can help them learn more about the experience of flying in an airplane. Yet they differ considerably in what they want to know. Because they ask different kinds of questions, they obtain different kinds of information. If they were to recount what they had learned about "flying in a jetliner," they would most likely describe very



different things. Although neither of them has the “wrong” story, neither does each one have the whole story.

In research, too, different questions yield different types of information. Different research problems lead to different research designs and methods, which in turn result in the collection of different types of data and different interpretations of those data.

Furthermore, many kinds of data may be suitable only for a particular methodology. To some extent, *the data dictate the research method*. As an example, consider historical data, those pieces of information gleaned from written records of past events. You cannot extract much meaning from historical documents by using a laboratory experiment. An experiment is simply not suited to the nature of the data.

Over the years, numerous research methodologies have emerged to accommodate the many different forms that data are likely to take. Accordingly, we must take a broad view of the approaches the term *research methodology* encompasses. Above all, we must not limit ourselves to the belief that only a true experiment constitutes “research.” Such an attitude prohibits us from agreeing that we can better understand Coleridge’s poetry by reading the scholarly research of John Livingston Lowes (1927, 1955) or from appreciating Western civilization more because of the historiography of Arnold Toynbee (1939–1961).

No single highway leads us exclusively toward a better understanding of the unknown. Many highways can take us in that direction. They may traverse different terrain, but they all converge on the same destination: the enhancement of human knowledge.

In Chapters 6 through 10 of this book we zero in on various research methodologies. But many researchers tend to categorize research studies into two broad categories: quantitative research and qualitative research. We now look at this distinction.

## Comparing Quantitative and Qualitative Approaches

As you might guess, **quantitative research** involves looking at amounts, or *quantities*, of one or more variables of interest. A quantitative researcher typically tries to measure variables in some numerical way, perhaps by using commonly accepted measures of the physical world (e.g., rulers, thermometers, oscilloscopes) or carefully designed measures of psychological characteristics or behaviors (e.g., tests, questionnaires, rating scales).

In contrast, **qualitative research** involves looking at characteristics, or *qualities*, that cannot be entirely reduced to numerical values. A qualitative researcher typically aims to examine the many nuances and complexities of a particular phenomenon. You are most likely to see qualitative research in studies of complex human situations (e.g., people’s in-depth perspectives about a particular issue, the behaviors and values of a particular cultural group) or complex human creations (e.g., television commercials, works of art). Qualitative research is not limited to research problems involving human beings, however. For instance, some biologists study, in a distinctly qualitative manner, the complex social behaviors of other animal species; Dian Fossey’s work with gorillas and Jane Goodall’s studies of chimpanzees are two well-known examples (e.g., see Fossey, 1983; Goodall, 1986).

Quantitative and qualitative approaches involve similar processes—for instance, they both entail identifying a research problem, reviewing related literature, and collecting and analyzing data. Yet these processes are often combined and carried out in different ways, leading to distinctly different research methods. For instance, quantitative researchers often start with one or more specific hypotheses to be tested. They isolate the variables they want to study, use a standardized procedure to collect some form of numerical data, and use statistical procedures to analyze and draw conclusions from the data. In contrast, qualitative researchers often start with general research questions rather than specific hypotheses, collect an extensive amount of verbal data and/or nonverbal artifacts, organize those data and artifacts into some form that gives them coherence, and use verbal descriptions to portray the situation they have studied.

To some extent, quantitative and qualitative research designs are appropriate for answering different kinds of questions. As a result, we learn more about the world when we have both

quantitative and qualitative methodologies at our disposal than when we are limited to only one approach or the other.

Let's consider how the two approaches might look in practice. Suppose two researchers are interested in investigating the "effectiveness of the case-based method of teaching business management practices." The first researcher asks the question, "How effective is case-based instruction in comparison with lecture-based instruction?" She finds five instructors who are teaching case-based business management classes; she finds five other instructors who are teaching the same content using lectures. At the end of the semester, the researcher administers an achievement test to students in all 10 classes. Using statistical analyses, she compares the scores of students in case-based and lecture-based courses to determine whether the achievement of one group is significantly higher than that of the other group. When reporting her findings, she summarizes the results of her statistical analyses. This researcher has conducted a *quantitative* study.

The second researcher is also interested in the effectiveness of the case method but asks the question, "What makes case-based instruction effective or ineffective?" To answer this question, he sits in on a case-based business management course for an entire semester. He spends an extensive amount of time talking with the instructor and some of the students in an effort to try to understand the participants' perspectives on case-based instruction. He carefully scrutinizes his data for patterns and themes in the responses. He then writes an in-depth description and interpretation of what he has observed in the classroom setting. This researcher has conducted a *qualitative* study.

Table 4.3 presents a summary of differences between quantitative and qualitative approaches. We briefly discuss these differences in the next few paragraphs—not to persuade you that one approach is better than the other, but to help you make a more informed decision about which approach might be better for your own research question.

**Purpose.** Quantitative researchers seek explanations and predictions that will generalize to other persons and places. The intent is to establish, confirm, or validate relationships and to develop generalizations that contribute to existing theories.

Qualitative researchers seek a better understanding of complex situations. Their work is sometimes (although not always) exploratory in nature, and they may use their observations to build theory from the ground up.

**TABLE 4.3**

Distinguishing characteristics of quantitative and qualitative approaches

What is the purpose of the research?	<ul style="list-style-type: none"> <li>● To explain and predict</li> <li>● To confirm and validate</li> <li>● To test theory</li> </ul>	<ul style="list-style-type: none"> <li>● To describe and explain</li> <li>● To explore and interpret</li> <li>● To build theory</li> </ul>
What is the nature of the research process?	<ul style="list-style-type: none"> <li>● Focused</li> <li>● Known variables</li> <li>● Established guidelines</li> <li>● Predetermined methods</li> <li>● Somewhat context-free</li> <li>● Detached view</li> </ul>	<ul style="list-style-type: none"> <li>● Holistic</li> <li>● Unknown variables</li> <li>● Flexible guidelines</li> <li>● Emergent methods</li> <li>● Context-bound</li> <li>● Personal view</li> </ul>
What are the data like, and how are they collected?	<ul style="list-style-type: none"> <li>● Numeric data</li> <li>● Representative, large sample</li> <li>● Standardized instruments</li> </ul>	<ul style="list-style-type: none"> <li>● Textual and/or image-based data</li> <li>● Informative, small sample</li> <li>● Loosely structured or nonstandardized observations and interviews</li> </ul>
How are data analyzed to determine their meaning?	<ul style="list-style-type: none"> <li>● Statistical analysis</li> <li>● Stress on objectivity</li> <li>● Deductive reasoning</li> </ul>	<ul style="list-style-type: none"> <li>● Search for themes and categories</li> <li>● Acknowledgment that analysis is subjective and potentially biased</li> <li>● Inductive reasoning</li> </ul>
How are the findings communicated?	<ul style="list-style-type: none"> <li>● Numbers</li> <li>● Statistics, aggregated data</li> <li>● Formal voice, scientific style</li> </ul>	<ul style="list-style-type: none"> <li>● Words</li> <li>● Narratives, individual quotes</li> <li>● Personal voice, literary style (in some disciplines)</li> </ul>

**Process.** Because quantitative studies represent the mainstream approach to research, carefully structured guidelines exist for conducting them. Concepts, variables, hypotheses, and methods of measurement tend to be defined before the study begins and remain the same throughout. Quantitative researchers choose methods that allow them to objectively measure the variable(s) of interest. They also try to remain detached from the phenomena and participants they are observing so that they can draw unbiased conclusions.

The qualitative research process is often more holistic and “emergent,” with the specific focus, design, measurement instruments (e.g., observations, interviews), and interpretations developing and possibly changing along the way. Researchers enter the setting with open minds, prepared to immerse themselves in the complexity of the situation and interact with their participants. Categories (variables) emerge from the data, leading to “context-bound” information, patterns, and/or theories that help explain the particular phenomenon under study.

**Data collection.** Quantitative researchers identify one or a few variables that they intend to study and then collect data specifically related to those variables. Methods of measuring each variable are identified, developed, and standardized, with considerable attention to the validity and reliability of the measurement instruments. Data are collected from a population—or from one or more large samples that represent the population—in forms that are easily converted to numbers.

