TYPES OF VARIABLES

Independent and Dependent Variables

With definitional background behind us, we return to variables. Variables can be classified in several ways. In this book three kinds of variables are very important and will be emphasized: (1) independent and dependent variables, (2) active and attribute variables, and (3) continuous and categorical variables.

The most useful way to categorize variables is as independent and dependent. This categorization is highly useful because of its general applicability, simplicity, and special importance in conceptualizing and designing research and in communicating the results of research. An independent variable is the presumed cause of the dependent variable, the presumed effect. The independent variable is the antecedent; the dependent is the consequent. When we say: If $A$, then $B$, we have the conditional conjunction of an independent variable ($A$) and a dependent variable ($B$).

The terms "independent variable" and "dependent variable" come from mathematics, where $X$ is the independent and $Y$ the dependent variable. This is probably the best way to think of independent and dependent variables, because there is no need to use the touchy word "cause" and related words, and because such use of symbols applies to most research situations. There is no theoretical restriction on numbers of $X$'s and $Y$'s. When, later, we consider multivariate thinking and analysis, we will deal with several independent and several dependent variables.

In experiments the independent variable is the variable manipulated by the experimenter. When, for example, educational investigators study the effects of different teaching methods, they may manipulate method, the independent variable, by using different methods. In nonexperimental research, where there is no possibility of experimental manipulation, the independent variable is the variable that "logically" has some effect on a dependent variable. For example, in the research on cigarette-smoking and lung cancer, cigarette-smoking, which has already been done by many subjects, is the independent variable.

The dependent variable, of course, is the variable predicted to, whereas the independent variable is predicted from. The dependent variable, $Y$, is the presumed effect, which varies concomitantly with changes or variation in the independent variable, $X$. It is the variable that is not manipulated. Rather, it is observed for variation as a presumed result of variation in the independent variable. In predicting from $X$ to $Y$, we can take any value of $X$ we wish. whereas the value of $Y$ we predict to is "dependent on" the value of $X$ we
have selected. The dependent variable is ordinarily the condition we are trying to explain. The most common dependent variable in education, for instance, is achievement or "learning." We want to account for or explain achievement. In doing so we have a large number of possible X's or independent variables to choose from.

When the relation between intelligence and school achievement is studied, intelligence is the independent and achievement the dependent variable. (Is it conceivable that it might be the other way around?) Other independent variables that can be studied in relation to the dependent variable, achievement, are social class, methods of teaching, personality types, types of motivation (reward and punishment), attitudes toward school, class atmosphere, and so on. When the presumed determinants of delinquency are studied, such determinants as slum conditions, broken homes, lack of parental love, and the like, are independent variables and, naturally, delinquency (more accurately, delinquent behavior) is the dependent variable. In the frustration-aggression hypothesis mentioned earlier, frustration is the independent variable and aggression the dependent variable. Sometimes a phenomenon is studied by itself, and either an independent or a dependent variable is implied. This is the case when teacher behaviors and characteristics are studied. The usual implied dependent variable is achievement or child behavior in general. Teacher behavior can of course be a dependent variable.

The relation between an independent variable and a dependent variable can perhaps be more clearly understood if we lay out two axes at right angles to each other, one axis representing the independent variable and the other axis the dependent variable. (When two axes are at right angles to each other, they are called orthogonal axes.) Following mathematical custom, X, the independent variable, is the horizontal axis and Y, the dependent variable, the vertical axis. (X is called the abscissa and Y the ordinate.) X values are laid out on the X axis and Y values on the Y axis. A very common and useful way to "see" and interpret a relation is to plot the pairs of XY values, using the X and Y axes as a frame of reference. Let us suppose, in a study of child development, that we have two sets of measures: the X measures chronological age, the Y measures reading age:

<table>
<thead>
<tr>
<th>X: Chronological Age (in Months)</th>
<th>Y: Reading Age (in Months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>48</td>
</tr>
<tr>
<td>84</td>
<td>62</td>
</tr>
<tr>
<td>96</td>
<td>69</td>
</tr>
<tr>
<td>108</td>
<td>71</td>
</tr>
<tr>
<td>120</td>
<td>100</td>
</tr>
<tr>
<td>132</td>
<td>112</td>
</tr>
</tbody>
</table>

These measures are plotted in Figure 3.3.

The relation between chronological age (CA) and reading age (RA) can now be "seen" and roughly approximated. Note that there is a pronounced tendency, as might be expected, for more advanced CA to be associated with higher RA, medium CA with medium RA, and less advanced CA with lower RA. In other words, the relation between the independent and dependent variables, in this case between CA and RA, can be seen from a graph such as this. A straight line has been drawn in to "show" the relation. It is a rough average of all the points of the plot. Note that if one has knowledge of independent variable measures and a relation such as that shown in Figure 3.3, one can predict with considerable accuracy the dependent variable measures. Plots like this can of course be used with any independent and dependent variable measures.

*Reading age is a so-called growth age. Saturated measurements of individuals' growths—in height, weight, intelligence, and so forth—are expressed as the average chronological age at which they appear in the standard population.
The student should be alert to the possibility of a variable being an independent variable in one study and a dependent variable in another, or even both in the same study. An example is economic development. McClelland, for instance, used it as a dependent variable in studying the relation between Protestantism and economic development. Cutright, on the other hand, used it as an independent variable predicting to political development as a dependent variable. Anxiety has been studied as an independent variable affecting the dependent variable achievement. But it can readily be conceived and used as a dependent variable—for example, if we wished to study the effectiveness of types of teaching or types of teacher supportive behavior, or types of tests, in reducing anxiety. In other words, the independent and dependent variable classification is really a classification of uses of variables rather than a distinction between different kinds of variables.

