Evidence and Inference in Social Research*

PAUL F. LAZARSFELD

IN RECENT decades a major development of quantitative research in real social situations has come about; in a very broad sense it may be labeled "survey research." The following discussion is concerned with several recurring and very general problems connected with this type of work.

Concentration on survey research does not underrate the importance of other techniques. The same period has seen a similar development in the use of experimental methods in the social sciences; but the logical problems of experiments here and in the natural sciences are essentially the same. There is also a long tradition of social analysis using historical data to study problems of large-scale social change—the relation of the Protestant Reformation to the rise of capitalism is the most famous example. This kind of study raises interesting logical problems of its own that need careful analysis, but the methodology of "historical sociology" is not yet developed enough to be discussed here. Finally there are the attempts to derive social inferences from institutional data. It has been said that man is a data-producing animal. Wherever he goes he leaves certain kinds of data—court records, tax records, school records, birth and death records, and the like. This leads to the possibility of using existing institutional data as indicators of complex social trends and relationships. Durkheim's use of suicide rates to study problems of social norms and social cohesion is the classic example here.

Studies of the type with which we shall be concerned, however, have three distinctive features: they are quantitative rather than qualitative; the researcher designs and uses his own data-gathering devices rather than depending on available historical or institutional records; and they concern people's behavior and attitudes in real-life situations rather than constructed experimental situations.

* Publication No. A-276 of the Bureau of Applied Social Research, Columbia University. I am indebted to Prof. Allen Barton and Mr. Herbert Menzel for their contributions to this review.

From the point of view of this group of papers, two topics deserve special attention. How are broad conceptual ideas converted into instruments of empirical research to provide evidence on a topic of inquiry? And how can the "variables" so developed be manipulated to lead to broader generalizations? Both these problems need much additional specification. For the first we shall pay special attention to two issues: What happens if we have a choice between several instruments? And what happens if our evidence pertains to both individuals and collectives—a topic of traditional interest to the social scientist? As far as inference goes, social research has brought two topics to the foreground: how can we come near to causal relations, if we have no experiments but descriptive data only? And what hope can we place on the role of mathematics in the social sciences? The following five sections will briefly discuss these topics.

I. THE FLOW FROM CONCEPTS TO EMPIRICAL INDICES

No science deals with its objects of study in their full concreteness. It selects certain of their properties and attempts to establish relations among them. The finding of such laws is the ultimate goal of all scientific inquiries. But in the social sciences the singling out of relevant properties is in itself a major problem. No standard terminology has yet been developed for this task. The properties are sometimes called aspects or attributes, and often the term "variable" is borrowed from mathematics as the most general category. The attribution of properties is interchangeably called description, classification, or measurement.

When social scientists use the term "measurement," it is in a much broader sense than the natural scientists do. For instance, if we are able to say that one department in a company has higher morale than another, we would be very pleased with ourselves and we would say that we had performed a "measurement." We would not worry that we cannot say that it is twice as high or only 20 per cent higher. This does not mean that we make no efforts to arrive at measurements in the traditional sense, with a precise metric. Some success has been achieved, but these efforts are only beginning, and they represent merely a small part of measurement activities in the broader sense.

Keeping in mind this generalized idea of measurement, let us see how social scientists establish devices by which to characterize the objects of empirical investigations. There appears to be a typical

process which recurs regularly when we establish "variables" for measuring complex social objects. This process by which concepts are translated into empirical indices has four steps: an initial imagery of the concept, the specification of dimensions, the selection of observable indicators, and the combination of indicators into indices.

1. Imagery. The flow of thought and analysis and work which ends up with a measuring instrument usually begins with something which might be called imagery. Out of the analyst's immersion in all the detail of a theoretical problem, he creates a rather vague image or construct. The creative act may begin with the perception of many disparate phenomena as having some underlying characteristic in common. Or the investigator may have observed certain regularities and is trying to account for them. In any case, the concept, when first created, is some vaguely conceived entity that makes the observed relations meaningful.

Suppose we want to study industrial firms. We naturally want to measure the management of the firm. What do we mean by management and managers? Is every foreman a manager? Somewhere the notion of management was started, within a man's writing or a man's experience. Someone noticed that, under the same conditions, sometimes a factory is well run and sometimes it is not well run. Something was being done to make men and materials more productive. This "something" was called management, and ever since students of industrial organization have tried to make this notion more concrete and precise.

The same process happens in other fields. By now the development of intelligence tests has become a large industry. But the beginning of the idea of intelligence was that, if you look at little boys, some strike you as being alert and interesting and others as dull and uninteresting. This kind of general impression starts the wheels rolling for a measurement problem.

2. Concept specification. The next step is to take this original imagery and divide it into components. The concept is specified by an elaborate discussion of the phenomena out of which it emerged. We develop "aspects," "components," "dimensions," or similar specifications. They are sometimes derived logically from the over-all concept, or one aspect is deduced from another, or empirically observed correlations between them are reported. The concept is shown to consist of a complex combination of phenomena, rather than a simple and directly observable item.

Suppose you want to know if a production team is efficient. You have a beginning notion of efficiency. Somebody comes and says, "What do you really mean? Who are more efficient—those who work quickly and make a lot of mistakes, so that you have many rejections, or those who work slowly but make very few rejects?" You might answer, depending on the product, "Come to think of it, I really mean those who work slowly and make few mistakes." But do you want them to work so slowly that there are no rejects in ten years? That would not be good either. In the end you divide the notion of efficiency into components such as speed, good product, careful handling of the machines—and suddenly you have what measurement theory calls a set of dimensions.

The development of dimensions can go quite far. One university in California has made a study under a Navy contract of an airplane factory, aimed at determining what is really efficient management on the lowest level. The notion of efficient management was divided into nineteen components, some of which were: absence of dissensions in the group, good communication downward, not too much compulsion, consistency of command, the size of command, and so on.

This can probably be overdone. I have rarely seen a concept that needed nineteen dimensions. But as a general principle, every concept we use in the social sciences is so complex that breaking it down into dimensions is absolutely essential in order to translate it into any kind of operation or measurement.

3. *Selection of indicators.* After we have decided on these dimensions, there comes the third step: finding indicators for the dimensions. Here we run into a number of problems. First of all, how does one "think up" indicators? The problem is an old one.

William James has written in *The Meaning of Truth:*

> . . . Suppose, e.g., that we say a man is prudent. Concretely, that means that he takes out insurance, hedges in betting, looks before he leaps . . . As a constant habit in him, a permanent tone of character, it is convenient to call him prudent in abstraction from any one of his acts. . . . There are peculiarities in his psychophysical system that make him act prudently. . . .

