Chapter 4

Planning your Research Project

Before constructing a home, a builder acquires or develops a detailed set of plans—how to frame the walls and roof, where to put doors and windows of various sizes, where to put pipes and electrical wiring, what kinds of materials to use, and the like. These plans enable the builder to erect a strong, well-designed structure. Researchers should pay similar attention to detail in planning a research project.

Learning Outcomes

4.1 Distinguish between primary data and secondary data, and describe a variety of forms that data for a research project might take.

4.2 Compare quantitative versus qualitative research methodologies in terms of their typical purposes, processes, data collection strategies, data analyses, and nature of the final reports.

4.3 Explain the difference between the internal validity and external validity of a research study. Also explain how you might use different strategies to determine the validity of a quantitative study versus that of a qualitative study.

4.4 Differentiate between substantial and insubstantial phenomena, as well as among nominal, ordinal, interval, and ratio scales.

4.5 Describe several different types of validity and reliability related to specific measurement techniques. Also, describe various strategies you might use to either determine or enhance the validity and/or reliability of a measurement technique.

4.6 Discuss ethical issues related to protection from harm, voluntary and informed participation, right to privacy, and honesty with professional colleagues. Also, explain the roles of internal review boards and professional codes of ethics in minimizing potential ethical problems in a research study.

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 time and expense are 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 your data will take—always with the central goal of solving your research problem in mind—at the very beginning of your project.

PLANNING A GENERAL APPROACH

In planning a research design, a researcher in quest of new knowledge and understandings cannot be shackled by discipline-specific methodological restraints. The course of a research project will frequently lead the researcher 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 data for resolving the research problem.

Instead of limiting their thinking to departmentalized knowledge, researchers might better 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 research problems relating to children, senior citizens, families, communities, cultural groups, ancestors, employees, mental and physiological processes, learning, motivation, social and educational problems, crime, rehabilitation, medical treatments, nutrition, language, and religion.
- ***Things.*** In this category are research problems relating to animal and vegetable life, viruses and bacteria, inanimate objects (rocks, soil, buildings, machines), matter (molecules, atoms, subatomic matter), stars, and galaxies.
- ***Records.*** In this category are research problems relating to newspapers, personal journals, letters, Internet websites, registers, speeches, minutes, legal documents, mission statements, census reports, archeological remains, sketches, paintings, and music.
- ***Thoughts and ideas.*** In this category are research problems relating to concepts, theories, perceptions, opinions, beliefs, reactions, issues, semantics, poetry, and political cartoons.
- ***Dynamics and energy.*** In this category are 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, and 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 a 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 data 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 is probably better that you abandon this project in favor of one for which you have the appropriate knowledge, skills, and resources. Your research should be *practical* research, built on precise and realistic *planning* and executed within the framework of a clearly conceived and feasible *design.*

THE NATURE AND ROLE OF DATA IN RESEARCH

Research is a viable approach to a problem only when data can be collected 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.

Researchers must always remember that data are not absolute reality or truth—if, in fact, any single "realities" and "truths" can ever be determined. (Recall the discussions of *postpositivism* and *constructivism* in Chapter 1.) Rather, data are merely *manifestations* of various physical, social, or psychological phenomena that we want to make better sense of. For example, we often see what other people do—the statements they make, 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!

Data Are Transient and Ever Changing

Data are rarely permanent, unchanging entities. Instead, they are transient—they may have validity for only a split second. Consider, for 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 to the next block, the data they have collected are already out of date. Some people in the previous block who voiced a particular opinion may have seen a television program or heard a discussion that changed their opinion. Some people may have moved away, and others may have moved in; some may have died, and others may 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

For now, let's take a *positivist* perspective and assume that out there—somewhere—is a certain Absolute Truth waiting to be discovered. A researcher's only perceptions of this 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.