Qualitative researchers operate under the assumption that reality is not easily divided into discrete, measurable variables. Some qualitative researchers describe themselves as *being* the research instrument because the bulk of their data collection is dependent on their personal involvement in the setting. Rather than sample a large number of participants with the intent of making generalizations, qualitative researchers tend to select a few participants who might best shed light on the phenomenon under investigation. Both verbal data (interview responses, documents, field notes) and nonverbal data (drawings, photographs, videotapes) may be collected.

**Data analysis.** All research requires logical reasoning. Quantitative researchers tend to rely more heavily on deductive reasoning, beginning with certain premises (e.g., hypotheses, theories) and then drawing logical conclusions from them. They also try to maintain objectivity in their data analysis, conducting predetermined statistical procedures and using objective criteria to evaluate the outcomes of those procedures.

In contrast, qualitative researchers make considerable use of inductive reasoning: They make many specific observations and then draw inferences about larger and more general phenomena. Furthermore, their data analysis is more subjective in nature: They scrutinize the body of data in search of patterns—subjectively identified—that the data reflect.

It is important to note, however, that quantitative research is not exclusively deductive, nor is qualitative research exclusively inductive. Researchers of all methodological persuasions typically use both types of reasoning in a continual, cyclical fashion. Quantitative researchers often formulate a theory by inductive reasoning (e.g., by observing a few situations), engage in the theory-building process described in Chapter 1, and then try to support their theory by drawing and testing the conclusions that follow logically from it. Similarly, after qualitative researchers have identified a theme in their data using an inductive process, they typically move into a more deductive mode to verify or modify it with additional data.

**Reporting findings.** Quantitative researchers typically reduce their data to means, medians, correlations, and other summarizing statistics. It is not necessary to look at individual performances; rather, the *averages* of those performances are of greater interest. The results are usually presented in a report that employs a formal, scientific style using impersonal language.

Qualitative researchers often construct interpretive narratives from their data and try to capture the complexity of a particular phenomenon. Especially in certain disciplines (e.g., anthropology), qualitative researchers may use a more personal, literary style than quantitative researchers do, and they often include the participants' own language and perspectives. Although all researchers must be able to write clearly, effective qualitative researchers must be especially skillful writers.

## Combining Quantitative and Qualitative Designs

We have drawn the preceding quantitative–qualitative distinctions, in part, as a way of pointing out the relative strengths and weaknesses of the two approaches. For example, a common weakness of quantitative research with human beings is that it is sometimes conducted in a laboratory—and thus in a somewhat artificial setting. Although contrived circumstances can give the researcher considerable control over the events that occur, the results obtained may in some cases not generalize to more naturalistic settings. In contrast, qualitative research occurs within natural contexts and so, in this respect, is more “true to life.” Yet the findings of qualitative studies may be so specific to a particular context that they do not apply (generalize) to other contexts.

Fortunately, quantitative and qualitative research designs are not necessarily mutually exclusive. Many researchers successfully combine elements of both approaches in what is sometimes called a **mixed-methods design**. For example, it is not unusual for researchers to *count* (and therefore quantify) certain kinds of data in what is, for all intents and purposes, a qualitative investigation. Nor is it unusual for quantitative researchers to report participants’ perceptions of or emotional reactions to various experimental treatments. Especially in studies of human behavior, mixed-methods designs with both quantitative and qualitative elements often provide a more complete picture of a particular phenomenon than either approach could do alone. We explore mixed-methods designs in more detail in Chapter 10.

## PRACTICAL APPLICATION Choosing a General Research Approach

Although we believe that research studies are sometimes enhanced by combining both quantitative and qualitative methods, we also realize that many novice researchers may not have the time, resources, or expertise to effectively combine approaches for their initial forays into research. Furthermore, good research does not necessarily have to involve a complex, multifaceted design. For example, in an article reviewing classic studies in his own discipline, psychologist Christopher Peterson had this to say in his abstract:

Psychology would be improved if researchers stopped using complicated designs, procedures, and statistical analyses for the sole reason that they are able to do so. . . . [S]ome of the classic studies in psychology [are] breathtakingly simple. . . . More generally, questions should dictate research methods and statistical analyses, not vice versa. (Peterson, 2009, p. 7)

As you choose your own general approach to addressing your research problem—whether to use a quantitative approach, a qualitative approach, or a combination of the two—you should base your decision on the research problem you want to address and the skills you have as a researcher, *not* on what tasks you wish to avoid. For example, disliking mathematics and wanting to avoid conducting statistical analyses are not good reasons for choosing a qualitative study over a quantitative one. The guidelines we offer here can help you make a reasonable decision.

### GUIDELINES Deciding Whether to Use a Quantitative or Qualitative Approach

Table 4.4 can help guide you in your choice between quantitative and qualitative approaches. Keep in mind that the items in the table are not necessarily ordered from most to least important. Each item should factor into your decision. Consider each component carefully before making your final selection.

Qualitative studies have become increasingly popular in recent years, even in some disciplines that have historically placed heavy emphasis on quantitative approaches. Yet we have met many students who have naively assumed that qualitative studies are easier or in some other way more “comfortable” than quantitative designs. Be forewarned: Qualitative studies require as much effort and rigor as quantitative studies, and data collection alone often stretches over the

TABLE 4.4

Which approach should I use?

	Quantitative	Qualitative
1. You believe that:	There is an objective reality that can be measured	There are multiple possible realities constructed by different individuals
2. Your audience is:	Familiar with/supportive of quantitative studies	Familiar with/supportive of qualitative studies
3. Your research question is:	Confirmatory, predictive	Exploratory, interpretive
4. The available literature is:	Relatively large	Limited
5. Your research focus:	Covers a lot of breadth	Involves in-depth study
6. Your time available is:	Relatively short	Relatively long
7. Your ability/desire to work with people is:	Medium to low	High
8. Your desire for structure is:	High	Low
9. You have skills in the areas of:	Deductive reasoning and statistics	Inductive reasoning and attention to detail
10. Your writing skills are strong in the areas of:	Technical, scientific writing	Literary, narrative writing

course of many months. In the following paragraphs, we briefly discuss each of the components in Table 4.4 from the perspective of someone who might be inclined to “go qualitative.”

1. *Consider your own comfort with the assumptions of the qualitative tradition.* If you believe that no single reality underlies your research problem but that, instead, different individuals may have constructed different, possibly equally valid realities relevant to your problem, then qualitative research is more appropriate.

2. *Consider the audience for your study.* If your intended audience (e.g., a dissertation committee, a specific journal editor, or colleagues in your field) is not accustomed to or supportive of qualitative research, it makes little sense to spend the time and effort needed to do a good qualitative study (e.g., see Miller, Nelson, & Moore, 1998).

3. *Consider the nature of the research question.* Qualitative designs can be quite helpful for addressing exploratory or interpretive research questions. But they may be of little use in testing specific hypotheses about cause-and-effect relationships.

4. *Consider the extensiveness of the related literature.* If the literature base is weak, underdeveloped, or altogether missing, a qualitative design can provide the researcher with the freedom and flexibility needed to explore a specific phenomenon so that important variables might be identified.

5. *Consider the depth of what you wish to discover.* If you want to examine a phenomenon in depth with a relatively small number of participants, a qualitative approach is ideal. But if you are skimming the surface of a phenomenon and wish to do so using a large number of participants, a quantitative study will be more efficient.

6. *Consider the amount of time you have available for conducting the study.* Qualitative studies typically involve an extensive amount of time both on site and off site. If your time is limited, you may not be able to complete a qualitative study satisfactorily.

7. *Consider the extent to which you are willing to interact with the people in your study.* Qualitative researchers who are working with human beings must be able to establish rapport and trust with their participants and interact with them on a fairly personal level. Also, gaining initial entry into a research site involves much advance planning and numerous preliminary contacts.

8. *Consider the extent to which you feel comfortable working without much structure.* Qualitative researchers typically work with few specific rules and procedures; their work is exploratory in

many respects. Thus, they must think creatively about how best to address various aspects of a research problem, and they need a high tolerance for ambiguity.

9. *Consider your ability to organize and draw inferences from a large body of information.* Qualitative research often involves the collection of a great many field notes, interview responses, and so on, that are not clearly organized at the beginning of the process. Working with extensive amounts of data and reasoning inductively about them require considerable self-discipline and organizational ability. In comparison, conducting a few statistical analyses—even for those who have little affection for mathematics—is a much easier task.

10. *Consider your writing skills.* Qualitative researchers must have excellent writing skills. Communicating findings is the final step in all research projects; the success of your research ultimately will be judged by how well you accomplish this final component of the research process.