Here James proceeds from an image to a series of indicators suggested directly by common experience. Today we would be rather more specific about the relation of these indicators to the underlying quality. We would not expect a prudent man always to hedge in betting, or to take out insurance on all possible risks; instead we would talk about the probability that he will perform any specific

act as compared with a less prudent individual. And we would know that the indicators might vary considerably, depending on the social setting of the individual. Among students in a Protestant denominational college, for instance, we might find little betting and rare occasions for taking out insurance. Still a measure of prudence could be devised which was relevant to the setting. We might use as indicators whether a student always makes a note before he lends a book, whether he never leaves his dormitory room unlocked, etc.

The fact that each indicator has not an absolute but only a probability relation to our underlying concept requires us to consider a great many possible indicators. The case of intelligence tests furnishes an example. First, intelligence is divided into dimensions of manual intelligence, verbal intelligence, and so on. But even then there is not just one indicator by which imaginativeness can be measured. We must use many indicators to get at it.

There is hardly any observation which has not at one time or another been used as an indicator of something we want to measure. We use a man's salary as one of the indicators of his ability; but we do not rely on it exclusively, or we would have to consider most businessmen more able than even top-ranking university professors. We take the number of patients a doctor has cured as another indicator of ability in that setting; but we know that a good surgeon is more likely to lose a patient than is a good dermatologist. We take the number of books in a public library as an indicator of the cultural level of the community; but we know that quality of books matters as much as quantity.

When a battery of indicators is being drawn up, one difficult problem is to decide where to stop. Which indicators are considered "part of" the concept, and which are considered independent of or external to it? If we start listing indicators of the "integration" of a community, is the crime rate a part of the conception of integration, or is it an external factor which we might try to predict from our measure of integration? Here again, as with the problem of projective indices, knowing the laws which relate indicators to one another is of great importance. Even if we exclude crime rates from our image of an "integrated" city, they might be so highly correlated, as a matter of empirical generalization, that we could use them as a measure of integration in situations where we could not get data on the indicators which we "really" want to call integration. To do this, of course, we must first have "validating studies" where we cor-

relate crime rate with the other indicators of integration and establish that it is generally closely related. We should also know whether there are other factors besides integration influencing crime rate which might confuse our measurements if we used it alone to measure integration, so that we can check on these other factors, or add enough other indicators so as to cancel out their influence.

4. *Formation of indices.* The fourth step is to put Humpty Dumpty together again. After the efficiency of a team or intelligence of a boy has been divided into six dimensions, and ten indicators have been selected for each dimension, we have to put them all together, because we cannot operate with all those dimensions and indicators separately.

For some situations we have to make one over-all index out of them. If I have six students and only one fellowship to give, then I must make an over-all rating of the six. To do this I must in some way combine all the information I have about each student into an index. At another time we may be more interested in how each of several dimensions is related to outside variables. But, even so, we must find a way of combining the indicators, since by their nature the indicators are many, and their relations to outside variables are usually both weaker and more unstable than the underlying characteristic which we would like to measure.

To put it in more formal language, each individual indicator has only a probability relation to what we really want to know. A man might maintain his basic position, but by chance shift on an individual indicator; or he might change his basic position, but by chance remain stable on a specific indicator. But if we have many such indicators in an index, it is highly unlikely that a large number of them will all change in one direction, if the man we are studying has in fact not changed his basic position.

To put the matter in another way, we need a lot of probings if we want to know what a man can really do or where he really stands. This, however, creates great difficulties in the fourth step of the measurement sequence which we described above. If we have many indicators and not all of them move in the same direction, how do we put them together in one index? Only recently have we raised the question: can you really develop a theory to put a variety of indicators together? The subject is a large one, and it is impossible to go into details here. The aim always is to study how these indicators are interrelated with each other, and to derive from these

interrelations some general mathematical ideas of what one might call the power of one indicator, as compared with another, to contribute to the specific measurement one wants to make.

In the formation of indices of broad social and psychological concepts, we typically select a relatively small number of items from a large number of possible ones suggested by the concept and its attendant imagery. It is one of the notable features of such indices that their correlation with outside variables will usually be about the same, regardless of the specific "sampling" of items which goes into them from the broader group associated with the concept. This rather startling phenomenon has been labeled "the interchangeability of indices."

II. THE INTERCHANGEABILITY OF INDICES

To present an example, we chose an index of "conservatism" used in a recent study of the response of college teachers of social sciences to the difficult years of the "McCarthy period," with its frequent attacks against colleges and professors for "leftist leanings."

One of our problems in this Teacher Apprehension study was to sort out those teachers who, because of their own convictions, could not possibly be the objects of such attacks: the men and women who hereafter, using the favorite term of their own spokesmen, will be called the conservatives.

From the beginning of our study we sought to find an acceptable way to locate this conservative group correctly. How was that to be done in a relatively short interview, in which the bulk of the questions necessarily was concerned with the nonconservatives who were the ones mainly involved in the controversies? This is a problem of classification common to all survey research. What indicators should we select?

In our study we could have submitted to our respondents certain conservative writings and asked them whether they approved of them. Or we could have selected the organizations they belonged to or the magazines they read as indicators. We preferred, as a result of much previous experience, to choose indicators more closely connected with the rest of the interview. We submitted to each respondent a series of rights and prohibitions, most of them taken from academic life, and asked whether they were for or against them. Out of this material an index of conservatism was formed. Since we

were aware that quite different material would have been equally suitable, we tested our index against a series of other possibilities.

Two questions had to do with the respondent's attitude towards student activities. "If there are students who want to join it, do you think that a Young Socialist League ought to be allowed on this campus, or not?" The attitude toward socialists seemed a good indicator because whether they should be classified with communists or not is an issue on which educated conservatives and their opponents are likely to disagree. Fourteen per cent, or 355 professors, reported they would be definitely against such a policy. Characteristically enough, the second question, also pertaining to student activities, gave almost the same number of conservative replies. We asked our respondents to suppose that they were faculty advisers to a student organization on the campus that "proposed inviting Owen Lattimore, Far Eastern expert (now under indictment in Washington) to speak at a public meeting here." Again, about 14 per cent of the sample, in this case 342 professors, put themselves on record that this "ought not to be allowed."

To both questions we get practically the same number of conservative answers: 342 and 355, respectively. One might expect that practically the same professors furnish these replies. This, however, is not completely the case. Table 1 shows how the answers are related.