Imagine, for a moment, that you live in a dungeon, where you can never see the sun—the Truth. Instead, you see a beam of sunlight on the dungeon floor. This light might give you an idea of what the sun is like. The direct beam of sunlight is *primary data.* Although the shaft is not the sun itself, it has come directly from the sun.[1]

But now imagine that, rather than seeing a direct beam of light, you see a diffused pattern of shimmering light on the floor. The sunlight (primary data) has fallen onto a shiny surface and then been reflected—distorted by imperfections of the shiny surface—onto the floor. The pattern is in some ways similar but in other ways dissimilar to the original shaft of light. This pattern of reflected 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 idea that an accident might occur until the car went out of control. Neither you nor the driver will ever be able to determine the 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 and then writes an account of the accident for the local paper. When your sister reads the account the next morning, she gets, as it were, the reflected-sunlight-on-the-floor version of the event. The newspaper article provides secondary data. The data are inevitably 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 others actually observed.

Figure 4.1 represents what we have been saying about data and their relation to any possible Truth that might exist. 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, including the limitations of the human senses, the weaknesses of instrumentation, the inability of language to communicate people's thoughts precisely, and the inability of two human beings to witness the same event and report it in exactly the same way.

Researchers must never forget the overall idea underlying Figure 4.1. Keeping it in mind can prevent them from making exaggerated claims or drawing unwarranted conclusions. No researcher can ever glimpse Absolute Truth—if such a thing exists at all—and researchers can perceive data that reflect that Truth only through imperfect senses and imprecise channels of communication. Such 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. . . .*

Planning for Data Collection

Basic to any research project are several fundamental questions about the data. To avoid serious trouble later on, the researcher must answer them 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 a ridiculously simple one, but in fact a specific, definitive answer to it is fundamental to any research effort. To resolve the

[1]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*.

How will you get them? Careful attention to this question marks the difference between a viable research project and a pipe dream.

4. ***What limits will be placed on the nature of acceptable data?*** Not all gathered data will necessarily be acceptable for use in a research project. Sometimes 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. The restrictions identified are sometimes called the **criteria for the admissibility of data**.

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. 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?

Now imagine a sociologist who plans to conduct a survey to determine people's attitudes and beliefs about a controversial issue in a particular area of the country. The sociologist constructs a 10-item survey that will be administered and collected at various shopping malls, county fairs, and other public places over a 4-week period. Some people will respond to all 10 items, but others may respond to only a subset of the items. Should the sociologist include data from surveys that are only partially completed, with some items left unanswered? And what about responses such as "I don't want to waste my time on such stupid questions!"—responses indicating that a person was not interested in cooperating?

The agronomist and the sociologist should be specific about such things—ideally, in sufficient detail that another researcher might reasonably replicate their studies.

5. ***How will the data be interpreted?*** This is perhaps the most important question of all. The four former hurdles have been overcome. You have the data in hand. But you must also spell out precisely what you intend to do with them 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 rethink the nature of your problem. If, instead, both answers are *yes,* a next important step is to consider an appropriate methodology.

LINKING DATA AND RESEARCH METHODOLOGY

Data and methodology are inextricably intertwined. For this reason, the methodology chosen 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 airliner. No one else in his village has ever ridden in an airplane, so after he returns home, his friends ask him about his trip. One friend asks, "How fast did you move?" "How far did you go?" and "How high did you fly?" A second one 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, but because they ask different kinds of questions, they obtain different kinds of information. Although neither of them gets the "wrong" story, neither does each one get the whole story.

In research, too, different questions yield different kinds 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 desired data dictate the research method.* As an example, consider historical data, those pieces of information gleaned from written records of past events. You can't extract much meaning from historical documents by conducting 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 and understandings.

Comparing Quantitative and Qualitative Methodologies

On the surface, quantitative and qualitative approaches involve similar processes—for instance, they both entail identifying a research problem, reviewing related literature, and collecting and analyzing data. But by definition, they are suitable for different types of data: Quantitative studies involve numerical data, whereas qualitative studies primarily make use of nonnumerical data (e.g., verbal information, visual displays). And to some degree, quantitative and qualitative research designs are appropriate for answering different kinds of questions.

