Ninth Edition

MASS MEDIA RESEARCH

An Introduction

ROGER D. WIMMER
Wimmer Research

JOSEPH R. DOMINICK
University of Georgia

WADSWORTH
CENGAGE Learning

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

WADSWORTH
CENGAGE Learning™

Mass Media Research: An Introduction, Ninth Edition
Roger D. Wimmer and Joseph R. Dominick

Senior Publisher: Lyn Uhl
Publisher: Michael Rosenberg
Developmental Editor: Stacey Sims
Assistant Editor: Jillian D'Urso
Editorial Assistant: Erin Pass
Media Editor: Jessica Badiner
Marketing Manager: Bryant Chrzan
Marketing Coordinator: Darlene Macanan
Marketing Communications Manager: Christine Dobberpuhl
Senior Content Project Manager: Rosemary Winfield
Art Director: Linda Helcher
Production Service: Pre-PressPMG, Sini Sivaraman
Manufacturing Buyer: Justin Palmeiro
Manufacturing Manager: Denise Powers
Rights Acquisitions Manager, Text: Margaret Chamberlain-Gaston
Rights Acquisitions Manager, Image: Mandy Groszko
Cover Image: © Corbis
Compositor: Pre-PressPMG

Library of Congress Control Number: 2009936675

ISBN-13: 978-1-4390-8274-4

ISBN-10: 1-4390-8274-X

Wadsworth
20 Channel Center Street
Boston, MA 02210
USA

Cengage Learning is a leading provider of customized learning solutions with office locations around the globe, including Singapore, the United Kingdom, Australia, Mexico, Brazil, and Japan. Locate your local office at **www.cengage.com/global**

Cengage Learning products are represented in Canada by Nelson Education, Ltd.

To learn more about Wadsworth, visit **www.cengage.com/wadsworth**

Purchase any of our products at your local college store or at our preferred online store **www.ichapters.com**

Printed in the United States of America
1 2 3 4 5 6 7 13 12 11 10 09

CHAPTER 2

ELEMENTS OF RESEARCH

CHAPTER OUTLINE

Chapter 1 presented an overview of the research process. In this chapter, we define and discuss four basic elements of this process: concepts and constructs, measurement, variables, and scales. A clear understanding of these elements is essential to conduct precise and meaningful research.

CONCEPTS AND CONSTRUCTS

A **concept** is a term that expresses an abstract idea formed by generalizing from particulars and summarizing related observations. For example, a researcher might observe that a public speaker becomes restless, starts to perspire, and fidgets with a pencil just before giving a speech. The researcher might summarize these observed patterns of behavior and label them "speech anxiety." On a more ordinary level, the word *table* is a concept that represents a wide variety of observable objects, ranging from a plank supported by concrete blocks to a piece of furniture commonly found in dining rooms. Typical concepts in mass media research include terms such as *advertising effectiveness*, *message length*, *media usage*, and *readability*.

Concepts are important for at least two reasons. First, they simplify the research process by combining particular characteristics, objects, or people into general categories. For example, a researcher may study families that own computers, modems, MP3 players, cell phones, and DVD or Blu-Ray machines. To make it easier to describe these families, the researcher calls them "Taffies" and categorizes them under the concept of "technologically advanced families." Instead of describing each of the characteristics that make these families unique, the researcher has a general term that is more inclusive and convenient to use.

Second, concepts simplify communication among those who have a shared understanding of them. Researchers use concepts to organize their observations into meaningful summaries and to transmit this information to others. Researchers who use the concept of "agenda setting" to describe a complicated set of audience and media activities find that their colleagues understand what is being discussed. Note that people must share an understanding of a concept for the concept to be useful. For example, when teenagers use the word *emo* to describe a person, most of their peers understand perfectly what is meant by the concept, although adults may not.

A **construct** is a concept that has three distinct characteristics: First, it is an abstract idea that is usually broken down into dimensions represented by lower-level concepts; a construct is a combination of concepts. Second, because of its abstraction, a construct usually cannot be observed directly. Third, a construct is usually designed for a specific research purpose so that its exact meaning relates only to the context in which it is found. For example, the construct "involvement" has been used in many advertising studies (search the Internet for "*advertising involvement*"). Advertising involvement is a construct that is difficult to see directly, and it includes the concepts of attention, interest, and arousal. Researchers can observe only its likely or presumed manifestations. In some contexts, involvement means a subject's involvement with the product; in others, it refers to involvement with the message or even with the medium. Its precise meaning depends on the research context.

Another example in mass communication research is the term *authoritarianism*, which *represents* a construct defined to describe a certain type of personality; it involves nine different concepts, including conventionalism, submission, superstition, and cynicism. Authoritarianism itself cannot be seen, so some type of questionnaire or standardized test is used to determine its presence. The results of such tests indicate what authoritarianism might be and whether it is present under given conditions, but the tests do not provide exact definitions for the construct itself.

The empirical counterpart of a construct or concept is called a **variable**. Variables are important because they link the empirical world with the theoretical; they are the phenomena and events that are measured or manipulated in research. Variables can have more than one value along a continuum. For example, the variable "satisfaction with pay-per-view TV programs" can take on different values—a person can be satisfied a lot, a little, or not at all—reflecting in the empirical world what the concept "satisfaction with pay-per-view TV programs" represents in the theoretical world.

Researchers try to test a number of associated variables to develop an underlying meaning or relationship among them. After suitable analysis, the most important variables are kept and the others are discarded. These important variables are labeled **marker variables** because they tend to define or highlight the construct under study. After additional studies, new marker variables may be added to increase understanding of the construct and to allow for more reliable predictions.

Concepts and constructs are valuable tools in theoretical research, but, as noted in Chapter 1, researchers also function at the observational, or empirical, level. To understand how this is done, it is necessary to understand variables and how they are measured.

INDEPENDENT AND DEPENDENT VARIABLES

Variables are classified in terms of their relationship with one another. It is customary to talk about independent and dependent variables.

Independent variables are systematically varied by the researcher; **dependent variables** are observed and their values are presumed to depend on the effects (influence) of the independent variables. In other words, the *dependent variable is what the researcher wishes to explain.* For example, assume a researcher is interested in determining how the angle of a camera shot affects an audience's perception of the credibility of a television newscaster. Three versions of a newscast are recorded: one shot from a very low angle, another from a high angle, and a third from eye level. Groups of subjects are randomly assigned to view one of the three versions and complete a questionnaire to measure the newscaster's credibility. In this experiment, the camera angle is the independent variable. The experimenter, who selects only three of the camera angles possible, systematically varies its values. The dependent variable is the perceived credibility of the newscaster as measured by the questionnaire. If the researcher's assumption is correct, the newscaster's credibility will vary according to the camera angle. (The values of the dependent variable are not manipulated; they are simply observed or measured.)