TABLE 1

A Cross-Tabulation of Answers to the Two Questions on What Students Should Be Permitted To Do

	Invite Lattimore			
Form a Socialist Club	*Approved*	*Undecided*	*Disapproved*	*Total*
Approved	1686	95	124	1905
Undecided	118	27	46	191
Disapproved	152	31	172	355
Total	1956	153	342	2451

We see that the great statistical similarity of replies to each question is really the result of a considerable amount of "turnover." Of the people who approved of a Socialist Club, 124 would be against an invitation to Lattimore; conversely, 152 people who approved this invitation would not want students to form a Young Socialist League. This is neither surprising nor disturbing. Any single indicator has a

specific element and can never be taken as fully representative for the classification we are striving to achieve—here, the classification of conservative respondents. Many of the interviewees make qualitative comments on their answers, and they do it most often when they see that on a specific point their response is somewhat out of line with their whole attitude pattern. We know, therefore, fairly well what explains the position of the people in the right upper and left lower corners of Table 1. Some of the respondents who were against inviting Lattimore dislike him personally. Others feel that a legal matter is at issue—a man who is under indictment should not be permitted to talk on a college campus. Inversely, the professors who would let Lattimore talk but who are against a Young Socialist League sometimes comment that on their campus there is a general policy against political student organizations or that they feel that a socialist organization could be especially open to subversive infiltration.

Suppose we decided to use one of the items in Table 1 as a crude index to conservatism. A serious discussion could start over which of the two questionnaire items is a better "measure" of conservatism. The Lattimore question is tinged with personal idiosyncrasies and legal implications. The Socialist League item has an element of ambiguity: do those who would forbid such an organization express their own opinion or the policy of their college? Neither of the two items is a very pure "measure," and arguments could therefore continue for and against each of them. Actually, however, it would make very little difference which one is used. And this is a point which needs to be driven home. Classifications in social research are mainly used to establish relations between a number of variables. The crucial question, therefore, is whether these relations, the empirical findings we are looking for, are much affected if we interchange one reasonable index with another.

To exemplify the matter we need an "outside variable." For it we chose the answer to an item which forced the respondent to make a hypothetical choice between the rights of the individual and the claims of the institution:

> If you had to make a choice, in a case in which a member of the faculty is accused of being subversive or of engaging in un-American activities, which do you think is *more* important for the college (university) administration to protect—the reputation of the college (university) or the rights of the faculty members?

What is the relation of conservatism to the concern for individual rights? This concern is the outside variable which we want to relate to conservatism. For the latter we have two measures available. Each of them can be tabulated against the choice between the protection of individual rights or the reputation of the college. What difference is there in the choice between the two indicators, namely the Lattimore or the Socialist League item?

The essential fact is that we get practically the same result irrespective of which of the two indicators we use to separate the conservative respondents from the others (Table 2).

TABLE 2

Proportion Giving Priority to Faculty Rights Related to Two Measures of "Conservatism"

Attitude on Lattimore Speech	*% Giving Priority to Faculty Rights*	*Attitude on Socialist League*	*% Giving Priority to Faculty Rights*
Conservative	46%	Conservative	43%
Neutral	50%	Neutral	51%
Permissive	70%	Permissive	70%

Among the "conservatives" to be found in the first line of either column less than half would feel that the faculty rights are paramount. Among the "permissives" in the bottom line more than two-thirds feel this way. The whole numerical trend in the two columns is about the same irrespective of which indicator has been used for classificatory purposes.

In actual research practice, a larger number of items rather than one item alone is used for the purpose of classification. This has a variety of reasons. For example, indices based on more items permit finer distinctions, and they tend to cancel out the peculiarities of any single item. These are details which need not be elaborated here. Even if we use several items for classificatory purposes, we have always a selection out of a much larger pool of reasonable possibilities.

This, then, is the general rule based on very diversified research practice. If we are dealing with a rather broad concept like conservatism, and if we want to "translate" it into an empirical research instrument, a large number of indicators will always be eligible to

form an index for classificatory purposes. Only a relatively small number of such items is practically manageable in most field research situations. If we choose two sets of such reasonable items to form two alternative indices, the following two facts will usually be found:

a. The two indices will be related, but they will not classify all the people in a study in precisely the same way; Table 1 exemplifies this.

b. The two indices will usually lead to very similar empirical results if they are cross-tabulated against a third outside variable; Table 2 exemplifies this.

One pays a serious but unavoidable price for the practical advantages of the interchangeability of indices. We can never reach "pure" classifications. Whatever index we use, the items will have "peculiarities" which result in some cases being misclassified, and therefore the empirical relationships which we find are lower than they would be if we had more precise measures of the variables with which the study is concerned.

The tentative character of the rule should also be stressed. There are some variables which are of great and general significance, and therefore over the years ever better instruments have been developed. Intelligence tests, for example, use a very large number of carefully selected items. If we were to use two such tests to classify the same group of people, the number of contradictions would be much smaller than that found in our Table 1. If a long series of studies over many years were intended to see whether the number of conservatives in the population increases, or how conservatism is related to a great many other variables, it would be worthwhile to develop a very refined classification device. But in a study like that of Teacher Apprehension, where a large number of variables had to be introduced for the first time, the only practical course for the researcher is to use fairly simple indices, and to make sure that he does not deceive himself or his readers about the remaining uncertainties.

III. ON THE RELATION BETWEEN INDIVIDUAL AND COLLECTIVE PROPERTIES

Social scientists often make use of variables to describe not only individual persons, but also groups, communities, or other "collectives." Thus, one reads of "racially mixed census tracts," of "highly bureaucratized voluntary organizations," or of a "centrally

located rooming house district." At other times the variables, although describing individuals, are based on data about some collectives, as in a comparison of "graduates of top-ranking medical schools" with "graduates of other medical schools." I shall try to clarify some of the operations involved in the construction and use of such variables in empirical research, and to provide a nomenclature for the different ways in which information about individuals and about collectives may be interwoven in these variables.

1. *Some features of generalizing propositions.* Because the intended meaning of such variables often remains ambiguous if they are not examined in the context of certain kinds of propositions in which they are used, it is necessary at the outset to highlight certain features of these propositions:

a. They say something about a set of *elements* ("cases," "units of observation").
b. For the research purposes at hand, these elements are considered comparable. This means that the same set of *properties* is used to describe each of the elements.
c. Each element has a certain value on each property (these values may be quantitative or qualitative).
d. The *propositions* assert interrelationships between the properties.

These features are, of course, common to all empirical or hypothetical generalizations. The propositions with which we are here concerned have the additional characteristic that their elements either are collectives, or are described by reference to collectives. Typical examples of the first case are these: the lower the average income in a precinct, the higher the proportion of Democratic votes cast in a presidential election; tank platoons composed of friends perform better than those composed otherwise. The precinct or the platoon are the elements of these propositions. The properties they relate are average income and voting rate in one example, and some measure of social relation and performance ratings by observers in the other.

2. *Special meaning of "collective" and "member."* The term "collective" is used here in a specific sense which needs clarification. A collective is any element in a proposition composed of *members*, i.e., constituent parts, which are regarded as comparable.

A Boy Scout troop, for example, is a collective, and the Boy Scouts who belong to it are its members. In the same sense, a city can be

treated as a collective, with the inhabitants as members. However, the members of a collective are not necessarily individual persons. A city, for example, can be described not only as a collective with the inhabitants as members, but also as a collective with the voting precincts as members. It follows that what appears as a collective in one context (e.g., precincts), can appear as a member in another. But a city could also be introduced as a collective of buildings. In any analysis of a piece of writing in which some of the elements in a proposition are collectives, it is always necessary to specify clearly of what members the collectives are composed (for the purposes at hand).