Once you have decided whether to take a quantitative or qualitative approach, you need to pin down your research method more precisely. Table 4.5 lists some common research methodologies and the types of problems for which each is appropriate.

**TABLE 4.5** Methodology and concomitant research goals

Methodology	Research Goals the Method Attempts to Achieve
Action research	A type of applied research that focuses on finding a solution to a local problem in a local setting. For example, a teacher might investigate whether a new spelling program she has adopted leads to improvement in her students' achievement scores. (See Mertler [2009] or Mills [2011] for guidance on conducting action research.)
Case study	A type of qualitative research in which in-depth data are gathered relative to a single individual, program, or event, for the purpose of learning more about an unknown or poorly understood situation. (See Chapter 7.)
Content analysis	A detailed and systematic examination of the contents of a particular body of material (e.g., television shows, magazine advertisements, Internet websites, works of art) for the purpose of identifying patterns, themes, or biases within that material. (See Chapter 7.)
Correlational research	A statistical investigation of the relationship between two or more variables. Correlational research looks at surface relationships but does not necessarily probe for causal reasons underlying them. For example, a researcher might investigate the relationships among high school seniors' achievement test scores and their grade point averages a year later when they are first-year college students. (See Chapter 9.)
Developmental research	An observational-descriptive type of research that either compares people in different age groups (a <i>cross-sectional study</i> ) or follows a particular group over a lengthy period of time (a <i>longitudinal study</i> ). Such studies are particularly appropriate for looking at developmental trends. (See Chapter 9.)
Ethnography	A type of qualitative inquiry that involves an in-depth study of an intact cultural group in a natural setting. (See Chapter 7.)
Experimental research	A study in which participants are randomly assigned to groups that undergo various researcher-imposed treatments or interventions, followed by observations or measurements to assess the effects of the treatments. (See Chapter 10.)
Ex post facto research	An approach in which one looks at conditions that have already occurred and then collects data to investigate a possible relationship between these conditions and subsequent characteristics or behaviors. (See Chapter 10.)
Grounded theory research	A type of qualitative research aimed at deriving theory through the use of multiple stages of data collection and interpretation. (See Chapter 7.)
Historical research	An effort to reconstruct or interpret historical events through the gathering and interpretation of relevant historical documents and/or oral histories. (See Chapter 8.)
Observation study	A type of quantitative research in which a particular aspect of behavior is observed systematically and with as much objectivity as possible. (See Chapter 9.)
Phenomenological research	A qualitative method that attempts to understand participants' perspectives and views of social realities. (See Chapter 7.)
Quasi-experimental research	A method similar to experimental research but without random assignment to groups. (See Chapter 10.)
Survey research	A study designed to determine the incidence, frequency, and distribution of certain characteristics in a population; especially common in business, sociology, and government research. (See Chapter 9.)

---

## Considering the Validity of Your Method

No matter what research methodology you choose, you must think about the *validity* of your approach. We have already described the importance of validity in measurement instruments. But here we are talking about the validity of *the research project as a whole*—its accuracy, meaningfulness, and credibility. Your research effort will be worth your time and effort only to the extent that it allows you to draw meaningful and defensible conclusions from your data.

When researchers consider the validity of a research study, they need to ask themselves two basic questions. First, does the study have sufficient controls to ensure that the conclusions drawn are truly warranted by the data? And second, can the results obtained reasonably be used to make generalizations about the world beyond that specific research context? The answers to these two questions address the issues of *internal validity* and *external validity*, respectively.

### Internal Validity

The *internal validity* of a research study is the extent to which its design and the data it yields allow the researcher to draw accurate conclusions about cause-and-effect and other relationships within the data. To illustrate, we present three situations in which the internal validity of a study is suspect:

1. A marketing researcher wants to study how humor in television commercials affects sales in the United States and Canada. To do so, the researcher studies the effectiveness of two commercials that have been developed for a new soft drink called Zowie. One commercial, in which a well-known but humorless television actor describes how Zowie has a zingy and refreshing taste, airs during the months of March, April, and May. The other commercial, a humorous scenario in which several teenagers spray one another with Zowie on a hot summer day, airs during the months of June, July, and August. The researcher finds that in June through August, Zowie sales are almost double what they were in the preceding three months. "Humor boosts sales," the researcher concludes.
2. An industrial psychologist wants to study the effects of soft classical music on the productivity of a group of typists in a typing pool. At the beginning of the month, the psychologist meets with the typists to explain the rationale for the study, gets their consent to play the music during the working day, and then begins to have music piped into the office where the typists work. At the end of the month, the typists' supervisor reports a 30% increase in the number of documents completed by the typing pool that month. "Classical music increases productivity," the psychologist concludes.
3. An educational researcher wants to study the effectiveness of a new method of teaching reading to first graders. The researcher asks all 30 of the first-grade teachers in a particular school district whether they would like to receive training in the new method and then use it during the coming school year. Fourteen teachers volunteer to learn and use the new method; 16 teachers say that they would prefer to use their current approach. At the end of the school year, students who have been instructed with the new method have, on average, significantly higher average scores on a reading achievement test than students who have received more traditional reading instruction. "The new method is definitely better than the old one," the researcher concludes.

Did you detect something wrong with the conclusions these researchers drew? If not, go back and read the three descriptions again. *None of the conclusions is warranted from the study conducted.*

In the first research study, the two commercials apparently differed from each other in many ways (e.g., the presence of teenagers, the amount of action) in addition to humor. And, of course, we should not overlook the fact that the humorous commercial aired during the summer months. People are more likely to drink soft drinks (including Zowie) when they are hot!

In the second study, the typists *knew* they were participating in a research study; they also knew the nature of the researcher's hypothesis. Sometimes the participants in a research study

change their behavior simply because they know they are in a research study and are getting extra attention as a result. This effect, known as the **Hawthorne effect**,<sup>3</sup> is an example of **reactivity**, a more general phenomenon in which people change their behavior when they are aware that they are being observed. But other explanations for the second study's results are possible as well. Perhaps the typists typed more because they liked the researcher and wanted to help him support his hypothesis. Perhaps the music perked up the typists for a few weeks simply because it was a change in their environment—a phenomenon known as the **novelty effect**. (In such a situation, reverting back to *no* music after a month or two might *also* lead to an increase in productivity.) Furthermore, the researcher did not consider the number of people who were working before and after the music started. Perhaps productivity increased simply because two people in the typing pool had just returned from vacation!

In the third study, notice that the researcher looked for *volunteers* to use the new method for teaching reading. Were the volunteer teachers different in some way from the nonvolunteers? Were they better educated or more motivated? Did they teach with more enthusiasm and energy because they *expected* the new method to be more effective? Or did the volunteer teachers happen to teach in areas of the school district where children had a better head start in reading skills before coming to school? Perhaps the children in the volunteers' classrooms performed better on the achievement test not because the instructional method was more effective, but because, as a group, they had been read to more frequently by their parents, gone to better preschools, and so on.

To ensure the internal validity of a research study, researchers need to take whatever precautions they can to *eliminate other possible explanations for the results observed*. Following are several strategies researchers sometimes use to increase the probability that *their explanations are the most likely ones* for the observations they have made:

- **A controlled laboratory study.** An experiment is conducted in a laboratory setting so that environmental conditions can be carefully regulated.
- **A double-blind experiment.** In a situation where two or more different interventions are being compared, neither the participants in the study nor the people administering the interventions (e.g., teachers, research assistants) know which intervention each participant is receiving. Such lack of knowledge ("blindness") decreases the likelihood that people's expectations for outcomes might influence the *actual* outcomes.
- **Unobtrusive measures.** People are observed in such a way that they do not know their actions are being recorded. We offer two real-life examples to illustrate. In one case, a university library measured student and faculty use of different parts of the library by looking at wear-and-tear patterns on the carpet. In another situation, researchers for the U.S. National Park Service looked at hikers' frequency of using different hiking trails by installing electronic counters in hard-to-notice locations beside the trails (R. K. Ormrod & Trahan, 1982). (Note that ethical issues sometimes arise when we observe people without their permission; we discuss ethics a bit later in this chapter.)
- **Triangulation.** Multiple sources of data are collected with the hope that they will all converge to support a particular hypothesis or theory. This approach is especially common in qualitative research; for instance, a researcher might engage in many informal observations in the field *and* conduct in-depth interviews, then look for common themes that appear in the data gleaned from both methods. Triangulation is also common in mixed-methods designs, in which both quantitative and qualitative data are collected to answer a single research question.

