MODULE 5

Defining, Measuring, and Manipulating Variables

LEARNING OBJECTIVES

- Explain and give examples of an operational definition.
- Explain the four properties of measurement and how they are related to the four scales of measurement.
- Identify and describe the four types of measures.

DEFINING VARIABLES

An important step when beginning a research project is to define the variables in the study. Some variables are fairly easy to define, manipulate, and measure. For example, if a researcher is studying the effects of exercise on blood pressure, she can manipulate the amount of exercise either by varying the length of time that individuals exercise or by varying the intensity of the exercise (and monitoring target heart rates). She can also periodically measure blood pressure during the course of the study; a machine already exists that takes this measurement in a consistent and accurate manner. Does the fact that a machine exists to take this measurement mean that the measurement is always accurate? No. (We discuss this issue in Module 6 when we address measurement error.)

Now let's suppose that a researcher wants to study a variable that is not as concrete or as easily measured as blood pressure. For instance, many people study abstract concepts such as aggression, attraction, depression, hunger, or anxiety. How would a researcher either manipulate or measure any of these variables? One person's definition of what it means to be hungry may be vastly different from someone else's. If the researcher decides to measure hunger simply by asking participants in an experiment if they are hungry, the measurement is not accurate because each individual may define hunger in a different way.

operational definition: A definition of a variable in terms of the operations (activities) a researcher uses to measure or manipulate it.

What such a study lacks is an **operational definition** of hunger, that is, a definition of the variable in terms of the operations the researcher uses to measure or manipulate it. Because this definition is somewhat circular, let's reword it in a way that may make more sense. An operational definition specifies the activities of the researcher in measuring and/or manipulating a variable (Kerlinger, 1986). In other words, the investigator might define hunger in terms of specific activities such as not having eaten for 12 hours. Thus one operational definition of hunger could be that simple: Hunger occurs when 12 hours have passed with no food intake. Notice how much more concrete this definition is than simply saying hunger is that "gnawing feeling" that you get in your stomach. Specifying hunger in terms of the number of hours without food is an operational definition; defining hunger as that "gnawing feeling" is not.

Researchers must operationally define all variables: those measured (dependent variables) and those manipulated (independent variables). One reason for so doing is to ensure that the variables are measured or manipulated consistently during the course of the study. Another reason is to help communicate ideas to others. As a consequence, if a researcher says he measured anxiety in his study, the question becomes how did he operationally define anxiety because it can be defined and therefore measured in many different ways? Anxiety can be defined as the number of nervous actions displayed in a 1-hour time period, as heart rate, or as a person's score on a GSR (galvanic skin response) machine or on the Taylor Manifest Anxiety Scale. Some measures are better than others, "better" meaning more reliable and valid (concepts we discuss in Module 6). Once other investigators understand how a researcher has operationally defined a variable, they can replicate the study if they so desire. They can better understand the study and whether it has problems. They can also better design their own studies based on how the variables were operationally defined.

PROPERTIES OF MEASUREMENT

In addition to operationally defining independent and dependent variables, we must consider the level of measurement of the dependent variable. There are four levels of measurement, each based on the characteristics, or properties, of the data: identity, magnitude, equal unit size, and absolute zero. When a measure has the property of **identity**, objects that are different receive different scores. Thus if participants in a study have different political affiliations, they receive different scores. Measurements have the property of **magnitude** (also called *ordinality*) when the ordering of the numbers reflects the ordering of the variable. That is, numbers are assigned in an order such that different numbers reflect more or less of the variable being measured.

Measurements have an **equal unit size** when a difference of 1 is the same amount throughout the entire scale. As an example, the height difference between people who are 64 inches tall and 65 inches tall is the same as the difference between those who are 72 inches tall and 73 inches tall. The difference in each situation (1 inch) is identical. Notice how this measurement differs from the property of magnitude. If we simply lined up and ranked a group of individuals based on their height, our scale would have the properties of identity and magnitude but not equal unit size. Why is this so? We would not actually measure people's height in inches but simply order them according to how tall they appear, from the shortest (the person receiving a score of 1) to the tallest (the person receiving the highest score). Therefore our scale would not meet the criterion of equal unit size because the difference in height between the two people receiving scores of 1 and 2 might not be the same as the difference in height between the two people receiving scores of 3 and 4.

Finally, measures have an **absolute zero** when assigning a score of zero indicates an absence of the variable being measured. For instance, time spent studying has the property of absolute zero because a score of zero means an individual spent no time studying. However, a score of zero is not always equal to the property of absolute zero. An example is the Fahrenheit temperature scale. Although that measurement scale has a score of zero (the thermometer can read 0 degrees), does the score indicate an absence of temperature? No, instead it indicates a very cold temperature. Hence it does not have the property of absolute zero.