In some studies, more than two levels appear; for example, inhabitants, precincts, and cities may all be elements of the same study. This whole matter could, therefore, be elaborated by pointing out the various relationships which can exist between inhabitants, precincts, and cities. In the next few pages we restrict ourselves to collectives which have only one kind of members; the members in most illustrations will be individual persons, but we will also present some cases in which the members themselves are larger units (e.g., "communities" considered as members of a state).

3. *Properties of collectives.* It is often useful to distinguish three types of properties which describe collectives: analytical properties based on data about each member; structural properties based on data about the relations among members; and global properties, not based on information about the properties of individual members. The following examples may clarify these distinctions:

a. Analytical properties. These are properties of collectives which are obtained by performing some mathematical operation upon some property of each single member.

The average income in a city is an example of an analytical property of a collective (city) made up of individuals. Another example of an analytical property is the proportion of the communities of a given state that have their own high school; this is a property of a collective (state) the members of which are communities.

The standard deviation of incomes in a nation appears as an analytical property in the following proposition: when incomes in a nation are more equally distributed, people will save more, because they will spend less money on display consumption which might help them be socially acceptable in the higher strata.

Correlations are sometimes used to characterize collectives, and then also constitute analytical properties. The individual correlation of age and prestige in a given community, for example, has been used as a measure of its norms regarding old age. Sometimes more indirect inferences are involved. For example, in urban areas voting is highly correlated with occupation, while this is not the case in rural districts. One may conclude from this that in rural districts there is a stronger spirit of community and cohesion.

b. Structural properties. These are properties of collectives which are obtained by performing some operation on data about the relations of each member to some or all of the others.

Assume for example, that a sociometrist has recorded the "best-liked classmate" of each student in a number of classes. He can then classify the classes according to the degree to which all choices are concentrated upon a few "stars." Or he might, alternately, classify them according to their cliquishness, the latter being defined as the number of subgroups into which a class can be divided so that no choices cut across subgroup lines. In these examples the collective is the class, and the members are the individual students; "concentration of choices" and "cliquishness" are structural properties of the classes.

For an example in which the members are larger units, consider a map of the precincts of a city, which indicates the number of Negroes residing in each. Let a "Negro enclave" be defined as a precinct in which some Negroes live, but which is completely surrounded by precincts without Negroes. The proportion of the precincts of a city which are Negro enclaves would then be a structural property of the city.

c. Global properties. Often collectives are characterized by properties which are not based on information about the properties of individual members.

Nations, for example, may be characterized by the ratio of the national budget alloted to education and to armaments. Army units may be characterized by the cleanliness of their mess equipment. American Indian tribes may be characterized by the frequency with which themes of "achievement motive" make their appearance in their folk tales.

The cultural level of a city might be measured by the presence or absence of certain "cultural" institutions (theatres, libraries, etc.)

or by the proportion of its buildings which are used for cultural purposes.

Having a city manager form of government is a global property of a city. The insistence on specified initiation rites as a prerequisite to membership is a global property of a religious cult or of a college fraternity. Accessibility from the nearest traffic artery is a global property of a village.

"Emergent," "integral," "syntalic" and other terms have been used in meanings very similar to that of our term "global." It is by no means certain which term is most useful, nor have all the logical problems been resolved. Thus, the number of members of a collective (population size, etc.) is classified here as a global property, although one might argue that it is an analytic property, obtained by the operation of counting performed upon the individual property of "existence." Even more ambiguous is the classification of rates based on the behavior of members, e.g., suicide rates.

4. *A subsidiary distinction among analytical properties of collectives.* An interesting distinction may be made among the analytical properties. The first two examples given above were the average income of a city and the proportion of the communities of a given state that have their own high school. These properties of collectives have what one might call a psychological similarity to the properties of members on which they are based. The wealth of a city seems to be the same sort of thing as the wealth of an inhabitant. The endowment of a community with a high school and the proportion of high-school-endowed communities in a state have a parallel meaning.

This is not true for the remaining examples of analytical properties given above—the standard deviation of incomes in a nation, or correlations like that between age and prestige in a given community. Correlations and standard deviations can apply only to collectives and have no parallel on the level of members. The standard deviation of incomes in a city, for example, denotes something quite different—lack of homogeneity, perhaps—from individual income, the datum from which it is computed. Some economists have surmised that the average proportion of income saved in a group is smaller the higher the dispersion of income; a lack of homogeneity supposedly makes for more consumption expenditures at the service of status striving. Average rate of saving in a group is "similar" to individual saving; but homogeneity of income has no "parallel" on the individual level,

at least at one point of time. One might speak of a "genuine" analytical variable in the latter case, although the distinction is of a somewhat intuitive nature.

Another variable of this sort is "degree of consensus." When a Democrat and a Republican are competing for the mayoralty, the degree of political consensus in a particular club might be measured by the extent of the club's deviation from a 50-50 split. In this instance the analytic property is measured by a proportion, but it is not the simple proportion of adherents of either party; clubs which are 80 per cent Democratic and those which are 20 per cent Democratic are regarded as equal in consensus.

5. *Properties of individual members of collectives.* Propositions in which the elements are individuals make up the main stock of empirical research findings: rich people vote more Republican, women read more fiction, etc. When people are considered as members of a collective, then their properties can be classified according to whether the rest of the collective enters into the characterization of its members or not. This leads to a classification which to a certain extent corresponds to the one just discussed.

a. Absolute properties. These are characteristics of members which are obtained without making any use either of information about the characteristics of the collective, or of information about the relationships of the member being described to other members. Thus in the preceding examples income and sex were absolute properties.

b. Relational properties. These properties of members are computed from information about the substantive relationships between the member described and other members.

Sociometric popularity-isolation (member of choices received) is a relational property. Many other sociometric indices fall into this category. For example, if each member of a small group has rated each other member on a 5-point scale of acceptance-rejection, each member can be characterized by the total score he received (popularity), by the total score he expressed (active sociability), by the average deviation of the scores he accorded the others (discrimination in his acceptance of other members), etc. In a study of the diffusion of the use of a new drug through a community of doctors, the physicians were classified according to whether or not they had a friend who had already used the new drug on a certain date.

c. Comparative properties. These characterize a member by a comparison between his value on some (absolute or relational) property with the distribution of this property over the entire collective of which he is a member.

Sibling order is a comparative property of individuals in the proposition, "First-born children are more often maladjusted than intermediate and last-born children." Note that each individual is characterized by comparison with the age of the other individuals in his family; in the resulting classification, many of the "last-born" will be older in years than many of the "first-born."