The distinction between types of variables depends on the purposes of the research. An independent variable in one study may be a dependent variable in another. Also, a research task may involve examining the relationship of more than one independent variable to a single dependent variable. For example, the researcher in the previous example could investigate the effects of camera angles and of how the newscaster's manner, or style, in closing the program affects his or her credibility (as perceived by the viewers). In many instances, multiple dependent variables are measured in a single study, which is called a **multivariate analysis**.

Discrete and Continuous Variables

Two forms of variables are used in mass media investigation. A **discrete variable** includes only a finite set of values; it cannot

A CLOSER LOOK

Mass Media Variables

Analysis of why people like certain movies, magazines, newspapers, or radio or television shows has historically been difficult because of the number of variables to consider. Even when researchers develop a relatively stable set of variables to measure, assessing popularity of the media is difficult because respondents say something like, "It depends on my mood." As a media researcher, how would you address this problem?

be divided into subparts. For instance, the number of children in a family is a discrete variable because the unit is a person. It does not make much sense to talk about a family size of 2.24 because it is hard to conceptualize 0.24 of a person. Political affiliation, population, and gender are other discrete variables.

A **continuous variable** can take on any value, including fractions, and can be meaningfully broken into smaller subsections. Height is a continuous variable. If the measurement tool is sophisticated enough, it is possible to distinguish between one person 72.12 inches tall and another 72.13 inches tall. Time spent watching television is another example; it is perfectly meaningful to say that Person A spent 3.12 hours viewing while Person B watched 3.13 hours. The *average* number of children in a family is a continuous variable; thus, in this context, it may be perfectly meaningful to refer to 0.24 of a person.

When dealing with continuous variables, researchers should keep in mind the distinction between the variable and the measure of the variable. If a child's attitude toward television violence is measured by counting his or her positive responses to six questions, then there are only seven possible scores: 0, 1, 2, 3, 4, 5, and 6. However, it is entirely likely that the underlying variable is continuous even though the measure is discrete. In fact, even if a fractionalized scale were developed, it would still be limited to a finite number of scores. As a generalization, most of the measures in mass media research tend to be discrete approximations of continuous variables.

Variables measured at the nominal level are always discrete variables. Variables measured at the ordinal level are generally discrete, although some underlying continuous measurement dimension may exist. (Nominal and ordinal levels are discussed later in this chapter.) Variables measured at the interval or ratio level can be either discrete (number of magazine subscriptions in a household) or continuous (number of minutes per day spent reading magazines). Both the level of measurement and the type of variable under consideration are important in developing useful measurement scales.

Other Types of Variables

In nonexperimental research, where there is no active manipulation of variables, different terms are sometimes substituted for independent and dependent variables. The variable that is used for predictions or is assumed to be causal (analogous to the independent variable) is sometimes called the **predictor, or antecedent, variable**. The variable that is predicted or assumed to be affected (analogous to the dependent variable) is sometimes called the **criterion variable**.

Researchers often wish to control certain variables to eliminate unwanted influences.

These **control variables** are used to ensure that the results of the study are due to the independent variables, not to another source. However, a control variable need not always be used to eliminate an unwanted influence. On occasion, researchers use a control variable such as age, gender, or socioeconomic status to divide subjects into specific, relevant categories. For example, in studying the relationship between newspaper readership and reading ability, researchers know that IQ will affect the relationship and must be controlled; thus, subjects may be selected based on IQ scores or placed in groups with similar IQ scores.

One of the most difficult steps in any type of research is to identify all the variables that may create spurious or misleading results. Some researchers refer to this problem as **noise**. Noise can occur in even simple research projects. For example, a researcher might design a telephone survey to ask respondents to name the local radio station they listened to most during the past week. The researcher uses an open-ended question—that is, provides no specific response choices—and the interviewer writes down each respondent's answer. When the completed surveys are tabulated, the researcher notices that several people mentioned radio station WAAA. However, if the city has a WAAA-AM and a WAAA-FM, which station gets the credit? The researcher cannot arbitrarily assign credit to the AM station or to the FM station, nor can credit be split because this may distort the description of the actual listening habits.

Interviewers could attempt callbacks to everyone who said "WAAA," but this is not suggested for two reasons: (1) the likelihood of reaching all the people who gave that response is low; and (2) even if the first condition is met, some respondents may not recall which station they originally mentioned. The researcher is therefore unable to provide a reliable analysis of the data because not all possible intervening variables were considered. (The researcher should have predicted this problem and instructed the interviewers to find out in each case whether "WAAA" meant WAAA-AM or WAAA-FM.)

People who unknowingly provide false information create another type of research noise. For example, people who keep diaries for radio and television surveys may err in recording the station or channel they tune in to; that is, they may listen to or watch station KAAA but incorrectly record "KBBB." (This problem is solved by the use of Nielsen's *people meters* and Arbitron's *portable people meters*; see Chapter 14.) In addition, respondents/subjects often answer a multiple-choice or yes/no research question at random (they make up answers) because they do not wish to appear ignorant or uninformed. To minimize this problem, researchers must take great care in constructing measurement instruments. Noise is always present, but a large and representative sample should decrease the effects of some research noise. (In later chapters, noise is referred to as "error.")

With experience, researchers learn to solve many simple problems in their studies. In many situations, however, researchers understand that total control over all aspects of the research is impossible, and they account for the impossibility of achieving perfect control in the interpretation of their results.

Defining Variables Operationally

In Chapter 1, we stated that an operational definition specifies the procedures to be followed to experience or measure a concept. Research depends on observations, and observations cannot be made without a clear statement of what is to be observed. An operational definition is such a statement. Operational definitions are indispensable in scientific research because they enable investigators to measure relevant variables. In any study, it is necessary to provide operational definitions for both independent variables and dependent variables. Table 2.1 lists examples of such definitions taken from research studies in mass communication.

Kerlinger (2010) identifies two types of operational definitions: *measured* and *experimental*. A measured operational definition specifies how to measure a variable. For instance, a researcher investigating dogmatism and media use might operationally define the term *dogmatism* as a subject's score on the *Twenty-Item Short Form Dogmatism Scale*. An experimental operational definition explains how an investigator has manipulated a variable. Obviously, this type of definition is used when the independent variable is defined in a laboratory setting. For example, in a study on the impact of television violence, the researcher might manipulate media violence by constructing two eight-minute films. The first film, labeled "the violent condition," could contain scenes from a boxing match. The second film, labeled "the nonviolent condition," could depict a swimming race. Similarly, source credibility might be manipulated by alternately attributing an article on health to the *New England Journal of Medicine* and to the *National Enquirer*.