Internal validity is especially of concern in experimental designs, where the specific intent is to identify cause-and-effect relationships; accordingly, we will revisit this issue in Chapter 9. But to some degree, internal validity is important in *any* research study. The researcher must have confidence that the conclusions drawn are warranted from the data collected.

<sup>3</sup>The effect owes its name to the Hawthorne Works, an industrial complex in Illinois where the effect was first observed.



## External Validity

The **external validity** of a research study is the extent to which its results apply to situations beyond the study itself—in other words, the extent to which the conclusions drawn can be *generalized* to other contexts. As a general rule, researchers contribute more to humanity's knowledge about the world when they conduct research that has implications that extend far beyond the specific situation studied.

Following are three commonly used strategies that enhance the external validity of a research project:

- **A real-life setting.** Earlier we mentioned that researchers sometimes use laboratory experiments to help them control the environmental conditions in which a study takes place. Laboratory studies have a downside, however: They provide an artificial setting that may be quite different from real-life circumstances. Research that is conducted in the outside world, although it may not have the tight controls of a laboratory project, may be more valid in the sense that it yields results with broader applicability to other real-world contexts.<sup>4</sup>
- **A representative sample.** Whenever researchers seek to learn more about a particular category of objects or creatures—whether they are studying rocks, salamanders, or human beings—they often study a *sample* from that category and then draw conclusions about the category as a whole. (Here is a classic example of inductive reasoning.) For example, to study the properties of granite, researchers might take pieces of granite from anywhere in the world and assume that their findings based on those pieces might be generalizable to granite found in other locations. The same might hold true for salamanders if researchers limit their conclusions to the particular species of salamander they have studied.

Human beings are another matter. The human race is incredibly diverse in terms of culture, childrearing practices, educational opportunities, personality characteristics, and so on. To the extent that researchers restrict their research to people with a particular set of characteristics, they may not be able to generalize their findings to people with a very different set of characteristics. Ideally, then, researchers want participants in a research study to be a *representative sample* of the population about which they wish to draw conclusions. In Chapter 8, we consider a number of strategies for obtaining representative samples.

- **Replication in a different context.** Imagine that one researcher draws a conclusion from a particular study in a specific context, and another researcher who conducts a similar study in a very different context reaches the same conclusion, and perhaps additional researchers also conduct similar studies in dissimilar contexts and, again, draw the same conclusion. Under such circumstances, these studies, taken together, provide evidence that the conclusion has validity and applicability across diverse contexts and situations.

Any researcher should consider both internal validity and external validity when designing a research project. One's conclusions are valid and meaningful only to the extent that they are warranted based on the data collected *and* have applicability beyond the specific research situation itself.

In Chapter 2 we introduced you to the distinction between basic and applied research. Well-designed basic research—research conducted under tightly controlled (and possibly artificial) conditions—ensures internal validity; that is, it allows the researcher to rule out other possible explanations for the results obtained. Applied research—research conducted in more naturalistic but invariably more complex environments—is more useful for external validity; that is, it increases the chances that a study's findings are generalizable to other real-life situations and

---

<sup>4</sup>The artificial nature of laboratory research has been a concern in psychology for many years. Anderson, Lindsay, and Bushman (1999) compared the results of laboratory and field studies related to a wide variety of psychological phenomena. They discovered that the two kinds of studies typically yield similar results—that studies conducted in the lab and those conducted in natural settings lead to the same conclusions about human nature. In their own words, "The psychological laboratory has generally produced psychological truths, rather than trivialities" (p. 3).

problems. Keep in mind, however, that the basic-versus-applied distinction is really a continuum rather than a dichotomy: Research studies can have varying degrees of artificiality versus real-world authenticity.

## Validity in Qualitative Research

The concepts of internal and external validity originated in discussions of quantitative research (Campbell & Stanley, 1963), and some qualitative researchers have questioned their relevance to qualitative designs. For instance, Lincoln and Guba (1985) and Creswell (2007) have suggested that, for qualitative research, such words as *credibility*, *trustworthiness*, *confirmability*, and *validation* be used instead of the term *validity*.

As noted earlier, qualitative researchers frequently use triangulation—comparing multiple data sources in search of common themes—to support the validity of their findings. Following are several additional strategies they employ:

- **Extensive time in the field.** The researcher may spend several months, perhaps even a year or more, studying a particular phenomenon, forming tentative hypotheses, and continually looking for evidence that either supports or disconfirms those hypotheses.
- **Negative case analysis.** The researcher actively looks for cases that contradict existing hypotheses, then continually revises his or her explanation or theory until all cases have been accounted for.
- **Thick description.** The situation is described in sufficiently rich, “thick” detail that readers can draw their own conclusions from the data presented.
- **Feedback from others.** The researcher seeks the opinion of colleagues in the field to determine whether they agree or disagree that the researcher has made appropriate interpretations and drawn valid conclusions from the data.
- **Respondent validation.** The researcher takes his or her conclusions back to the participants in the study and asks quite simply, Do you agree with my conclusions? Do they make sense based on your own experiences?

As you can see, then, researchers use a wide variety of approaches to support the validity of their findings. Different approaches are appropriate in different situations, depending on the nature of the data and the specific methodologies used.

Regardless of the kind of study you decide to conduct, you must address the validity of your study at the very beginning of your project—that is, *at the planning stage*. If you put off validity issues until later in the game, you may end up conducting a study that has little apparent validity, either in terms of minimizing alternative explanations for the results obtained (internal validity) or in terms of being generalizable to the world “out there” (external validity). As a result, you are almost certainly wasting your time and effort on what is, for all intents and purposes, a trivial enterprise.

## Ethical Issues in Research

Within certain disciplines—the social sciences, education, medicine, and similar areas of study—the use of human beings in research is, of course, quite common. And in biology, subjects of investigation are often nonhuman animals. Whenever human beings or other creatures with the potential to think, feel, and experience physical or psychological distress are the focus of investigation, researchers must look closely at the ethical implications of what they are proposing to do.

Most ethical issues in research fall into one of four categories: protection from harm, voluntary and informed participation, right to privacy, and honesty with professional colleagues. In the following sections we raise issues related to each of these categories. We describe the internal review boards and professional codes of ethics that provide guidance to researchers.

## Protection from Harm

Researchers should not expose research participants—whether they be human beings or non-human animals—to unnecessary physical or psychological harm. When a study involves human beings, the general rule of thumb is that the risk involved in participating in a study should not be appreciably greater than the normal risks of day-to-day living. Participants should not risk losing life or limb, nor should they be subjected to unusual stress, embarrassment, or loss of self-esteem.

In thinking about this issue, researchers must be particularly sensitive to and thoughtful about potential harm they might cause participants from potentially vulnerable populations (Sieber, 2000). For example, some participants may have allergies or health conditions that place them at greater-than-average risk in certain environments or with certain foods or medications. Participants of a particular gender, cultural background, or sexual orientation might feel embarrassed or otherwise uncomfortable when asked to answer some kinds of questions or engage in some kinds of activities. Special care must be taken with participants who cannot easily advocate for their own needs and desires—children, the elderly, people with disabilities, and the like.

Especially when working with human participants, a researcher should ideally also think about potential *benefits* that participation in a study might offer. At a minimum, the researcher should treat all participants in a courteous and respectful manner. A researcher can also consider how people might gain something useful from participating in a study—perhaps unique insights about a topic of personal interest or perhaps simply a sense of satisfaction about contributing in a small way to advancements in society's collective knowledge about the world. In some cases a researcher can offer an incentive for participating (e.g., money or course credit), provided that it isn't so excessive that it's essentially a form of disguised coercion (Scott-Jones, 2000).<sup>5</sup>

In cases where the nature of a study involves creating a small amount of psychological discomfort, participants should know this ahead of time, and any necessary debriefing or counseling should follow immediately after their participation. A debriefing can simultaneously accomplish several things (Sales & Folkman, 2000):

- It can help alleviate any uncomfortable reactions—either anticipated or unanticipated—to certain questions, tasks, or activities.
- It can alert the researcher to necessary follow-up interventions for any participants experiencing extreme reactions.
- It provides an opportunity for the researcher to correct any misinformation participants might have gotten during the study.
- It provides a time during which participants can learn more about the nature and goals of the study, about how its results may fit in with what is already known about a topic, and about the nature of research more generally.