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 for 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 others 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 factors make case-based instruction more effective or less effective?" 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 learn 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 4.1 presents typical 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 tend to seek explanations and predictions that will generalize to other persons and places. The intent is to identify relationships among two or more variables and then, based on the results, to confirm or modify existing theories or practices.

Qualitative researchers tend to seek better understandings 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.

and then drawing logical conclusions from them. They also try to maintain objectivity in their data analysis, conducting predetermined statistical procedures and using relatively 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 might formulate a preliminary theory through 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 summarizing statistics (e.g., means, medians, correlation coefficients). In most cases, *average* performances are of greater interest than the performances of specific individuals (you will see exceptions in the single-subject designs described in Chapter 7). Results are typically presented in a report that uses a formal, scientific style with 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

Given that quantitative and qualitative methodologies are useful in answering somewhat different kinds of questions and solving somewhat different kinds of research problems, we can gain better understandings of our physical, social, and psychological worlds when we have both methodologies at our disposal. Fortunately, the two approaches aren't necessarily mutually exclusive; many researchers successfully combine them in a *mixed-methods design.* For example, it isn't 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 12.

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 doesn't 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)

CONSIDERING THE VALIDITY OF YOUR METHOD

No matter what research methodology you choose, you must think about the general *validity* of your approach for your purpose—the likelihood that it will yield accurate, meaningful, and credible results that can potentially help you address your research problem. 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.

Researchers use a variety of strategies to support the validity of their findings. Different strategies are appropriate in different situations, depending on the nature of the data and the specific methodologies used. In the following sections, we examine two concepts—internal validity and external validity—that originated in discussions of quantitative research (Campbell & Stanley, 1963). However, some qualitative researchers have questioned the relevance of these two concepts to qualitative designs; thus, in a subsequent section, we present validation strategies that qualitative researchers often use.

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 3 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 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 anything 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 differed from each other in several ways (e.g., the presence of teenagers, the amount of action) in addition to humor. And we shouldn't overlook the fact that the humorous commercial aired during the summer months. People are more likely to drink soft drinks (including Zowie) when they're hot.

Internal validity is especially of concern in experimental designs, where the specific intent is to identify cause-and-effect relationships; accordingly, we revisit this issue in Chapter 7. But to some degree, internal validity is important in *any* research study. Researchers and those who read their research reports must have confidence that the conclusions drawn are warranted from the data collected.

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. 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 might 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.[3]
- ***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 the same kinds of 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 6 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. Taken together, these studies provide evidence that the conclusion has validity and applicability across diverse situations.

You have previously encountered the distinction between *basic research* and *applied research* in Chapter 2. Well-designed basic research—research conducted under tightly controlled (and possibly artificial) conditions—ensures internal validity; that is, it allows the researcher to rule

[3]The artificial nature of laboratory research has been a concern in psychology for many years. In most cases, however, studies conducted in a laboratory and those conducted in real-world settings lead to the same conclusions about human nature, especially when lab-based studies reveal large differences among treatment groups (e.g., see C. A. Anderson, Lindsay, & Bushman, 1999; G. Mitchell, 2012).

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 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

Qualitative researchers don't necessarily use the term *validity* in describing their research; they may instead use such words as *quality, credibility, trustworthiness, confirmability,* and *interpretive rigor* (Creswell, 2013; Lincoln & Guba, 1985; O'Cathain, 2010; Teddlie & Tashakkori, 2010). Nevertheless, they do take certain precautions to substantiate their methods, findings, and conclusions. As noted earlier, they often use *triangulation*—comparing multiple data sources in search of common themes—to give credence to their findings. Following are several additional strategies they employ:

- ***Extensive time in the field.*** A 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.
- ***Analysis of outliers and contradictory instances.*** A researcher actively looks for examples that are inconsistent with existing hypotheses, then continually revises his or her explanation or theory until all examples have been accounted for.
- ***Thick description.*** A researcher who uses **thick description** describes a situation in sufficiently rich, "thick" detail that readers can draw their own conclusions from the data presented.
- ***Acknowledgment of personal biases.*** Rather than claim to be an objective, impartial observer, a researcher describes personal beliefs and attitudes that may potentially be slanting observations and interpretations.
- ***Respondent validation.*** In **respondent validation**, a researcher takes 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?
- ***Feedback from others.*** A 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.