Operationally defining a variable forces a researcher to express abstract concepts in concrete terms. Occasionally, after unsuccessfully struggling with the task of making a key variable operational, the researcher may conclude that the variable as originally conceived is too vague or ambiguous and must be redefined. Because operational definitions are expressed so concretely, they can communicate exactly what the terms represent. For instance, a researcher might define "political knowledge" as the number of correct answers on a 20-item true/false test. Although it is possible to argue about the validity (does the test actually measure political knowledge) of the definition, there is no confusion as to what the statement "Women possess more political knowledge than men" actually means.

Finally, there is no single foolproof method for operationally defining a variable.

Table 2.1 Examples of Operational Definitions

Study	Variable	Operational Definition
Henning and Vorderer (2001)	Need for cognition	Summated scores on a five-point Likert Scale to eight cognition items
Wu (2000)	Press freedom	Scale of press freedom ranging from 1 to 100 taken from yearly evaluations by the Freedom House organization
Angelini (2008	Arousal	Measure of galvanic skin response
Buijen and Valkenburg (2000)	Children's gift ideas	Children were asked to write down their two most favorite Christmas wishes
Kamhawi and Grabe (2008)	Appreciation of news stories	Semantic differential scale with six bi-polar adjective pairs

No operational definition satisfies everybody. The investigator must decide which method is best suited for the research problem at hand. The numerous articles and examples available from an Internet search of "*operational definition*" illustrate the various methods.

QUALITATIVE AND QUANTITATIVE RESEARCH

Mass media research, like all research, can be qualitative or quantitative. **Qualitative research** involves several methods of data collection, such as focus groups, field observation, in-depth interviews, and case studies. In all of these methods, the questioning approach is varied. In other words, although the researcher enters the project with a specific set of questions, follow-up questions are developed as needed. The variables in qualitative research may or may not be measured or quantified.

In some cases, qualitative research has certain advantages. The methods allow a researcher to view behavior in a natural setting without the artificiality that sometimes surrounds experimental or survey research. In addition, qualitative techniques can increase a researcher's depth of understanding of the phenomenon under investigation. This is especially true when the phenomenon has not been investigated previously. Finally, qualitative methods are flexible and allow the researcher to pursue new areas of interest. A questionnaire is unlikely to provide data about questions that were not asked, but a person conducting a field observation or focus group might discover facets of a subject that were not considered before the study began.

However, some disadvantages are associated with qualitative methods. First, sample sizes are sometimes too small (sometimes as small as one) to allow the researcher to generalize the data beyond the sample selected for the particular study. For this reason, qualitative research is often the preliminary step to further investigation rather than the final phase of a project. The information collected from qualitative methods is often used to prepare a more elaborate quantitative analysis, although the qualitative data may in fact be all the information needed for a particular study.

Data reliability can also be a problem, since single observers are describing unique events. Because a person conducting qualitative research must become closely involved the respondents, it is possible to lose objectivity when collecting data. A researcher who becomes too close to the study may lose the necessary professional detachment.

Finally, if qualitative research is not properly planned, the project may produce nothing of value. Qualitative research appears to be easy to conduct, but projects must be carefully designed to ensure that they focus on key issues. Although this book is primarily concerned with quantitative research, we discuss several qualitative methods in Chapter 5.

Quantitative research also involves several methods of data collection, such as telephone surveys, mail surveys, and Internet surveys. In these methods, the questioning is static or standardized—all respondents are asked the same questions and there is no opportunity for follow-up questions.

In the past, some researchers claimed that the difference between qualitative and quantitative research related to only two things:

1. Qualitative research uses smaller samples of subjects or respondents.
2. Because of the small sample size, results from qualitative research could not be generalized to the population from which the samples were drawn.

While these two points may affect some qualitative research, the fact is that sample sizes in both qualitative and quantitative can be the same.

Quantitative research requires that the variables under consideration be measured. This form of research is concerned with how often a variable is present and generally uses numbers to communicate this amount. Quantitative research has certain advantages. One is that the use of numbers allows greater precision in reporting results. For example, the Violence Index (Gerbner, Gross, Morgan & Signorielli, 1980), a quantitative measuring device, makes it possible to report the exact increase or decrease in violence from one television season to another, whereas qualitative research could report only whether there was more or less violence.

For the past several years, some friction has existed in the mass media field and in other disciplines between those who favor quantitative methods and those who prefer qualitative methods. Most researchers have now come to realize that both methods are important in understanding any phenomenon. In fact, the term *triangulation*, commonly used by marine navigators, frequently emerges in conversations about communication research. If a ship picks up signals from only one navigational aid, it is impossible to know the vessel's precise location. However, if signals from more than one source are detected, elementary geometry can be used to pinpoint the ship's location. In this book, the term **triangulation** refers to the use of both qualitative methods and quantitative methods to fully understand the nature of a research problem.

Although most of this book is concerned with skills relevant to quantitative research, we do not imply that quantitative research is in any sense better than qualitative research. It is not. Each approach has value, and the decision to use one or the other depends on the goals of the research.

THE NATURE OF MEASUREMENT

The importance of mathematics to mass media research is difficult to overemphasize. As pointed out by measurement expert J. P. Guilford (1954, p. 1):

> The progress and maturity of a science are often judged by the extent to which it has succeeded in the use of mathematics. . . . Mathematics is a universal language that any science or technology may use with great power and convenience. Its vocabulary of terms is unlimited. . . . Its rules of operation . . . are unexcelled for logical precision.

The idea behind **measurement** is simple: A researcher assigns numerals to objects, events, or properties according to certain rules. Examples of measurement are everywhere: "She or he is a 10" or "Unemployment increased by 1%" or "The earthquake measured 5.5 on the Richter scale." Note that the definition contains three central concepts: numerals, assignment, and rules. A numeral is a symbol, such as V, X, C, or 5, 10, 100. A *numeral* has no implicit quantitative meaning. When it is given quantitative meaning, it becomes a number and can be used in mathematical and statistical computations. *Assignment* is the designation of numerals or numbers to certain objects or events. A simple measurement system might entail assigning the numeral 1 to the people who obtain most of their news from television, the numeral 2 to those who get most of their news from a newspaper, and the numeral 3 to those who receive most of their news from some other source.

Rules specify the way that numerals or numbers are to be assigned. Rules are at the heart of any measurement system; if they are faulty, the system will be flawed. In some situations, the rules are obvious and straightforward. To measure reading speed, a stopwatch and a standardized message may

A CLOSER LOOK

Qualitative and Quantitative Research

The only difference between qualitative and quantitative research is the style of questioning. Qualitative research uses flexible questioning; quantitative uses standardized questions. Assuming that the sample sizes are large enough and that the samples are properly selected, the results from both methods can be generalized to the population from which the sample was drawn.

be sufficient. In other instances, the rules are not so apparent. Measuring certain psychological traits, such as "source credibility" or "attitude toward violence," calls for carefully explicated measurement techniques.

