

WHAT IS A VARIABLE?

In Chapter 1 we said that a hypothesis is your (informed) best guess about the relationship between two or more variables. But what is a variable, what types of variables are we dealing with, and how are they measured? In this chapter we will first consider a definition of *variable* and then discuss how variables can be classified according to how they are measured and according to their functions in our research.

A variable can be defined as *an attribute of a person or of an object which "varies" from person to person or from object to object*. For example, you may be left-handed. That is an attribute, and it varies from person to person. There are also right-handed people. Height, sex, nationality, and language group membership are all variables commonly assigned to people. Variables often attributed to objects include weight, size, shape, and color. In some of these examples, variables may appear to be of the all-or-nothing type. That is, the *Ss* either are or they aren't foreign students; they either are or they aren't speakers of French; they either are or they aren't left-handed. In such cases, we are simply asking whether a *S* does or does not possess the attribute, the variable. However, this is not always the case. There are much more general variables. The syntactic, semantic, and phonological elements of language are also variables. They are attributes of language and they are also something which people may possess (to some varying degree of proficiency). Even language acquisition is a variable, a very broad variable to be sure, but a variable all the same. The more specific a variable is, the easier it will be to locate and measure.

VARIABLE SCALES

As you can see from the examples, some variables are of an all-or-nothing sort. In such cases we cannot measure *how much* of the variable to attribute to a person but only whether or not it is present. In other cases we can measure very accurately how much of the variable subjects possess. Variables will be quantified on different scales depending on whether we want to know how much of the variable a person has or only about the presence or absence of the variable.

Nominal scale

You can probably guess that nominal refers to naming variables. For example, suppose it is important to your research to know whether your *Ss* know a third language. The plus or minus on knowledge of a third language is a variable which could be scaled in nominal form. While you could use a plus or minus, it is much more practical, since you are likely to use a computer, to assign a 1 to those *Ss* who know a third language and a 2 to those who don't. These numbers don't mean anything except that they identify the *S* as either possessing or not possessing that variable.

If you are doing research overseas, it may be important to know whether your *Ss* have been to the United States or not. Again, you could assign a 1 to those who have and a 2 to those who have not. Thus, the assignment of numbers on a nominal scale is not related to the characteristics of the *S* in any mathematical way. You could as well assign 1 to those who know two languages and a 2 to those who know three. Or you could assign a 1 to those who have not traveled to the United States and a 2 to those who have. The assignment of numbers simply allows you to separate the *Ss* into those who do and those who do not possess the variable.

It is also possible to have more than two levels for a variable. For example, if your variable is "foreign student," you could measure this as plus or minus on a nominal scale, but you could also decide to treat it as a variable with many levels. The variable +/- foreign student could be subdivided into levels, for example, Brazilian, Malaysian, Korean, and Mexican. Then your nominal scale would consist of five numbers—1 for nonforeign students, 2 for Brazilians, 3 for Malaysians, 4 for Koreans, and 5 for Mexicans. School grade is also a nominal variable which could have 1 to 12 levels (or 13, if you include kindergarten). Again the numbers have no mathematical value; they only serve to identify which *Ss* belong in which level of the variable. It's important to keep *levels* distinct from *variables*. It can be confusing, for a variable such as "foreign student" may be a multilevel variable in one study but a level within another variable—say "university student"—in another study.

Ordinal scale

We often want to measure how much of a variable is present when we can't measure the amount on an equal-interval scale. No ruler or thermometer, for example, can measure anomie, a child's interest in reading, culture shock, or a person's reactions to Suggestopedia. We can say, however, that a student is *very happy—happy—unhappy* or her performance in English is *poor—fair—good—excellent*, provided that we operationally define these terms. It is possible to measure the variable *culture shock* along a high to low scale with intermediate points identified in some rational way.

It is also possible to use an ordinal scale that ranks *Ss*. For example, you may interview a group of *Ss* and rank them from most to least proficient, assigning numbers from 1 and so on. Again, we assume you had some specific criteria in mind when making your rank-order judgments.

Interval scale

Interval measurement also tells us how much of something is present, but the measurement of the interval can be described. The interval should mean the same thing wherever it appears—for example, a nickel is a nickel (more or less), a year is a year, a millisecond is a millisecond. The interval unit is the same no matter when or where it occurs. The units can be added and subtracted. You can't do that with ordinal measurement. You can't say "fair + fair = good," not even if you've assigned numbers to fair and to good.

