CHAPTER 3

Constructs, Variables, and Definitions

From F. Kerlinger (1973) Foundations of Behavioral Research. NY: Holt, Rinehart Winston.

Scientists operate on two levels: theory-hypothesis-construct and observation. More accurately, they shuttle back and forth between these levels. A psychological scientist may say, "Early deprivation produces learning deficiency." This statement is a hypothesis consisting of two concepts, "early deprivation" and "learning deficiency," joined by a relation word, "produces." It is on the theory-hypothesis-construct level. Whenever the scientist utters relational statements and whenever he uses concepts, or constructs, as we shall call them, he is operating, so to speak, at this level.

The scientist must also operate at the level of observation. He must gather data that test his hypotheses. In order to do this, he must somehow get from the construct level to the observation level. He cannot simply make observations of "early deprivation" and "learning deficiency." He must so define these constructs that observations are possible. The problem of this chapter is to examine and clarify the nature of scientific concepts or constructs and the way in which behavioral scientists get from the construct level to the observation level, how they shuttle from one to the other.

Concepts and Constructs

The terms "concept" and "construct" have similar meanings. Yet there is an important distinction. A *concept* expresses an abstraction formed by generalization from particulars. "Weight" is a concept: it expresses numerous observations of things that are more or less "heavy" or "light." "Mass," "energy," and "force" are concepts used by physical scientists. They are, of course, much more abstract than concepts such as "weight," "height," and "length."

A concept of more interest to readers of this book is "achievement." It is an abstraction formed from the observation of certain behaviors of children. These behaviors are associated with the mastery or "learning" of school tasks—reading words, doing arithmetic problems, drawing pictures, and so on. The various observed behaviors are put together and expressed in a word—"achievement." "Intelligence," "aggressiveness," "conformity," and "honesty" are all concepts used to express varieties of human behavior of interest to behavioral scientists.

A *construct* is a concept. It has the added meaning, however, of having been deliberately and consciously invented or adopted for a special scientific purpose. "Intelligence" is a concept, an abstraction from the observation of presumably intelligent and nonintelligent behaviors. But as a scientific construct, "intelligence" means both more and less than it may mean as a concept. It means that scientists consciously and systematically use it in two ways. One, it enters into theoretical schemes and is related in various ways to other constructs. We may say, for example, that school achievement is in part a function of intelligence and motivation. Two, "intelligence" is so defined and specified that it can be observed and measured. We can make observations of the intelligence of children by administering X intelligence test to them, or we can ask teachers to tell us the relative degrees of intelligence of their pupils.

Variables

Scientists somewhat loosely call the constructs or properties they study "variables." Examples of important variables in sociology, psychology, and education are: sex, income, education, social class, organizational productivity, occupational mobility, level of aspiration, verbal aptitude, anxiety, religious affiliation, political preference, ego strength, task orientation, authoritarianism, conformity, intelligence, achievement. It can be said that a variable is a property that takes on different values. Putting it redundantly, a variable is something that varies. While this way of speaking gives us an intuitive notion of what variables are, we need a more general and yet more precise definition.

A *variable* is a symbol to which numerals or values are assigned. For instance, x is a variable: it is a *symbol* to which we assign numerical values. The variable x may take on any justifiable set of values—for example, scores on an intelligence test or an attitude scale. In the case of intelligence we assign to x a set of numerical values yielded by the procedure designated in a specified test of intelligence. This set of values, often called IQ's, ranges from low to high, from, say, 50 to 150.

A variable, x, however, may have only two values. If sex is the construct under study, then x can be assigned 1 and 0, 1 standing for one of the sexes and 0 standing for the other. It is still a variable. Other examples of two-valued variables are: alive-dead, citizen-noncitizen, middle class-working class, teacher-nonteacher, Republican-Democrat, and so on. Such variables are often called dichotomies or dichotomous variables.

Some of the variables used in behavioral research are true dichotomies—that is, they are characterized by the presence or absence of a property: male-female, alive-dead, employed-unemployed. Some variables are polytomies. A good example is religious preference: Protestant, Catholic, Jew, Other.[1] Most variables, however, are theoretically capable of taking on continuous values. It has been common practice in behavioral research to convert continuous variables to dichotomies or polytomies. For example, intelligence, a continuous variable, has been broken down into high and low intelligence, or into high, medium, and low intelligence. Variables such as anxiety, introversion, and authoritarianism have been treated similarly. Note that while it is not possible to convert a truly dichotomous variable such as sex to a continuous variable, it is always possible to convert a continuous variable to a dichotomy or a polytomy. As we will see later, such conversion can serve a useful conceptual purpose. But it is poor practice in the analysis of data because it throws information away.

Constitutive and Operational Definitions of Constructs and Variables

The distinction made earlier between "concept" and "construct" leads naturally to another important distinction: that between kinds of definitions of constructs and variables. Words or constructs can be defined in two general ways. First, we can define a word by using other words, which is what a dictionary usually does. We can define "intelligence" by saying it is "operating intellect," "mental acuity," or "the ability to think abstractly." Note that such definitions use other concepts or conceptual expressions in lieu of the expression being defined.