## Voluntary and Informed Participation

When research involves public documents or records that human beings have previously created—birth certificates, newspaper articles, and so on—such documents and records are generally considered to be fair game for research investigation. But when people are specifically recruited for participation in a research study, they should be told the nature of the study to be conducted and given the choice of either participating or not participating. Furthermore, they should be told that, if they agree to participate, they have the right to withdraw from the study at any time. And under no circumstances should people feel pressure to participate from employers or other more powerful individuals. *Any participation in a study should be strictly voluntary.*

In general, research with human beings requires **informed consent**. That is, participants—or legal guardians in the case of children and certain other populations—must know the nature

<sup>5</sup> Two qualifications should be noted here. When working with children, enticing incentives should be offered *only after* parents have already given permission for their participation. And when offering course credit to college students, alternative ways to earn the same credit must be provided as well—for instance, reading and writing a review of a research article (Scott-Jones, 2000).

of the study and grant written permission. One common practice—and one that is *required* for certain kinds of studies at most research institutions—is to present an **informed consent form** that describes the nature of the research project, as well as the nature of one's participation in it. Such a form should contain the following information:

- A brief description of the nature and goal(s) of the study, written in language that its readers can readily understand
- A description of what participation will involve, in terms of activities and duration
- A statement indicating that participation is voluntary and can be terminated at any time without penalty
- A description of any potential risk and/or discomfort that participants may encounter
- A description of potential benefits of the study, including those for participants, science, and/or human society as a whole
- A guarantee that all responses will remain confidential and anonymous
- The researcher's name, plus information about how the researcher can be contacted
- An individual or office that participants can contact if they have questions or concerns about the study
- An offer to provide detailed information about the study (e.g., a summary of findings) upon its completion
- A place for the participant to sign and date the letter, indicating agreement to participate (when children are asked to participate, their parents must read and sign the letter)

An example of such a form, used by Rose McCallin in a research project for her doctoral dissertation, is presented in Figure 4.6. The form was used to recruit college students who were enrolled in a class in a teacher preparation program. It is missing one important ingredient: an offer to provide information about the study after its completion. Instead, McCallin appeared in class a few weeks after she had collected data to give a summary of the study and its implications for teachers.

A dilemma sometimes arises as to *how informed* potential participants should be. If people are given too much information—for instance, if they are told the specific research hypothesis being tested—they may behave differently than they would under more normal circumstances (recall the earlier description of a study involving classical music and typists' productivity). A reasonable compromise is to give potential participants a general idea of what the study is about (e.g., "This study is investigating the effects of a physical exercise program on people's overall mental health") and to describe what specific activities their participation will involve—in other words, to give them sufficient information to make a reasonable, informed judgment about whether they wish to participate.

On rare occasions (e.g., in some studies of social behavior), telling participants the true nature of a study might lead them to behave in ways that would defeat the purpose of the study. In general, deception of any kind is frowned on and should be used only when the study cannot meaningfully be conducted without it. Even then, the degree of deception should be as minimal as possible, and participants should be told the true nature of the research as soon as their involvement is over. (An internal review board, to be described shortly, can give you guidance regarding this matter.)

Earlier in the chapter we mentioned the use of unobtrusive measures as a strategy for measuring behavior. Strictly speaking, unobtrusive measures violate the principle of informed consent. But if people's behaviors are merely being recorded in some way during their normal daily activities—if people are not being asked to do something they ordinarily would not do—and if they are not being scrutinized in any way that might be potentially invasive or embarrassing, then unobtrusive measures are quite appropriate. Recall our two earlier examples: examining the frequency with which people use different parts of the library and the frequency with which they hike along certain trails in a national park. Both of these examples involved behaviors within the scope of participants' normal activities.

### Understanding How Students Organize Knowledge

You are being asked to participate in a study investigating ways in which students organize their knowledge.

We are interested in determining how students organize their knowledge in memory and use that knowledge. It is hoped that the results of this study can be useful in helping teachers understand why students perform differently from one another in the classroom.

As a future teacher, you will most likely have to use your knowledge in a variety of situations. However, relatively little is known about relationships among factors involved in knowledge application. Your participation may help to clarify some of these relationships so that we can better identify why students perform differently. And, although you may not directly benefit from this research, results from the study may be useful for future students, both those you teach and those who, like yourself, plan to be teachers.

If you agree to participate, you will complete two activities. In addition, we need to use your anonymous grade point average (GPA) as a control variable in order to account for initial differences among students. To ensure anonymity, we will submit only your social security number to the UNC Registrar, who will use this number to locate your GPA. The Registrar will black out the first three digits of your social security number before giving us this information, and the remaining six-digit number will be used only to keep track of your performance on the other activities. You will not be putting your name on anything except this form. And, there will be no attempt to link your name with the last six digits of your social security number because individual performance is not of interest in this study. Only group results will be reported.

In the first activity, you will be asked to complete a 15-minute Self-Rating Checklist. This checklist consists of statements about knowledge application that you will judge to be true or false according to how each statement applies to you. In the second activity (which will be administered two days later), you will be given a list of concepts and asked to organize them on a sheet of paper, connect concepts you believe to be related, and describe the type of relationship between each connected pair of concepts. This activity should take about 30 minutes.

Although all studies have some degree of risk, the potential in this investigation is quite minimal. All activities are similar to normal classroom procedures, and all performance is anonymous. You will not incur any costs as a result of your participation in this study.

Your participation is voluntary. If at any time during this study you wish to withdraw your participation, you are free to do so without prejudice.

If you have any questions prior to your participation or at any time during the study, please do not hesitate to contact us.

**AUTHORIZATION:** I have read the above and understand the nature of this study. I understand that by agreeing to participate in this study I have not waived any legal or human right and that I may contact the researchers at the University of Northern Colorado (Dr. Jeanne Ormrod or Rose McCallin, 303-555-2807) at any time. I agree to participate in this study. I understand that I may refuse to participate or I may withdraw from the study at any time without prejudice. I also grant permission to the researchers to obtain my anonymous grade point average from the UNC Registrar for use as a control variable in the study. In addition, I understand that if I have any concerns about my treatment during the study, I can contact the Chair of the Internal Review Board at the University of Northern Colorado (303-555-2392) at any time.

Participant's signature: \_\_\_\_\_ Date: \_\_\_\_\_

Researcher's signature: \_\_\_\_\_ Date: \_\_\_\_\_

#### FIGURE 4.6

##### Example of an informed consent form

Adapted from *Knowledge Application Orientation, Cognitive Structure, and Achievement* (pp. 109–110) by R. C. McCallin, 1988, unpublished doctoral dissertation, University of Northern Colorado, Greeley. Adapted with permission.

## Right to Privacy

Any research study involving human beings must respect participants' right to privacy. Under no circumstances should a research report, either oral or written, be presented in such a way that other people become aware of how a particular participant has responded or behaved—unless, of course, the participant has specifically granted permission *in writing* for this to happen.

In general, a researcher must keep the nature and quality of individual participants' performance strictly confidential. For instance, the researcher might give each participant a unique, arbitrary code number and then label any written documents with that number rather than with the person's name. And if a particular person's behavior is described in depth in the research

report, he or she should be given a pseudonym—and other trivial, irrelevant details that might give away the person's identity should be changed—to ensure anonymity.

In this age of the Internet, researchers must also take precautions that computer hackers cannot access participants' individual data. Our advice here is simple: Don't post raw data or easily decodable data about individual participants online in any form. If you use the Internet to share your data with co-researchers living elsewhere, use e-mail and well-encoded attachments to transmit your data set; send your coding scheme in a separate e-mail message at another time.

Occasionally employers or other powerful individuals in a research setting might put considerable pressure on a researcher to reveal participants' individual responses. The researcher must not give in to such pressure. In general, knowledge about participants' individual performances should be revealed *only* to any co-researchers who have a significant role in the research investigation—unless, of course, participants have specifically granted permission *in writing* that it be shared with certain other individuals. There is one important exception to this rule: Researchers are legally obligated to report to the proper authorities any information that suggests present or imminent danger to someone (e.g., child abuse, a planned terrorist act).

## Honesty with Professional Colleagues

Researchers must report their findings in a complete and honest fashion, without misrepresenting what they have done or intentionally misleading others about the nature of their findings. And under no circumstances should a researcher fabricate data to support a particular conclusion, no matter how seemingly "noble" that conclusion might be. Such an action constitutes scientific fraud, plain and simple.