Test scores are considered as though they were interval scales. In a test score, if two *Ss* obtain scores of 95 and 100, respectively, we say that one is better than the other to the extent of the value of those five points. In other words, we assume that those five points are of equal value and that all 100 points are based on intervals of equal value. Assigning numbers to the scores from one to one hundred is based on the assumption that the intervals between 1 and 2 or 5 and 6 or 99 and 100 are of equal value.

Ideally, when we measure how much of a variable is present on an interval scale, we expect the intervals to be of equal value. In some cases there can be little argument about the value of intervals; in other cases we may not be so sure. For example, if age is the variable we are researching, we can assume that a year is a year. But the value of a year may differ along the scale. For second language learning, the difference between each year in the 10- to 20-year range may be much larger in value than the year intervals between 60 and 70.

Whether a variable is placed on a nominal, ordinal, or interval scale is sometimes determined by the type of variable, but more frequently the researcher must decide on the most appropriate scale for the variable. For example, if *bilingualism* is the variable you wish to research, you could place it on a nominal scale—your *Ss* either are or are not bilingual—and assign a 1 or 2 value to the variable. You could also assign it to an ordinal scale and either rank-order your *Ss* in relation to each other on how bilingual you think they are or assign them to a scale of extremely to not very bilingual. Or you could give them a test to measure proficiency in each language and thus obtain data on an interval scale for the variable.

We will discuss the measurement of variables in much more detail in later chapters. At this point, however, it should be clear that a variable is an attribute of a person or of an object. How it varies from person to person or from object to object can be measured by placing it on a scale.

Whatever the variables are that you want to investigate and whatever scale you select for the variables, you will need to define them further in terms of your research design. You must be clear about the function of each variable in your investigation.

THE FUNCTIONS OF VARIABLES

In order to assess the relationship between variables in research, we must be able to identify each according to the type of relationship we expect to

investigate. Variables can be classified as dependent, independent, or moderator variables. It is also possible to have intervening and control variables as well. Each of these will now be defined in turn.

Independent variable

The *independent variable* is the major variable which you hope to investigate. It is the variable which is selected, manipulated, and measured by the researcher. For example, if you want to investigate the effect of your instruction on reading scores of your students, then instruction is the independent variable because that is what you want to investigate. You will select some special instruction method, control the amount of this instruction, and manipulate it to meet your definition of instructional method. It is the independent variable, instruction, which you believe will affect the other variable (Ss' reading ability).

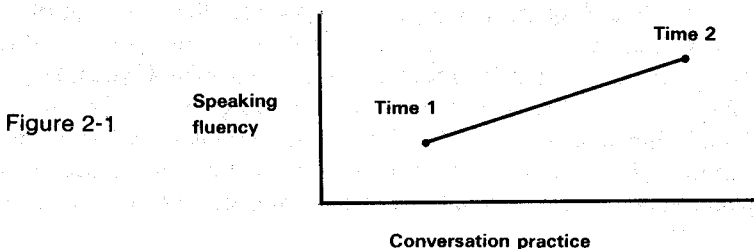
Dependent variable

The *dependent variable*, on the other hand, is the variable which you observe and measure to determine the effect of the independent variable. In our example, the reading scores of your Ss would be the dependent variable. Those scores, you believe, are *dependent* on the independent variable of instruction. (In reality, they may depend on many other things as well.) Usually dependent variables are represented by the letter *Y*, and independent variables by *X*, *Z*, or other letters.

Moderator variable

A *moderator variable* is a special type of independent variable which you may select for study in order to investigate whether it modifies the relationship between the dependent and the major independent variables. Suppose you were investigating the effect of conversation practice on the speaking fluency of foreign students. Conversation practice, then, would be the independent variable that you are interested in investigating. Fluency, operationally defined, is the dependent variable. However, you may have a hunch that conversation practice works better for your Spanish students than for your Chinese students. Or you may have a hunch that it works better for men than for women or vice versa. Thus, language and/or sex could be moderator variables.