SCALES OF MEASUREMENT

The level, or scale, of measurement depends on the properties of the data. There are four scales of measurement: nominal, ordinal, interval, and ratio. Each of these scales has one or more of the properties described in the previous section. We discuss the scales in order, from the one with the fewest properties to the one with the most, that is, from the least to the most sophisticated. As we see in later modules, it is important to establish the scale of data measurement in order to determine the appropriate statistical test to use when analyzing the data.

Nominal Scale

nominal scale: A scale in which objects or individuals are assigned to categories that have *no numerical properties.*

In a **nominal scale** objects or individuals are assigned to categories that have no numerical properties. Nominal scales have the characteristic of identity but lack the other properties. Variables measured on a nominal scale are often referred to as *categorical variables* because the data are divided into categories. However, the *categories carry no numerical weight. Some examples of categorical variables, or* data measured on a nominal scale, are ethnicity, gender, and political affiliation.

We can assign numerical values to the levels of a nominal variable. Take ethnicity for example: we could label Asian-Americans as 1, African-Americans as 2, Latin Americans as 3, and so on. Yet these scores do not carry numerical weight; they are simply labels for the categories. In other words, the scores are used for identity but not for magnitude, equal unit size, or absolute value. We cannot order the data and claim that 1s are more or less than 2s in any way. We cannot analyze these data mathematically. It would not be appropriate to report that the mean ethnicity was 2.56, and we cannot say that there is a true zero, that is, that someone has no ethnicity.

Ordinal Scale

ordinal scale: A scale in which objects or individuals are categorized and the categories form a rank order along a continuum.

In an **ordinal scale** objects or individuals are categorized, and the categories form a rank order along a continuum. Data measured on an ordinal scale have the properties of identity and magnitude but lack equal unit size and absolute zero. Ordinal data are often referred to as *ranked data* because they are ordered from highest to lowest or from biggest to smallest. For example, reporting how students did on an examination based simply on their rank (highest score, second highest, and so on) involves an ordinal scale. This variable carries identity and magnitude because each individual receives a rank (a number) that carries identity, and the rank also conveys information about order or magnitude (how many students performed better or worse in the class). However, the ranking score does not have equal unit size: The difference in performance on the examination between the students ranked 1 and 2 is not necessarily the same as the difference between those ranked 2 and 3. Nor does this scale have an absolute zero that indicates an absence of the variable being measured.

Interval Scale

interval scale: A scale in which the units of measurement (intervals) between the numbers on the scale are all equal in size.

In an **interval scale** the units of measurement (intervals) between the numbers on the scale are all equal in size. An interval scale meets the criteria for identity, magnitude, and equal unit size. For example, the Fahrenheit temperature scale is an interval scale of measurement. A given temperature carries:

- Identity—days with different temperatures receive different scores on the scale.
- Magnitude—cooler days receive lower scores, whereas hotter days receive higher scores.
- Equal unit size—the difference between 50 and 51 degrees is the same as that between 90 and 91 degrees.

The Fahrenheit scale, however, does not have an absolute zero. Because of this lack, we are not able to form ratios based on this scale (for example,

a temperature of 100 degrees is not twice as hot as a temperature of 50 degrees). We can still perform mathematical computations on interval data, as we will see in later modules.

Ratio Scale

ratio scale: A scale in which in addition to order and equal units of measurement there is an absolute zero that indicates an absence of the variable being measured.

In a **ratio scale** in addition to order and equal units of measurement, there is an absolute zero that indicates an absence of the variable measured. Ratio data have all four properties of measurement (identity, magnitude, equal unit size, and absolute zero). Examples of ratio scales of measurement include weight, time, and height. Each of these scales has (1) identity (individuals with different weights receive different scores), (2) magnitude (those who weigh less receive lower scores than those who weigh more), and (3) equal unit size (1 pound is the same unit of weight anywhere along the scale). Ratio scales also have an absolute zero, meaning that a score of zero reflects an absence of the variable. For instance, all bathroom scales start at a weight of zero, and, although a person obviously cannot weigh zero, when the scale reads zero it reflects an absence of the variable.

Also, ratios can be formed. Thus a weight of 100 pounds is twice as much as a weight of 50 pounds. As with interval data mathematical computations can be performed on ratio data. Because interval and ratio data are very similar, many psychologists simply refer to the category as *interval-ratio data* and typically do not distinguish between the types. You should be familiar with the difference between interval and ratio data, but you should also be aware that because they are so similar, they are often referred to as one type of data.