Additionally, in mass media research and in much of social science research, investigators usually measure indicators of the properties of individuals or objects rather than the individuals or objects themselves. Concepts such as "authoritarianism" or "motivation for reading the newspaper" cannot be observed directly; they must be inferred from presumed indicators. Thus, if a person endorses statements such as "Orders from a superior should always be followed without question" and "Law and order are the most important things in society," it can be deduced that he or she is more authoritarian than someone who disagrees with the same statements.

Measurement systems strive to be isomorphic to reality. **Isomorphism** means *identity or similarity of form or structure*. In some research areas, such as the physical sciences, isomorphism is not a problem because the objects being measured and the numbers assigned to them usually have a direct relationship. For example, if an electric current travels through Substance A with less resistance than it does through Substance B, it can be deduced that A is a better conductor than B. Testing more substances can lead to a ranking of conductors, where the numbers assigned indicate the degrees of conductivity. The measurement system is isomorphic to reality.

In mass media research, the correspondence is seldom that obvious. For example, imagine that a researcher is trying to develop a scale to measure the "persuasibility" of people in connection with a certain type of advertisement. A test is developed and given to five people. The scores are displayed in Table 2.2. Now imagine that an omniscient being is able to disclose the "true" persuasibility of the same five people. These scores are also shown in Table 2.2. For two people, the test scores correspond exactly to the "true" scores. The other three scores miss the true scores, but there is a correspondence between the rank orders. Also note that the true persuasibility scores range from 0 to

Table 2.2 Illustration of Isomorphism

Person	Test Score	"True" Score
A	1	0
B	3	1
C	6	6
D	7	7
E	8	12

12, and the measurement scale ranges from 1 to 8. To summarize, there is a general correspondence between the test and reality, but the test is far from an exact measure of what actually exists.

Unfortunately, the degree of correspondence between measurement and reality is rarely known in research. In some cases, researchers are not even sure they are actually measuring what they are trying to measure (**validity**). In any event, researchers must carefully consider the degree of isomorphism between measurement and reality. This topic is discussed in detail later in the chapter.

LEVELS OF MEASUREMENT

Scientists have distinguished four different ways to measure things, or four different levels of measurement, depending on the rules that are used to assign numbers to objects or events. The operations that can be performed with a given set of scores depend on the level of measurement achieved. The four levels of measurement are nominal, ordinal, interval, and ratio.

The **nominal level** is the weakest form of measurement. In nominal measurement, numerals or other symbols are used to classify people, objects, or characteristics. For example, in the physical sciences, rocks can generally be classified into three categories: igneous, sedimentary, and metamorphic. A geologist who assigns a 1 to igneous, a 2 to sedimentary, and a 3 to metamorphic has formed a nominal scale. Note that the numerals are simply labels that stand for the respective categories; they have no mathematical significance. A rock that is placed in Category 3 does not have more "rockness" than those in Categories 1 and 2. Other examples of nominal measurement are the numbers on football jerseys, license plates, and Social Security numbers. An example of nominal measurement in mass media is classifying respondents according to the medium they depend on most for news. Those depending most on TV may be in Category 1, those depending most on newspapers in Category 2, those depending on magazines in Category 3, and so on.

The nominal level, like all levels, possesses certain formal properties. Its basic property is equivalence. If an object is placed in Category 1, it is considered equal to all other objects in that category. Suppose a researcher is attempting to classify all the advertisements in a magazine according to primary appeal. If an ad has economic appeal, it is placed in Category 1; if it uses an appeal to fear, it is placed in Category 2; and so on. Note that all ads using "fear appeal" are equal even though they may differ on other dimensions such as product type or size, or use of illustrations.

Another property of nominal measurement is that all categories are *exhaustive* and *mutually exclusive*. This means that each measure accounts for every possible option and that each measurement is appropriate to only one category. For instance, in the example of primary appeals in magazine advertisements, all possible appeals need to be included in the analysis (exhaustive): economic, fear, morality, religion, and so on. Each advertisement is placed in one and only one category (mutually exclusive). Nominal measurement is frequently used in mass media research, and several are available on the Internet by searching for *"nominal measurement" examples*.

Even a variable measured at the nominal level may be used in higher-order statistics if it is converted into another form. The results of this conversion process are known as **dummy variables**. For example, political party affiliation could be coded as follows:

Republican	1
Democrat	2
Independent	3
Other	4

This measurement scheme could be interpreted incorrectly to imply that a person classified as "Other" is three units "better" than a person classified as a "Republican." To measure political party affiliation and use the data in higher-order statistics, a researcher must convert the variable into a more neutral form.

One way to convert the variable to give equivalent value to each option is to recode it as a dummy variable that creates an "either/or" situation for each option; in this example, a person is either a "Republican" or something else. For example, a binary coding scheme could be used:

Republican	001
Democrat	010
Independent	100
Other	000

This scheme treats each affiliation equivalently and allows the variable to be used in higher-order statistical procedures. Note that the final category "Other" is coded using all zeros. A complete explanation for this practice is beyond the scope of this book; basically, however, its purpose is to avoid redundancy, since the number of individuals classified as "Other" can be found from the data on the first three options. If, in a sample of 100 subjects, 25 belong in each of the first three options, then it is obvious that there are 25 in the "Other" option.

Objects measured at the **ordinal level** are usually ranked along some dimension, such as from smallest to largest. For example, one might measure the variable "socioeconomic status" by categorizing families according to class: lower, lower middle, middle, upper middle, or upper. A rank of 1 is assigned to lower, 2 to lower middle, 3 to middle, and so forth. In this situation, the numbers have some mathematical meaning: Families in Category 3 have a higher socioeconomic status than families in Category 2. Note that nothing is specified with regard to the distance between any two rankings. Ordinal measurement often has been compared to a horse race without a stopwatch. The order in which the horses finish is relatively easy to determine, but it is difficult to calculate the difference in time between the winner and the runner-up.

An ordinal scale possesses the property of *equivalence*. Thus, in the previous example, all families placed in a category are treated equally, even though some might have greater incomes than others. It also possesses the property of order among the categories. Any given category can be defined as being higher or lower than any other category. Common examples of ordinal scales are rankings of football or basketball teams, military ranks, restaurant ratings, and beauty pageant results.

Ordinal scales are frequently used in mass communication research, and several can be found on the Internet by searching for *"ordinal scales" communication examples*. There are also several studies on the Internet about media content, known as **agenda setting**.