Active and Attribute Variables

A classification that will be useful in our later study of research design is based on the distinction between experimental and measured variables. It is important when planning and executing research to distinguish between these two types of variables. Manipulated variables will be called active variables; measured variables will be called attribute variables.

Any variable that is manipulated, then, is an active variable. "Manipulation" means, essentially, doing different things to different groups of subjects, as we will see clearly in a later chapter where we discuss in depth the differences between experimental and non-experimental research. When a researcher does one thing to one group—for example,
positively reinforces a certain kind of behavior—and does something different to another
group, or has the two groups follow different instructions, this is manipulation. When one
uses different methods of teaching, or rewards the subjects of one group and punishes
those of another, or creates anxiety through worrisome instructions, one is actively manip-
ulating the variables methods, reinforcement, and anxiety.

Variables that cannot be manipulated are attribute variables. It is impossible, or at
least difficult, to manipulate many variables. All variables that are human characteristics—
intelligence, aptitude, sex, socioeconomic status, conservatism, field dependence, need
for achievement, and attitudes, for example—are attribute variables. Subjects come to our
studies with these variables (attributes) ready-made. Early environment, heredity, and
other circumstances have made individuals what they are.14 The word “attribute” moreover,
is accurate enough when used with inanimate objects or referents. Organizations,
institutions, groups, populations, homes, and geographical areas have attributes. Organiz-
ations are variably productive: institutions become outmoded; groups differ in cohesiv-
eness; geographical areas vary widely in resources.

This active-attribute distinction is general, flexible, and useful. We will see that some
variables are by their very nature always attributes, but other variables that are attributes
can also be active. This latter characteristic makes it possible to investigate the “same”
relations in different ways. A good example, again, is the variable anxiety. We can
measure the anxiety of subjects. Anxiety is in this case obviously an attribute variable.
But we can manipulate anxiety too. We can induce different degrees of anxiety, for
example, by telling the subjects of one experimental group that the task they are about to
do is difficult, that their intelligence is being measured, and that their futures depend on
the scores they get. The subjects of another experimental group are told to do their best but
to relax; the outcome is not too important and will have no influence on their futures.
Actually, we cannot assume that the measured (attribute) and the manipulated (active)
“anxieties” are the same. We may assume that both are “anxiety” in a broad sense, but
they are certainly not the same.

---

CONSTRUCTS, OBSERVABLES, AND LATENT VARIABLES

In much of the previous discussion of this chapter it has been implied, though not explicit-
lly stated, that there is a sharp difference between constructs and observed variables. In
fact, we can say that constructs are nonobservables, and variables, when operationally

14 Such variables have been called “dummy variables.” Since they are highly useful and powerful, even
indispensable, in modern research data analysis, they should be clearly understood. See F. Kerlinger and
1973, chaps. 6 and 7, and Chapter 34 of this volume. A *polyvany* is a division of the members of a group into
three or more subdivisions.
defined. are observables. The distinction is important, because if we are not always keenly aware of the level of discourse we are on when talking about variables, we can hardly be clear about what we are doing.

An important and fruitful expression, which we will encounter and use a good deal later in the book, is "latent variable." A latent variable is an unobserved "entity" presumed to underlie observed variables. The best-known example of an important latent variable is "intelligence." We note, say, that three ability tests, verbal, numerical, and spatial, are positively and substantially related. This means, in general, that persons high on one tend to be high on the others; similarly, persons low on one tend to be low on the others. We believe that something is common to the three tests or observed variables and name this something "intelligence." It is a latent variable.

We have encountered many examples of latent variables in previous pages: achievement, creativity, social class, anti-Semitism, conformity, and so on. Indeed, whenever we utter the names of phenomena on which people or objects vary, we are talking about latent variables. In science our real interest is more in the relations among latent variables than it is in the relations among observed variables because we seek to explain phenomena and their relations. When we enunciate a theory, we enunciate in part systematic relations among latent variables. We are not too interested in the relation between observed frustrated behaviors and observed aggressive behaviors, for example, though we must of course work with them at the empirical level. We are really interested in the relation between the latent variable frustration and the latent variable aggression.

We must be cautious, however, when dealing with nonobservables. Scientists, using such terms as "hostility," "anxiety," and "learning," are aware that they are talking about invented constructs the "reality" of which has been inferred from behavior. If they want to study the effects of different kinds of motivation, they must know that "motivation" is a latent variable, a construct invented to account for presumably "motivated" behavior. They must know that its "reality" is only a postulated reality. They can only judge that youngsters are motivated or not motivated by observing their behavior. Still, in order to study motivation, they must measure it or manipulate it. But they cannot measure it directly because it is an "in-the-head" variable, an unobservable entity, a latent variable, in short. The construct was invented for "something" presumed to be inside individuals, "something" prompting them to behave in such-and-such manners. This means that researchers must always measure presumed indicants of motivation and not motivation itself. They must, in other words, always measure some kind of behavior, be it marks on paper, spoken words, or meaningful gestures, and then make inferences about presumed characteristics—or latent variables.