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 credibility and worth, 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.

IDENTIFYING MEASUREMENT STRATEGIES

Especially if you are planning a quantitative research project, you must also determine how you will *measure* the variables you intend to study. 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.

Appropriate measurement procedures provide a solid basis on which any good quantitative study 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.

We should note here that *some* measurement is almost inevitable in qualitative research as well. At a minimum, qualitative researchers are apt to *count* things—perhaps the members of certain groups or the frequencies of certain events. And during data analyses, many of them *code* their observations to reflect various categories into which different observations fall. Because their measurement strategies are often specific to certain qualitative designs and may continue to be refined over the course of a study (recall our earlier point that qualitative designs are often *emergent* in nature), we postpone discussion of such strategies until Chapter 11.

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 zoom in on various parts of this definition. The first five words are *measurement is limiting the data.* When we measure something, we constrain 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.* In some cases, observable physical entities are measured. These are **substantial phenomena**; 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 phenomena**, 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.[4]

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.

[4]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).

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.

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 their 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. For example, imagine that we want to examine—and also to *measure*—the interpersonal dynamics within a small group of people. Let's take a group of nine people who work together in the human resources department of a large corporation. They attend a recognition dinner at an exclusive hotel and 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 4.2.

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 group member a slip of paper on which to write three sets of names, one set each for (a) one or more individuals in 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 4.3. We might also assign "weights" that place the data into

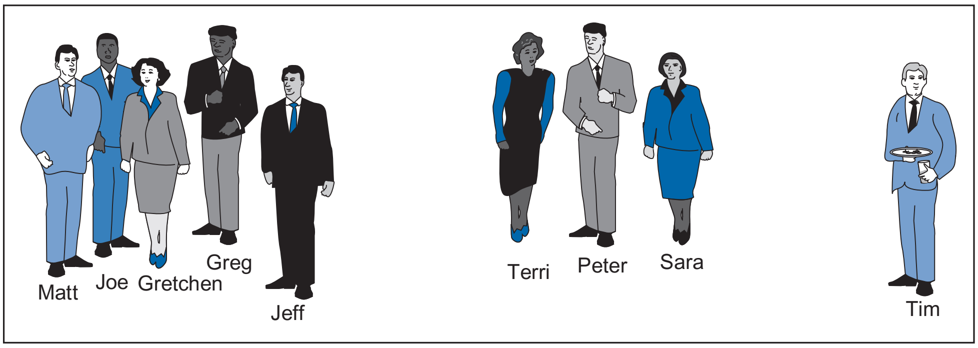

FIGURE 4.2 Conversation Groups in a Hypothetical Human Resources Department

TABLE 4.3 Data from Figure 4.3 Presented as a Sociometric Matrix

How Each Person Rated the Others	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

- 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.
- Tim is apparently 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 didn't just observe the behaviors of nine individuals at a social event; we also looked below the surface to identify possible hidden social forces at play. 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**: nominal, ordinal, interval, and ratio (Stevens, 1946). The scale of measurement will ultimately dictate the statistical procedures (if any) that can be used in processing the data.

Nominal Scales

The word *nominal* comes from the Latin *nomen,* meaning "name." Hence we might "measure" data to some degree simply by assigning a name to each data point. 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 where each of them lives. 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, people measured the population of the entire nation in just such a manner.

TABLE 4.4 ■ 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 make proportional comparisons; allows one to conduct virtually any inferential statistical analysis

Table 4.4 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 8), 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.