Another example is contained in the following proposition: "Students who had the highest I.Q. in their respective high school classes have greater difficulty in adjusting in college than students who were not quite at the top in high school, even when their actual I.Q. score is equally high." Here the comparative property (being at the top in high school or not) is established in terms of the I.Q. distribution in each student's respective high school; the proposition pertains to a set of college students which includes boys from several high schools (collectives).

d. Contextual properties. These describe a member by a property of his collective.

Consider the example: "Graduates of large law schools are more likely to earn high incomes at age 40 than graduates of small law schools." In this proposition, "being a member of a large law school" is a contextual property of individuals.

Contextual properties are also used in the following propositions: "Union members in closed shops are less militant than union members in open shops." "Residents of racially mixed districts show more racial prejudice than those of racially homogeneous districts." "Soldiers in units where many promotions have been granted are less satisfied with the promotion policy than those in units where few promotions have been granted." In these propositions, being a member of a closed shop, residing in a mixed district, or being a soldier in a unit with frequent promotions are all examples of contextual properties.*

* The contextual properties can of course be divided once more according to the distinctions made in the previous section. Observe the difference between the poor man raised in a neighborhood "where everyone else was rich" and the American lawyer who was trained in a country "where Roman law pre-

Note that a contextual property, unlike a comparative property, has the same value for all members of a given collective. It is not meaningful to speak of contextual or comparative properties when the elements under study are all members of the same collective, for instance, when only graduates of one law school are being studied. The reason is that any *contextual* property would, in that case, have the same value for all the elements; hence nothing could be said about the interrelationship of this property to any other property. Any *comparative* property would, under these circumstances, classify the elements in exactly the same way as the absolute property from which it was derived, except that the calibration may be grosser.

An interesting situation comes about as follows. We have a set of collectives, and take as elements of a proposition all their members (or a representative sample of the members of each); one property of the proposition is a contextual one. Suppose, for instance, we take all the graduates of fifty law schools in the United States as of a certain year, and see how large their income is ten years later. The resultant finding may be "the income of law school graduates is correlated with the size of the school they graduated from;" this is a proposition about students, relating their income (an absolute property) to the size of their law school (a contextual property). The same proposition could be interpreted also as one where the elements are the law schools: the average income of the students would then be an analytical property of each law school; its size would be a global property of these collectives.

At first sight one may feel that the distinctions made in this section are not much more than an exercise in elementary logic. Actually they have a bearing on major discussions which have continued for years. In the debate between Durkheim and Tarde, for instance, the former waved the flag of the "social fact" against his opponent the psychologist. The issue, properly formulated, is whether a system of propositions can be built up, the elements of which are exclusively collectives. Durkheim would not have insisted on using only global

---

vailed." It seems not worthwhile to follow these combinations into further detail. It is, however, useful to emphasize the difference between a relational characteristic of a member and a contextual characteristic based on a structural property of the collective. An example of the former would be the sociometric isolate; an example of the latter would be a man who comes from an atomized group (containing many isolates, irrespective of whether he is an isolate himself or not).

characteristics, such as tribal customs or laws; his notion of density of contact is a structural, and the suicide rate conceivably an analytical, property of a collective. The real issue is whether it is necessary to introduce propositions the elements of which are individuals in order to develop a coherent system of social theory. A radical Durkheim position denying this necessity is today rather unlikely.

Another controversy can be clarified with the help of the preceding analysis. The objection is often raised that social research is atomistic and therefore does not take into account the complexities of social reality. This is sometimes put into the form that "wholes" or "structures" cannot be described by combinations of "separate" variables. Using the word structure in this general sense we can see that its meaning is caught in the following two ways:

a. by propositions in which the elements are collectives, if the latter are characterized by structural properties in the narrow sense of the preceding pages;

b. by propositions in which the elements are individuals, if the latter are partly characterized by contextual properties.

This case is important enough to deserve a special example provided by a study of college students. Norman Miller has analyzed tests measuring attitudes toward labor. He showed that students who came from a middle class background (a contextual property) became relatively more anti-labor between Freshman and Senior year while students from labor background became progressively more pro-labor. The correlation between age and attitude is different according to the background of the students.

The main finding in this last example is due to combining more than two variables into a proposition. What can we say in general about this procedure of multi-variable analysis?

IV. THE ANALYSIS OF STATISTICAL RELATIONS

In the preceding sections we have referred to many relations between two variables. For our present purpose we can assume that they were all dichotomous attributes and that we are interested in one question only: are the two attributes related or not? For our answer we have to form a cross-tabulation of the following form:

attribute y	attribute x: present	attribute x: absent	
present	P_{xy}	$P_{\bar{x}y}$	P_y
absent	$P_{x\bar{y}}$	$P_{\bar{x}\bar{y}}$	$P_{\bar{y}}$
	P_x	$P_{\bar{x}}$	1

The symbolism is simple: $p_{\bar{x}y}$ for instance is the proportion of people who are y but not x. If x and y are unrelated then

$$\frac{P_{xy}}{P_{\bar{x}y}} = \frac{P_{x\bar{y}}}{P_{\bar{x}\bar{y}}}$$

This means that the presence of x is relatively equally frequent among those people who are and those who are not y. The preceding condition can be expressed in terms of the so-called *crossproduct*

$$[xy] = P_{xy}\,P_{\bar{x}\bar{y}} - P_{\bar{x}y}\,P_{x\bar{y}}\,.$$

If $[xy] = 0$ (vanishes) then x and y are unrelated. (Sampling problems are not relevant in the present context.)

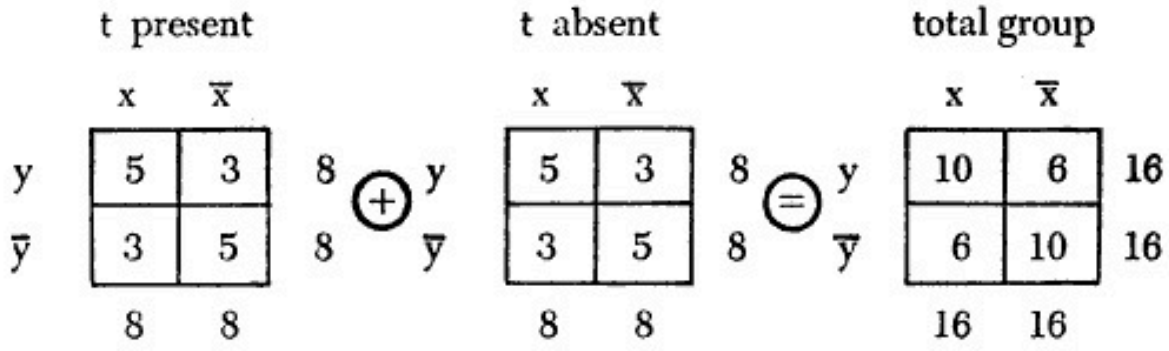
Now suppose that a third attribute t is introduced. Then we can develop the two correlative operations of *mixture* and of *elaboration*. A numerical example of mixture is given in the following scheme:

	t present: x	t present: $\bar{x}$			t absent: x	t absent: $\bar{x}$			total group: x	total group: $\bar{x}$	
y	9	3	12	⊕ y	1	3	4	= y	10	6	16
$\bar{y}$	3	1	4	$\bar{y}$	3	9	12	$\bar{y}$	6	10	16
	12	4			4	12			16	16	

In the two tables on the left the variables x and y are unrelated. But if we mix the two groups originally separated according to t we do find that people who are y are also more likely to be x than those who are not y.