Second, we can define a word by telling what actions or behaviors it expresses or implies. Defining "intelligence" this way requires that we specify what behaviors of children are "intelligent" and what behaviors are "not intelligent." We may say that a child of seven who successfully reads a story we give him to read is "intelligent." If the child cannot read the story, we may say he is "not intelligent." In other words, this kind of definition can be called a behavioral or observational definition. Both "other word" and "observational" definitions are used constantly in everyday living.

There is a looseness about this discussion that would disturb a scientist. Though he uses the types of definition just described, he does so in a more precise and articulated manner. We express this usage by defining and explaining Margenau's distinction between constitutive and operational definitions.[2]

A *constitutive definition* defines a construct with other constructs. For instance, we can define "weight" by saying that it is the "heaviness" of objects. Or we can define "anxiety" as "subjectified fear." In both cases we have substituted one concept for another. Some of the constructs of a scientific theory may be defined constitutively. Torgerson, borrowing from Margenau, says that

[1]Such dichotomies and polytomies have been called "qualitative variables." The questionable nature of this designation will be discussed later.

[2]H. Margenau, *The Nature of Physical Reality*. New York: McGraw-Hill, 1950, chaps. 4, 5, and 12. The present discussion leans heavily on Margenau and on Torgerson's excellent presentation of the same distinction: W. Torgerson, *Theory and Methods of Scaling*. New York: Wiley, 1958, pp. 2–5.

all constructs, in order to be scientifically useful, must possess constitutive meaning.[3] This means that they must be capable of being used in theories.

An *operational definition* assigns meaning to a construct or a variable by specifying the activities or "operations" necessary to measure it. Alternatively, an operational definition is a specification of the activities of the researcher in measuring a variable or in manipulating it. An operational definition is a sort of manual of instructions to the investigator. It says, in effect, "Do such-and-such in so-and-so a manner." In short, it defines or gives meaning to a variable by spelling out what the investigator must do to measure it.

A well-known, if extreme, example of an operational definition is: Intelligence (anxiety, achievement, and so forth) is scores on X intelligence test, or intelligence is what X intelligence test measures. Notice that this definition tells us what to do to measure intelligence. It says nothing about how well intelligence is measured by the specified instrument. (Presumably the adequacy of the test was ascertained prior to the investigator's use of it.) In this usage, an operational definition is an equation where we say, "Let intelligence equal the scores on X test of intelligence." We also seem to be saying, "The meaning of intelligence (in this research) is expressed by the scores on X intelligence test." definitional operationism

There are, in general, two kinds of operational definitions: (1) *measured* and (2) *experimental.* The definition given above is more closely tied to measured than to experimental definitions. A *measured* operational definition describes how a variable will be measured. For example, achievement may be defined by a standardized achievement test, by a teacher-made achievement test, or by grades. Hare defined the consensus of a group as follows: "The amount of consensus in the group is measured by having each individual rate the ten pieces of camping equipment before discussion, the group rate the equipment during the discussion, and the individuals again rate them after discussion."[4] A study may include the variable *consideration*. It can be defined operationally by listing behaviors of children that are presumably considerate behaviors and then requiring teachers to rate the children on a five-point scale. Such behaviors might be when the children say to each other, "I'm sorry," or "Excuse me," when one child yields a toy to another on request (but not on threat of aggression), or when one child helps another with a school task.

An *experimental* operational definition spells out the details (operations) of the investigator's manipulations of a variable. Reinforcement can be operationally defined by giving the details of how subjects are to be reinforced (rewarded) and not reinforced (not rewarded) for specified behaviors. In the Hurlock study discussed earlier, for example, some children were praised, some blamed, and some ignored. Dollard et al. define frustration as prevention from reaching a goal, or "... interference with the occurrence of an instigated goal response at its proper time in the behavior sequence ..."[5] This definition contains clear implications for experimental manipulations. Barker, Dembo, and Lewin, apparently

[3] *Ibid.*, p. 5.

[4] A. Hare, "A Study of Interaction and Consensus in Different Sized Groups," *American Sociological Review*, XVII (1952), 261–267.

[5] J. Dollard et al., *Frustration and Aggression*. New Haven: Yale University Press, 1939, p. 7.

influenced by the definition of Dollard et al., operationally defined frustration by describing children put into a playroom with "a number of highly attractive, *but inaccessible*, toys."[6] (The toys were put behind a wire-net partition; the children could see them but not touch them.) Other examples of both kinds of operational definitions will be given later.

Scientific investigators must sooner or later face the necessity of measuring the variables of the relations they are studying. Sometimes measurement is easy, sometimes difficult. To measure sex or social class is easy; to measure creativity, anxiety, or organizational effectiveness is difficult. The importance of operational definitions cannot be overemphasized. They are indispensable ingredients of scientific research because they enable researchers to measure variables and because they are bridges between the theory-hypothesis-construct level and the level of observation. There can be no scientific research without observations, and observations are impossible without clear and specific instructions on what and how to observe. Operational definitions are such instructions.

Though indispensable, operational definitions yield only limited meanings of constructs. No single operational definition can ever express all of a variable. No operational definition of intelligence can ever express the rich and diverse meaning of human intelligence. This means that the variables measured by scientists are always limited and specific in meaning. The "creativity" studied by psychologists is not the "creativity" referred to by artists, though there will of course be common elements.