IN REVIEW Features of Scales of Measurement

	Scales of Measurement			
	Nominal	**Ordinal**	**Interval**	**Ratio**
Examples	Ethnicity	Class rank	Temperature	Weight
	Religion	Letter grade	(Fahrenheit and Celsius)	Height
	Sex		Many psychological tests	Time
Properties	Identity	Identity	Identity	Identity
		Magnitude	Magnitude	Magnitude
			Equal unit size	Equal unit size
				Absolute zero
Mathematical operations possible	None	Rank order	Add	Add
			Subtract	Subtract
			Multiply	Multiply
			Divide	Divide

CRITICAL THINKING CHECK 5.1

1. Provide several operational definitions of anxiety. Include nonverbal measures and physiological measures. How would your operational definitions differ from a dictionary definition?
2. Identify the scale of measurement for each of the following variables:
   a. zip code
   b. grade of egg (large, medium, small)
   c. reaction time
   d. score on the SAT
   e. class rank
   f. number on a football jersey
   g. miles per gallon

DISCRETE AND CONTINUOUS VARIABLES

discrete variables: Variables that usually consist of whole number units or categories and are made up of chunks or units that are detached and distinct from one another.

Another means of classifying variables is in terms of whether they are discrete or continuous in nature. **Discrete variables** usually consist of whole number units or categories. They are made up of chunks or units that are detached and distinct from one another. A change in value occurs a whole unit at a time; decimals do not make sense in discrete scales. Most nominal and ordinal data are discrete. For example, gender, political party, and ethnicity are discrete scales. Some interval or ratio data can be discrete. For instance, the number of children someone has is reported as a whole number (discrete data), yet it is also ratio data (you can have a true zero and form ratios).

continuous variables: Variables that usually fall along a continuum and allow for fractional amounts.

Continuous variables usually fall along a continuum and allow for fractional amounts. The term *continuous* means that it "continues" between the whole number units. Examples of continuous variables are age (22.7 years), height (64.5 inches), and weight (113.25 pounds). Most interval and ratio data are continuous in nature. Discrete and continuous data will have increased importance in later modules when we discuss research design and data presentation.

TYPES OF MEASURES

When psychology researchers collect data, the types of measures used can be classified into four basic categories: self-report measures, tests, behavioral measures, and physical measures. We discuss each category, noting its advantages and possible disadvantages.

Self-Report Measures

self-report measures: Usually questionnaires or interviews that measure how people report that they act, think, or feel.

Self-report measures are typically administered as questionnaires or interviews to measure how people report that they act, think, or feel. Thus self-report measures aid in collecting data on behavioral, cognitive, and affective events (Leary, 2001).

Behavioral self-report measures typically ask people to report how often they do something such as how often they eat a certain food, eat out at a restaurant, go to the gym, or have sex. The problem with this and the other types of self-report measures is that we are relying on the individuals to

report on their own behaviors. When collecting data in this manner, we must be concerned with the veracity of the reports and with the accuracy of the individual's memory. Researchers much prefer to collect data using a behavioral measure, but direct observation of some events is not always possible or ethical.

Cognitive self-report measures ask individuals to report what they think about something. You have probably participated in a cognitive self-report measure of some sort. You may have been stopped on campus and asked what you think about parking, food services, or residence halls. Once again, we are relying on the individual to make an accurate and truthful report.

Affective self-report measures ask individuals to report how they feel about something. You may have participated in an affective self-report measure if you ever answered questions concerning emotional reactions such as happiness, depression, anxiety, or stress. Many psychological tests are affective self-report measures. These tests also fit into the category of measurement tests described in the next section.

Tests

tests: Measurement instruments used to assess individual differences in various content areas.

Tests are measurement instruments used to assess individual differences in various content areas. Psychologists frequently use two types of tests: personality tests and ability tests. Many *personality tests* are also affective self-report measures; they are designed to measure aspects of an individual's personality and feelings about certain things. Examples of such tests include the MMPI II or the Beck Depression Inventory.

Ability tests, however, are not self-report measures and generally fall into two different categories: aptitude tests and achievement tests. *Aptitude tests* measure an individual's potential to do something, whereas *achievement tests* measure an individual's competence in an area. In general, intelligence tests are aptitude tests, and school exams are achievement tests.

Most tests used by psychologists have been subjected to extensive testing themselves and are therefore considered an objective, unbiased means of collecting data. Keep in mind, however, that any measuring instrument has the potential for problems, which may range from the state of the participant on a given day to scoring and interpretation.

Behavioral Measures

behavioral measures: Measures taken by carefully observing and recording behavior.

Psychologists take **behavioral measures** by carefully observing and recording behavior. Behavioral measures are often referred to as *observational measures* because they involve observing what a participant does. Briefly (because we discuss observational research studies in detail in the next chapter), behavioral measures can be applied to anything a person or an animal does—a pigeon pecking a disk, the way men and women carry their books, or how many cars actually stop at a stop sign. The observations can be direct (while the participant is engaging in the behavior) or indirect (via audio- or videotape).