Within this context, we ask you to recall our discussion in Chapter 3 about giving appropriate credit where credit is due. Any use of another person's ideas or words demands full acknowledgment; otherwise, it constitutes plagiarism and documentary theft. Full acknowledgment of all material belonging to another person is mandatory. To appropriate the thoughts, ideas, or words of another without acknowledgment—even if you paraphrase the borrowed ideas in your own language—is dishonest, unethical, and highly circumspect. Honest researchers do not hesitate to acknowledge their indebtedness to others.

## Internal Review Boards

In the United States, any college, university, or research institution must have an **internal review board (IRB)**<sup>6</sup> that scrutinizes all proposals for conducting human research under the auspices of the institution. This board, which is made up of scholars and researchers across a broad range of disciplines, checks proposed research studies to be sure that the procedures are not unduly harmful to participants, that appropriate procedures will be followed to obtain participants' informed consent, and that participants' privacy and anonymity are ensured.

It is important to note that the research is reviewed at the proposal stage. *A proposal must be submitted to and approved by the IRB before a single datum is collected.* Depending on the extent to which the study intrudes in some way on people's lives and imposes risk to participants, the board's chairperson may quickly declare it *exempt from review*, give it an *expedited review*, or bring it before the board for a *full review*. In any case, the researcher cannot begin the study until (a) the board has given its seal of approval to the study as originally designed or (b) the researcher has made modifications that the board requests.

The criteria and procedures of an IRB vary slightly from one institution to another. For examples of institutional policies and procedures, you might want to visit the websites of Duke University ([www.ots.duke.edu](http://www.ots.duke.edu)), Tufts University ([tnemcirb.tufts.edu](http://tnemcirb.tufts.edu)), or University of New Hampshire ([www.unh.edu/osr](http://www.unh.edu/osr)). You can find other helpful sites on the Internet by using a search engine (e.g., Google or Yahoo!) and such keywords as "IRB," "human participants," and "human subjects."

<sup>6</sup> You might also see this committee called something along the lines of "Committee for Protection of Human Subjects."

Universities and other research institutions have review boards for animal research as well. Any research that may potentially cause suffering, distress, or death to animals must be described and adequately justified to an **institutional animal care and use committee (IACUC)**. Furthermore, the researcher must minimize or prevent such suffering and death to the extent it is possible to do so. For examples of research institutions' IACUC policies and procedures, we refer you to the University of Maryland ([www.umresearch.umd.edu/IACUC](http://www.umresearch.umd.edu/IACUC)) and the University of Arizona ([www.uac.arizona.edu](http://www.uac.arizona.edu)).


## Professional Codes of Ethics

Many disciplines have their own codes of ethical standards governing research that involves human subjects and, when applicable, research involving animal subjects as well. One good source of discipline-specific ethical codes is, of course, the Internet. Following are examples of organizational websites with ethical codes related to research in their disciplines:

- American Anthropological Association ([www.aaanet.org](http://www.aaanet.org))
- American Association for Public Opinion Research ([www.aapor.org](http://www.aapor.org))
- American Educational Research Association ([www.aera.net](http://www.aera.net))
- American Psychological Association ([www.apa.org](http://www.apa.org))
- American Society of Criminology ([www.asc41.com](http://www.asc41.com))
- American Sociological Association ([www.asanet.org](http://www.asanet.org))
- Society for Conservation Biology ([www.conbio.org](http://www.conbio.org))

## PRACTICAL APPLICATION Planning an Ethical Research Study

Ethical practices in research begin *at the planning* stage. The following checklist can help you scrutinize your own project for its potential ethical implications.

 **CHECKLIST**

**Determining Whether Your Proposed Study Is Ethically Defensible**

\_\_\_\_\_ 1. Might your study present any physical risks or hazards to participants? If so, list them here.

\_\_\_\_\_ 2. Might your study incur any psychological harm to all or some participants (e.g., offensive stimulus materials, threats to self-esteem)? If so, identify the specific forms of harm that might result.

\_\_\_\_\_ 3. Will participants incur any significant financial costs (e.g., transportation costs, mailing expenses)? If so, how might you minimize or eliminate those costs?

\_\_\_\_\_ 4. What benefits might your study have for (a) participants, (b) your discipline, and (c) society at large?

- \_\_\_\_\_ 5. Do you need to seek informed consent from participants? Why or why not?
- \_\_\_\_\_ 6. If you need to seek informed consent, how might you explain the nature and goals of your study to potential participants in a way that they can understand? Write a potential explanation here.
- \_\_\_\_\_ 7. What specific steps will you take to ensure participants' privacy? List them here.
- \_\_\_\_\_ 8. If applicable, what format might a post-participation debriefing take? What information should you include in your debriefing?

## Critically Scrutinizing Your Overall Plan

At this point, you have presumably attended to the nature and availability of the data you need in order to address your research problem, identified potentially appropriate ways of measuring your observations, chosen an overall approach to your research methodology, and examined the ethical implications of what you intend to do. But ultimately, you must step back a bit and look at the overall forest—the big picture—rather than at the specific, nitty-gritty trees. And, of course, you must be realistic and practical regarding what you can reasonably accomplish. Remember the title of this book: *Practical Research*.

## PRACTICAL APPLICATION Judging the Feasibility of a Research Project

Many beginning researchers avoid looking closely at the practical aspects of a research endeavor. Envisioning an exotic investigation or a solve-the-problems-of-the-world study sometimes keeps a researcher from making an impartial judgment about practicability. Completing the following checklist can help you wisely plan and accurately evaluate the research you have in mind. After you have finished, review your responses. Then answer this question: Can you reasonably accomplish this study? If your answer is *no*, determine which parts of the project are not terribly practical, and identify things you might do to make it more realistically accomplishable.

### ✓ CHECKLIST

#### Determining Whether a Proposed Research Project Is Realistic and Practical

##### THE PROBLEM

- \_\_\_\_\_ 1. With what area(s) will the problem deal?
- \_\_\_\_\_ People
- \_\_\_\_\_ Things
- \_\_\_\_\_ Records



- Thoughts and ideas  
 Dynamics and energy
2. Are data that relate directly to the problem available for each of the categories you've just checked?  Yes  No
3. What academic discipline is primarily concerned with the problem?  
 \_\_\_\_\_
4. What other academic disciplines are possibly also related to the problem?  
 \_\_\_\_\_
5. What special qualifications do you have as a researcher for this problem?  
 Interest in the problem  
 Experience in the problem area  
 Education and/or training  
 Other (specify): \_\_\_\_\_

## THE DATA

6. How available are the data to you?  
 Readily available  
 Available with permission  
 Available with great difficulty or rarely available  
 Unavailable
7. How often are you personally in contact with the source of the data?  
 Once a day       Once a week       Once a month  
 Once a year       Never
8. Will the data arise directly out of the problem situation?  
 Yes       No  
 If your answer is no, where or how will you obtain the data?  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_
9. How do you plan to gather the data?  
 Observation       Questionnaire       Test       Rating scale  
 Photocopying of records       Interview and audio recording  
 Specialized machine/device       Computer technology  
 Other (explain): \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_
10. Is special equipment or are special conditions necessary for gathering or processing the data?  
 Yes       No  
 If your answer is yes, specify: \_\_\_\_\_  
 \_\_\_\_\_
11. If you will need special equipment, do you have access to such equipment and the skill to use it?  
 Yes       No

If your answer is no, how do you intend to overcome this difficulty?

---



---

- \_\_\_\_\_ 12. What is the estimated cost in time and money to gather the data?
- 
- \_\_\_\_\_ 13. What evidence do you have that the data you gather will be valid and reliable indicators of the phenomena you wish to study?
- 
- 

#### CRITERIA-BASED EVALUATION

- \_\_\_\_\_ 14. Does your research project meet the four criteria applicable to all research? (For a refresher on these criteria, see the section "General Criteria for a Research Project.")

Universality            \_\_\_\_\_ Yes            \_\_\_\_\_ No

Replication            \_\_\_\_\_ Yes            \_\_\_\_\_ No

Control                \_\_\_\_\_ Yes            \_\_\_\_\_ No

Measurement           \_\_\_\_\_ Yes            \_\_\_\_\_ No

- \_\_\_\_\_ 15. As you review your responses to this checklist, might any of the factors you've just considered, or perhaps any other factors, hinder a successful completion of your research project?