When a scale has all the properties of an ordinal scale and the intervals between adjacent points on the scale are of equal value, the scale is at the **interval level**. The most obvious example of an interval scale is temperature. The same amount of heat is required to warm an object from 30 to 40 degrees as to warm it from 50 to 60 degrees. Interval scales incorporate the formal *property of equal differences*; that is, numbers are assigned to the positions of objects on an interval scale in such a way that one may carry out arithmetic operations on the differences between them.

One disadvantage of an interval scale is that it lacks a true zero point, or a condition of nothingness. For example, it is difficult to conceive of a person having zero intelligence or zero personality. The absence of a true zero point means that a researcher cannot make statements of a proportional nature;

for example, someone with an IQ of 100 is not twice as smart as someone with an IQ of 50, and a person who scores 30 on a test of aggression is not three times as aggressive as a person who scores 10. Despite this disadvantage, interval scales are frequently used in mass communication research.

Scales at the **ratio level** of measurement have all the properties of interval scales plus one more: the existence of a true zero point. With the introduction of this fixed zero point, ratio judgments can be made. For example, since time and distance are ratio measures, one can say that a car traveling at 50 miles per hour is going twice as fast as a car traveling at 25 miles per hour. Ratio scales are relatively rare in mass media research, although some variables, such as time spent watching television or number of words per story, are ratio measurements.

As discussed in Chapter 12, researchers who use interval or ratio data can use parametric statistics, which are specifically designed for these data. Procedures designed for use with "lower" levels of measurement can also be used with data at a higher level of measurement. Statistical procedures designed for higher-level data, however, are generally more powerful than those designed for use with nominal or ordinal levels of measurement. Thus, if an investigator has achieved the interval level of measurement, parametric statistics should generally be used. Statisticians disagree about the importance of the distinction between ordinal scales and interval scales and about the legitimacy of using interval statistics with data that may in fact be ordinal. Without delving too deeply into these arguments, we suggest that the safest procedure is to assume interval measurement unless there is clear evidence to the contrary, in which case ordinal statistics should be used. For example, ordinal statistics should be used for a research task in which a group of subjects ranks a set of objects. On the other hand, parametric procedures are justified if subjects are given an attitude score constructed by rating responses to various questions.

Most statisticians seem to feel that statistical analysis is performed on the numbers yielded by the measures, not on the measures themselves, and that the properties of interval scales belong to the number system (Nunnally & Bernstein, 1994; Roscoe, 1975). Additionally, there have been several studies in which various types of data have been subjected to different statistical analyses. These studies suggest that the distinction between ordinal data and interval data is not particularly crucial in selecting an analysis method (McNemar, 1969).

MEASUREMENT SCALES

A *scale* represents a *composite measure* of a variable; it is based on more than one item. Scales are generally used with complex variables that do not easily lend themselves to single-item or single-indicator measurements. Some items, such as age, newspaper circulation, or number of radios in the house, can be adequately measured without scaling techniques. Measurement of other variables, such as attitude toward TV news or gratification received from going to a movie theater, generally requires the use of scales. Several scaling techniques have been developed over the years. This section discusses only the better-known methods. Search the Internet for additional information about all types of measurement scales.

Simple Rating Scales

Rating scales are common in mass media research. Researchers frequently ask respondents to rate a list of items such as a list of programming elements that can be included in a radio station's weekday morning show, or to rate how much respondents like radio or TV on-air personalities.

The researcher's decision is to decide which type of scale to use: 1 to 3? 1 to 5? 1 to 7? 1 to 10? 1 to 100? Or even a 0 to 9 scale, which is commonly used by researchers who don't have computer software to accept double-digit numbers (like 10). Selecting a type of scale is largely a matter of personal preference, but there are a few things to consider:

1. *A scale with more points rather than fewer points allows for greater differentiation on the item or items being rated.* For example, assume we are rating the importance of programming elements contained in a radio station's weekday morning show. Let's say the respondents are told, "The higher the number, the more important the element is to you." Will a 1–3 scale or 1–10 scale provide more information? Obviously, the 1–10 scale provides the broadest differentiation.

Broad differentiation in opinions, perceptions, and feelings is important because it gives the researcher more information. Artificially restricting the range of ratings is called **factor fusion**, which means that opinions, perceptions, and feelings are squeezed into a smaller space. It's better for the respondents and the researcher to have more rating points than fewer rating points. Restricting respondents' responses by using too few scale points always hides the range of potential responses and restricts the potential of any research study.

2. *Our experience shows that males and females of all age groups, and all races and nationalities like to use a 1–10 scale.* This is true because the 1–10 scale is universally used, particularly in sporting events like the Olympics. Virtually everyone understands the 1–10 rating scale. A 10 is best or perfect, a 1 is worst or imperfect. Our experience also shows that researchers should not use a 0–9 or 1–9 rating scale because, quite frankly, respondents do not associate well with a 9 as the highest number.

3. *When using simple rating scales, it is best to tell respondents that "The higher the number, the more you agree," or "The higher the number, the more you like."* Over thousands of research studies, we have found this approach better than telling respondents, "Use a scale of 1 to 10, where '1' means *Dislike* and '10' means *Like a lot*."

Transforming Scales

On occasion, a researcher will conduct a study using one scale and then later want to compare those data to other data using a different rating scale. For example, let's say that a researcher uses a 1–7 rating scale and wants to convert the results to a 1–100 scale. What can be done?

The procedure is always the same: Divide the smaller rating scale into the larger to produce a multiplier to transform the scale. For the transformation of 1–7 to 1–100, first divide 100 by 7, which is 14.2857, and then multiply this number times each of the 1–7 elements to compute the converted 1–100 scale numbers. The new, transformed (rounded) ratings are:

1 = 14
2 = 29
3 = 43
4 = 57
5 = 71
6 = 86
7 = 100

What about transforming a 5-point scale to a 7-point scale? The procedure is the same: Divide 7 by 5, which produces a multiplier of 1.4. This number is multiplied times each of the numbers in the 5-point scale to produce a transformed scale:

1 = 1.4
2 = 2.8
3 = 4.2
4 = 5.6
5 = 7.0

If you transform scores the other way, such as a 10-point scale to a 5-point scale, simply divide each of the numbers in the scale by the multiplier.

SPECIALIZED RATING SCALES

Thurstone Scales

Thurstone scales are also called *equal-appearing interval scales* because of the technique used to develop them and are typically used to measure the attitude toward a given concept or construct. To develop a Thurstone scale, a researcher first collects a large number of statements (Thurstone recommends at least 100) that relate to the concept or construct to be measured.

Next, judges rate these statements along an 11-category scale in which each category expresses a different degree of favorableness toward the concept. The items are then ranked according to the mean or median ratings assigned by the judges and are used to construct a questionnaire of 20 to 30 items that are chosen more or less evenly from across the range of ratings. The statements are worded so that a person can agree or disagree with them. The scale is then administered to a sample of respondents whose scores are determined by computing the mean or median value of the items agreed with. A person who disagrees with all the items has a score of zero.