CONCEPTUAL ANALYSIS EXERCISE Identifying Scales of Measurement

USING TECHNOLOGY

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 1 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 market researcher is studying the relationship between (a) a country's average annual relative humidity levels and (b) the number of raincoats sold in the country during the year. What scales underlie the two variables in this study?
3. A sports journalist in Spain wants to determine whether a football club's fan club membership correlates to the frequency with which the club has won in the Spanish football league during the past five years. The researcher can easily obtain information about fan club membership and wins from the club's records. To simplify data collection, the researcher uses the following coding scheme for the fan club membership: 1 = FC Barcelona, 2 = Real Madrid, and 3 = Atletico de Madrid. What measurement scale(s) underlie (a) fan club membership and (b) frequency of wins?
4. A new audio system was installed in a theatre that has a seating capacity of 500 people. The seating area has been divided into three distinct areas: platinum, gold, and silver.

At the end of every show in one week, the audience was asked to fill out a questionnaire to determine the effectiveness of the new audio system. What measurement scale does the researcher's coding scheme for the seating area represent?

5. An advertiser is studying the impact of a chocolate commercial on different age groups. He puts the age groups in four categories: Group 1 includes 3- to 12-year-olds, Group 2 includes teenagers aged 13 to 19, Group 3 includes adults aged 20 to 49, and Group 4 includes people aged 50 and above. What kind of scale is the classification of age groups in this study?
6. A car manufacturer is studying sales patterns over five years of four different types of car that vary in size and available utilities: hatchbacks (small cars with no trunk space), midsized sedans (cars with limited trunk space and more utilities compared to hatchbacks), sedans (proper trunk space and all utilities), and luxury sedans (big trunk space and luxurious interiors). Based on size and available utilities, which measurement scale does the type of car reflect?
7. A child psychologist is developing an instrument designed to measure the social etiquette of small children. The instrument includes 25 statements; for example, "wishes elders when they meet them," "courteous to their sibling," and "says please when asking for something or thank you when someone has helped." Parents must rate each of these statements on a 5-point scale as follows:

   0 never
   1 rarely
   2 sometimes
   3 often
   4 always

   Parents who answer "never" to each of the 25 questions get the lowest possible score of 0 on the instrument. Parents 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

Earlier in the chapter we discussed the importance of determining that your chosen method will have validity for your purpose—that it will yield meaningful, credible results. When used to describe a measurement tool, however, the term *validity* has a somewhat different meaning. Regardless of the type of scale a measurement instrument involves, the instrument must have both validity and another, related characteristic—*reliability*—for its intended purpose. The validity and reliability of measurement instruments influence the extent to which a researcher can legitimately 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.

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 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 phenomenon 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. As noted earlier, head size is *not* a good indication of intelligence level.

Creative researchers use a 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 modifying 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 examine 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

USING TECHNOLOGY

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
- Content validity
- Criterion validity
- Construct validity
- Interrater reliability
- Test–retest reliability
- Equivalent forms reliability
- Internal consistency reliability

8. A novice film critic attempted to find out how good a certain movie is by interviewing viewers on the first day of its release. He passed around a questionnaire to the viewers with certain standard queries such as "What made you decide to watch this movie?" and "Is there good humor in the movie?" Afterwards, the film critic shared his research results with the magazine he was working for. The chief editor analyzed the questions and answers, and pointed out that the questions did not ask about how good the movie was but focused on the individual's preference for watching that movie. The questions also tended to bracket the movie within a certain class.

ETHICAL ISSUES IN RESEARCH

In 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 the 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—during the *planning* stage—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 then describe the internal review boards and professional codes of ethics that provide guidance for researchers.

Protection from Harm

Researchers should not expose research participants—whether they be human beings or nonhuman 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 especially 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 to engage in some kinds of activities. Special care must be taken with participants who cannot easily advocate for their own needs and desires—such as children, elderly individuals, and people with significant physical or mental disabilities.

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).[5]

[5] 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).

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—such as birth certificates, newspaper articles, and Internet websites—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 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 might 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 4.4. 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.