Elaboration is mixture in reverse order. It consists in decomposing the group on the right into two subgroups and studying the relation of x and y separately for people who are and are not t. Elaboration

is clearly not a unique operation. We could obtain the right side, e.g., by mixing the two following subgroups:



There is one important difference between this and the preceding scheme. Both attributes x and y were formerly related to t. Neither $[xt]$ nor $[yt]$ were zero. In the new scheme we find that the proportion of x and y is the same for people who are and are not t:

$$[xt] = 0 \text{ and } [yt] = 0.$$

We shall call the relations on the left side of the schemes *partial relations* (depending on t) and symbolize them by $[xy;t]$ and $[xy;\bar{t}]$ respectively. Elaboration then consists of studying how $[xy]$ depends upon $[xy;t]$ and $[xy;\bar{t}]$ under varying conditions of $[xt]$ and $[yt]$. An algebraic development of the problems leads to the general formula.

Formula (1):

$$[xy] = \frac{[xy;t]}{P_t} + \frac{[xy;\bar{t}]}{P_{\bar{t}}} + \frac{[xt] \cdot [ty]}{P_t \cdot P_{\bar{t}}}.$$

The original relationship can be described as the sum of the two partial relationships and an additional factor, which is the product of what are called the marginal relationships between the test factor and each of the two original variables.

This elaboration leads to two major forms. Either the two partial relations may vanish; then formula (1) reduces to

Formula (2a):

$$[xy] = \frac{[xt] \cdot [ty]}{P_t \cdot P_{\bar{t}}}.$$

Or the test factor t might be unrelated to x (this means that $[xt] = 0$), and we then have

Formula (2b):

$$[xy] = \frac{[xy;t]}{P_t} + \frac{[xy;\bar{t}]}{P_{\bar{t}}},$$

a form which will turn out of interest only if one of these two partial relations is markedly stronger than the other. We shall call this the P-form (emphasis on partials), while formula (2a) will be called the M-form (emphasis on what the statistician calls "marginals").

To this formal distinction we now add a substantive one: the time order of the 3 variables. Assuming that x is prior to y, then t either can be located between x and y in time, or it can precede both. In the former case we talk of an intervening test variable, in the latter case of an antecedent one. We thus have 4 major possibilities.

	Statistical Form	
	P	M
Position of t		
Antecedent	PA	MA
Intervening	PI	MI

We are now ready to present the decisive point. It is claimed that there are essentially four operations which can be performed with two original and one test variable. It makes no difference whether this is done with actual data or whether they take the form of theoretical analyses. If a relation between two variables is analyzed in the light of a third, only these four operations or combinations thereof will occur irrespective of whether they are called interpretation, understanding, theory, or anything else.

We now turn to some concrete examples which will clarify what these four types of elaboration stand for. In cases of the type PA, we usually call the test variable t a "condition." General examples easily come to mind, although in practice they are fairly rare and are a great joy to the research man when they are found. For example, the propaganda effect of a film is greater among less-educated than among highly educated people. The depression had worse effects on authoritarian families than on other types.

Three general remarks can be made about this type of finding or reasoning: (a) It corresponds to the usual stimulus-disposition-response sequence, with x being the stimulus and the antecedent t being the disposition. (b) The whole type might best be called one of *specification*. One of the two partials will necessarily be larger than the original relationship. We specify, so to speak, the circum-

stances under which the original relationship holds true more strongly. (c) Usually we will go on from there and ask why the relationship is stronger on one side of the test dichotomy. This might then lead into one of the other types of analysis. Durkheim uses types PA in discussing why married people commit suicide less than unmarried people. He introduces as a test variable "a nervous tendency to suicide, which the family, by its influence, neutralizes or keeps from developing." This is type PA exactly. We do not take it to be a convincing explanation because the introduction of the hypothetical test variable (tendency to suicide) sounds rather tautological. We rather want to know why the family keeps this tendency from developing, which leads to type MI, as we shall see later.

The type PI is also easily exemplified. We study the relationship between job success *(y)* and whether children did or did not go to progressive schools *(x)*. We find that if the progressively educated children come into an authoritarian job situation *(t)* they do less well in their work than the others; on the other hand, if they come into a democratic atmosphere, their job success is greater.

The relation between type of education and job success is elaborated by an intervening test factor, the work atmosphere. We call such a test factor a "contingency." In many prediction studies, the predicted value depends upon subsequent circumstances which are not related to the predictor. An example is the relation between occupational status and participation in the life of a housing community. White-collar people participate more if they are dissatisfied, whereas manual workers participate more if they are satisfied with their jobs.

Type MA is used mainly when we talk of rectifying what is usually called a *spurious relationship*. It has been found that the more fire engines that come to a fire *(x)*, the larger is the damage *(y)*. Because fire engines are used to reduce damage, the relationship is startling and requires elaboration. As a test factor *(t)* the size of the fire is introduced. The partials then become zero and the original result appears as the product of two marginal relationships; the larger the fire, the more engines—and also the more damage.

When we encounter a relationship which is psychologically puzzling, we usually stop at that point; but this same mode of elaboration is also used under different psychological circumstances. More people commit suicide during the summer than during the

winter. Durkheim suggests, as a t factor for elaboration, that increased social activities are going on during the summer. Our general knowledge tells us that x (the season) is indeed related to t in this way.

Our interest immediately shifts to the *[ty]* relationship, namely: the presumed fact that greater intensity of social life leads to more suicides. Actually, of course, whether this explanation which comes from Durkheim is correct would depend upon a disappearance of the partials. Durkheim would have to show that if intensity of social life is kept constant, the season does not make any difference in suicides. Because he has no data on this point, he looks for other situations where he can presume the intensity of social life to vary. He finds that there are more suicides during the day as compared with the number during the night, which he again explains with the help of the same test factor. This leads into the whole question of probability of inference, which we do not follow up here.

We now turn to type MI, for which we shall use the term "*interpretation.*" The difference between "explanation" and "interpretation" in this context is related to the time sequence between x and t. In an interpretation, t is an intervening variable situated between x and y in the time sequence.