One advantage of the Thurstone method is that it is an interval measurement scale. On the downside, this method is time consuming and labor intensive. Thurstone scales are not often used in mass media research, but they are common in psychology and education research.

Guttman Scaling

Guttman scaling, also called *scalogram* analysis, is based on the idea that items can be arranged along a continuum in such a way that a person who agrees with an item or finds an item acceptable will also agree with or find acceptable all other items expressing a less extreme position. For example, here is a hypothetical four-item Guttman scale:

1. Indecent programming on TV is harmful to society.
2. Children should not be allowed to watch indecent TV shows.
3. Television station managers should not allow indecent programs on their stations.
4. The government should ban indecent programming from TV.

Presumably, a person who agrees with Statement 4 will also agree with Statements 1–3. Furthermore, if we assume the scale is valid, a person who agrees with Statement 2 will also agree with Statement 1 but will not necessarily agree with Statements 3 and 4. Because each score represents a unique set of responses, the number of items a person agrees with is the person's total score on a Guttman scale.

A Guttman scale requires a great deal of time and energy to develop. Although they do not appear often in mass media research, Guttman scales are common in political science, sociology, public opinion research, and anthropology.

Likert Scales

Perhaps the most commonly used scale in mass media research is the Likert scale, also called the *summated rating approach*. A number of statements are developed with respect to a topic, and respondents can strongly agree, agree, be neutral, disagree, or strongly disagree with the statements (see Figure 2.1). Each response option is weighted, and each subject's responses are added to produce a single score on the topic.

Figure 2.1 Sample of Likert Scale Items

1. Only U.S. citizens should be allowed to own broadcasting stations.

Response	Score Assigned
________ Strongly agree	5
________ Agree	4
________ Neutral	3
________ Disagree	2
________ Strongly disagree	1

2. Prohibiting foreign ownership of broadcasting stations is bad for business.

Response	Score Assigned
________ Strongly agree	1
________ Agree	2
________ Neutral	3
________ Disagree	4
________ Strongly disagree	5

Note: To maintain attitude measurement consistency, the scores are reversed for a negatively worded item. Question 1 is a positive item; Question 2 is a negative item.

This is the basic procedure for developing a Likert scale:

1. Compile a large number of statements that relate to a specific dimension. Some statements are positively worded; some are negatively worded.
2. Administer the scale to a randomly selected sample of respondents.
3. Code the responses consistently so that high scores indicate stronger agreement with the attitude in question.
4. Analyze the responses and select for the final scale those statements that most clearly differentiate the highest from the lowest scorers.

Semantic Differential Scales

Another commonly used scaling procedure is the **semantic differential** technique. As originally conceived by Osgood, Suci, and Tannenbaum (1957), this technique is used to measure the meaning an item has for an individual. Research indicated that three general factors—activity, potency, and evaluation—were measured by the semantic differential. Communication researchers were quick to adapt the evaluative dimension of the semantic differential for use as a measure of attitude.

To use the technique, a name or a concept is placed at the top of a series of seven-point scales anchored by bipolar attitudes. Figure 2.2 shows an example of

Figure 2.2 Sample Form for Applying the Semantic Differential Technique

***Time* Magazine**

Biased	____ : ____ : ____ : ____ : ____ : ____ : ____	Unbiased
Trustworthy	____ : ____ : ____ : ____ : ____ : ____ : ____	Untrustworthy
Valuable	____ : ____ : ____ : ____ : ____ : ____ : ____	Worthless
Unfair	____ : ____ : ____ : ____ : ____ : ____ : ____	Fair

this technique as used to measure attitudes toward *Time magazine*.

The bipolar adjectives that typically "anchor" such evaluative scales are *pleasant/ unpleasant, valuable/worthless, honest/ dishonest, nice/awful, clean/dirty, fair/unfair,* and *good/bad*. However, we recommend that a unique set of anchoring adjectives be developed for each particular measurement situation.

Strictly speaking, the semantic differential technique attempts to place a concept in semantic space using an advanced multivariate statistical procedure called **factor analysis**. When researchers borrow parts of the technique to measure attitudes, or images or perceptions of objects, people, or concepts, they are not using the technique as originally developed. Consequently, perhaps a more appropriate name for this technique is bipolar rating scales.

RELIABILITY AND VALIDITY

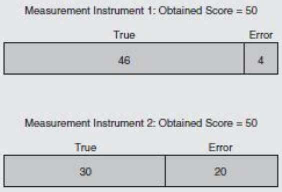
Using any scale without preliminary testing is poor research. At least one pilot study should be conducted for any newly developed scale to ensure its reliability and validity. To be useful, a measurement must possess these two related qualities. A measure is reliable if it consistently gives the same answer. **Reliability** in measurement is the same as reliability in any other context. For example, a reliable person is one who is dependable, stable, and consistent over time. An unreliable person is unstable and unpredictable and may act one way today and another way tomorrow. Similarly, if measurements are consistent from one session to another, they are reliable and can be believed to some degree.

In understanding measurement reliability, you may think of a measure as containing two components. The first represents an individual's "true" score on the measuring instrument. The second represents random error and does not provide an accurate assessment of what is being measured. Error can slip into the measurement process from several sources. Perhaps a question was worded ambiguously or a person's pencil slipped when completing a measuring instrument.

Whatever the cause, all measurements are subject to some degree of random error, as shown in Figure 2.3. As is evident, Measurement Instrument 1 is highly reliable because the ratio of the true component of the score to the total score is high. Measurement Instrument 2 is unreliable because the ratio of the true component to the total is low.

A completely unreliable measurement measures nothing at all. If a measure is repeatedly given to individuals and each person's responses at a later session are unrelated to his or her earlier responses, the measure is useless. If the responses are identical or nearly identical each time the measure is given, the measure is reliable; it at least measures something, though not necessarily what the researcher intended. (This problem is discussed later in this chapter.)

Figure 2.3 Illustration of "True" and "Error" Components of a Scale



The importance of reliability should be obvious now. Unreliable measures cannot be used to detect relationships between variables. When the measurement of a variable is unreliable, it is composed mainly of random error, and random error is seldom related to anything else. Reliability is not a unidimensional concept. It consists of three different components: stability, internal consistency, and equivalency.

Stability refers to *the consistency of a result or of a measure at different points in time*. For example, suppose that a test designed to measure proofreading ability is administered during the first week of an editing class and again during the second week. The test possesses stability if the two results are consistent.

Caution should be exercised whenever stability is used as a measure of reliability, since people and things can change over time. In the proofreading example, it is possible for a person to score higher the second time because some people might actually improve their ability from Week 1 to Week 2. In this case, the measure is not unstable—actual change occurred.