Some scientists say that such limited operational meanings are the only meanings that "mean" anything, that all other definitions are metaphysical nonsense. They say that discussions of anxiety are metaphysical nonsense, unless adequate operational definitions of anxiety are available and are used. This view is extreme, though it has healthy aspects. To insist that every term we use in scientific discourse be operationally defined would be too narrowing, too restrictive, and, as we shall see, scientifically unsound.[7]

Despite the dangers of extreme operationism, it can be safely said that operationism has been and still is a healthy influence because, as Skinner puts it, "The operational attitude, in spite of its shortcomings, is a good thing in any science but especially in psychology because of the presence there of a vast vocabulary of ancient and nonscientific origin."[8] When the terms used in education are considered, it is clear that education, too, has a vast vocabulary of ancient and nonscientific terms. Consider these: the whole child, horizontal and vertical enrichment, meeting the needs of the learner, core curriculum, emotional adjustment, and curricular enrichment.

[6]R. Barker, T. Dembo, and K. Lewin, "Frustration and Regression." In R. Barker, J. Kounin, and H. Wright, *Child Behavior and Development*. New York: McGraw-Hill, 1943, p. 443.

[7]For a good discussion of this point, see F. Northrop, *The Logic of the Sciences and the Humanities*. New York: Macmillan, 1947, chaps VI and VII. Northrop says, for example, "The importance of operational definitions is that they make verification possible and enrich meaning. They do not, however, exhaust scientific meaning" (p. 130). Margenau makes the same point in his extended discussion of scientific constructs. (See Margenau, *op. cit.*, pp. 232ff.)

[8]B. Skinner, "The Operational Analysis of Psychological Terms." In H. Feigl and M. Brodbeck, eds., *Readings in the Philosophy of Science*. New York: Appleton, 1953, p. 586.

To clarify constitutive and operational definitions—and theory, too—look at Fig. 3.1, which has been adapted after Margenau and Torgerson. The diagram is supposed to illustrate a well-developed theory. The single lines represent theoretical connections or relations between constructs. These constructs, labeled with lower-case letters, are defined constitutively; that is, c_4 is defined somehow by c_3, or vice versa. The double lines represent operational definitions. The C constructs are directly linked to observable data; they are indispensable links to empirical reality. But it is important to note that not all constructs in a scientific theory are defined operationally. Indeed, it is a rather thin theory that has all its constructs so defined.

Let us build a "small theory" of underachievement to illustrate these notions. Suppose an investigator believes that underachievement is, in part, a function of pupils' self-concepts. He believes that pupils who perceive themselves "inadequately," who have negative self-percepts, also tend to achieve less than their potential capacity and aptitude indicate they should achieve. He further believes that ego-needs (which we will not define here) and motivation for achievement (call this *n*-ach, or need for achievement) are tied to underachievement. Naturally, he is also aware of the relation between aptitude and intelligence and achievement in general. A diagram to illustrate this "theory" might look like Fig. 3.2.

The investigator has no *direct* measure of self-concept, but he assumes that he can draw inferences about an individual's self-concept from a figure-drawing test. He operationally defines self-concept, then, as certain responses to the figure-drawing test. This is probably the most common method of measuring psychological (and educational) constructs. The heavy single line between c_1 and C_1 indicates the relatively direct nature of the presumed relation between self-concept and the test. (The double line between C_1 and the level of observation indicates an

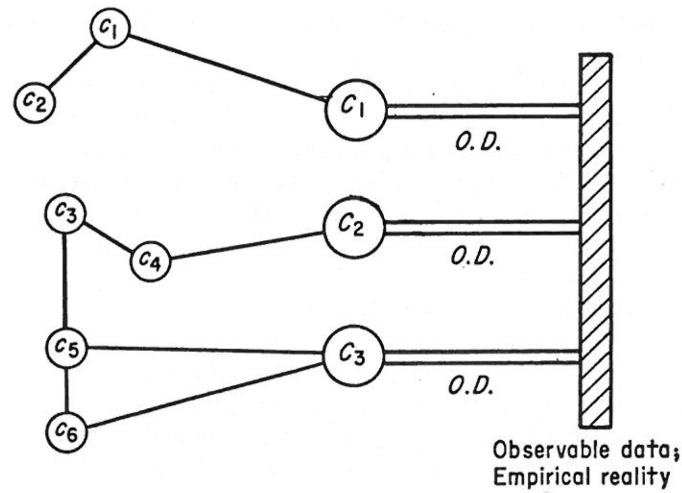

FIG. 3.1 Constructs defined operationally—that is, connected to observable data by operational definitions (*O. D.*'s): C_1, C_2, C_3. Constructs defined constitutively: $c_1, c_2, \ldots, c_6$