Examples of type MI are numerous. Living in a rural community rather than a city *(x)* is related to a lower suicide rate *(y)*. The greater intimacy of rural life *(t)* is introduced as an intervening variable. If we had a good test of cohesion, we should undoubtedly find that a community being a rural rather than an urban one *(x)* is positively correlated with its degree of cohesion *(t)*, and that greater cohesion *(t)* is correlated with lower suicide rates *(y)*. But obviously some rural communities will have less cohesion than some urban communities. If cohesion is kept constant as a statistical device, then the partial relationship between the rural-urban variable and the suicide rate would have to disappear.

It might be useful to illustrate the difference between type MA and type MI in one more example. It was found during the war that married women working in factories had a higher rate of absence from work than single women. Here are a number of possible elaborations:

a. Test factor: more responsibilities at home. This is an intervening variable. If it is introduced and the two partial relationships—

between marital status and absenteeism—disappear, we have an elaboration of type MI. We interpret the relation by showing what intervening variable connects the original two variables.

b. Test factor: physical infirmity as crudely measured by age; the *older* women are more likely to be married, and to have less physical strength, both as a result of their age. Age is an antecedent variable. If it turns out, when age is kept constant, that the relation between marital status and absenteeism disappears, we would have an explanation, and probably call it a spurious effect: type MA.

The latter case suggests again an important point. After having explained the original relationship, our attention might shift to *[ty]*, the fact that older people show a higher absentee rate. This, in turn, might lead to new elaborations: is it really the case that older women have less physical resistance, be they married or single? Or, is it that older women have been born in a time when work is not as yet important for women and, therefore, they have a lower work morale. In other words, after one elaboration is completed, we will, as good scientists, immediately turn to a new one; but the basic analytical processes will always be the same.

One final point can be cleared up, at least to a certain degree, by this analysis. We can suggest a clear-cut definition of the *causal* relation between two attributes. If we have a relationship between x and y, and if for any *antecedent* test factor the partial relationships between x and y do not disappear, then the original relationship should be called a causal one. It makes no difference here whether the necessary operations are actually carried through or made plausible by general reasoning. In a controlled experiment we have two matched groups: the experimental exposure corresponds to the variable x, and the observed effect to y. Matching makes sure that for any antecedent t we shall have $[xt] = 0$. If then $[xy] \neq 0$ we can in the light of the preceding analysis always be sure that there is a causal relation between exposure x and effect y.

This has special bearing on the following kind of discussion. It is found that in densely populated areas the crime rate is higher than in sparsely populated areas. Some authors state that this could not be considered a true causal relationship, but such a remark is often used in two very different ways. Assume an intervening variable, for instance, the increased irritation which is the result of crowded

conditions. Such an interpretation does not detract from the causal character of the original relationship. On the other hand, the argument might go this way: crowded areas have cheaper rents and, therefore, attract poorer, partly demoralized people. Here the character of the inhabitants is antecedent to the characteristics of the area. In this case the original relationship is indeed explained as a spurious one and should not be called causal.

We can, finally, link the present discussion with some observations made in the previous section. Explanation consists of the formal aspect of elaboration and some substantive ordering of variables. We have here focused on ordering by time sequence. But not all variables can be ordered this way. We saw that we can distinguish orders of complexity, e.g., variables characterizing persons, collectives, and sets of collectives. Other ordering principles could be introduced, e.g., degree of generality, exemplified by the instance of a specific opinion, a broader attitude, and a basic value system. What is needed is to combine the formalism of elaboration with a classification of variables according to different ordering principles. This would cover a great part of what needs to be known about the logic of explanation and inference in contemporary survey analysis.

V. MATHEMATICAL MODELS—SOME USES AND LIMITATIONS

So far, we have discussed how observations are organized into concepts and indices, and how relationships among several variables are analyzed. There is, however, a small group of social scientists which tries to go even further and builds mathematical models of human behavior.

It is useful to divide the area of models into what we may call 1) static or measurement models, and 2) dynamic or behavioral models. Anyone who develops a theory of scaling works in a static area. That is to say, he relates manifest data to underlying utility or probability. One who deals, let us say, with learning theory, or with the kind of work which Rashevsky does, where the element of time enters, works in the second field. There is, of course, some overlapping. For example, the problem of introducing scales and measurements in dynamic models is very important. One of the main reasons for the criticisms of Rashevsky is that he uses dynamic models that introduce characteristics of people without concerning himself about the measurement of the characteristics. But it is worth while

to remember that most of our normal work conveniently divides into those models which are used to develop variables, such as attitudes, utility, or scales, and those models which are used to analyze dynamic behavior, such as learning or changes in social systems.

Most of these remarks will deal with the second type, the dynamic models, because of their relevance to problems which interest the majority of non-mathematical sociologists. They are not interested in the formation of units or scales; they take that for granted. They call in "George" to do it, and they believe in his technical ability. But after George has developed a "scale," what they want to know is how it can be used for the analysis of social processes.

There are two major problems in this area on which we need greater clarification:

1) What are the scientific tasks which mathematical models can perform in the broad area of social science?

2) How is the choice of the specific area in which we develop models made?

As to the uses of models, there is no doubt that one important function is to help predict behavior. But there is another function which might be called the *linguistic* function of models in the social sciences. There may be a great value in developing models on their own merits, irrespective of whether we have any data or of whether they can at this moment predict anything at all.

The crux is, of course, the meaning of the term "data." Experiments might be the only source of data for the physicist, and they might be useful for the social scientist, but they are for us certainly not the main material we are working with today. We have data which have been accumulating for about 2,500 years. Plato developed very interesting propositions on human behavior in society, and if we now developed models which somehow organized or clarified what Plato had to say, we should be doing an important job.

Some interesting illustrations of the linguistic function of mathematical models could be drawn from efforts to "translate" humanistic writings. However, here we shall draw our examples from recent empirical studies, and show that even in this area there is an important task of clarification which does not involve prediction.

The linguistic function of mathematical models can be divided into three parts: 1) the organizing function, 2) the analytical function, and 3) the mediating function.

As an example of the *organizing function*, consider the many studies which are now available on how people vote and why they vote. One finding which has been well corroborated in different studies is that Catholics for many decades in this country have been voting much more frequently Democratic than have Protestants. A number of recent studies refined this result. We made studies during election campaigns, interviewing people every month, asking them how they intended to vote, and classifying voters as Democratic Catholics, Republican Catholics, etc. In August, most Catholics said that they intended to vote Democratic, and most Protestants said that they planned to vote Republican. In October, when the same people were asked the same questions, there were some changes. The Catholic Democrats and the Republican Protestants kept their conviction, but the other two groups showed changes. Republican Catholics started to say, "I changed my mind; I think I had better vote Democratic." And many of the Protestants who had wanted to vote Democratic began to change their minds and veer to the Republicans. This is true not only for Catholics and Protestants, but for urban and rural groups as well; urban dwellers are more likely to be Democrats. This leads to a generalization on a somewhat higher level. People who want to vote in agreement with their own group are much more likely to be stable in their intention than the people who intend to vote against the prevailing trend of their group. As time goes on, the people who first want to vote against the prevailing trend in their group slowly "return to the fold."