An assessment of reliability is necessary in all mass media research and should be reported along with other facets of the research as an aid in interpretation and evaluation. One commonly used statistic for assessing reliability is the correlation coefficient, denoted as r_{xx}. Chapter 12 provides a more detailed examination of the correlation coefficient.

For now let's say only that r_{xx} is a number ranging from −1.00 to +1.00 and is used to gauge the strength of a relationship between two variables. When r_{xx} is high—that is, approaching +1.00—the relationship is strong. A negative number indicates a negative relationship (high scores on one variable are associated with low scores on the other), and a positive number indicates a positive relationship (a high score goes with another high score). In measuring reliability, a high positive r_{xx} is desired.

One method that uses correlation coefficients to compute reliability is the test-retest method. This procedure measures the stability component of reliability. The same people are measured at two different points in time, and a coefficient between the two scores is computed. An r_{xx} that approaches +1.00 indicates that a person's score at Time A was similar to his or her score at Time B, showing consistency over time. There are two limitations to the test-retest technique. First, the initial administration of the measure might affect scores on the second testing. If the measuring device is a questionnaire, a person might remember responses from session to session, thus falsely inflating reliability. Second, the concept measured may change from Time A to Time B, thus lowering the reliability estimate.

Internal consistency involves examining the consistency of performance among the items that compose a scale. If separate items on a scale assign the same values to the concept being measured, the scale possesses internal consistency. For instance, suppose a researcher designs a 20-item scale to measure attitudes toward newspaper reading. For the scale to be internally consistent, the total score on the first half of the test should correlate highly with the score on the second half of the test. This method of determining reliability is called the **split-half technique**.

Only one administration of the measuring instrument is made, but the test is split into halves and scored separately. For example, if the test is in the form of a questionnaire, the even-numbered items might constitute one half and the odd-numbered items the other half. A correlation coefficient is then computed between the two sets of scores. Since this coefficient is computed from a test that is only half as long as the final form, it is corrected by using the following formula:

$$r_{xx} = \frac{2(r_{oe})}{1 + r_{oe}}$$

where r_{oe} is the correlation between the odd items and the even items.

Another common reliability coefficient is alpha (sometimes referred to as **Cronbach's alpha**), which uses the analysis of variance approach to assess the internal consistency of a measure (see Chapter 12).

The **equivalency** component of reliability, sometimes referred to as **cross-test reliability**, assesses the relative correlation between two parallel forms of a test. Two instruments that use different scale items or different measurement techniques are developed to measure the same concept. The two versions are then administered to the same group of people during a single time period, and the correlation between the scores on the two forms of the test is taken as a measure of the reliability. The major problem with this method, of course, is developing two forms of a scale that are perfectly equivalent. The less parallel the two forms, the lower the reliability.

A special case of the equivalency component occurs when two or more observers judge the same phenomenon, as is the case in content analysis (see Chapter 6). This type of reliability is called **intercoder reliability** and is used to assess the degree to which a result can be achieved or reproduced by other observers. Ideally, two individuals who use the same operational measure and the same measuring instrument should reach the same results. For example, if two researchers try to identify acts of violence in television content based on a given operational definition of violence, the degree to which their results are consistent is a measure of intercoder reliability. Disagreements reflect a difference either in perception or in the way the original definition was interpreted. Special formulas for computing intercoder reliability are discussed in Chapter 6.

In addition to being reliable, a measurement must have *validity* if it is to be of use in studying variables. A valid measuring device measures what it is supposed to measure. Or, to put it another way, determining validity requires an evaluation of the congruence between the operational definition of a variable and its conceptual or constitutive definition. Assessing validity requires some judgment on the part of the researcher. In the following discussion of the major types of measurement validity, note that each one depends at least in part on the judgment of the researcher. Also, validity is almost never an all-or-none proposition; it is usually a matter of degree. A measurement rarely turns out to be totally valid or invalid. Typically it winds up somewhere in the middle.

Concerning measurement, there are four major types of validity, and each has a corresponding technique for evaluating the measurement method: *face validity, predictive validity, concurrent validity*, and *construct validity*.

The simplest and most basic kind of validity, **face validity**, is achieved by examining the measurement device to see whether, on the face of it, it measures what it appears to measure. For example, a test designed to measure proofreading ability could include accounting problems, but this measure would lack face validity. A test that asks people to read and correct certain paragraphs has more face validity as a measure of proofreading skill. Whether a measure possesses face validity depends to some degree on subjective judgment. To minimize subjectivity, the relevance of a given measurement should be judged independently by several experts.

Checking a measurement instrument against some future outcome assesses **predictive validity**. For example, scores on a test to predict whether a person will vote in an upcoming election can be checked against actual voting behavior. If the test scores allow the researcher to predict, with a high degree of accuracy, which people will actually vote and which will not, then the test has predictive validity. Note that it is possible for a measure to have predictive validity and at the same time lack face validity. The sole factor in determining validity in the predictive method is the measurement's ability to forecast future behavior or events correctly. The concern is not with what is being measured but with whether the measurement instrument can predict something. Thus, a test to determine whether a person will become a successful mass media researcher could conceivably consist of geometry problems. If it predicts the ultimate success of a researcher reasonably well, the test has predictive validity but little face validity. The biggest problem associated with predictive validity is determining the criteria against which test scores are to be checked. What, for example, constitutes a "successful mass media researcher"? One who obtains an advanced degree? One who publishes research articles? One who writes a book?

Concurrent validity is closely related to predictive validity. In this method, however, the measuring instrument is checked against some present criterion. For example, it is possible to validate a test of proofreading ability by administering the test to a group of professional proofreaders and to a group of nonproofreaders. If the test discriminates well between the two groups, it can be said to have concurrent validity. Similarly, a test of aggression might discriminate between one group of children who are frequently detained after school for fighting and another group, the members of which have never been reprimanded for antisocial behavior.

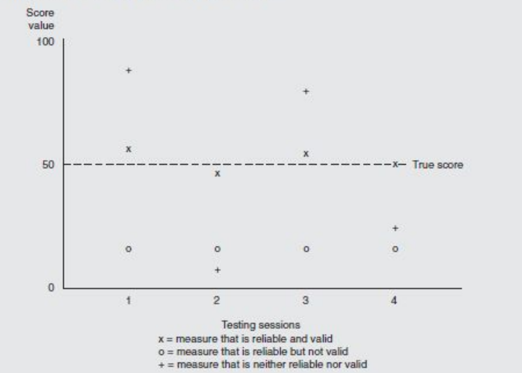
The fourth type of validity, **construct validity**, is the most complex. In simplified form, construct validity involves relating a measuring instrument to some overall theoretic framework to ensure that the measurement is logically related to other concepts in the framework. Ideally, a researcher should be able to suggest various relationships between the property being measured and the other variables. For construct validity to exist, the researcher must show that these relationships are in fact present. For example, an investigator might expect the frequency with which a person views a particular television newscast to be influenced by his or her attitude toward that program. If the measure of attitudes correlates highly with the frequency of viewing, there is some evidence for the validity of the attitude measure. Similarly, construct validity is present if the measurement instrument under consideration does not relate to other variables when there is no theoretic reason to expect such a relationship. Therefore, if an investigator finds a relationship between a measure and other variables that is predicted by a theory and fails to find other relationships that are not predicted by a theory, there is evidence for construct validity. Figure 2.4 summarizes the four types of validity.