\_\_\_\_\_ Yes            \_\_\_\_\_ No

If your answer is yes, list those factors.

---



---

## When You Can't Anticipate Everything in Advance: The Value of a Pilot Study

Did you have trouble answering some of the questions in the checklist? For instance, did you have difficulty estimating how much time it would take you to gather or process your data? Did you realize that you might need to develop your own questionnaire, test, or other measurement instrument but then wonder how valid and reliable the instrument might be for your purpose?

Up to this point, we have been talking about planning a research project as something that occurs all in one fell swoop. In reality, a researcher may sometimes need to do a brief exploratory investigation, or *pilot study*, to try out particular procedures, measurement instruments, or methods of analysis. *A brief pilot study is an excellent way to determine the feasibility of your study.* Furthermore, although it may take some time initially, it may ultimately save you time by letting you know—after only a small investment on your part—which approaches will and will not be effective in helping you solve your overall research problem.

## PRACTICAL APPLICATION Establishing Research Criteria and Justifying Your Research Methodology

In this exercise, you will continue developing your research proposal, for which you have already conducted a review of the related literature (see Chapter 3). In writing proposals, we leave nothing unspecified. Here we explore how to treat each subproblem in terms of its data, the criteria

for the admission of those data into your study, and a justification of the methodology you propose to employ. In the form presented in Figure 4.7 (extend it as needed on either a sheet of paper or word processing document), we suggest the following approach to justify your research methods:

1. Write your principal problem at the top of the page, and divide the rest of the page into the two columns.
2. In the left-hand column, write the first subproblem.
3. Immediately below the subproblem, write a description of the data you will need to resolve that subproblem.
4. In the right-hand column, write the criteria you will establish for the admissibility of those data into your research design. In other words, what kinds of data will be acceptable and what kinds will *not* be acceptable for inclusion in your study? Be very specific. For instance, if you are planning to administer a questionnaire to determine people's attitudes about a controversial issue, will you include data from questionnaires that are only partially completed, with some items left unanswered?
5. Repeat steps 2–4 for the remaining subproblems.

Now refer to the earlier sections "Comparing Quantitative and Qualitative Approaches" and "Guidelines: Deciding Whether to Use a Quantitative or Qualitative Approach." Then do the following:

1. Describe the characteristics that the data in your research project will exhibit.
2. For the data you have just described, identify the approach—quantitative, qualitative, or some combination of the two—that would be most appropriate for collecting, analyzing, and interpreting those data. Justify your choice.

## PRACTICAL APPLICATION Developing a Plan of Attack

Once you have determined that your research project is feasible, you can move ahead. Yet especially for a novice researcher, all the things that need to be done—writing and submitting the proposal, getting IRB or IACUC approval, arranging for access to one or more research sites, setting up any experimental interventions you have planned, collecting the data, analyzing and

**FIGURE 4.7**  
Justifying your research approaches

The problem: (Write it as you stated it in Chapter 3.)	
<p>A statement of each subproblem together with a description of those data needed for its solution.</p> <p>Subproblem 1:</p> <p>The data needed for its solution:</p>	<p>The criteria that will govern the admissibility of the data into the research design.</p> <p>Criterion 1:</p> <p>Criterion 2:</p> <p>Criterion 3:</p>

interpreting it, and writing the final research report (almost always in multiple drafts)—may, in combination, seem like a gigantic undertaking. We authors recall, with considerable disappointment and sadness, the many promising doctoral students we have known who took all required courses, passed their comprehensive exams with flying colors, and then never earned their doctoral degrees because they couldn't persevere through the process of completing a dissertation.

You must accept the fact that *your project will take time—lots of time*. All too often, we have had students tell us that they anticipate completing a major research project (e.g., a thesis or dissertation) in a semester or less. In the vast majority of cases, such a belief is unrealistic. Consider the many steps we listed in the preceding paragraph. If you think you can accomplish all of these things within two or three months, you are almost certainly setting yourself up for failure and great disappointment. We would much rather you think of any research project—and especially your first project—as something that is a valuable learning experience in its own right. As such, it is worth however much of your time and effort it takes to do the job well.

The most effective strategy we can suggest here is to *develop a research and writing schedule and try to stick to it*. Figure 4.8 provides a workable format for your schedule. In the left-hand column, list all of the specific tasks you need to accomplish for your research project (writing the proposal, getting approval from the IRB and any other relevant faculty committees, conducting any needed pilot studies, etc.) in the order in which you need to accomplish them. In the second column, estimate the number of weeks or months it will take you to complete each task, always giving yourself a little more time than you think you will need. In the third column, establish appropriate target dates for accomplishing each task, taking into account any holidays, vacations, business trips, and other breaks in your schedule that you anticipate. Also include a little bit of slack time for unanticipated illnesses or family emergencies. Use the right-hand column to check off each step as you complete it.

## Using Project Management Software and Electronic Planners



Project management software is commonplace in the business world. Such software is available both commercially (e.g., FastTrack Schedule, ManyMoon, Milestones, ToDoList) and in freeware available for download from the Internet (e.g., go to [www.ganttproject.biz](http://www.ganttproject.biz) or [download.cnet.com](http://download.cnet.com)). You can use such software to organize and coordinate the various aspects of a research project. For example, it lets you outline the different phases of the project, the dates by which those phases need to be completed, the ways in which they are interconnected, and the person who is responsible for completing each task. This information can be displayed in graphic form with due dates and milestones highlighted.

Project management software is especially helpful when a research project has many separate parts that all need to be carefully organized and coordinated. For example, suppose a large research effort is being conducted in a local school district. The effort requires a team of observers and interviewers to go into various schools and observe teachers in class, interview students during study halls, and discuss administrative issues with the principals. Coordinating the efforts of the many observers, teachers, students, and administrators is a complex task that can be easily laid out and scheduled by project management software.

You might consider computerizing your schedule even if you don't expect your research project to be as multifaceted as the one just described. For instance, computerized versions of day planners (e.g., Daily Planner, DevPlanner, Now Up-to-Date) are widely available. These customized databases can store and organize the various tasks you need to accomplish. You can insert electronic reminders that you need to do certain things on such-and-such a date. And you can easily revise your long-term schedule if unforeseen circumstances occur.

## Keeping an Optimistic and Task-Oriented Outlook

In our own experiences, we have found that a schedule goes a long way in helping us complete a seemingly humongous task. In fact, this is exactly the approach we took when we wrote various editions of this book. Make no mistake about it, writing a book such as this one can be even more overwhelming than conducting a research project!

Task to Complete	Estimated Amount of Time Needed	Target Date for Completion	Task Completed (indicate with a ✓)

**FIGURE 4.8**

Establishing a schedule for your project

A schedule in which you break your project into small, easily doable steps accomplishes several things for you simultaneously. First, it gives you the confidence that you *can* complete your project if you simply focus on one piece at a time. Second, it helps you persevere by giving you a series of target dates that you strive to meet. And last (but certainly not least!), checking off each task as you complete it provides a regular reminder that you are making progress toward your final goal of solving the research problem.

## For Further Reading

### Planning Your Research Design

- Bordens, K. S., & Abbott, B. B. (2010). *Research design and methods: A process approach* (8th ed.). New York: McGraw-Hill.
- Butler, D. L. (2006). Frames of inquiry in educational psychology: Beyond the quantitative-qualitative divide. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 903–927). Mahwah, NJ: Erlbaum.
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- Ercikan, K., & Roth, W.-M. (2006). What good is polarizing research into qualitative and quantitative? *Educational Researcher*, 35(5), 14–23.
- Ethridge, D. (2004). *Research methodology in applied economics: Organizing, planning, and conducting economic research* (2nd ed.). New York: Wiley.
- Firestone, W. A. (1987). Meaning in method: The rhetoric of quantitative and qualitative research. *Educational Researcher*, 16(7), 16–21.
- Hedrick, T. E., Bickman, L., & Rog, D. J. (1993). *Applied research design: A practical guide*. Thousand Oaks, CA: Sage.
- Herr, K., & Anderson, G. L. (2005). *The action research dissertation: A guide for students and faculty*. Thousand Oaks, CA: Sage.
- Jacob, H. (1984). *Using published data: Errors and remedies*. Thousand Oaks, CA: Sage.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14–26.
- Malhotra, N. K. (2010). *Marketing research: An applied orientation* (6th ed.). Upper Saddle River, NJ: Prentice Hall.
- Maxfield, M. G., & Babbie, E. R. (2011). *Research methods for criminal justice and criminology* (6th ed.). Belmont, CA: Wadsworth/Cengage Learning.
- Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Upper Saddle River, NJ: Pearson.
- Singleton, R. A., Jr., & Straits, B. C. (2009). *Approaches to social research* (5th ed.). New York: Oxford University Press.
- Tashakkori, A., & Teddlie, C. (Eds.) (2010). *SAGE handbook of mixed methods in social and behavioral research* (2nd ed.). Thousand Oaks, CA: Sage.
- Wood, M. J., & Ross-Kerr, J. C. (2011). *Basic steps in planning nursing research: From question to proposal* (7th ed.). Sudbury, MA: Jones & Bartlett.