Suppose you conducted your study and concluded that conversation practice increased the speaking fluency of your students. The outcome can be shown as in Figure 2-1. However, when you included sex as a moderator variable, you



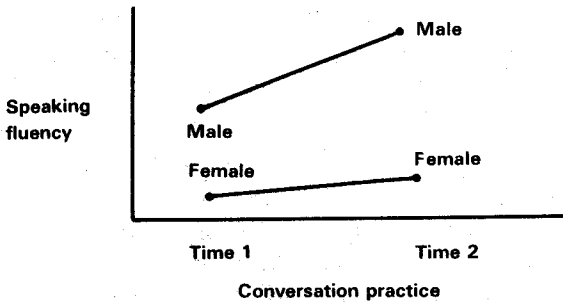


Figure 2-2

found a different pattern. That is, conversation practice turned out to be more effective for males than for females. Then your results would, perhaps, look something like Figure 2-2. (Please remember this is all hypothetical, *not* a prediction that men really would benefit more from conversation practice!) The effectiveness of conversation practice in promoting fluency is moderated by the sex of the learner.

Control variable

A *control variable* is a variable which is held constant in order to neutralize the potential effect it might have on behavior. For example, you might select all your Ss from Mexico to control the nationality factor. If you selected all your Ss from the beginning-level ESL class, then you have controlled the language proficiency factor. If your Ss are all right-handed, you have controlled for the laterality factor. Notice that you also limit the generalizability of your research by using control variables.

Intervening variables

Usually, the effect of the independent variable on the dependent variable is shown in terms of scores, counts, time measurement, etc. That is, the dependent variable is measured in some way to determine the effect of the independent variable. However, there is a process underlying the behavior we are measuring which is usually neither observable nor measurable. For example, in the study of oral fluency, oral fluency is measured. We have not, however, said anything about the *process* underlying the acquisition of fluency. A number of variables have not been measured which may or may not be part of that process—learning, intelligence, frustration. These have not been measured or manipulated. These are called *intervening variables*.

From this chapter, it may seem to you that there is a lot of jargon to be learned. But it is important that you be able to identify your variables by name even if you do not intend to analyze your data yourself. If you take your data to a consultant, you may be sure that the first thing you will be asked to do is to identify the dependent and independent variables. As soon as you have identified the variables, the next question will be on the scales you used in quantifying the data. If you hope to communicate with someone else who will

analyze your data for you, you must know these terms as well as someone who plans to carry out the entire procedure without help.

ACTIVITIES

1. Label the following variables as either (a) nominal, (b) ordinal, or (c) interval and give your justification for assigning them to the scale you select.

Examples: IQ score. Interval (because the score is obtained from a test which I assume has equal-interval measurement). Left-handed. Nominal (because Ss either are or are not left-handed).

Variables: male, bilingual, applied linguist, articulation accuracy, happy, EMR (educable mentally retarded), reading speed, audiolingual, placement test score, dyslexic, lip reader, English language proficiency.

2. Your research concerns the English language proficiency of students enrolled in English classes at the largest adult school in Los Angeles. There are many ways that you might want to subdivide these ESL learners into groups and levels. How many ways can you think of? Select two of these ways and explain why a researcher would want to divide the foreign students into these levels.

3. You are completing a study on the effect of participation in a volunteer aiding program on later performance in ESL practice teaching. What is the dependent variable? What is the independent variable? You also believe that type of class in which aiding was done (elementary school, adult school, university class) might have some relationship to success in student teaching if it was or wasn't in the same type of school. Identify this variable type.

4. You have asked each of your students to go out in the real world and make five complaints during the next week. They will judge their success on a five-point "success" to "failure" scale. During the week, half of these foreign students watched a videotape of an American woman returning a watch to a store, complaining that it didn't work properly. You want to know if the model helped. What is the dependent variable and what is the independent variable? All your students are adult women, and some of them work in factories, some in stores, and some in offices. How would you identify these variables? You decide that this might be a good "pilot" research project on foreign student success in small speech events. Give one other variable which might be important in such a study. How would you measure it?

5. You have worked out a whole series of role-play activities for your adult-school ESL class because you believe that practice in role-play promotes overall language proficiency. Your study shows that you are correct. What was the independent variable? The dependent?

6. You believe that the speed with which students go through the SRA reading kits in your ESL reading lab is negatively related to their reading comprehension. If you tested this, what would your variables be? Identify them, and say how you would measure them.

7. There are many characteristics identified within foreigner talk—the language addressed to language learners. Slow speech rate (WPM) is one of these. You believe that it is pause length rather than slower articulation of the message which makes the message easier for beginning language learners to understand. You decide to investigate this idea. State the question, the variables, and the hypotheses, and give your scales.

Suggested further reading for this chapter: Campbell and Stanley, and Tuckman.