Before closing this discussion, we should point out that reliability and validity are related. Reliability is necessary to establish validity, but it is not a sufficient condition; a reliable measure is not necessarily a valid one. Figure 2.5 shows this relationship. An X represents a test that is both reliable and valid; the scores are consistent from session to session and lie close to the true value. An

Figure 2.4 Types of Validity

Judgment-based	Criterion-based	Theory-based
Face validity	Predictive validity Concurrent validity	Construct validity

Figure 2.5 Relationship of Reliability and Validity



O represents a measure that is reliable but not valid; the scores are stable from session to session but they are not close to the true score. A + represents a test that is neither valid nor reliable; scores vary widely from session to session and are not close to the true score.

SUMMARY

Understanding empirical research requires a basic knowledge of concepts, constructs, variables, and measurement. Concepts summarize related observations and express an abstract notion that has been formed by generalizing from particulars. Connections among concepts form propositions that, in turn, are used to build theories. Constructs consist of combinations of concepts and are also useful in building theories.

Variables are phenomena or events that take on one or more different values. Independent variables are manipulated by the researcher, whereas dependent variables are what the researcher attempts to explain. All variables are related to the observable world by operational definitions. Researchers frequently use scales to measure complex variables. Thurstone, Guttman, Likert, and semantic differential scales are used in mass media research.

Measurement is the assignment of numerals to objects, events, or properties according to certain rules. The four levels of measurement are nominal, ordinal, interval, and ratio. To be useful, a measurement must be both reliable and valid.

Key Terms

Agenda setting
Antecedent variable
Concept
Concurrent validity
Construct
Construct validity
Continuous variable
Control variable
Criterion variable
Cronbach's alpha
Cross-test reliability
Dependent variable
Discrete variable
Dummy variable
Equivalency
Face validity
Factor analysis
Factor fusion
Guttman scale
Independent variable
Intercoder reliability
Internal consistency
Interval level
Isomorphism
Likert scale
Marker variable
Measurement
Multivariate analysis
Noise
Nominal level
Ordinal level
Predictive validity
Predictor variable
Rating scale
Ratio level
Reliability
Semantic differential
Split-half technique
Stability
Thurstone scale
Triangulation
Validity
Variable

Using the Internet

1. Search the Internet for:
   - operational definition
   - rating scales
   - reliability
   - validity
   - dummy variable coding
   - "frequency scales" communication examples
   - "ratio scales" communication examples
   - "split-half reliability"
   - "measurement validity" communication examples
   - "semantic differential" communication examples

Questions and Problems for Further Investigation

1. Provide conceptual and operational definitions for the following items:
   - Artistic quality
   - Objectionable song lyrics
   - Programming appeal
   - Sexual content
   - Violence

Compare your definitions to those of others in the class. Would there be any difficulty in conducting a study using these definitions? Have you demonstrated why so much controversy surrounds the topics, for example, of sex and violence on television? What can you find on the Internet about these terms?

2. What type of data (nominal, ordinal, interval, or ratio) is associated with each of the following concepts or measurements?
   - Baseball team standings
   - A test of listening comprehension
   - AC Nielsen's list of the top 10 television programs
   - Frequency of heads versus tails on coin flips
   - Baseball batting averages
   - A scale measuring intensity of attitudes toward violence
   - VHF channels 2–13
   - A scale for monitoring your weight over time
3. Develop a measurement technique to examine each of these concepts:
   - Newspaper reading
   - Aggressive behavior
   - Brand loyalty in purchasing products
   - Television viewing
4. Assume you are going to conduct a study that requires respondents to rate the importance of programming elements on a radio station. Would you use a semantic differential scale or a 1–10 scale? Why?
5. Provide three examples of variables that could be either an independent or dependent variable in different types of research studies.

For additional resources go to *www.wimmerdominick.com* and *www.cengage.com/masscomm/wimmer/mediaresearch9e*.

References and Suggested Readings

Anderson, J. A. (1987). *Communication research: Issues and methods*. New York: McGraw-Hill.

Babbie, E. R. (2010). *The practice of social research* (12th ed.). Belmont, CA: Cengage.

Bae, H. S. (2000). Product differentiation in national TV newscasts, *Journal of Broadcasting and Electronic Media*, 44(1), 62–77.

Buijen, M., & Valkenburg, P. (2000). The impact of television advertising on children's Christmas wishes. *Journal of Broadcasting and Electronic Media*, 44(3), 456–470.

Gerbner, G., Gross, L., Morgan, M., & Signorielli, N. (1980). The mainstreaming of America: Violence profile no. 11. *Journal of Communication*, 30(3), 10–29.

Guilford, J. P. (1954). *Psychometric methods*. New York: McGraw-Hill.

Henning, B., & Vorderer, P. (2001). Psychological escapism: Predicting the amount of television viewing by need for cognition. *Journal of Communication*, 51(1), 100–120.

Hindman, D. B. (2000). The rural-urban digital divide. *Journalism and Mass Communication Quarterly*, 77(3), 549–560.

Kerlinger, F. N. (2000). *Foundations of behavioral research* (4th ed.). New York: Holt, Rinehart & Winston.

Kerlinger, F. N., & Pedhazur, E. (1997). *Multiple regression in behavioral research* (3rd ed.). New York: Holt, Rinehart & Winston.

Mason, E. J., & Bramble, W. J. (1989). *Understanding and conducting research* (2nd ed.). New York: McGraw-Hill.

McNemar, Q. (1969). *Psychological statistics* (4th ed.). New York: John Wiley.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (4th ed.). New York: McGraw-Hill.

Osgood, C., Suci, G., & Tannenbaum, P. (1957). *The measurement of meaning*. Urbana: University of Illinois Press.

Robinson, J., & Shaver, P. (1973). *Measures of social psychological attitudes* (2nd ed.). Ann Arbor, MI: Institute for Social Research.

Roscoe, J. (1975). *Fundamental research statistics for the behavioral sciences*. New York: Holt, Rinehart & Winston.

Wu, H. D. (2000). Systemic determinants of international news coverage. *Journal of Communication*, 50(2), 110–130.