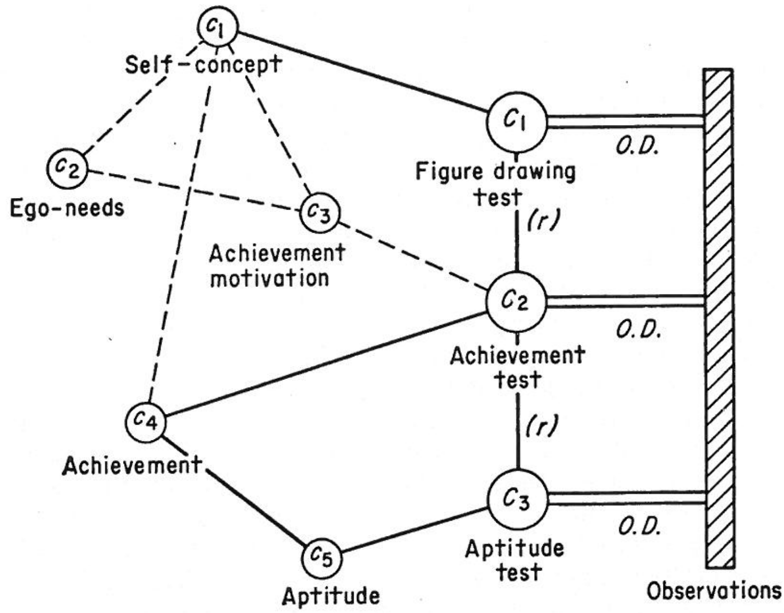

FIG. 3.2

operational definition, as it did in Fig. 3.1.) Similarly, the construct achievement (c_4) is operationally defined as the discrepancy between measured achievement (C_2) and measured aptitude (c_5). In this model the investigator has no direct measure of achievement motivation, no operational definition of it. In another study, naturally, he may specifically hypothesize a relation between achievement and achievement motivation, in which case he will try to define achievement motivation operationally.

A single solid line between concepts—for example, the one between the construct achievement (c_4) and achievement test (C_2)—indicates a relatively well-established relation between postulated achievement and what standard achievement tests measure. The single solid lines between C_1 and C_2 and between C_2 and C_3 indicate obtained relations between the test scores of these measures. (The lines between C_1 and C_2 and between C_2 and C_3 are labeled *r* for "relation," or "coefficient of correlation.")

The broken single lines indicate postulated relations between constructs that are not relatively well established, are more tenuous. A good example of this is the postulated relation between self-concept and achievement motivation. One of the aims of science is to make these broken lines solid lines by bridging the operational definition-measurement gap. In this case, it is quite conceivable that both self-concept and achievement motivation can be operationally defined and directly measured.

In essence, this is the way the behavioral scientist operates. He shuttles back and forth between the level of theory-constructs and the level of observation. He does this by operationally defining the variables of his theory that are amenable to such definition and then by estimating the relations between the operationally defined and measured variables. From these estimated relations he makes inferences as to the relations between the constructs. In the above example, he calculates the relation between C_1 (Figure-drawing test) and C_2 (Achievement test) and, if the relation is established on this observational level, he infers that a relation exists between c_1 (Self-concept) and c_4 (Achievement).

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 important and 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. Indeed, it can even be said that in scientific research the relations between X's and Y's are constantly pursued. And 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.

seems implicit nonetheless.

In experiments the independent variable is the variable manipulated by the experimenter. When, for example, an educational investigator studies the effect of different teaching methods, he may manipulate method, the independent variable, by using different methods. In nonexperimental research, where there is no possibility of manipulation, the independent variable is the variable that has presumably been "manipulated" before he got it. He may, for instance, study the presumed effects on achievement of a ready-made teaching situation in which different methods have already been used. Methods, here, is also the independent variable. Or he may study the effect on school achievement of parental attitudes. Here parental attitudes 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:[9]

X: Chronological Age (in Months)	Y: Reading Age (in Months)
72	48
84	62
96	69
108	71
120	100
132	112

These measures are plotted in Fig. 3.3.

[9]*Reading age* is so-called growth age. Seriatim 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 data reported above are from one of the author's studies: F. Kerlinger, "The Statistics of the Individual Child: The Use of Analysis of Variance with Child Development Data," *Child Development*, XXV (1954), 265–275. This article refers to the original sources of the growth data.

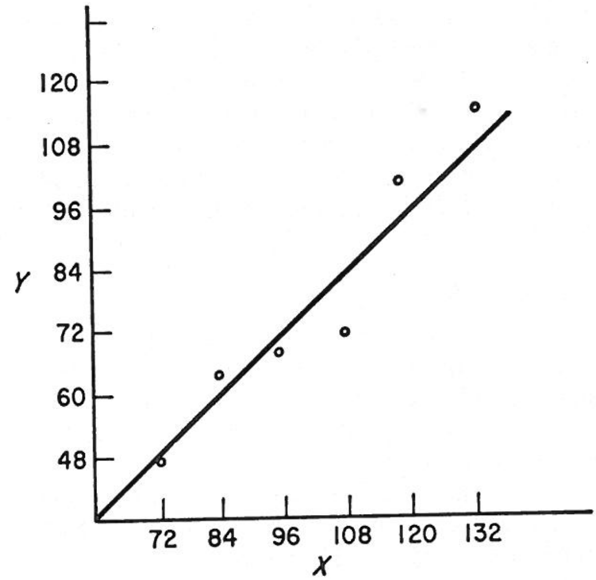

FIG. 3.3 Y: reading age (in months); X: chronological age (in months)

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 Fig. 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.