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 used different parts of the library and the frequency with which people hiked along certain trails in a national park. Both of these examples involved behaviors within the scope of participants' normal activities.

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. Knowledge about participants' individual performances should be revealed *only* to any co-researchers who have a significant role in the research investigation unless 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—to be blunt—makes you a thief. 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 don't hesitate to acknowledge their indebtedness to others.

Internal Review Boards

Historically, some researchers had few (if any) scruples about the harm they inflicted on certain people or animals. Among the most notorious were German doctors who conducted horrific experiments on concentration camp prisoners during World War II—experiments that sometimes resulted in death or permanent disabilities. Other researchers, too, exposed people or animals to conditions that created significant physical or psychological harm, with virtually no oversight by more ethical colleagues. Fortunately, safeguards are now in place in many countries to keep inappropriate experimentation in check.

In the United States, in Canada, and among members of the European Union, any college, university, or research institution must have an **internal review board (IRB)**[6] 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 (a) quickly declare it *exempt from review,* (b) give it an *expedited review,* or (c) bring it before the board for a *full review.* In any case, the researcher cannot begin the study until either the board has given its seal of approval or 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 Tufts University (tnemcirb.tufts.edu), the University of Northern Colorado (unco.edu/osp/ethics), or the University of Texas (utexas.edu/research/rsc). You can find other helpful sites on the Internet by using a search engine (e.g., Google, Bing, or Yahoo!) and such keywords as *IRB*, *human participants*, and *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 that it's possible to do so. For examples of research institutions' IACUC policies and procedures, we refer you to the University of Maryland (umresearch.umd.edu/IACUC) and the University of Arizona (uac.arizona.edu).

Many novice researchers view IRB and IACUC reviews as a bothersome hurdle to jump in their efforts to carry out a successful research project. We authors can assure you that members of these boards want to encourage and support research—*not* impede it—and typically work hard to make their proposal reviews as quick and painless as possible. Also, they can give helpful advice to ensure that your study does not needlessly jeopardize the welfare of participants in your study.

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 the Internet. Following are examples of organizational websites with ethical codes related to research in their disciplines:

American Anthropological Association (aaanet.org)

American Association for Public Opinion Research (aapor.org)

[6]Some institutions use a different label (e.g., *Institutional Review Board*, *Committee for Protection of Human Subjects*).

American Educational Research Association (aera.net)
American Psychological Association (apa.org)
American Sociological Association (asanet.org)
Society for Conservation Biology (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 (a) attended to the nature and availability of the data you need; (b) decided whether a quantitative, qualitative, or mixed-methods methodology is best suited to address your research problem; (c) possibly identified valid, reliable ways of measuring certain variables; and (d) 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 you must definitely 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 practicality. 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?

____

OVERALL ASSESSMENT

_____ 14. 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 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 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 data, analyzing and 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. Such a waste! we thought then . . . and continue to think now.

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 listed in the preceding paragraph. If you think you can accomplish all these things within 2 or 3 months, you are almost certainly setting yourself up for failure and 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 4.5 provides a workable format for your schedule. In the left-hand column, list all 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 TECHNOLOGY

Using Project Management Software and Electronic Planners

Project management 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 ganttproject.biz or freedcamp.com). You can use such software to organize and coordinate the various aspects of a research project. For example, it will let 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 school 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 electronically organizing your schedule even if you don't expect your research project to be as multifaceted as the one just described. For example, you might use the *calendar* application that comes with your laptop or smartphone, or you might download day-planning freeware from the Internet (e.g., My Daily Planner and Free Day Planner are two possibilities). With such applications 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 authors 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!

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. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th 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.

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.

Kerlinger, F. N., & Lee, H. B. (1999). *Foundations of behavioral research* (4th ed.). New York: Harcourt.

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.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Los Angeles: Sage.

Neuman, W. L. (2011). *Social research methods: Qualitative and quantitative approaches* (7th ed.). Upper Saddle River, NJ: Pearson.