### Measurement

- Aft, L. (1992). *Productivity measurement and improvement* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Earickson, R., & Harlin, J. (1994). *Geographic measurement and quantitative analysis*. Upper Saddle River, NJ: Prentice Hall.

- Fried, H. O., Knox Lovell, C. A., & Schmidt, S. S. (Eds.). (2008). *The measurement of productive efficiency and productivity growth*. New York: Oxford University Press.
- Miller, D. C., & Salkind, N. J. (2002). *Handbook of research design and social measurement* (6th ed.). Thousand Oaks, CA: Sage.
- Thorndike, R. M., & Thorndike-Christ, T. (2010). *Measurement and evaluation in psychology and education* (8th ed.). Upper Saddle River, NJ: Merrill/Pearson Education.

### Ethics

- American Educational Research Association. (1992). Ethical standards of the American Educational Research Association. *Educational Researcher*, 21(7), 23–36.
- American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist*, 57, 1060–1073.
- Bankowski, Z., & Levine, R. J. (Eds.). (1993). *Ethics and research on human subjects: International guidelines*. Albany, NY: World Health Organization.
- Brickhouse, N. W. (1992). Ethics in field-based research: Ethical principles and relational considerations. *Science Education*, 76(1), 93–103.
- Cheney, D. (Ed.). (1993). *Ethical issues in research*. Frederick, MD: University Publishing Group.
- Christians, C. G. (2000). Ethics and politics in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 133–155). Thousand Oaks, CA: Sage.
- Eiserman, W. C., & Behl, D. (1992). Research participation: Benefits and considerations for the special educator. *Teaching Exceptional Children*, 24, 12–15.
- Elliot, D., & Stern, J. E. (Eds.). (1997). *Research ethics: A reader*. Hanover, NH: University Press of New England.
- Erwin, E., Gendin, S., & Kleiman, L. (Eds.). (1994). *Ethical issues in scientific research: An anthology*. New York: Garland.
- Hemmings, A. (2009). Ethnographic research with adolescent students: Situated fieldwork ethics and ethical principles governing human research. *Journal of Empirical Research on Human Research Ethics*, 4(4), 27–38.
- King, N. M. P., & Churchill, L. R. (2000). Ethical principles guiding research on child and adolescent subjects. *Journal of Interpersonal Violence (Special Issue: The Ethical, Legal, and Methodological Implications of Directly Asking Children about Abuse)*, 15, 710–724.
- Loue, S., & Case, S. L. (2000). *Textbook of research ethics: Theory and practice*. New York: Plenum Press.
- Macrina, F. L. (2005). *Scientific integrity: Text and cases in responsible conduct of research* (3rd ed.). Washington, DC: American Society for Microbiology.

- Mertens, D. M., & Ginsberg, P. (Eds.). (2008). *The handbook of social research ethics*. Thousand Oaks, CA: Sage.
- Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Upper Saddle River, NJ: Pearson. [Provides an excellent discussion of ethical issues.]
- Panter, A. T., & Sterba, S. K. (Eds.). *Handbook of ethics in quantitative methodology*. New York: Routledge.
- Pimple, K. D. (2008). *Research ethics*. Aldershot, England: Ashgate.
- Pimple, K. D., Orlans, F. B., & Gluck, J. P. (1997). *Ethical issues in the use of animals in research*. Mahwah, NJ: Erlbaum.
- Roberts, L. W. (2006). Ethical principles and practices for research involving human participants with mental illness. *Psychiatric Services*, 57, 552–557.
- Sales, B. D., & Folkman, S. (Eds.). (2000). *Ethics in research with human participants*. Washington, DC: American Psychological Association.
- Sieber, J. E. (1992). *Planning ethically responsible research: A guide for students and internal review boards*. Thousand Oaks, CA: Sage.
- Yan, E. G., & Munir, K. M. (2004). Regulatory and ethical principles in research involving children and individuals with developmental disabilities. *Ethics & Behavior*, 14(1), 31–49.

### ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE "Identifying Scales of Measurement" on page 88:

1. This is a *ratio* scale, with an absolute zero point (i.e., no bacteria at all).
2. A country's average temperature is an *interval* scale, because an average temperature of 0—regardless of whether the temperature is reported in Fahrenheit or Celsius—still means *some* heat. (In the Kelvin scale, a temperature of 0 means no heat at all, but people typically don't use this scale in reporting climatic temperature.) Amount of tourist dollars is, of course, a *ratio* scale.
3. The party membership coding scheme is a *nominal* scale, because the numbers assigned indicate only category membership, not quantity or order. For example, a Republican (who is coded "2") does not have "twice as much" party membership as a Democrat (who is coded "1"). Meanwhile, voting frequency—how many times each person has voted in the past five years—is a *ratio* scale, with equal units of measurement (every trip to the polls is counted once) and an absolute zero point (a score of 0 means that a person hasn't voted even once in the past five years).
4. The zip code strategy for creating regions is a *nominal* scale, reflecting only category membership. Regions with higher zip codes don't necessarily have "more" of anything, nor are they necessarily "better" in some respect.
5. Don't be distracted by the absolute zero point here (an income of \$0 means no money at all). The ranges of income are different in each group: Group A has a \$20,000 range, Group B has a \$30,000 range, Group C has a \$50,000 range, and Group D—well, who knows how much the richest person in the study makes each year? Because of the unequal measurement units, this is an *ordinal* scale.
6. This is an *ordinal* scale that reflects varying levels of quality. There is no indication that the four categories each reflect the same range of quality, and a true zero point (no road at all!) is not represented by the categorization scheme.
7. This is a tricky one. Despite the 0, this is *not* a ratio scale because virtually all students have at least a tiny amount of anxiety about tests, even if they respond "never" to all 25 questions. But the scale does involve an amount of something, so this must be either an ordinal or interval scale. Many psychologists would argue that the scores reflect an *interval* scale and would treat it as such in their statistical analyses. We authors don't necessarily agree, for two reasons. First, some of the statements on the instruments might reflect higher levels of test anxiety than others, so a "4" response to one item isn't necessarily the equivalent of a "4" response to another. Second, the 5-point rating scale embedded within the instrument ("never" to "always") doesn't necessarily reflect equal intervals of frequency; for instance, perhaps a student thinks of "sometimes" as being a broad range of frequencies of test-anxiety occurrence but thinks of "often" as being a more limited range. Thus, we argue that, in reality, the scores on the test anxiety instrument reflect an *ordinal* scale.

### ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE

"Identifying Problems with Validity and Reliability in Measurement" on page 93:

1. The test lacks *content validity*: It doesn't reflect the content domain that the instruction covered—actual tennis *skills*.
2. This is a problem of *equivalent forms reliability*: The different versions of the same instrument yield different results.
3. The issue here is *interrater reliability*: The researchers are evaluating the same behaviors differently.
4. The problem in this case is *criterion validity*: Two measures of energy level—blood-test results and observer ratings—yield very different results.
5. This is a problem of *test-retest reliability*: The students responded differently within a very short time interval (two weeks at most), even though their general opinions about climate change probably changed very little, if at all, during that interval.
6. The questionnaire lacks *internal consistency reliability*: Different items in the instrument yield different results, even though all items are intended to measure a single characteristic: matrimonial harmony.
7. In this case the instrument's *construct validity* is suspect. Religious tolerance is a hypothesized internal characteristic that can be inferred and measured only indirectly through observable patterns in people's behaviors. Here the behaviors being observed are simply responses to a questionnaire.
8. *Face validity* is at stake here: Although many psychologists contend that intelligence involves reaction time to some degree—and thus the reaction-time task might be a valid measure—on the surface the task doesn't *appear* to be a good measure. Although face validity is not "true" validity, a lack of it can sometimes negatively impact participants' cooperation in a research project.