The student should be alert to the possibility of a variable being an independent variable in one study and a dependent variable in another. A good example is the variable "group cohesiveness." It is possible to study the determinants of cohesiveness and equally possible to study the presumed results of cohesiveness. In the former case cohesiveness is the dependent variable; in the latter it is the independent variable.

It is quite possible to consider school achievement as an independent variable, even though it is usually treated as a dependent variable. In studies of success in college, for instance, achievement in high school is often used as a predictor, or an independent variable. Here we have the interesting case of achievement being now the independent and now the dependent variable—achievement in different educational situations, naturally. Anxiety has been studied as an independent variable affecting the dependent variable achievement.[10] But anxiety

[10]See S. Sarason et al., *Anxiety in Elementary School Children*. New York: Wiley, 1960.

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.[11] It is also possible—but too rarely done—to treat a variable as independent in one study and dependent in another study. We can even give a variable this dual role in the same study. 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. Any variable that cannot be manipulated is an attribute variable. When one uses different methods of teaching, rewarding the children of one group and punishing those of another, creating anxiety through worrisome instructions, one is *actively* manipulating the variables methods, reinforcement, and anxiety.

On the other hand, it is impossible, or at least difficult, to manipulate many variables. All variables that are human characteristics—intelligence, aptitude, sex, socioeconomic status, field dependence, education, need for achievement, and attitudes, for example—are attribute variables. Subjects come to our studies with these variables (attributes) ready-made. They are, so to speak, already manipulated. Early environment, heredity, and other circumstances have made individuals what they are.[12] The word "attribute," moreover, is accurate enough when used with inanimate objects or referents. Organizations, institutions, groups, populations, homes, and geographical areas have attributes. Organizations are variably productive; institutions become outmoded; groups differ in cohesiveness; 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 is the variable anxiety. We can measure the anxiety of subjects. Anxiety is in this case

[11] *Ibid.*, p. 280.

[12] Such variables are also called *organismic* variables. Any property of an individual, any characteristic or attribute, is an organismic variable. It is part of the organism, so to speak. In other words, organismic variables are those characteristics that individuals have in varying degrees when they come to the research situation. The term individual differences implies organismic variables.

Another related classification, used mainly by psychologists, is *stimulus* and *response* variables, A *stimulus variable* is any condition or manipulation by the experimenter of the environment that evokes a response in an organism. A *response variable* is any kind of behavior of the organism. The assumption is made that for any kind of behavior there is always a stimulus. Thus the organism's behavior is a response. This classification is reflected in the well-known equation: $R = f(O, S)$, which is read: "Responses are a function of the organism and stimuli," or "Response variables are a function of organismic variables and stimulus variables."

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.

Continuous and Categorical Variables

A distinction especially useful in the planning of research and the analysis of data — that between continuous and categorical variables — has already been introduced. Its later importance, however, justifies more extended consideration.

A *continuous* variable is capable of taking on an ordered set of values within a certain range. This definition means, first, that the values of a continuous variable reflect at least a rank order, a larger value of the variable meaning more of the property in question than a smaller value. The values yielded by a scale to measure dogmatism, for instance, express differing amounts of dogmatism from high through medium to low. Second, continuous measures in actual use are contained in a range, and each individual obtains a "score" within the range. A scale to measure dogmatism may have the range 1 through 7. Most scales in use in the behavioral sciences also have a third characteristic: there are a theoretically infinite set of values within the range. (Rank-order scales are somewhat different; they will be discussed later in the book.) That is, a particular individual's score may be 4.72 rather than simply 4 or 5.

Categorical variables, as I will call them, belong to a kind of measurement called nominal. (It will be explained in Chapter 25.) In nominal measurement, there are two or more subsets of the set of objects being measured. Individuals are categorized by their possession of the characteristic that defines any subset. "To categorize" means to assign an object to a subclass or subset of a class or set on the basis of the object's having or not having the characteristic that defines the subset. The individual being categorized either has the defining property or does not have it; it is an all-or-none kind of thing. The simplest examples are dichotomous categorical variables: sex, Republican-Democrat, white-black. Polytomies, variables with more than two subsets or partitions, are fairly common, especially in sociology and economics: religious preference, education (usually), nationality, occupational choice, and so on.

Categorical variables — and nominal measurement — have simple requirements: all the members of a subset are considered the same and all are assigned the same name (nominal) and the same numeral. If the variable is religious preference, for instance, all Protestants are the same, all Catholics are the same, and all "others" are the same. If an individual is a Catholic — operationally defined in a suitable way — he is assigned to the category "Catholic" and also assigned a "1" in that category. In brief, he is counted as a "Catholic." Categorical variables are "democratic": there is no rank order or greater-than-and-less-than among the categories, and all members of a category have the same value: 1.