O'Cathain, A. (2010). Assessing the quality of mixed methods research: Toward a comprehensive framework. In A. Tashakkori & C. Teddlie (Eds.), *Mixed methods in social & behavioral research* (2nd ed., pp. 531–555). Thousand Oaks, CA: Sage.

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.

Vogt, W. P., Gardner, D. C., & Haeffele, L. M. (2012). *When to use what research design.* New York: Guilford Press.

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. (2000). *Work measurement and methods improvement.* New York: Wiley.

Campbell, D. T., & Russo, M. J. (2001). *Social measurement.* Thousand Oaks, CA: Sage.

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.

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.

Elliott, 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.

Israel, M., & Hay, I. (2006). *Research ethics for social scientists.* Thousand Oaks, CA: Sage.

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.) (2012). *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.

Rhodes, C. S., & Weiss, K. J. (Eds.) (2013). *Ethical issues in literacy research.* New York: Routledge.

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., & Tolich, M. B. (2013). *Planning ethically responsible research* (2nd ed.). 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":

1. This is a *ratio* scale, with an absolute zero point (i.e., no bacteria at all).
2. Relative humidity (RH) varies from 0% to 100% on a scale. For average RH levels, a range is chosen, which tends to make this an interval scale. Note that 0% RH is not an absolute value as there will always be traces of water in the air even at 0%. However, the number of raincoats can be zero, which is an absolute value; hence, this is a ratio scale.
3. The fan club membership coding scheme is a *nominal* scale, because the numbers assigned≈indicate only the fan club they are member of, not quantity or order. For example, a Real Madrid fan club member (who is coded "2") does not have "twice as much" fan club membership as a FC Barcelona fan club member (who is coded "1"). Frequency of wins (the number of times each club has won in the Spanish League during the last five years) is a *ratio* scale with an absolute zero point (i.e., no wins in the last five years).
4. The strategy for creating three distinct areas—platinum, gold, and silver—in the seating area reflects an *ordinal* scale because of the varying levels of quality associated with each category of seats.
5. Note that the ranges for all age groups are different, and the age of the oldest person is not identified. Since the measurement units are unequal, this is an *ordinal* scale.
6. This is an *ordinal* scale that reflects varying sizes and available utilities. An absolute zero or no car at all is not being considered here.
7. There is a zero available, and an equal interval is given. This is neither a ratio scale nor an interval scale; the ratings are based on the parent's perception and expectations of the child, which will vary from parent to parent. A parent with high expectations may rate "never" for their child on some areas of etiquette even if the child sometimes adheres to the requirement. Hence, a zero may not be an accurate indication of judgment and cannot be absolute. Therefore, it is not a ratio scale. The second possibility is an interval scale. Note that the terms used for identification of the score, such as "rarely," "sometimes," and "often," cannot be quantified exactly. How many times is "often"? Is it the same for Parent A and Parent B? Six times out of ten may be "often" for some, while nine would be for someone else. Such words do not reflect equal intervals. The ranges vary in all cases. Hence, this is an *ordinal* scale.

ANSWERS TO THE CONCEPTUAL ANALYSIS EXERCISE "Identifying Problems with Validity and Reliability in Measurement":

1. The questionnaire does not demonstrate *content validity* because the components do not measure the effectiveness of the shampoo in getting rid of dandruff.
2. As different versions of the performance evaluation form yielded different results, this is a problem of *equivalent forms reliability*.
3. Since the researchers are evaluating the same behaviors differently, this is an issue of *interrater reliability.*
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 blood tests indicated different values within a very short time interval (2 hours) during which no change is likely to have occurred in the preserved samples.
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. A happiness level can be inferred and measured only indirectly by observing people's behaviors. Since the behaviors being observed here are simply responses to a questionnaire, we can conclude that the instrument's *construct validity* is suspect.
8. The film critic was able to secure participants' cooperation with his questionnaire, but the questions do not appear to be a good measure of how good the film was. This is an issue of *face validity.*