When taking behavioral measures, a researcher usually employs some sort of coding system, which is a means of converting the observations to numerical data. A very basic coding system involves simply counting the number

of times that participants do something. How many times does the pigeon peck the lighted disk, or how many cars stop at the stop sign? A more sophisticated coding system involves assigning behaviors to categories. For example, a researcher might watch children playing and classify their behavior into several categories of play such as solitary, parallel, and cooperative. In the example of cars stopping at a stop sign, simply counting the number of stops might not be adequate. What is a stop? The researcher might operationally define a full stop as the car not moving for at least 3 seconds. Other categories might include a complete stop of less than 3 seconds, a rolling stop, and no stop. The researcher then has a more complex coding system consisting of various categories.

Also think about the problems of collecting data at a stop sign. If someone is standing there with a clipboard making notes, how might the presence of the data collector affect the behavior of drivers approaching the stop sign? Are researchers going to get a realistic estimate of how many cars actually stop at the sign? Probably not. For this reason measures are sometimes taken in an unobtrusive manner. Observers may hide what they are doing, hide themselves, or use a more indirect means of collecting the data (such as videotape). Using an unobtrusive means of collecting data reduces **reactivity**, that is, participants reacting in an unnatural way to being observed. This issue is discussed more fully in the next chapter.

reactivity: A possible reaction by participants in which they act unnaturally because they know they are being observed.

Finally, let's note some of the possible problems with behavioral measures. First, they rely on humans observing events. How do we know that the observers perceived the events accurately? Second, the observers must then code the events into a numerical format. There is tremendous potential for error in this coding. Finally, if the observers are visible, there is the possibility that participants may not be acting naturally because the latter know they are being observed.

Physical Measures

physical measures: Measures of bodily activity such as pulse or blood pressure that may be taken with a piece of equipment.

Most **physical measures**, or measures of bodily activity, are not directly observable. Physical measures are usually taken by means of equipment. Weight is measured with a scale, blood pressure with a sphygmomanometer, and temperature with a thermometer. Sometimes the equipment is more sophisticated. Psychologists, for instance, frequently use the galvanic skin response (GSR) to measure emotional arousal, electromyography (EMG) recordings to measure muscle contractions, and electroencephalogram (EEG) recordings to measure electrical activity in the brain.

Physical measures are much more objective than behavioral measures. A physical measure is not simply an observation (which may be subjective) of how a person or animal is acting. Instead, it is a measure of a physical activity that takes place in the brain or body. This is not to say that physical measures are problem free. Keep in mind that humans are still responsible for running the equipment that takes the measures and ultimately for interpreting the data provided by the measuring instrument. Thus even when using physical measures, a researcher needs to be concerned with the accuracy of the data.

IN REVIEW Feat

Description

Examples

Considerations

CRITICAL THINKING CHECK 5.

SUMMARY

REVIEW Features of Types of Measures

	Types of Measures			
	Self-Report	Tests	Behavioral	Physical
Description	Questionnaires or interviews that measure how people report that they act, think, or feel	A measurement instrument used to assess individual differences	Careful observations and recordings of behavior	Measures of bodily activity
Examples	Behavioral self-report Cognitive self-report Affective self-report	Ability tests Personality tests	Counting behaviors Classifying behaviors	Weight EEG recordings GSR recordings Blood pressure
Considerations	Are participants being truthful? How accurate are participants' memories?	Are participants being truthful? How reliable and valid are the tests?	Is there reactivity? How objective are observers?	Is the individual taking the measure skilled at using the equipment? How reliable and valid is the measuring instrument?

CRITICAL THINKING CHECK 5.2

1. Which types of measures are considered more subjective? Which are more objective?
2. Why might there be measurement error even when a researcher uses an objective measure such as a blood pressure cuff? What would you recommend to control or to minimize this type of measurement error?

SUMMARY

We discussed many elements important to measuring and manipulating variables. We learned the importance of operationally defining both the independent and the dependent variables in terms of the activities involved in measuring or manipulating each variable. It is also important to determine the scale, or level, of measurement of a variable based on its properties (identity, magnitude, equal unit size, and absolute zero). Once established, the level of measurement (nominal, ordinal, interval, or ratio) helps determine the appropriate type of statistics to be used. Data can also be classified as discrete (whole number units) or continuous (allowing for fractional amounts). We next described several types of measures, including self-report measures (reporting on how you act, think, or feel), tests (ability or personality), behavioral measures (observing and recording behavior), and physical measures (measurements of bodily activity).