The expression "qualitative variables" has sometimes been applied to categorical variables, especially to dichotomies, probably in contrast to "quantitative variables" (our continuous variables). Such usage reflects a somewhat distorted notion of what variables are. They are always quantifiable, or they are not variables. If x has only two subsets and can take on only two values, 1 and 0, these are still values, and the variable varies. If x is a polytomy, like religious preference, we quantify again by assigning 1's and 0's to individuals. If an individual, say, is a Catholic, then put him in the Catholic subset and assign him a 1. It is extremely important to understand this because, for one thing, it is the basis of quantifying many variables—even experimental treatments—for complex analysis. In multiple regression analysis, as we will see later in the book, all variables, continuous and categorical, are entered as variables into the analysis. Earlier, the example of sex was given, 1 being assigned to one sex and 0 to the other. We can set up a column of 1's and 0's just as we would set up a column of IQ's or dogmatism scores. The column of 1's and 0's is the quantification of the variable sex. There is no mystery here. The method is easily extended to polytomies.[13]

Constructs, Observables, and Intervening Variables

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

Constructs have been called intervening variables.[14] *Intervening variable* is a term invented to account for internal and directly unobservable psychological processes that in turn account for behavior. Tolman, using William James' picturesque expression, says, ". . . the sole 'cash-value' of mental processes lies . . . in this their character as a set of intermediate functional processes which interconnect between the initiating causes of behavior, on the one hand, and the final resulting behavior itself, on the other."[15] An intervening variable is an "in-the-head" variable. It cannot be seen, heard, or felt. It is inferred from behavior.

[13]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 E. Pedhazur, *Multiple Regression Analysis in Behavioral Research*. New York: Holt, Rinehart and Winston, 1973, chaps. 6 and 7, and Chapter 36 of this volume. A *polytomy* is a division of the members of a group into three or more subdivisions. The method of "coding" variables described above—and other methods—are explained in Chapter 36. See also Chapters 5 and 25.

[14]E. Tolman, *Behavior and Psychological Man*. Berkeley, Calif.: University of California Press, 1958, pp. 115–129.

[15]*Ibid.*, p. 116.

"Hostility" is inferred from presumably hostile or aggressive acts. "Learning" is inferred from, among other things, increases in test scores. "Anxiety" is inferred from test scores, from skin responses, from heart beat, and so on.

The scientist, using such terms, is always aware that he is talking about invented constructs the "reality" of which he has inferred from behavior. If he wants to study the effects of different kinds of motivation, he must know that "motivation" is an intervening variable, a construct invented by man to account for presumably "motivated" behavior. He must know that its "reality" is only a postulated reality. He can only judge that a youngster is motivated or not motivated by observing the youngster's behavior. Still, in order to study motivation, he must measure it. But he cannot measure it directly because it is an in-the-head variable, an intervening variable, an unobservable entity. Other men have invented the construct to stand for "something" *presumed to be* inside the individual, "something" prompting him to behave in such-and-such manner. This means that he must always measure presumed indicants of motivation and not motivation itself. He 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.

Examples of Constructs and Operational Definitions

A number of constructs and operational definitions have already been given. To illustrate and perhaps clarify the preceding discussion, especially that in which the distinction was made between experimental and measured variables and between constructs and operationally defined variables, several and varied examples of constructs and operational definitions are given below. If a definition is experimental, it is labeled (E). If it is measured, it is labeled (M).

The student should note that operational definitions differ in degree of specificity. Some are quite closely tied to observations. "Test" definitions like "Intelligence is defined as score on X intelligence test" are very specific. A definition like "Frustration is prevention from reaching a goal" is more general and requires further specification, such as that of the Barker, Kounin, and Wright definition, in order to be directly measurable.

Social Class "... two or more orders of people who are believed to be, and are accordingly ranked by the members of a community, in socially superior and inferior positions."[16] (M) (To be operational this definition has to be specified by questions aimed at people's beliefs about other people's positions.)

This is a subjective definition of social class. Social class, or social status, is also defined more objectively by using such indices as occupation, income, and education, or by combinations of such indices. For example, "To get an index of SES (socioeconomic status), we combined a measure of occupational level with one of income."[17] (Then details of the definition follow.) (M)

[16] W. Warner and P. Lunt, *The Social Life of a Modern Community*. New Haven: Yale University Press, 1941, p. 82.

[17] R. Sears, E. Maccoby, and H. Levin, *Patterns of Child Rearing*. New York: Harper & Row, 1957, pp. 423–424.

Anxiety (Test Anxiety) "Our measure of test anxiety was a revised and shortened version of the Test Anxiety (TA) questionnaire which is described elsewhere."[18] (M) (The authors refer to a research study that describes the construction and development of the TA Scale. It is always good practice to refer to the original sources of tests and scales.)

Achievement (School, Arithmetic, Spelling) Achievement is customarily defined operationally by citing a standardized test of achievement (for example, Iowa Every-Pupil Tests of Basic Skills, Elementary); by grade-point averages; or by teacher judgments.

"The criterion of school achievement, grade-point average . . . was generally obtained by assigning weights of 4, 3, 2, 1, and 0 to grades of A, B, C, D, and F, respectively. Only courses in the so-called 'solids,' that is, mathematics, science, social studies, foreign language, and English, were used in computing grade-point averages."[19] (M)

Social Organization (of Classes) ". . . has to do with the amount of social grouping and pupil autonomy in a class. A class scoring high was one in which it was relatively common to find the class broken up into two or more groups working independently, and in which the teachers talked relatively little."[20] (M) (Note the relative looseness of this definition, a looseness introduced by such words as "relatively common" and "relatively little." But it is not too difficult to tighten it.