Offhand it would seem easy to understand this by referring to the influence of the environment. And indeed we have information that people with similar social characteristics are more likely to meet one another, and that such contacts affect the vote intention of the participants. But why does the population not split into groups which are homogeneous within and completely different from each other? Obviously there are countertrends, and one of them is the random fluctuation of people who have little interest in politics. This uninterested group is larger between elections than during campaigns and guarantees a certain amount of reshuffling. At the same time the uninterested people introduce another complication. They are less likely to talk politics, but are more easily influenced when they do.

One could go on enumerating the many well-established but disconnected propositions relating attitude, social characteristics, con-

tact frequencies, etc. But no one knows what the minimum of variables and relations are from which the total set could be derived. Here is where the organizing function of mathematics enters. In the small-group field, H. Simon has shown how much clarity can thus be gained. In the area of voting, inventories of propositions have been made, but the final formalization is still lacking.

The *analytical function* of models is exemplified by studies showing that groups appear to solve problems better than individuals. In the election example we have an overabundance of data, and it is a function of the mathematical models to organize them. In the problem-solving case we have a deficiency of data, and the mathematical model points out where we should look further. The notion of group problem-solving is vague. Do we mean a kind of division of labor within the group? Or do we mean that in a group there is a greater probability that one outstanding person will appear and find the answer for all the rest? Empirical studies are available showing the probability of both individual and group solutions. It is then possible to set up two models: one assuming division of labor, and one assuming solution by a top individual in the group. Derivations from the two models can be tested against the data. The model helps us to infer what is going on within the group while it works, although we know only the final outcome. This analytical use of mathematical models often leads to ideas of what new empirical work is needed.

The *mediating function* of model construction can be illustrated as follows. Economists have asserted for eighty years that one cannot measure utilities. Social psychologists have, for the last thirty years, measured attitudes without worrying whether this was a legitimate operation. Perhaps economists did not measure utilities because they spent so much time on the problem of whether utilities could be measured. Only by putting the utility and the attitude problems into a formal language, and thereby showing that they are really the same problem, could there be a joining of effort between the people who have conceptual problems and the people who measured without realizing that the conceptual problem existed. We begin to have now a way of formulating the problem correctly; economists believe it is susceptible to measurement; and social psychologists realize that measurement is not as easy as they thought before.

The mediating function of the formalization appears in a variety

of other areas. For instance, almost any textbook in sociology has a final chapter on social change or the lack of it. But the chapters on social change make sense only if one has read a reasonable book on business cycle analysis or economic growth. Sociologists have discussed relations between the individual and society for a hundred and fifty years; physicists have considered the relations between thermodynamics and atomic theory; "aggregation" is a word familiar to economists. When problems are put into formal mathematical language, similarities become obvious—a very important cross-disciplinary mediating function. Only after problems have been formalized is it really possible to work on cross-disciplinary approaches and to make mutual contributions from one discipline to another.

Thus the linguistic function of a mathematical model helps to organize an abundance of material, helps to pin down deficiencies of data, and helps to mediate between procedures that are formally alike but terminologically different—a function of mathematical models which has been often underrated.

Now to the second major point. Are we judicious enough in the choice of the subject matter to which we apply our mathematical efforts? Those who contrive experiments to test mathematical models of human behavior tend to introduce gambling situations. The decision problems social scientists deal with are different: Why do people commit crimes? Why do they buy cars? Why did they vote for Eisenhower? Social scientists in general study decision problems. But when it comes to model construction, mathematicians are interested mainly in how people bet. Perhaps we are again facing a danger which we faced about forty years ago, at the time of the early Watsonian behaviorists. When the idea of learning experiments emerged, there was a general conviction that only in very primitive situations could one really experiment. The main learning effort was concentrated on rats. But social psychologists are more interested in people, and today in America we are no longer convinced that the study of animal psychology can solve the major problems of human psychology.

As mathematical models become important there is a real danger that we repeat the cycle, that we start again at the lowest rung of the ladder. One cannot study how rats bet, but one can study how human beings bet. The relative ease of these experiments seems to me to tempt mathematical model builders to repeat the error of the

behavioristic extremists. There are, of course, those who feel that Watson was right. They should at least agree that the concentration of model-building on betting is an analogous situation. The rat maze and the betting experiment are characterized by the same tendency to seek the simplest configuration.

Connected with the choice of area is another choice, that of experiment *versus* observation. I have not been convinced by accounts of the history of psychology that the strong emphasis on experimentation is justified. One can experiment with how people bet. But if one wants to know how people vote, and how people buy, one cannot really experiment. One has to make systematic analytical observations. In the voting studies we have made, we took a sample of people in a small town in New York, and interviewed them every month. We asked them for whom they wanted to vote, and then observed how they changed from month to month. In another study, we kept potential car buyers under observation for nine months, and tried to study the decision process by re-interviewing the same people in various stages of their approach to a final purchase. Somewhat more attention should be devoted to observational data taken over a time interval. The models to be looked for have a similarity to the business cycle analysis of economists. The whole new mathematics of economic growth and population growth are relevant here. In any problem which involves time-series analysis, in any complex problem that involves a comparison of a series of variables over a number of time periods, the relationship of the model to the data is somewhat different from that in the experimental situation.

* * *

THE TOPICS here chosen for review may seem arbitrary, but they have one feature in common: they are entirely characteristic of empirical social science research today. There is a danger that methodological discussions become very general and miss the relevant detail by trying to be all-embracing. In a Colloquium, the theme of which is the unity of the sciences, it is salutary to reverse the old French proverb: the more it is the same, the more one should stress the differences. It is the new step in the over-arching continuity which leads to progress.

BIBLIOGRAPHICAL NOTE

The general theory of index-formation is exemplified in Section I of the *Language of Social Research* edited by P. F. Lazarsfeld and

M. Rosenberg (Free Press, 1955). There one will also find additional material on the interchangeability of indices. An inventory of group properties as used in current sociological literature is provided by R. K. Merton on pages 311-325 of his *Social Theory and Social Structure* (second edition, Free Press, 1957); a good discussion of the general principles involved was given by P. M. Blau in the *American Journal of Sociology* (Volume 63, 1957, p. 58 ff.). Many concrete examples of the treatment of multivariant statistical relations will be found in Chapter VII of H. H. Hyman's *Survey, Design and Analysis* (Free Press, 1955). The organizing use of mathematics is shown at its best in Section 2, H. A. Simon, *Models of Man* (Wiley, 1957). The literature on group problem-solving models is reviewed by I. D. Lorge and H. Solomon in *Psychometrical* (Volume 20, 1955, p. 139 ff.).