Popularity Popularity is often defined operationally by the number of sociometric choices an individual receives from other individuals (in his class, play group, and so on). Individuals are asked: "With whom would you like to work?", "With whom would you like to play?", and the like. Each individual is required to choose one, two, or more individuals from his group on the basis of such criterion questions.

"The sociometric popularity score (choices received) based on a criterion of enjoyment of participation, and the expansiveness (choices made), confidence (choices expected) . . ."[21] (M)

Reinforcement Reinforcement definitions come in a number of forms. Most of them involve, one way or another, the principle of reward. But note that both positive and negative reinforcement may be used.

". . . statements of *agreement* or *paraphrase*."[22] (E) Then the author gives specific experimental definitions of "reinforcement." For example, "In the second 10 minutes, every opinion statement *S* made was recorded by *E* and reinforced. For two groups, *E* agreed with every opinion statement by saying: 'Yes, you're

[18] I. Sarnoff et al., "A Cross-Cultural Study of Anxiety among American and English School Children," *Journal of Educational Psychology*, XLIX (1958), 129–136.

[19] W. Holtzman and W. Brown, "Evaluating the Study Habits and Attitudes of High School Students," *Journal of Educational Psychology*, LIX (1968), 404–409.

[20] D. Medley and H. Mitzel, "Some Behavioral Correlates of Teacher Effectiveness," *Journal of Educational Psychology*, L (1959), 239–246.

[21] E. Borgatta, "Analysis of Social Interaction: Actual, Role-Playing, and Projective," *Journal of Abnormal and Social Psychology*, LI (1955), 394–405.

[22] W. Verplanck, "The Control of the Content of Conversation: Reinforcement of Statements of Opinion," *Journal of Abnormal and Social Psychology*, LI (1955), 668–676.

right,' 'That's so,' or the like, or by nodding and smiling affirmation if he could not interrupt."[23] (E)

"... the model and the child were administered alternately 12 different sets of story items ... To each of the 12 items, the model consistently expressed judgmental responses in opposition to the child's moral orientation ... and the experimenter reinforced the model's behavior with verbal approval responses such as 'Very good,' 'That's fine,' and 'That's good.' The child was similarly reinforced whenever he adopted the model's class of moral judgments in response to his own set of items."[24] (E) (This is called "social reinforcement.")

"Specified Comment students, regardless of teacher or student differences, all received comments designated in advance for each letter grade, as follows:

A. Excellent! Keep it up.
B. Good work. Keep it up.
C. Perhaps try to do still better?
D. Let's bring this up.
F. Let's raise this grade!

Teachers were instructed to administer the comments 'rapidly and automatically, trying not even to notice who the students are.'"[25] (E) Two other experimental conditions, Free Comment and No Comment, were also operationally (experimentally) defined by the author.

Acquisition (of Learning or Conditioning) "... probability of occurrence, expressed as the number of trials on which a given subject produces a CR, or the percentage of subjects giving a CR on a given trial."[26] (E-M) (CR means conditioned response.)

Extinction (of Learning or Conditioning) "... the decrease of response strength of nonreinforcement."[27] (E-M)

Response Set "... a general 'tendency to agree or disagree with questionnaire items, regardless of their content.'" (M) The authors later operationalize this definition in different ways. One of these is: "An Overall Agreement Score (OAS) was computed for each *S* by taking the mean of their responses to the 360 items."[28]

Community Reputation This is a variable of Newcomb's study of the attitudes of Bennington College girls. It was operationally defined by having knowledgeable students choose other students sociometrically. Twenty-eight criteria were used. For each of these the student judges nominated three individuals.

[23] *Ibid.*

[24] A. Bandura and F. McDonald, "Influence of Social Reinforcement and the Behavior of Models in Shaping Children's Moral Judgments," *Journal of Abnormal and Social Psychology*, LXVII (1963), 274–281.

[25] E. Page, "Teacher Comments and Student Performance: A Seventy-Four Classroom Experiment in School Motivation," *Journal of Educational Psychology*, XLIX (1958), 173–181.

[26] G. Kimble, *Hilgard and Marquis' Conditioning and Learning*, 2d ed. New York: Appleton, 1961, p. 82.

[27] *Ibid.*

[28] A. Couch and K. Keniston, "Yeasayers and Naysayers: Agreeing Response Set as a Personality Variable," *Journal of Abnormal and Social Psychology*, LX (1960), 151–174.

Examples of the criteria are:

1. *Most absorbed* in social life, weekends, etc.
2. *Most influenced* by faculty authority.
3. *Least likely* to engage actively in pursuits related to college interest.[29]

Values "'Rank the ten goals in the order of *their importance to you*.' (1) financial success; (2) being liked; (3) success in family life; (4) being intellectually capable; (5) living by religious principles; (6) helping others; (7) being normal, well adjusted; (8) cooperating with others; (9) doing a thorough job; (10) occupational success."[30] (M)

Approval-Disapproval (of Pupils by Teachers) ". . . time samples of classroom behavior were spread over an entire school year . . . interactions between teachers and pupils were classified into two categories: (a) praise contacts (teacher initiated interactions with a child in which she verbally expressed approval of some behavior which the child had displayed), and (b) blame contacts (teacher initiated interactions with a child in which she verbally expressed disapproval for some bit of behavior which the child had displayed)."[31] (M)

Honesty "A child was considered 'dishonest' on the *Clapp-Young Arithmetic Test* in the correction of his paper if the inside of the test booklet showed: (a) that he had changed his answer by drawing a circle around his wrong response and had made an X for the correct response; (b) that he had erased the wrong response and marked the correct one; (c) that he marked the correct response when checking his paper but his work on the outside of the test booklet did not agree with his answer."[32] (M)

Leadership "We have chosen to measure two specific dimensions of leader behavior, 'Initiating Structure' and 'Consideration.' Initiating Structure refers to the leader's behavior in delineating the relationship between himself and the members of the work-group, and in endeavoring to establish well-defined patterns of organization, channels of communication, and methods of procedure. Consideration refers to behavior indicative of friendship, mutual trust, respect, and warmth in the relationship between the leader and the members of his staff."[33] (M) (Later, the author defines the two dimensions more precisely and operationally.)

Cohesiveness ". . . group cohesiveness refers to the degree to which the members of a group desire to remain in the group."[34] (M) (The author goes on to elaborate and operationalize the definition.)

[29]T. Newcomb, *Personality and Social Change*. New York: Holt, Rinehart and Winston, 1943, pp. 66–67, 184–185.

[30]T. Newcomb, *The Acquaintance Process*. New York: Holt, Rinehart and Winston, 1961, pp. 40 and 83.

[31]W. Meyer and G. Thompson, "Sex Differences in the Distribution of Teacher Approval and Disapproval among Sixth-Grade Children," *Journal of Educational Psychology*, XLVII (1956), 385–396.

[32]Sister M. Gross, "The Effect of Certain Types of Motivation on the 'Honesty' of Children," *Journal of Educational Research*, XL (1946), 133–140.

[33]A. Halpin, *The Leadership Behavior of School Superintendents*. Chicago: Midwest Administration Center, University of Chicago, 1956, p. 4.

[34]D. Cartwright, "The Nature of Group Cohesiveness." In D. Cartwright and A. Zander, eds., *Group Dynamics*, 3d ed. New York: Harper & Row, 1968, 91–109.

Institutional Quality (Excellence) "... the average academic ability of the entering student body and the per-student expenditures for 'educational and general' purposes (meaning, primarily, salaries for faculty and staff)."[35] (M)

Before leaving operational definitions, we should discuss something that may puzzle the student. In reading the literature, as often as not the student will not encounter operational definitions as such. Investigators may or may not *explicitly* define their terms operationally. A good research report, of course, should be so written that another investigator, if he chooses, can repeat the research. This implies that he will be able to measure the variables or manipulate the experimental conditions in the same way the original investigator did. To do this, naturally, he must know clearly and explicitly how to measure the variables or how to manipulate the experimental conditions. Therefore report writers must include operational definitions directly or indirectly. The commonest practice seems to be to mention the variables of the study early in the report and later to discuss the instruments used to measure the variables, or to discuss the experimental procedures used to manipulate the independent variables. Obviously the variables are thus operationally defined. An investigator will state a hypothesis, for example, that contains his variables: Underachievement is a function of inadequate self-concept. He will then probably discuss the hypothesis and in so doing say something about the concepts "underachievement" and "self-concept." But he may not define them operationally at this point. Later, in his section on method or procedure he will explain how he intends to measure underachievement and self-concept.

The benefits of operational thinking can be great. Although operationism can be carried to extremes, and although extreme operationism is dangerous because it tends to shut out the recognition of the importance of constructs and constitutive definitions in scientific theory and research and also tends to narrow research to perhaps trivial problems, there can be little doubt that it is a healthy scientific influence. As Underwood has said, in one of his fine chapters on operational definitions:

> ... I would say that operational thinking makes better scientists. The operationist is forced to remove the fuzz from his empirical concepts ...
>
> ... operationism facilitates communication among scientists because the meaning of concepts so defined is not easily subject to misinterpretation.[36]

Study Suggestions

1. Make up operational definitions for the following constructs. When possible, write two such definitions: an experimental one and a measured one.

permissiveness	underachievement
reinforcement	leadership

[35] A. Astin, "Undergraduate Achievement and Institutional 'Excellence,'" *Science*, CLXI (1968), 661–668.

[36] B. Underwood, *Psychological Research*. New York: Appleton, 1957, p. 53.

reading ability	class atmosphere
achievement	delinquency
interests	organizational conflict
needs	self-other attitudes
transfer of training	conformity

Notice that some of these concepts or variables—for example, needs and transfer of training—are difficult to define operationally. Why?

2. It is instructive and broadening for specialists to read outside their fields. This is particularly true for students of behavioral research. It is suggested that the student of a particular field read two or three research studies in one of the best journals of another field. If you are in psychology, read a sociology journal, say the *American Sociological Review*. If you are in education or sociology, read a psychology journal, say the *Journal of Personality and Social Psychology*. Students not in education can sample the *Journal of Educational Psychology* or the *American Educational Research Journal*. When you read, jot down the names of the variables and compare them to the variables in your own field. Are they primarily active or attribute variables? Note, for instance, that psychology's variables are more "active" than sociology's. What implications does the nature of the variables of a field have for research in the field?